

Children's Acquisition of Assamese Split Ergativity

Pori Saikia

A thesis submitted for the degree of
Doctor of Philosophy

Department of Language and Linguistics
University of Essex

Date of Resubmission for Examination: November 2022

Abstract

The aim of this thesis is to provide a comprehensive study of the Assamese case alignment system and its acquisition by L1 child speakers. It comprises of three interrelated papers that collectively make the following contributions:

Paper 1 addresses the ongoing debate between *NOMINATIVE* and *ERGATIVE* alignments in the existing Assamese literature. Its objective is to establish a definitive account of the Assamese case alignment system, thereby laying the foundation for subsequent research on the acquisition of the split ergative system by children.

Paper 2 examines the acquisition of ergative splits by forty 2 to 6 year old children, based on two datasets obtained from an experiment and a semi-structured language game. The results demonstrate that children begin to use *ERGATIVE* morphology in adult-like structures as early as 2;6 years old, and their proficiency in this aspect increases with age. Additionally, our data indicates that children tend to grasp the Differential Object Marking (DOM) split in Assamese with relative ease, typically achieving competence by the age of 3.

Paper 3 further explores the case alignment in Assamese, examining it from both synchronic and diachronic perspectives. It endeavors to trace the historical development of the *ERGATIVE* marker, shedding light on its origins and evolution. Additionally, this paper offers a unique comparison of adult and child language data, revealing an ongoing shift in the language's alignment system.

As a whole, this thesis contributes to our deeper understanding of the Assamese split ergative system, unravels the intricate procedure through which children effortlessly acquire this complex system at a young age, and sheds light on the historical evolution of the ergative case marker within the language while focusing on an ongoing change in the split ergative alignment.

Dedication

This thesis is dedicated to the loving memories of my dear friends, J. R. Philemon Chiru and Bashabi Barua, without whose love, guidance, and motivation, I would not be here today. Though they are no longer with us, their impact on my life is eternal.

Declaration

Chapter 4 of this thesis, titled “Assamese Case Alignment Shifts in Progress” was authored in collaboration with Dr. Maris Camilleri and was originally published in the proceedings of the LFG 2019 Conference (Saikia & Camilleri, 2019). Subsequent to its publication, I have undertaken further revisions and expanded its scope by incorporating an additional study to provide deeper insights into the adult usage of the Assamese split ergative system. Additionally, I have reanalysed the data elicited from the child participants in order to incorporate valuable feedback and suggestions provided by my examiners.

Acknowledgements

I would like to take this opportunity to acknowledge my heartfelt gratitude to my current supervisor Dr Laurel Lawyer and my former supervisors Prof. Monika Schmid, Prof. Louisa Sadler, Dr Sonja Eisenbeiss, and Prof. Ayesha Kidwai for their unwavering guidance throughout this journey. They have not only provided me with invaluable inputs on my work, but have also supported me in overcoming personal challenges I faced during my PhD journey. I will particularly be forever grateful to Prof. Monika Schmid, who in spite of relocating to a different university, went above and beyond her responsibilities as my (former) supervisor to make my PhD experience more manageable. I must acknowledge that Prof. Monika Schmid's guidance and encouragement were instrumental in ensuring that I stayed on the program and successfully completed my PhD.

I would like to thank my examiners Prof. Miriam Butt and Dr Claire Delle Luche for their invaluable inputs that has greatly improved this thesis. They pointed out numerous flaws in the methodology, statistical analysis, and overall writing, and I will forever appreciate their constructive and compassionate approach. This thesis remains a work in progress, and I will continue to incorporate their suggestions to enhance its quality.

I am grateful to the Doctoral Training Centre - Economic and Social Research Council (ESRC) for funding this research. Also, I express my profound gratitude to British Federation of Women Graduates for offering me an Emergency Fund which helped me to continue my research without any financial concerns in the middle of a global pandemic.

My heartfelt appreciation goes to my PhD Supervisory Board Members, with special mention to my former chair, Dr Claire Delle Luche, and my advisor, Dr Doug Arnold, for their invaluable feedback and continued support. I'm also grateful to Dr Doug Arnold for always being just a text message away to help me with all my LaTeX issues. The department administrators, particularly Carrie and Carolina were utmost helpful throughout my doctoral program. I am also profoundly thankful to Felicity Kingsgate, the Postgraduate Research Education Officer at the University of Essex, for her ongoing

support during my entire Ph.D. journey.

I extend my heartfelt gratitude to my professors Mr. Rana Kr. Changmai, Ms. Mayuri Sharma Baruah, Mr. Rajeev Mohan, Dr Mriganka Choudhury, Dr Chaitanya Borah, Dr Nita Paul, and Prof. Gautam K. Borah for serving as role models and a constant source of inspiration in my life during my college and university days in Assam. Your dedication and commitment to education have had a profound impact on my personal and academic growth, and I am truly thankful for the knowledge and values you have instilled in me.

I would also like to express my gratitude to Dennis Somadula for creating the pictures for my experiment, to Rocktim Gohain and Poli Saikia for their assistance during the fieldwork, and to my participants and informants for generously sharing their valuable time. I extend my sincere thanks to Nabajyoti Borah and Nabajyoti Lakhkar for proofreading my Assamese documents and providing expert advice on the Assamese language.

I have to acknowledge the efforts my friend Sabina Yasmin Rahman went through in order to get me the initial financial assistance I needed to get here. She not only trusted me with her own credit card so that I could immediately have access to funds to apply for my student visa, she also appealed on Twitter so that I could raise the initial funding I needed. I will forever be grateful to Annapurna Verma and Pranveer Singh Rathore who responded to Sabina's tweet and came forward to help me despite not knowing me personally. I am deeply grateful to the members of Xoixik Xikhya Xongram Xomiti of Baruahola Village, particularly Ananda Chetia, as well as Anup Saikia, Ajay Senapati, and the Professors of Tinsukia College for providing financial assistance when I needed it the most. All of these people deserve a special mention for playing a pivotal role in making it possible for me to come to the UK.

I can never express in words what my former flatmate and dear friend, Maris Camilleri, has done for me since 2017. She has not only been my mentor throughout my Ph.D. but has also been there for me during some of the darkest days of my life. From co-authoring Chapter 4 of this thesis and offering her valuable input in the others, to forcing me to get out of bed during my depressive phase and spending entire nights at the library alongside me so that I have company, she went above and beyond her duties as a friend to help me complete this thesis.

I would like to take this opportunity to most gratefully acknowledge the love and support of Angkur Jyoti Dipanka Shaikeea, Anuraag Bordoloi, Nandana Saikia, Nancy Phukan, Pompei Baidya, Prasujya Gogoi, Rick Das, Sarvenaz Moradi, Sayan Banerjee, Siffi Singh, Shrikanth Mohan, and Zain Odho for being there for me whenever I needed someone to listen to me in a non-judgmental manner. You all have been a constant rock in my life and have helped me grow as a person throughout the years.

I wish to express my sincere gratitude for the unwavering love and support bestowed upon me by Ioannis Korkos, Lucas Shuttleworth, Julia Heimann, Khimpi Dutta, Swapnali Baruah, Dushyanthi Gooneratne, Rrituraj Sharma, Pinaki Sharma, Ishaan, and Pokhi. I may never be able to fully repay your kindness and generosity. You all were kind enough to invite me into your homes, where you have treated me with the love and care of a family, and have provided me with a sense of belonging and a home that I've always yearned for in this faraway country.

I would also like to acknowledge the efforts that Hasitha Wijeyeratne has put in during the last few months leading up to my submission. From buying me my favourite chicken mayo to accompanying me to the garden center whenever I found myself in a less-than-ideal mood (which happened more often than I'd care to admit), your support has been immeasurable. Moreover, I acknowledge that my four-legged baby, Buzz, played an indispensable role, and without him in my life, this work wouldn't have been possible.

Last but not least, I shall forever remain indebted to my parents, Nipon Saikia and Jyotsna Saikia, my sister, Poli Saikia, my grandfather, Dinesh Gohain, and my aunts, Monju Gogoi Saikia and Sukhama Saikia, for all that they have done for me.

Date: 14th November 2022

Pori Saikia

List of Abbreviations

ERG	Ergative
ACC	Accusative
NOM	Nominative
S _a	Unergative
S _o	Unaccusative
DAT	Dative
GEN	Genitive
LOC	Locative
SUB	Subject
OBJ	Object
IOBJ	Indirect Object
INST	Instrumental
PRES	Present
PST	Past
FUT	Future
PROG	Progressive
PERF	Perfect
SG	Singular
PL	Plural
M	Masculine
F	Feminine
CLF	Classifier
1	1 st PERSON
2	2 nd PERSON
3	3 rd PERSON
EM	Emphasis
HON	Honorific
COMPL	Completive
A	Agentive

List of Tables

2.1	Case-Marking on Assamese SUBJ Pronouns	24
2.2	Case-Marking Distribution on Assamese (non-pronominal) NPs	27
3.1	Intransitivity Based Split in Assamese	43
3.2	PERSON and NUMBER Based Split in Assamese	44
3.3	Summary of Participants	50
3.4	Stimuli Design for Contrastive Case Elicitation Task	52
3.5	Non-randomised Item List of The Contrastive Case Elicitation Task	53
3.6	Utterance Types Produced in the Contrastive Case Elicitation Task	54
3.7	Age-wise Distribution of Overt SUBJ Realisation in Child-data .	55
3.8	Age-wise Distribution of Overt OBJ Realisation in Child-data .	56
3.9	Age-wise Distribution of Overt IOBJ Realisation in Child-data .	57
3.10	Summary of the Full Utterances of the Child Participants	58
3.11	Summary of the Full Utterances of the Adult Participants . . .	58
3.12	Summary of ERG and NOM Omissions in Obligatory Contexts .	58
3.13	Observed and Expected Counts of Case Elicitation Task Data .	60
3.14	Stimuli Design for Final Destination Board Game	65
3.15	Summary of the Results from The Final Destination Game . . .	66
3.16	Age-wise distribution of the Results from The Final Destination Game	68
3.17	Age-wise distribution of the Production Errors in The Final Destination Game	68
3.18	Age-wise distribution of the Production Errors in The Final Destination Game	69
3.19	Age-wise distribution of the Results for OBJs from The Final Destination Game	71
3.20	Age-wise distribution of the Results for IOBJs from The Final Destination Game	71
4.1	Distribution of case-marking on SUBJ GFs	81
4.2	Distribution of case-marking on non-SUBJ GFs	90
4.3	Summary of Participants	91

4.4	The emergent ANIMACY-based split	96
4.5	Summary of Grammaticality Judgement Test of Unergative Verbs	96
4.6	Summary of Grammaticality Judgement Test of Unaccusative Verbs	97

List of Figures

2.1	Case Alignment	8
2.2	Tripartite Alignment	14
2.3	Four Way Distinction of Case Marking (Devi, 1986, p. 185)	19
2.4	Nominative-Accusative Case Alignment	21
2.5	Split Ergative Case Model for Assamese	22
2.6	Animacy Hierarchy Scale (Silverstein, 1976, p. 176)	23
3.1	Split Ergative Case Model for Assamese	46
3.2	Contrastive pictures for transitive ‘push’ and unaccusative ‘burn’	54
3.3	Case Elicitation Task: Error Types by Age Group	60
3.4	The handmade felt castle for the Final Destination Board Game	64
3.5	Summary of the Error Types by Age Group	69
4.1	Contrastive Elicitation Task for Testing Case Marking (based on Ruigendijk (2015))	92
4.2	Case Elicitation Task: Error Types by Age Group	94
4.3	GLM Output for Adult Data	94
4.4	Significance of Age in Child Data	95
4.5	Summary of Grammaticality Judgement Test for Intransitive Verbs	97

Contents

Abstract	iii
Dedication	v
Declaration	vii
Acknowledgements	ix
List of Abbreviations	xiii
1 Introduction	1
1.1 Motivation	2
1.2 Research Questions	2
1.3 Research Methods	2
1.4 Significance of this Study	3
1.5 The Structure of this Thesis	4
2 Descriptive Analysis of Split Ergativity in Assamese	7
2.1 Introduction	7
2.2 Background	10
2.2.1 Assamese in the CASE literature	12
2.3 Split Ergativity in Assamese	20
2.3.1 Person and Number Based Split	22
2.3.2 Intransitivity Based Split	26
2.4 Differential Object Marking	28
2.5 Non-canonical Subjects	32
2.6 Summary and Conclusion	34
3 Acquisition of Assamese Split Ergativity	41
3.1 Introduction	41
3.2 Split Ergativity in Assamese	42
3.2.1 Intransitivity Based Split	43
3.2.2 Person and Number Based Split	44
3.3 Earlier Studies on Case Acquisition by Children	46
3.4 Methods	48

3.4.1	Participants	49
3.4.2	Data Collection	50
3.4.3	Equipment Used	51
3.4.4	Procedure of Data Transcription and Analysis	51
3.5	Study 1: Contrastive Case Elicitation Task	52
3.5.1	Materials	52
3.5.2	Procedure	53
3.5.3	Data Analysis	54
3.5.4	Overt SUBJ NP Realisation in Child-data	55
3.5.5	Overt OBJ NP Realisation in Child-data	56
3.5.6	Overt IOBJ NP Realisation in Child-data	57
3.5.7	Results and Discussion	57
3.5.8	Results for Subjects	57
3.5.9	Results for Direct Objects	59
3.5.10	Results for Indirect Objects	62
	3.5.10.1 Development of Split Ergative Case Marking Among Two-year Olds	62
3.6	Study 2: Final Destination Board Game	63
3.6.1	Procedure	64
3.6.2	Data Analysis	66
3.6.3	Results and Discussion	66
3.6.4	Results for Objects	69
3.7	Summary and Conclusion	72
4	Assamese Case Alignment Shifts in Progress	75
4.1	Introduction	75
4.2	Previous Studies in Children-Led language Change	76
4.3	Distribution of Case in Assamese	78
4.3.1	Assamese as a Split Ergative Language	78
4.3.2	The Historical Development of the Ergative in Assamese	82
4.3.3	Non-SUBJ Case Marking in Assamese	86
4.3.4	Current Predictions of Change in Progress	90
4.4	The study	91
4.4.1	Methodology	91
4.4.2	Results	93
	4.4.2.1 Results of Contrastive Case Task	93
	4.4.2.2 Results of Grammaticality Judgement Test	96
4.5	Conclusion	97
5	Conclusion	103
5.1	Summary of the Main Findings	103
5.2	Limitations and scope of future research	105
A	Ethical Approval Application	107
B	Consent Form for Child Participants	109

<i>CONTENTS</i>	xxi
C Consent Form for Adult Participants	111
D Language Background Information: Child Participants	113
E Language Background Information: Adult Participants	117
F SUBJ Realisations for Child Participants	121

Chapter 1

Introduction

Child language acquisition is a fascinating and complex field of study that has captivated researchers for decades. It is well-established that language acquisition begins before birth, with research, such as DeCasper and Spence (1986) and Fifer and Moon (1994), examining how infants in the womb respond to the rhythmic patterns and sounds of their native language. Studies have also shown that shortly after birth newborns exhibit the ability to differentiate among a broad spectrum of phonetic distinctions ((Bertoncini et al., 1987; Mehler et al., 1988; Fifer and Moon, 1994)).

The process of language acquisition starts in the womb and continues throughout a child's early years, with a remarkable surge in linguistic development occurring by the age of 4 (Bates et al., 1992). During this intensive period, children reach their pinnacle of understanding grammatical and phonological rules, expand their vocabulary exponentially, and become increasingly adept at expressing their thoughts and feelings through language.

Building on the foundations of ongoing research in various aspects of child language acquisition, this thesis takes a closer look at the acquisition of the Assamese split ergative system, with a primary focus on a contemporary language change driven by children that seems to be reshaping the existing case marking system in Assamese. The current research centers around three distinct yet interrelated objectives: Firstly, to provide an in-depth descriptive analysis of the split ergative system in Assamese, including its historical development and a comprehensive account of how split ergativity functions in Assamese based on current literature and adult data. This descriptive account of Assamese split-ergativity forms the basis for addressing the research questions of this investigation. Secondly, the Assamese language is known for its complex split ergative system, which has not been extensively studied in the context of child language acquisition. This research seeks to bridge this gap by shedding light on how Assamese-speaking children acquire and internalize this intricate case marking system. By examining this aspect, we aim to contribute to our understanding of children's acquisition of ERGATIVE case markers. Thirdly, this research seeks to document and analyze the use of Assamese case markers by

participants across three different participant groups. Moreover, by comparing their linguistic production to the traditionally accepted norms of Assamese grammar, we aim to gain insights on an ongoing linguistic shift with respect to its case marking system.

1.1 Motivation

The motivation behind this research is multifold. Assamese, in spite of having a rich linguistic history still remains relatively underrepresented in the field of linguistics, particularly in the context of child language acquisition. Investigating the acquisition of a complex split ergative system in Assamese provides an opportunity to explore linguistic diversity while contributing to language acquisition research.

Furthermore, the documentation of the Assamese case alignment in its current state, and the investigation into the undergoing linguistic shift, holds significance not only from an academic perspective but also as a vital resource for future research endeavors.

1.2 Research Questions

This thesis will systematically address the following research questions in subsequent chapters, thereby contributing to our understanding of Assamese case alignment system and its acquisition:

- 1) Does Assamese conform to the NOMINATIVE-ACCUSATIVE language pattern, as commonly believed among researchers?
- 2) To what extent do children between the ages of 2 and 6 demonstrate adult-like case marking skills?
- 3) Are there any observable changes in the current case marking system?

1.3 Research Methods

In order to investigate the nature of the Assamese case alignment system and its acquisition, a dataset was created from both child and adult speakers using three primary methods: an experiment, a semi-structured language game, and a grammaticality judgement test.

In order to look at the use of the ergative split in the 3rd PERSON NPs by children and young adults a picture set was created, drawing inspiration

from Contrastive Elicitation Task (Ruigendijk, 2015). The resultant custom-designed picture set comprising of 11 minimally contrastive pairs were culturally and linguistically tailored to the Indian context. During the experiment, participants viewed these picture pairs in a predetermined sequence and were encouraged to provide descriptions before moving on to the next pair.

On the other hand, a semi-structured language game based on Final Destination (Al-Houti, 2013) was developed to study the use of ERGATIVE-ACCUSATIVE case markers by children in the 1st PERSON (SG/PL) and 2nd PERSON (SG/PL) scenarios for both subjects and objects. To play this game, we captured images of child participants, their caregivers, and the researcher engaged in four different actions. Participants later looked at these pictures while playing the game, wherein they had to describe a set of pictures to progress through each hurdle in order to rescue a princess trapped in a castle.

Additionally, a grammaticality judgement questionnaire was designed for adults aged 15 to 60 to gain deeper insights into the intransitivity-based split in Assamese. To maintain the integrity of the responses, we employed a 4-point Likert scale, thereby preventing participants from selecting a mid-scale option. In total, the test comprised 48 questions, with 11 designated test items and 37 fillers. To eliminate potential order effects, we randomized the presentation of these questions using a Latin Square design.

All the data collected was transcribed in ELAN and were subsequently analyzed using SPSS and R. This analytical approach facilitated the investigation into the case marking patterns, comparisons between child and adult speech, and an assessment of the observed shift in the case alignment system in the language.

1.4 Significance of this Study

As part of this research, I have accumulated a corpus of both child and young adult speakers using a production experiment and a semi-structured language game. Through the analysis and evaluation of this rich dataset, this thesis attempts to contribute to the theoretical perspectives in the fields of language change and first language acquisition, as well as descriptive linguistics of contemporary Assamese. I hope that the dataset of a relatively lesser known language will aid researchers in the future. Furthermore, I created a picture set for the Contrastive Case Elicitation Task that is based on Ruigendijk (2015), but is culturally and linguistically adapted to the Indian context and can be used to study the acquisition of case systems for a broad range of Indian languages.

This study also attempts to highlight the ongoing debate regarding the case alignment system in Assamese, and provide a definitive account of the align-

ment system by building on relevant earlier research. It also tries to capture the current state of a language that seems to be undergoing change. This contributes to the broader field of linguistics by enhancing our understanding of split ergativity in lesser-studied languages, especially one that is already a change in progress.

1.5 The Structure of this Thesis

This thesis is a compilation of three distinct but interrelated papers, complemented by an overall introduction and a conclusion. It is worth noting that the three papers, while self-contained, share a common focus on the grammatical phenomenon of split ergativity in Assamese. Consequently, some degree of repetition and overlap is present in terms of literature review, methodology, and data analysis across the individual papers.

The thesis is structured as follows:

Paper 1/Chapter 2 offers a comprehensive examination of the existing literature on Assamese case alignment and attempts to settle the ongoing NOMINATIVE-ERGATIVE debate by presenting a definitive account of the case alignment system based on the analysis of the existing literature.

Paper 2/Chapter 3 delves into the first language (L1) acquisition of the complex intransitivity and PERSON and NUMBER-based ergative split in Assamese. This investigation is substantiated by two sets of data obtained through a production experiment and a semi-structured language game.

Paper 3/Chapter 4, a version of which was co-authored with Dr. Maris Camilleri, focuses on the evolving change in the case alignment system in Assamese. This change is observed as a shift from an intransitive-based split to an animacy-based split. The chapter supports this shift through a comparative analysis of data from both children and adults.

Finally, the conclusion provides a cohesive overview of the findings from the three main chapters and highlights the potential avenues for future research while acknowledging the limitations of the current study.

Bibliography

Al-Houti, S. K. (2013). *The acquisitions of finiteness, nominal and verbal agreement by Kuwaiti children acquiring Arabic*. PhD thesis, University of Essex.

- Bates, E., Thal, D., and Janowsky, J. S. (1992). Early language development and its neural correlates. *Handbook of Neuropsychology*, 8:109–176.
- Bertoncini, J., Bijeljac-Babic, R., Blumstein, S. E., and Mehler, J. (1987). Discrimination in neonates of very short cvs. *The Journal of the Acoustical Society of America*, 82(1):31–37.
- DeCasper, A. J. and Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development*, 9(2):133–150.
- Fifer, W. P. and Moon, C. M. (1994). The role of mother's voice in the organization of brain function in the newborn. *Acta Paediatrica*, 83:86–93.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., and Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29(2):143–178.
- Ruigendijk, E. (2015). Contrastive elicitation task for testing case marking. In Armon-Lotem, S., de Jong, J., and Meir, N., editors, *Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment*, pages 38–54. Multilingual Matters, Bristol.

Chapter 2

Descriptive Analysis of Split Ergativity in Assamese

Abstract

The case alignment system in Assamese has been the subject of a longstanding debate. The alignment system has been primarily described as a NOMINATIVE-ACCUSATIVE (Kakati, 1941; Goswami and Tamuli, 2003; DeLancey, 1981; Nath, 2003) or a (fully) ERGATIVE system (Butt and Deo, 2001; Devi, 1986; Zakharyin, 2015). The only two papers that categorise Assamese as a SPLIT ERGATIVE system have either focused on the split being conditioned by intransitivity (Amritavalli and Sarma, 2002) or by PERSON and NUMBER (Saha and Patgiri, 2013). Such literature, however, has provided a classification informed by only a partial view of the data. Consequently, the primary goal of this paper is to address the conflicting views and thereby present a definitive account of the case alignment system in Assamese. By analysing the case alignment in different clause types, I propose that Assamese is a morphologically split ergative language. Further, I discuss the animacy hierarchy in Assamese, and how it gives rise to a Differential Object Marking.

Keywords: Indo-Aryan Languages, Split Ergative, Differential Object Marking

2.1 Introduction

Natural languages employ different mechanisms to indicate grammatical relations within a clause, including constituent order, case-marking and verbal agreement or cross-referencing. Languages with a fixed word order convey distinct pragmatic or semantic interpretations through the position of its constituents, while the languages with a comparatively free word order use case marking to indicate grammatical relations. Some languages also indicate the relationship between arguments and their predicate through verbal agreement, where the head verb reflects grammatical properties of its NP arguments. If

we consider a language with a fixed word order like English, interchanging the order of the NPs would convey an opposite meaning, as in *The hunter killed the tiger* vs. *The tiger killed the hunter*. However, if we look at the Assamese examples in (1), we see that changing the order of the constituents *sikari* ‘hunter’ and *bag^h* ‘tiger’, does not result in any change in meaning.

- (1) a. sikari-zon=**e** bag^h-tu=(k) mar-il-e
 hunter-CLF=ERG tiger-CLF=(ACC) kill-PST-3
 ‘The hunter killed the tiger.’
- b. bag^h-tu=(k) sikari-zon=**e** mar-il-e
 tiger-CLF=(ACC) hunter-CLF=ERG kill-PST-3
 ‘The hunter killed the tiger.’

In instances such as (1), we also see that the subject of the transitive clause (A) is overtly marked with the ERGATIVE case marker *-e*¹, whereas the object of the transitive clause (O) is optionally marked with the ACCUSATIVE case marker *-(o)k*. More generally, case alignment in languages depends on the markings realised on the subject of an intransitive clause (S). Therefore, if a language groups its A and S together, and differentiates its O through distinct ways, with the most common being via a morphological case inflection, we end up with a nominative-accusative alignment. However, if a language distinguishes its A from both S and O, the result is an ergative-absolutive case alignment, as seen in Figure 2.1:

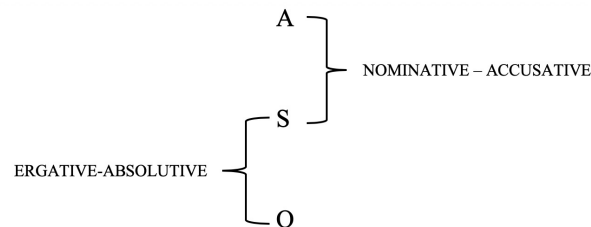


Figure 2.1: Case Alignment

Comrie (2013) includes four other alignments in his main alignment types: 1. The neutral case marking system where S, A, and O are coded in a similar way as in Mandarin. 2. A tripartite case alignment system where S, A and O are marked differently, as in Nez Perce and Semelai. 3. A NOM-marking with overtly marked S and O and unmarked accusative as in Harar Oromo. 4.

¹*-e* has been described as an ergative post-position (Zakharyin, 2015; Kakati (1941) and a suffix (Goswami & Kakati, 2003). However, I consider it as an enclitic based on Zwicky and Pullum’s (1983) test that clitics can attach to hosts already consisting of a clitic as observed in (1), where both the ergative *-e* and accusative *-(o)k* case enclitic gets attached to a host that already has the classifiers *-zon* and *-tu*. Since, it is beyond the scope of this paper, I will not discuss this in further details.

Active-inactive languages where the agent-like S arguments are marked, while patient-like arguments are unmarked as in Georgian. Zwart and Lindenbergh (2015), on the other hand propose a new alignment typology with eighteen alignment patterns to account for all those languages that do not fit into the traditional alignment typology. However, for the purpose of this paper, I will focus only on the ERGATIVE-ABSOLUTIVE system and its splits, and show how Assamese corresponds to this alignment pattern. We have already mentioned that ergativity, in its broader sense, refers to a phenomenon where the S and O are distinguished from A. This notion was introduced by Dirr as early as 1912 (Dixon, 1994, p. 3), and only gained popularity in the 1970s when Dixon (1972, 1979), Anderson (1976), Silverstein (1976), Comrie (1978), and Plank (1979) made extensive theoretical contributions on this grammatical pattern.

Languages can either display syntactic or morphological ergativity, i.e. whether all is simply a matter of a distinct morphological form, or whether a given language’s syntax considers absolutive-marked arguments as most prominent, affecting agreement behaviours, and the like. Syntactic or inter-clausal ergativity exists in languages such as the Western Nilotic language Pãri, where S and O are treated similarly through constituent ordering, coordination and relativization (Andersen, 1988; Dixon, 1994; Dryer, 1997). Although, a language with morphological ergativity usually does not show any syntactic ergativity, a language with syntactic ergativity uses morphological means to code ergativity, as seen in the following examples from Pãri. Intransitive clauses in Pãri have an SV order, as seen in (2a), while independent transitive clauses have an OVA order, as illustrated in (2b). ERGATIVE-ABSOLUTIVE alignment is further marked with $-i \sim \varepsilon$ on A, and \emptyset in S and O. When A is topicalised, as in (2c), it loses the ergative inflection, but shows agreement on the verb.

- (2) a. ùbúr á-túuk
 Ubur COMPL-play
 ‘Ubur played.’
- b. jòobì á-kèel ùburr-ì
 buffalo COMPL-shoot Ubur-ERG
 ‘Ubur shot the buffalo.’
- c. ùbúr jòobì á-kèel-é
 Ubur buffalo COMPL-shoot-3SG.A
 ‘Ubur shot the buffalo.’ (Pãri: Dixon, 1994, p. 51)

Relatively few languages display syntactic ergativity, and when they do, they additionally exhibit morphological realisation and display constituent order distinctions, as just exemplified by Pãri above. Morphological distinctions that express an ERGATIVE behaviour usually demonstrate a semantic basis behind the assignment of case to the functions. For example, the function of the constituent associated with A must “initiate or control an activity”, and can be an agent, initiator, donor, speaker, and so on. The O, on the other

hand, plays the role that is affected by the actions initiated or controlled by A. The S, however, could either take the roles associated with A, and align with a nominative-accusative system, or can take the role of O, triggering an ERGATIVE-ABSOLUTIVE alignment. Alternatively, it can associate with both roles, and form a mixed alignment system. This third type of alignment pattern is often referred to as split ergative, split intransitive, split S, active-inactive, agentive-patientive and active-stative (Dixon, 1979, 1994; Bickel, 1995; Comrie, 2013; Mithun, 1991; Handschuh, 2008).

These languages make a clear distinction between their S arguments on the basis of their semantic roles. For example, the NP referent (S_a) of unergative verbs like *jump*, *dance* and *swim* controls an activity. The NP referent (S_o) of unaccusative verbs like *fall*, *sink* and *burn* have no control over the activity, and similar to referents of an O function, they can be affected by the event. It is usually rare to find a fully ergative language, given that most languages have some sort of split, either in their syntax, or the morphology that gives rise to a split ergative alignment. In such a system, a construction can either use ergative syntax and morphology, or show another pattern, usually nominative-accusative, based on certain conditions, such as intransitivity, varied tense/aspect/mood values, the semantic nature of the NP, and other differences attributed to differences between main vs. subordinate clause. It is well attested that languages exhibit more than one alignment across their grammatical domains (Comrie, 2013; Handschuh, 2008; Harris, 1982, 2008; Aissen, 2003; Dixon, 1979, 1994; Silverstein, 1976). This is evidenced in Assamese, Hindi, Basque, Warlpiri, Dyirbal and many more split ergative languages. In this chapter, I will discuss the splits that condition ergativity in Assamese in Section 2.3. However, before doing that, I will give a broad overview of the prominent literature on Assamese case alignment in Section 2.2, and discuss animacy hierarchy and differential object marking (DOM) in Section 2.4. Further, I will discuss the non-canonically marked GEN and DAT subjects in section 2.5. Section 2.6, then summarises the analysis.

2.2 Background

Assamese is an Indo-Aryan language spoken by 14 million native speakers² in the North-East Indian state of Assam. It additionally serves as a lingua-franca among different speech communities in the state and in the neighbouring states of Nagaland and Arunachal Pradesh (Goswami & Tamuli, 2003, p. 394). Assamese has three varieties: Eastern (spoken in Tinsukia, Dibrugarh, Sivasagar, Jorhat and Lakhimpur district), Central (Morigaon district) and Western (Guwahati). In this paper, only the eastern variety, which is considered to be the standard variety and used as the official language of the state, will be taken into account.

²As per The Census of India (2011), there are 1,53,11,351 Assamese speakers in India out of which 1,48,16,414 identify themselves as L1 speakers

Assamese takes both head marking and dependant marking, and is an SOV language that is syntactically NOMINATIVE-ACCUSATIVE, for agreement purposes. In other words, the finite verbs in Assamese always cross-references information about PERSON exclusively for S and A, as illustrated in (3a) and (3b), respectively. However, non-canonical subjects in Assamese, expressed via non-ERG/NOM morphology, represented via the GEN-marked 1st PERSON subject *mur* in (3c), trigger a default 3P agreement. PERSON agreement on unaccusative verbs is optional in the past tense for 3rd PERSON subjects as in (3d). Finite verbs in Assamese also inflect for tense and aspect, while, (pro)nominal stems inflect for (numeral) classifiers, number and case. Case markers always follow the (numeral) classifiers or plural affixes (if any). However, when there is a numeral or quantifier modifying the head noun, the classifier attaches to it, while the case enclitic appears with the head as illustrated through (3e)-(3g). It is important to note that case enclitics appear only on head nouns, and lack any form of concordance with their modifiers.

- (3) a. tai xu-l-e
 3.F.SG.NOM sleep-PST.3
 ‘She slept.’
- b. kukur-tu-e muk kamur-il-e
 dog-CLF=ERG 1.ACC bite-PST-3
 ‘The dog bit me.’
- c. mur b^huk lag-is-e
 1.GEN hunger get-PERF-3
 ‘I am hungry.’
- d. rugi-zon mor-il
 patient.CLF.NOM die-PST
 ‘The patient died.’
- e. gari e-k^kɔn=e amak agbheti dhor-is-e
 car one-CLF=ERG 1.PL.ACC block catch-PERF-3
 ‘A car has blocked us.’
- f. e-k^kɔn gari=e amak agbheti dhor-is-e
 one-CLF car=ERG 1.PL.ACC block catch-PERF-3
 ‘A car has blocked us.’
- g. manuh-bur=e amak agusi dhor-is-e
 personPL=ERG 1.PL.ACC block catch-PERF-3
 ‘The people have blocked us.’

2.2.1 Assamese in the CASE literature

There are only a handful of scholarly works that specifically focus on case marking in Assamese, and within this limited literature, no clear consensus emerges regarding whether Assamese exhibits a NOMINATIVE-ACCUSATIVE, ERGATIVE-ABSOLUTIVE, or a split alignment. While a majority of the previous studies have emphasised that Assamese is a NOMINATIVE-ACCUSATIVE language (Morey, 2013; Bhattacharjya, nd; Chowdhary, 2014; Kakati, 1941; Sarma, 2015) yet, a closer inspection of the available literature suggests that this is mostly a mischaracterization of the terminology, rather than a misalignment.

Brown (1848)³ mentions the use of *-e* as an emphatic marker and states that there are four degrees of emphasis in the language that are represented through \emptyset , *-e*, *-i*, and *-he* (p. 4). He points out that the NOM-marking in Assamese is identical to its vocative, and cites the noun *manuh* ‘person’ in its citation form as an example of such. He further says that *-e* is used as a suffix for nouns marked as NOM to show a slight degree of emphasis or control. In fact, case alternations from NOMINATIVE to ERGATIVE driven by semantic factors such as volitionality or control is a common phenomenon in many other ERGATIVE languages, as highlighted through the following Hindi examples from Butt (2006).

- (4) a. ram k^hās-a
 Ram.M.SG.NOM cough-PFV.M.SG
 ‘Ram coughed.’
- b. ram=ne k^hās-a
 Ram.M.SG=ERG cough-PFV.M.SG
 ‘Ram coughed (purposefully).’ (Hindi/Urdu: Butt, 2006, p.71)

Kakati (1941), in his seminal work uses the ‘NOM’ terminology to describe the subjects of transitive verbs that ‘must always’ take the *-e* marker. Despite acknowledging the emphatic role of the *-e* enclitic, he contends that the presence of the *-e* marker on the subjects of transitive verbs (A) does not constitute an instance of emphasis. Kakati (1941) maintains that there is no distinct ‘agent case’ in Assamese, even though he asserts that the *-e* marker is influenced by the instrumental case marker *-(er)e* which is mandatory in passive constructions of transitive verbs, as exemplified in: *hat-e buwa kapur* ‘cloth woven by hand’ (p. 286). He further posits that it is this mandatory usage that leads to the habitual use of *-e* on the NPs expressing the meaning of A. This argument is in agreement with one of the popular views that ergativity in Indo-Aryan languages developed because the passivisation of transitive verbs attracted the instrumental case-marker in Classical Sanskrit (Sen, 1973; Masica, 1982); also see contra. (Butt and Deo, 2001; Beames, 1872; Kellogg,

³The first grammar of Assamese was written in 1839 by William Robinson, however, he considered Assamese a part of the Bengali dialect instead of giving it an independent language status as done by Brown later (Bhattacharyya (2011)).

1972).

Kakati explains that the case enclitic on A is necessary to distinguish it from the object NPs that do not take the ‘dative-accusative postposition’ *-ok* as shown in (5b). As Assamese is a pro-drop language with a flexible constituent order, constructions like (5) can lead to ambiguity if one of the NPs is not marked for case. Kakati argues that the *-e* marking on the noun phrase *manuh* ‘person’ in (5a) signifies its role as the subject of the sentence⁴, while the same noun phrase in (5b) serves as the object. Ambiguity would have arisen since the ACC marker on animate common nouns is optional in Assamese, as illustrated through (5b) and (5c), while inanimate nouns are always unmarked. When an animate object NP is marked, it accounts for the semantic feature of specificity⁵

- (5) a. *manuh=e mar-e*
 person=ERG beat-3
 ‘A man beats.’ (Kakati, 1941, p. 286)
- b. *manuh mar-e*
 person beat-3
 ‘(He/she/it) beats a man.’ (Lit: ‘(He/she/it) beats a person.’) (Kakati, 1941, p. 286)
- c. *manuh=ok mar-e*
 person=ACC beat-3
 ‘Beats the man.’ (Lit: ‘(He/she/it) beats the person.’)
- d. *zud^hɔ=t manuh=e manuh=ok mar-e*
 war=LOC person=ERG person=ACC beat-3
 ‘People kill people in war.’

Contrary to Kakati’s claim that the subjects of intransitive verbs are unmarked and appear in their ‘bare stem’, except when *-e* is used to put special emphasis, some S NPs can take the same case marking as A, such as the *-e* on *kesuwa* ‘baby’ as seen in (6b), even when there is no notion of emphasis. Across (6), we observe that the agent of the transitive verb *pulis* ‘police’ and the subject of the intransitive verb *kesuwa* ‘baby’ bears a similar marker, that is, *-e*, whereas the object of the transitive verb *sur* ‘thief’ is coded with the ACC-marker *-(o)k*.⁶

- (6) a. *pulis=e sur-tu=k d^hor-il-e*
 police=ERG thief-CLF=ACC hold-PST-3
 ‘Police caught the thief.’ (Lit. ‘Police held the thief.’)

⁴This is an intransitive sentence, hence *manuh* ‘person’ is technically the subject.

⁵We will look into differential object marking in Assamese later in §4

⁶The ACC marker here can be optional, given the animacy hierarchy in Assamese. More to follow in §4.

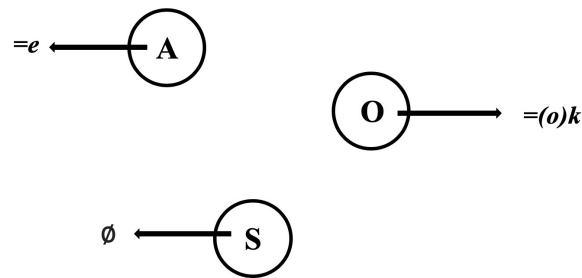


Figure 2.2: Tripartite Alignment

- b. *kesuwa-tu=e kand-il-e*
 baby-CLF=ERG cry-PST-3
 ‘The baby cried.’

Furthermore, if we consider Kakati’s (1941) analysis that the subjects of transitive verbs (A) are obligatorily marked with *-e*, and the objects of transitive verbs (O) are marked with *-(o)k*, while the subjects of intransitive verbs (S) remain unmarked, except for instances where *-e* is added to put special emphasis on it, then we would have a tripartite alignment in Assamese⁷, as illustrated in Figure 2.2.

However, adopting Kakati’s perspective and description of case in Assamese raises two important issues. First, as demonstrated in (7a), it becomes evident that only animate objects of transitive verbs receive the accusative marker *-(o)k*, while inanimate objects do not require marking. Furthermore, even among animate objects, there exists a hierarchy whereby the use of the accusative marker is optional, particularly with common nouns, as illustrated by the noun phrase *kesuwa* ‘baby’ in (7b).

- (7) a. *ma=e gari-k^hɔn d^hu-l-e*
 mother=ERG car-CLF wash-PST-3
 ‘Mother washed the car.’
- b. *mak=e kesuwa-tu(=k) ga d^hu-wa-l-e*
 mother=ERG baby-CLF(ACC) body wash-CAUS-PST-3
 ‘The mother washed the baby.’

The second issue involves the fact that the *-e* in overtly marked intransitive subjects is not solely intended for emphasis. This is evident from the fact that it appears as an obligatory marker with the subjects of unergative verbs in the

⁷Devi’s (1986, p. 196) also indicates that the split in nouns trigger a tripartite alignment in Assamese

language (refer to (6b and 9). Nevertheless, it holds true that the agentive *-e* in Assamese serves a dual purpose and is sometimes used to convey emphasis. This aspect has also been recognized by Chowdhary (2014) and Goswami and Tamuli (2003). The use of *-e* in this context is highlighted through the contrast on the SUBJ *lora* ‘boy’ in (8). In both examples, the unaccusative verb *por* ‘fall’ is utilized. However, in (8a), it is accompanied by a *-e*-marked subject, while in (8b), the subject is not marked with *-e*. The difference lies in the deliberate nature of the action of falling in (8a). On the other hand, omitting the *-e* marker on ‘lora’ in *lora* ‘boy’ in (8b) implies that the subject fell down unintentionally or accidentally:⁸

- (8) a. *lora-tu=e por-i di-l-e*
 boy-CLF=ERG fall-NF give-PST-3
 ‘The boy (deliberately/purposefully) fell down.’ (Choudhary, 2014, p. 111)
- b. *lora-tu.∅ por-il(-e)*
 boy-CLF.NOM fall-PST(-3)
 ‘The boy fell down.’

Chowdhary (2014) and Goswami and Tamuli (2003) also observe that the intransitive subjects of ‘conjunct’ or ‘potential conjunct verbs’ can prompt the use of an ‘overt nominative case’, as illustrated through the verb *xãtur* ‘swim’ in the following example:

- (9) *ram=e xãtur-il-e*
 Ram=ERG swim-PST-3
 ‘Ram swam.’ (Goswami & Tamuli, 2003, p. 432)

Chowdhary (2014) goes on to argue that intransitive verbs such as *dance*, *cry*, *run*, *fight*, *laugh*, and *cough* exhibit some ‘property of transitivity’, since the subject of such predicates is in full control over the situation. Hence, the presence of these actual or underlying arguments consequently leads to the appearance of the agentive case marker, originally intended for the agent A, on S, as we see in the following examples:

- (10) *suwali-zoni=e b^hal nas-e*
 girl-CLF=ERG good dance-3
 ‘The girl dances well.’

Chowdhary (2014) highlights two others instances where the subjects do not trigger an ERGATIVE case marker. First, in the context of complex predicates with a ‘patientive light verb’, such as *d^hor-a por-il* ‘get caught’ in 11, where the subject receives a patient or theme interpretation and consequently remains unmarked for case. Second, in the case of a subject of a copular predicate that lacks an object, as demonstrated by the unmarked 3rd PERSON pronoun *tek^het* ‘he’ in 12:

⁸The 3rd PERSON marker in (8b) is optional on certain unaccusative verbs.

- (11) sur-tu.∅ d^hor-a por-il(-e)
 thief-CLF.NOM catch-NOML fall-PST-3
 ‘The thief got caught.’ (Choudhary, 2014, p. 113)
- (12) tek^het.∅ hikk^hk as-il
 3SG.HON.NOM teacher be-PST-3
 ‘He was a teacher.’ (Choudhary, 2014, p. 113)

Chowdhary (2014) reasons that the subject of intransitive verbs (S) in Assamese aligns with the subjects of transitive verbs (A), and not with the objects of transitive verbs (O), and since the verbs agree with S and A, hence Assamese is a NOMINATIVE-ACCUSATIVE language. She further argues that Assamese exhibits a mixed surface ergativity, given how the unmarked (pro)nominals in both subject as well object positions are in absolutive case. The existence of the ‘absolute’ case for the unmarked subjects of intransitive verbs, as well as the unmarked objects of transitive verbs was proposed for Assamese by Goswami (1982). In fact, it is this postulation of ‘absolute’ case that gave rise to the proposal of ERGATIVE case in Assamese (Nath, 2003; Bhattacharja, nd). Apart from Chowdhary (2014), Haddad (2007) and Bhattacharjya (nd) both follow Goswami’s (1982) terminology and identify the unmarked subjects of intransitive verbs as instances of an as ‘absolute’ case.

In contrast, Nath (2003) uses the ‘nominative’ terminology to refer to the case marking on both A and S (S_a, S_o). He contradicts Kakati’s (1941) claim and suggests that, on occasions, the subjects of intransitive verbs can trigger an *-e* marking. However, he disputes Amritavalli and Sarma’s (2002) hypothesis that Assamese features an intransitivity-based split, based on his explanation in (13), where the external arguments/unergatives in the language do not consistently bear the *-e* case ending. Yet, he however fails to realise that the PP *bozar-oloi* ‘to the market’ functions as an oblique, and the verb *gol* ‘go’ is hence classified as an unaccusative verb, and therefore, the lack of *-e* marking, i.e. a zero marked (∅) subject is expected. Thus, in contrast to Nath’s assertion, Amritavalli and Sarma’s (2002) hypothesis that unaccusative verbs in Assamese are unmarked, stands correct.

- (13) ram/*ram=e bozar-oloi go-l
 Ram/*Ram=ERG market-ALL go-PST
 ‘Ram went to the market.’ (Nath, 2003, p. 14)

Nath (2003) further builds on Kulkarni’s (1988) examples of ERGATIVE alignment in Marathi, and states that a similar phenomenon of unmarked S (14a) and marked A (14b) is also observed in Assamese. Although he correctly identifies the lack of any ‘agentive NP’ in (14a), he considers the sole argument *pani* ‘water’ to be an O, instead of an S. If we look at his examples we see that in (14b) we have a valence increasing construction for the intransitive verb *utol* ‘boil’. The new causative construction is realised by adding the causer *Ram* as the ERGATIVE subject (A), while the original subject *pani* ‘water’ gets demoted to an object position:

- (14) a. *pani.∅ utol-il*
 water-NOM boil-PST
 ‘Water boiled.’
- b. *ram=e pani utol-a-l-e*
 Ram=ERG water boil-CAUS-PST-3
 ‘Ram boiled the water.’ (Nath, 2003, p. 15)

In comparing the data in (15) with (14), he argues that since the \emptyset marked NP argument *kobita* ‘poem’ in (15a) gives an ungrammatical reading, unlike what *pani* ‘water’ does in (14a), the hypothesis that Assamese is an ERGATIVE-ABSOLUTE language is false.

Examples (15a) and (14a) might seem as if they are parallel constructions, however, if we analyse it further, it is clear that *kobita* ‘poem’ in (15a) is the object of the verb, unlike *pani* ‘water’ in (14a). As Assamese is a pro-drop language, the structural NP in subject position can be unexpressed. Nevertheless this dropped subject still remains the implied subject of the construction with the agreement on the verb also demonstrating reference to an implied/understood, even if non-overt subject, which is the case in (15a). Moreover, based on the semantics of the verb *parh* ‘read’ and *utol* ‘boil’, the agent cannot distance itself from the process in (15a) as a poem cannot be read on its own. In contrast, water will continue to boil on its own, even if the agent leaves after putting the kettle on.

- (15) a. *kobita porh-il*
 poem read-PST
 ‘Read (the) poem.’ (Intended: ‘Poem was read.’)
- b. *ram=e kabitaa porh-il-e*
 ram=ERG poem read-PST-3
 ‘Ram read the poem.’ (Nath, 2003, p. 16)

Nath (2003) later compares some Assamese and Hindi examples in the perfect aspect⁹ but ends up deciding that since Assamese retains the same case ending *-e* in both perfect and simple past clauses, Assamese cannot be considered a split ergative language. Nath, however, fails to realise that ergativity in Assamese is not conditioned by aspect (refer to example (29)) as will be shown in §3, in contrast to the prevalence of such a split in most Indo-Aryan languages. Rather, the split in Assamese is based on an entirely different system, as will be discussed in details in §3.

Two of the most recent studies on Assamese maintain that Assamese has a NOMINATIVE-ACCUSATIVE alignment but do not make any substantial contributions to the ongoing case alignment debate on the language. Deb (2012)

⁹Although his observation that Assamese does not show a split on the basis of aspect is correct, his examples are in the simple past, rather than in the perfect.

states at the beginning of his paper that his study reaffirms the status of Assamese as a NOMINATIVE ACCUSATIVE language, in contrast to the claims made by Amritavalli and Sarma (2002). However, throughout his paper, he primarily present some examples illustrating case marking differences in Kamrupi Assamese, Sylheti Bengali, and Baleswari Oriya, without essentially making any contributions to his claim about NOMINATIVE-ACCUSATIVE case alignment in Assamese. Similarly, Sarma (2015) states that the nominative case in Assamese has two forms: *-e* and \emptyset , but does not expand further on this phenomenon.

Bhattacharjya (n.d.), on the other hand, does identify the split in S, and states that only a “small group” of intransitive verbs trigger an overt *-e* marking on their subjects. Following Goswami’s (1982) terminologies, she refers to the unmarked subjects of intransitive verbs as well as the unmarked objects of transitive verbs as ‘absolute’. However, she continues to use the nominative term for A and adds that if *-e* in Assamese is indeed an ERGATIVE marker, the language would then lack a nominative case marker. However, this would only be true if Assamese is a (fully) ERGATIVE language, as claimed by Zakharyin (2015), Butt and Deo (2005) and Devi (1986). In a (fully) ERGATIVE alignment, overtly marked A is distinguished from zero marked S and O. The following examples by Zakharyin (2015), do in principle show such a pattern, where both S and O are marked with \emptyset , whereas A is marked with *-e*.

- (16) a. ram=**e** kam- \emptyset kor-ib-o
 Ram=ERG work-ABS do-fut-3
 ‘Ram will do the work.’
- b. ram- \emptyset ah-ib-o
 ram.NOM come-FUT-3
 ‘Ram will come.’ (Zakharyin, 2015)

While Zakharyin’s (2015) provides evidence to argue that Assamese is a fully ERGATIVE language, his characterisation of the alignment system is flawed as he fails to provide any examples involving unergative verbs. Even Butt and Deo’s (2005) assessment, which largely relies on the analysis of Devi (1986), tends to overlook unergative verbs in Assamese that require their S_a to be overtly marked with *-e*.

Devi (1986), on the other hand, not only suggests the possibility of Assamese being a split ergative language, but also mentions about an underlying ‘tripartite subsystem’ (p. 196). According to her analysis, Assamese is predominantly a neutral and ERGATIVE language. However, the split in pronouns results in the emergence of an accusative sub-system, while, the split in nouns triggers a tripartite subsystem. This analysis leads her to propose a four-fold case alignment system for Assamese: ergative, accusative, neutral, and tripartite, as demonstrated in figure 2.3:

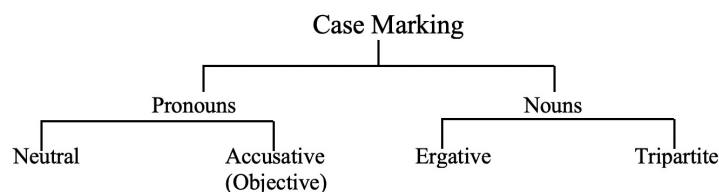


Figure 2.3: Four Way Distinction of Case Marking (Devi, 1986, p. 185)

Devi (1986) argues that the \emptyset -marked ‘free form pronouns’¹⁰ in either A or O function, and the \emptyset marked inanimate NPs in O reflect a neutral case marking system. Whereas, in the ERGATIVE system, the bound pronouns behave like other nominals and are overtly marked with *-e*. In the tripartite alignment, which she considers as a sub-system of the ERGATIVE system, the S remains unmarked, the A takes the *-e*, and the animate O takes *-(o)k*. She further argues that, in an accusative alignment, a sub-type of the neutral alignment, the pronouns remain unmarked, while the animate O takes the *-(o)k*. She adds that although there is no difference in the DOs of the tripartite and the accusative system, the former is a sub-set of the ERGATIVE alignment that takes only nominals in account, whereas, the latter is a sub-set of the neutral alignment that only encompasses pronouns.

Although Devi (1986, p.176) mentions the split within intransitive verbs in Assamese, she argues that it is a case of some sort of grammaticalization, i.e. an attribute that results out of a grammaticalisation process. She does not expand further on the process of grammaticalisation, but argues that as a consequence of it, the subjects of a cognate objects end up displaying the *-e* marker even when it is not necessary. She justifies this observation by claiming that although these constructions do not have an overt object as in any prototypical cognate verb constructions, they have an implied object. For example, under her analysis, (17a) implies the existence of the object *mat* ‘speech/words’ which would thereby mean something in the lines of *mora mat mat matise* ‘The peacock crowed a crow.’ Similarly, (17b) actually implies *rame hāhi eta hāhise* ‘Ram laughed a laugh.’ While this explanation seems convincing, one cannot determine whether this hypothesis is able to account for all the marked S NPs in Assamese, unless all intransitive verbs were to be tested for this possible behaviour.

- (17) *mora=e mat-is-e*
 peacock=ERG speak-PERF-3
 ‘The peacock has crowed’ (Devi, 1986, p. 131)

- (18) *ram=e hāh-is-e*
 ram=ERG laugh-PERF-3

¹⁰The single morpheme pronouns such as 1SG *moi* ‘I’ and 1PL *ami* ‘we’, 2.SG *tumi* ‘you’ and 3.SG *xi/tai* ‘he/she’.

‘Ram has laughed.’

(Devi, 1986, p. 171)

At present, the only hypothesis that seems viable and can account for all the splits observed in the language is a split ergativity conditioned by intransitivity (Amritavalli & Sarma, 2002), and PERSON and NUMBER (Saha & Patgiri, 2013) as these are two of the most common splits in any split ergative languages. Amritavalli and Sarma (2002) builds on the description of Kakati (1941) and Goswami (1982) that the instrumental case in Assamese is syncretic with the case ending on A, and argues that it is linked to agentivity. They, however, argue that Assamese does not have any nominative case, and the unmarked subjects of unaccusative verbs are in absolutive case (Goswami, 1982; Haddad, 2015; Bhattacharjya, n.d.). They also do not touch upon the pronoun split in Assamese, except when stating that this is observed on 2nd PERSON and 3rd PERSON pronouns, while 1st PERSON pronouns exhibit no morphological differentiation with respect to the way they are marked. This observation is erroneous, as the singular 2nd PERSON and 3rd PERSON pronouns in Assamese get zero coding. This split is also highlighted in Kakati (1941) and Goswami (1982) and discussed in detail by Devi (1981) and Saha and Patgiri (2013).

Saha and Patgiri (2013) further offer a phonological explanation for the \emptyset marked 1st PERSON, 2nd PERSON and 3rd PERSON singular pronouns and claim that since these pronouns end in a high front vowel /i/, they cannot take the -e marking except when required for emphasis. If this argument is true, then we need to understand why the same /i/ vowel accepts the -e enclitic for emphasis, and why this rule does not apply to nominals, such as proper names ending in the same high vowel as is the name *Meghali* in (19), which obligatorily takes the ERGATIVE case ending¹¹:

- (19) meghali=e bhat randh-i as-e
 meghali=ERG rice cook-PROG be.-PRES-3
 ‘Meghali is cooking rice.’

All in all, as noted earlier, the literature that claim Assamese as a NOMINATIVE-ACCUSATIVE language accurately describes all the intricacies of the case system, but uses the wrong terminology. On the other hand, Amritavalli and Sarma (2002) and Saha and Patgiri (2013) are the only available literatures that use the accurate terminologies and firmly claim the existence of a split ergative case alignment in Assamese. Therefore, building on the split ergativity accounts highlighted in these two papers, I will attempt to present a definitive account of the case alignment system in Assamese in the next sections.

2.3 Split Ergativity in Assamese

Assamese is not a NOMINATIVE-ACCUSATIVE language, in spite of this being the most popular view among linguists. As demonstrated in the previous

¹¹see footnote 12 on p.25

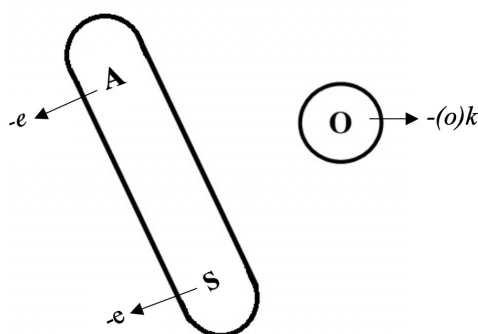


Figure 2.4: Nominative-Accusative Case Alignment

section, this alignment would have only accounted for: A/S vs O observations across NPs, and would not have taken into considerations observations about pronouns, or the split that occurs in S.

Although the examples in (6), which are repeated here in (20) for convenience, show that the A in Assamese aligns with S, and have a different marker for O, as demonstrated by figure (2.4).

- (20) a. pulis=**e** sur-tu=**k** d^hor-il-e
 police=ERG thief-CLF=ERG hold-PST-3
 ‘Police caught the thief.’ (Lit. ‘Police held the thief.’)
- b. kesuwa-tu=**e** kand-il-e
 baby-CLF=ERG cry-PST-3
 ‘The baby cried.’

However, this violates one of the main rules of nominative case marking, that is: nominative case is morphologically unmarked¹² (Blake, 1994; Dixon, 1979; Silverstein, 1976). Moreover, the case alignment in Assamese cannot be characterised as a (fully) ERGATIVE system, given the person and number-based splits for subjects, prevalent in the language.

Based on the discussion in the previous section, and drawing from the analysis of split ergativity by (Amritavalli & Sarma, 2002) and (Saha & Patgiri, 2013), I propose a model which is designed to capture the two splits in the case alignment in Assamese. As illustrated in Figure (2.5), the model shows that the agents of transitive verbs (A) and the subjects of unergative verbs (S_a) are obligatorily marked with the ERGATIVE case enclitic in Assamese, while the subject of unaccusative verbs (S_o) remains unmarked. This distinctive split

¹²There are a few rare marked-nominative languages, such as the Oromo, Dasenech and Kambata languages of the Cushitic family. In such an alignment system, the S and A functions are overtly marked, while O remains in its citation form (Dixon, 1979; Handschuh, 2008). In contrast, Assamese does not fit within this category. In Assamese, +human +animate objects are consistently marked with the accusative case marker, while the subjects of unaccusative verbs do not receive any obligatory marking.

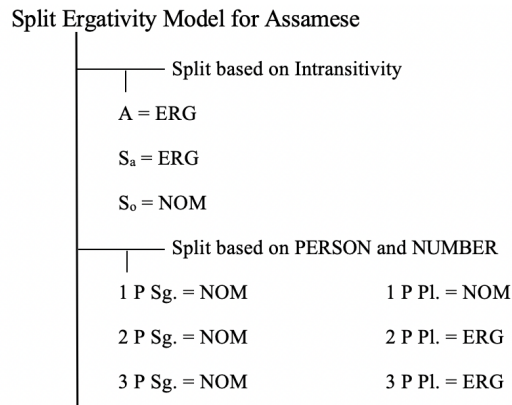


Figure 2.5: Split Ergative Case Model for Assamese

extends to pronominals as well, with the 1st PERSON singular and plural, 2nd PERSON singular, and 3rd PERSON singular pronominal subjects remaining unmarked in NOMINATIVE case, while the 2nd and 3rd PERSON plural pronominal subjects are obligatorily marked with ERGATIVE case enclitics.

2.3.1 Person and Number Based Split

An NP-based split where the (pro)nominals get different types of morphological case markings is a common crosslinguistic phenomenon as observed in Dyrbal, Punjabi, Thargari, Aranda, and Gumbaynggir, among others (Legate, 2014; Bjorkman, 2018; Butt, 2005; Garrett, 1990; Handschuh, 2008; Silverstein, 1976). In the Indo Aryan ergative language Punjabi, for example, the 1st and 2nd PERSON pronouns are not marked for ergative case, whereas 3rd PERSON pronouns as well as nominals get ERGATIVE marking, as demonstrated in 21 where the 1st PERSON singular ‘maiN’ is unmarked while the 3rd PERSON ‘oh’ gets the ERGATIVE case marking:

- (21) a. maiN kamm kita
 1.NOM work.M.SG.NOM do.PST.M.SG
 ‘I did some/the work.’
- b. o=ne kamm kita
 3.SG=ERG work.M.SG.NOM do.PST.M.SG
 ‘He/She did some/the work.’ (Punjabi: Bhatia, 1993 as cited in Butt and Deo, 2005)

Silverstein (1976) argues that ‘natural agents’, such as 1st and 2nd PERSON pronouns are higher on the animacy hierarchy scale, and hence do not require an overt agentive marking. Thus, an overtly marked ergative-absolutive alignment is found on the lower end of Silverstein’s hierarchy scale, while zero marked nominative-accusative on the upper end as shown in Figure 2.6.

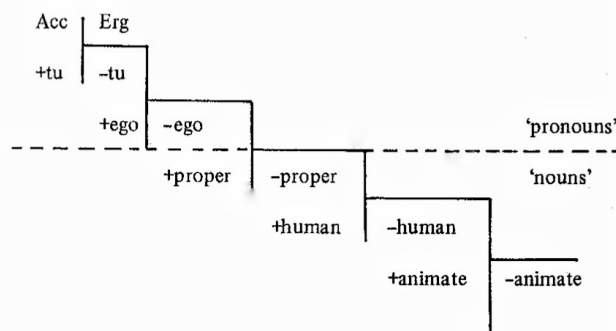


Figure 2.6: Animacy Hierarchy Scale (Silverstein, 1976, p. 176)

Table 2.1 below demonstrates that the PERSON and NUMBER based split in Assamese consistent with Silverstein’s Overt Marking Hypothesis. The pronouns at the lower end get ERG-marked, while the ones on the upper end, with the highest animacy features remain unmarked. However, this hierarchy scale has several limitations, as the different alignment systems do not show consistent agreement with the scale. For example, while in many systems the 1st PERSON pronoun is higher on the animacy scale than the 2nd PERSON, there are languages where the latter is considered higher than the first (Dixon, 1994; McGregor, 2009). In contrast, the ERG marking on full NPs in Assamese does not fall within Silverstein’s animacy hierarchy scale, since [+proper], [+human] and [+animate] nouns, which are higher on the scale, also trigger an ERG-marking (apart from being ACC-marked, when functioning as O (refer to §2.4)).

Assamese has 1st, 2nd, and 3rd PERSON personal pronouns like English with additional honorific pronouns for 2nd and 3rd PERSON singular and plural forms. This language also marks personal pronoun agreement on the verbs as shown in Table 2.1. Assamese, like most other Indo-Aryan languages has a threefold distinction in personal pronouns that is regulated by various sociolinguistic norms, such as age, social, educational and relationship status for the 2nd PERSON (Kakati, 1941; Goswami and Tamuli, 2003). The speaker can use the lowest 2nd PERSON singular/plural *toi/tohôt* to address someone younger or lower in social status, and siblings. Moreover, this pronoun is also used in a reciprocal manner among close friends to show intimacy. The familiar and polite singular/plural forms *tumi/tumaluk* are used in a reciprocal way among friends, siblings and in some circumstances with parents and partners. The 3rd PERSON honorific singular/plural forms *apuni/apunaluk* are usually used with strangers whose social status or age is unknown, elders, and for people in professions that are regarded as socially esteem. However, what is important here is that irrespective of the degree of honorificity, the PERSON and NUMBER based split is maintained in the language.

In other words, only the plural forms of all the three degrees of the 2nd PERSON pronouns take the ergative case marking. Furthermore, the distinct 2nd PERSON pronouns get cross-referenced on the verb irrespective of the different case endings, as shown in (23). The examples in (22)-(24) further demonstrate

Person	Singular	Singular Honorific	Plural	Plural Honorific	Tense	Agreement Markers
1	moi.NOM		ami.NOM		Present	-u
					Past	-u
					Future	-m/-im
2	toi.NOM		tohõt.ERG		Present	-a
	tumi.NOM		tumaluk.ERG		Past	-a
		apuni.NOM		apunaluk.ERG	Future	-ba/-bo
3	i/ei.NOM	tẽu.NOM	ihõt.ERG	tẽu-luk.ERG	Present	-e
	PROXIMITY	tek ^h et.NOM	PROXIMITY	tek ^h et-xokol.ERG	Past	-e
	xi/tai.NOM		xihot.ERG		Future	-bo
	DISTANCE		DISTANCE			

Table 2.1: Case-Marking on Assamese SUBJ Pronouns

the PERSON and NUMBER based split highlighted in table (2.1).

- (22) a. moi xãtur-il-u
1SG.NOM swim-PST-1
'I swam.'
- b. ami xãtur-il-u
1PL.NOM swim-PST-1
'We swam.'
- (23) a. toi xãtur-il-i
2SG.NOM swim-PST-2
'You swam.'
- b. tohot=**i** xãtur-il-i
2PL=ERG swim-PST-2
'You swam.'
- c. tumi xãtur-il-a
2SG.NOM swim-PST-2
'You swam.'
- d. tumaluk=**e** xãtur-il-a
2PL=ERG swim-PST-2
'You swam.'
- e. apuni xãtur-il-e
2SG.NOM swim-PST-2
'You swam.'

- f. apunaluk=e xātur-il-e
 2PL=ERG swim-PST-2
 ‘You swam.’

- (24) a. xi/tai/teu xātur-il-e
 3SG.NOM swim-PST-3
 ‘S/he swam.’

- b. xihōt/tāhat=e xātur-il-e
 3PL=ERG swim-PST-3
 ‘They swam.’

Although Saha and Patgiri (2013) and Choudhary (2014) claim that only those pronouns that end in consonants take the overt case marker, while those with a vowel ending remain \emptyset marked, we saw through the data in (19), repeated here as (25), that this cannot be true¹³. Hence, a PERSON and NUMBER split based on Silverstein’s hierarchy scale seems more acceptable.

- (25) meghali=e bhat randh-i as-e
 meghali=ERG rice cook-PROG be.-PRES-3
 ‘Meghali is cooking rice.’

Furthermore, the PERSON and NUMBER split in Assamese is independent of the intransitivity split.¹⁴ The PERSON and NUMBER-based split takes supremacy over any other behaviour that has to do with the intransitivity split in unergative verbs¹⁵, such that we end up with a non-ERG-marked SUBJ of an unergative verb such as *kāh* ‘cough’, in (26), when the subject is a 3rd PERSON pronoun, as in *tai* ‘she’.

- (26) a. suwali-zoni=e kāh-is-e
 girl-CLF=ERG cough-PERF-3
 ‘The girl has coughed.’

¹³Miriam Butt has pointed out that the pronouns without an overt ERGATIVE marker are old obliques while the ones with an overt ERGATIVE marking are innovated forms (usually the 3rd PERSON). Alternatively, these forms may also result from the incorporation of a form with nominal features, subsequently adopting the ergative case. Interestingly, Kakati (1941, p.293) has mentioned the term ‘oblique bases’ that were derived from the ‘old genitive’ in Middle Indo Aryan. He states that the Assamese pronouns have ‘oblique bases’ to which different case endings can be added to create its final form, such as the 1st PERSON singular *moi*. He states that *moi* consists of the oblique base *mo* to which the ACCUSATIVE *-ok* or the DATIVE *-or* can be added to form *muk* ‘to me’ or *mur* ‘mine’. However, in the subject position, the 1st PERSON singular *moi* does not bear any overt marking.

¹⁴This will be discussed in more details in the following section.

¹⁵This is not applicable for unaccusative verbs, which would still remain unmarked even for 2nd PERSON pronoun or 3rd PERSON pronoun SUBJs.

- b. tai.∅ kāh-is-e
 3.F.NOM cough-PERF-3
 ‘She has caughed.’

The PERSON and NUMBER-based split also takes supremacy in (di)transitive verbs. This is observed in an example such as (27), where the 3rd PERSON singular subject *xi* ‘he’ of the transitive verb *t^hel* ‘push’ remains zero marked. This example is in contrast to example (29) where the full NP subject *manuh-zon* ‘the person’ of the same transitive verb is overtly marked with the ERG case marker *-e*.

- (27) xi soki-k^hon t^hel-is-e
 he.NOM chair-CLF push-PERF-3
 ‘He has pushed the chair.’

- (28) a. manuh-zon=**e** soki-k^hon t^hel-il-e
 person-CLF=ERG chair-CLF push-PST-3
 ‘The man pushed the chair.’

- b. manuh-zon=**e** soki-k^hon t^hel-e
 person-CLF=ERG chair-CLF push-3
 ‘The man pushes the chair.’

2.3.2 Intransitivity Based Split

Case marking distributions on the basis of the verb’s valency, along the nature of thematic role associated with subject has been reported in several languages from all over the world (Li, 2007a; Takeuchi et al., 1995; Johns, 1992; Legate, 2006; Coon, 2012, 2013; Harris, 1982, 2008; Dixon, 1972, 1994). In many languages, the subjects of unergative (S_a) verbs perform, initiate and control an event, such as the subjects of the verbs, ‘talk’, ‘jump’, and ‘dance’. On the other hand, the subjects of unaccusative (S_o) verbs are affected by uncontrollable events such as ‘drown’, ‘fall’, ‘be lost’, and ‘burn’. However, the distinction between S_a and S_o verbs can be fuzzy. For example, the Siouan language Hidatsa marks intransitive verbs like ‘have hiccups’, ‘forget’ and ‘die’ just like its other intransitive volitional verbs ‘run’ and ‘sing’, while verbs like ‘roll over’ and ‘stand up’ are categorised as S_o verbs (Dixon, 1994, p.74). Similarly, Acehnese (western Austronesian) treats ‘cough’, ‘vomit’ and ‘dream’ as S_a even when the subject does not have any control over these actions. In a similar manner, Assamese also happens to treat these verbs as unergatives.

There are additionally certain languages that have fluid intransitive verbs, whose subjects can either get marked like an A or an O depending on the context. For example, the verb ‘go’ in Crow (Siouan family) can either behave like S_a or S_o ‘depending on whether or not volition is involved’ (Dixon, 1994, p. 81). The same phenomenon is also observed in Tsova-Tush, a North-east

Caucasian language and Indo-Aryan languages such as Hindi and Assamese. However, if we keep volitionality, emphasis, and context aside, Assamese and Sylheti are the only Indo-Aryan languages that displays a split on the basis of intransitivity. Nepali, on the other hand, has both an intransitivity and aspect based split (Li, 2007b). While most other Indo-Aryan split-ergative languages such as, Hindi/Urdu, Marathi, Gujarati and Punjabi exhibit a split based on aspect, this is not the case for Assamese. Having said this, however, Bhatt (2007, p. 4) claims that Assamese does have an aspect based split where ergative-marked subjects are obligatory in perfective tenses and optional for non-perfective tenses. This, however, we note that is not the case. This is illustrated by the following examples, where the subject NP *manuh* ‘person’ gets ergative marking in both perfective and imperfective aspectual contexts.

- (29) a. *manuh-zon=e soki-k^hon t^hel-il-e*
 person-CLF=ERG chair-CLF push-PST-3
 ‘The man pushed the chair.’
- b. *manuh-zon=e soki-k^hon t^hel-e*
 person-CLF=ERG chair-CLF push-3
 ‘The man pushes the chair.’

The full NP subjects of (di)transitive verbs always take an overt ergative case marker in Assamese. Among the intransitive verbs, the unergatives trigger an overt marker, while the patient-like unaccusative verbs remain unmarked for case. More specifically, unergative subjects exhibit a pattern similar to that of (di) transitive subjects, while unaccusative subjects exhibit a pattern akin to transitive inanimate objects as seen in Table (2.2):

Constituent	Subject	Object (Animate)	Object (Inanimate)
A	-e	-(o)k	∅
S _a	-e	—	—
S _o	∅	—	—

Table 2.2: Case-Marking Distribution on Assamese (non-pronominal) NPs

Although this phenomena in Assamese has been previously discussed in Amritavalli and Sarma (2002), they have been of the opinion that the ∅ marked subjects of the unaccusative verbs are in absolutive case and there is no nominative case in the language. Comrie (2013) uses the same terminology and refers to the case that encodes S and P/O as the absolutive. In an active-inactive (active-stative) system, the S argument of an active predicate behaves like an A argument and take ergative marking. On the other hand, the S argument of an inactive/stative argument aligns with an O argument, and hence takes absolutive case and remains zero-marked.

I, however, prefer to maintain the traditional view that nominative and absolutive are unmarked cases for subjects and objects, and choose to use the

term nominative for unmarked S_o . This is in accordance to Mohanan's (1994) description that if a language has two distinct cases associated with subjects, where one is inflected and the other uninflected, the ergative refers to the case of the former and nominative to the latter, as seen in (30a) and (30b), respectively. Here, the 2nd PERSON plural pronoun *tumaluk* 'you' takes the ergative case ending while the 2nd PERSON singular pronoun *tumi* 'you' remain unmarked for the unergative verb *dour* 'run':

- (30) a. *tumaluk=e dour-il-a*
 2PL=ERG run-PST-2
 'You ran.'
- b. *tumi.∅ dour-il-a*
 2.NOM run-PST-2
 'You ran.'

DeLancey (1981) and Mithun (1991), on the other hand, object to treating such split intransitive systems as instances of split ergativity and argue that these are a constituent of the active system¹⁶.

2.4 Differential Object Marking

The animacy hierarchy accounts for a good deal of the cross-linguistic variation in split ergative systems, and that this can differ on the basis of the nature of the noun type (McGregor, 2009). The same premise can be applied to objects in Assamese as the occurrence of the accusative marker depends on where the object NPs stand on the hierarchy scale. For example, in (31a) the animate DO *rita* 'Rita' of the transitive verb *d^hor* 'hold' takes the accusative case marker *-(o)k*, while in (31b), the inanimate DO *bol* 'ball', associated with the same transitive verb, remains unmarked.

- (31) a. *nitu=e rita=k d^hor-il-e*
 nitu=ERG rita=ACC hold-PST-3
 'Nitu caught Rita.' (Lit. 'Nitu held Rita.')
- b. *nitu=e bol-tu d^hor-il-e*
 nitu=ERG ball-CLF hold-PST-3
 'Nitu caught the ball.' (Lit. 'Nitu held the ball.')

Hindi, however, can optionally mark its inanimate NPs with the accusative case as illustrated through the examples in (32) (Mohanan, 1994, p. 104). In (32a) the animate DO *bacce* 'child' is obligatorily marked with the accusative case marker, while the inanimate NP *haar* is marked with the accusative in (32b), to give it a definite, as well as a specific reading, while its unmarked form gives a non-specific reading in (32c).

¹⁶Bernard Comrie suggested that Assamese is also an instance of such active languages (P.C.)

- (32) a. *Ilaa-ne bacce-ko (*baccaa) uTaayaa*
 Ila-ERG child-ACC (*child) lift.PERF
 ‘Ila lifted a/the child’
- b. *Ilaa-ne haar-ko uTaayaa*
 Ila-ERG necklace-ACC lift.PERF
 ‘Ila lifted the necklace.’
- c. *Ilaa-ne haar uTaayaa*
 Ila-ERG necklace lift.PERF
 ‘Ila lifted a/the necklace.’ (Hindi: Mohanan, 1994, p. 104)

Similarly, in the following Hindi example, we see that the [+human] proper name in (33a) *Dilīp* ‘Dilip’ is obligatorily marked with the accusative marker, while *Kalkattā*, ‘Calcutta’ which is the proper name of a city, and thereby an inanimate entity, remains unmarked.

- (33) *maiñe Dilīp ko (*Dilīp) dekhā/ maiñe Kalkattā*
 1SG.ERG Dilip ACC (*Dilip) see.PFV.1SG/ 1SG.ERG Calcutta
dekhā
 see.PFV.1SG
 ‘I saw Dilip / I saw Calcutta.’ (Hindi: Montaut, 2018)

In Assamese, animals, birds and trees are higher in the hierarchy than non-living things and hence can be optionally marked as shown through the NP *tamul* ‘areca nut’ in the example below:

- (34) *tamul-zupa(=k) ne-kat-ib-i*
 areca-nut-CLF(=ACC) NEG-cut-FUT-2
 ‘Do not cut the areca nut (tree).’

Furthermore, when an object that is lower on the hierarchy scale, such as a common noun, gets optional marking, the (non-)markedness accounts for definite/specificity information. For example, the unmarked object NP *sur* ‘thief’ in (35a) refers to a [+definite] [-specific] object, while the accusative marked NP in (35b) implies a [+definite] [+specific] instance.

- (35) a. *nitu=e sur-tu d^hor-il-e*
 nitu=ERG thief-CLF hold-PST-3
 ‘Nitu caught the thief.’ (Lit. ‘Nitu held the thief.’)
- b. *pulis=e sur-tu=k d^hor-il-e*
 police=ERG thief-CLF=ERG hold-PST-3
 ‘Police caught the thief.’ (Lit. ‘Police held the thief.’)

Further, it should be noted here that DOM in Assamese only applies to objects in neutral contexts. If the objects are placed in a non-neutral context, as in TOPIC positions, the inanimate indefinite NP can be accusative-marked. This is highlighted through the inanimate, indefinite NP *zibon* ‘life’ in the following example.

- (36) *zibon-ok adər-ok d^hopat-ok nə-hə-i*
 life-ACC welcome-IMP tobacco-ACC NEG-be-3
 ‘Welcome life, not tobacco’ (Chowdhary, p. 118)

Although the majority of the literature (e.g. Gundel et al. (1993), Enç (1991), but see Prince (1990)) suggest that definiteness also implies specificity, this does not hold true for Assamese, given that the numeral classifier *-tu* attached to the NP *sur* ‘thief’ implies that it already takes a definite reference. While DOM is determined by definiteness and (or) specificity in both Hindi and Assamese, there are other Indo-Aryan languages like Sinhala/Sinhalese where an animate O can be optionally marked by accusative, and where DOM is determined by animacy and not definiteness (Thampoe, 2017).

Outside Indo-Aryan languages, this grammatical phenomenon is also observed in many other languages such Turkish, Spanish and Maltese. In Turkish, only a definite/specific DO gets marked with the accusative suffix as shown in the em -1 marked NP *kitab* ‘book’ in (37a) and (37c). The indefinite/non-specific reference of the same object NP in (37b) remains unmarked (Spencer, 2008; Kornfilt, 2009; Von Heusinger and Kornfilt, 2005). definite, indefinite non-specific, indefinite specific

- (37) a. (Ben) *kitab-ı oku-du-m*
 I book-ACC read-PAST-1SG
 ‘I read the book.’ (Definite-(von Heusinger & Kornfilt, 2005, p. 7))
- b. (Ben) *bir kitab-ı oku-du-m*
 I a book-ACC read-PAST-1SG
 ‘I read a certain book.’ (Indefinite specific)
- c. (Ben) *bir kitap oku-du-m*
 I a book read-PAST-1SG
 ‘I read a book.’ (Indefinite, Non-specific-Turkish:(von Heusinger & Kornfilt, 2005, p. 8))

In Spanish, the accusative preposition *a* appears in front of animate NPs, such as the male proper name *Juan* ‘John’ (38a). The inanimate NP *el libro* ‘the book’ on the other hand, remains unmarked, as illustrated in (38b). Further, similar to Turkish (and Maltese), the accusative marker infers definiteness and specificity in this language.

- (38) a. María vio a Juan.
 Mary see.AOR.3SG ACC John
 ‘Mary saw John.’
- b. María vio el libro
 Mary see.AOR.3SG the book
 ‘Mary saw the book.’ (Spanish: Comrie, 2013)

In Maltese, the accusative marker *lil* which is also homophonous with the dative marker, triggers differential object marking given that it obligatory marks proper names, yet it usually also tends to mark definite human NPs (39a). Indefinite human NPs may also be marked. Non-human NPs, on the other hand are usually unmarked as seen in the case of the NP object ‘the road’ in (39b) (Camilleri and Sadler, 2012).

- (39) a. Raj-t (l)it-tifel
 saw.PV-1SG ACC.DEF-boy
 ‘I saw the boy.’
- b. Wera t-triq lil Pawlu
 showed.PV.3SGM DEF-road DAT Paul
 ‘He showed the road to Paul.’ Maltese: (Camilleri & Sadler, 2012, p. 121)

While, the animacy and definiteness/specificity hierarchy scale can differ from language to language, it is a common grammatical phenomenon seen in most languages of the world. Croft (2003, p. 132)’s ‘animacy hierarchy’ scale tells us that the referents higher on the scale, such as 1st/2nd PERSON pronouns are more likely to receive overt case marking than inanimate common nouns that are lower on the hierarchy. Croft’s hierarchy is provided in (40):

- (40) **first/second person pronoun > third person pronoun > proper names > human common noun > non-human animate common noun > inanimate common noun**

Montaut (2018) integrates specificity into Croft’s (2003) animacy scale to show that the definite specific NPs that are higher on the scale will be obligatorily marked, as opposed to non-specific NP counterparts referring to inanimate objects. Montaut’s hierarchy is provided in (41) below:

- (41) **Personal pronoun/Proper name > Definite NP > Indefinite specific NP > Non-specific NP**

DOM in Assamese corroborates with the above animacy hierarchy and specificity scale as, only the personal pronouns and proper names get obligatory marking. Human common nouns and non-human animate common nouns

are optionally marked. Inanimate common nouns, such as *gāri* ‘vehicle’ or *g^hor* ‘house’ remain unmarked in the object position. Devi (1986) mentions that a similar animacy hierarchy also exists internal to the subject NPs. She argues that the inanimate nouns of motion are higher on the animacy scale and therefore get coded with the ergative case marker. Static inanimate NPs are, however, on the lower end of the scale, and therefore remain unmarked.

2.5 Non-canonical Subjects

Majority of languages employ a different strategy to code its core arguments: S, A and O. The split marking in S that have been discussed in the previous sections is also a non-canonical form of coding as it does not follow a rigid pattern. However, in this section, we will only focus on the non-canonical marking of S and A that are either encoded in the dative or genitive case. Although these subjects show similar grammatical properties to its canonical subjects, they exhibit a different agreement pattern. The non-canonical subjects always take a default third person agreement as opposed to a standard NOMINATIVE-ACCUSATIVE agreement pattern followed by the ERGATIVE/NOMINATIVE subjects in the language as shown by the 1st PERSON marker *-u* in (42a) for the 1st PERSON nominative subject *moi* in contrast to the 3rd PERSON marker *-e* in (42b) for the 1st PERSON genitive¹⁷ subject *mur*:

- (42) a. *moi b^hoi k^ha-is-u*
 I.NOM fear eat-PERF-1
 ‘I am scared’ (Lit: ‘I have eaten fear.’)
- b. *mur b^hoi lag-is-e*
 I-GEN fear get-PERF-3
 ‘I am scared’ (Lit: ‘I have got fear.’)

A similar agreement pattern is also observed in Nepali and Bengali where the non-canonical subjects do not trigger the person agreement which is otherwise obligatory for canonical subjects (Bhatt, 2007; Onishi, 2001). In example (43a), the Nepali verb *dekhin* ‘appear’ agrees with the object when the subject is in dative case. Similarly, the Bengali light verb *bhalo lag* ‘like’ gets a third person agreement instead of agreeing with the first person genitive subject *amar* ‘I’ in (43b)

- (43) a. *budhi manche-lai chara dekhin-cha*
 old woman.F-DAT bird.M appear-PRES.M.SG
 ‘The bird appears to the old woman.’ (Bhatt, 2007, p. 15)

¹⁷Other uses of the Genitive, specifically in the oblique object position will be discussed in Chapter 4.

- b. ama-r ca bhalo lag-e
 1.SG-GEN tea good be.attached.PRES-3
 ‘I like tea.’ (Lit: ‘I feel good about tea.’) (Onishi, 2001, p. 118)

In most of the South Asian languages an experiencer subject that expresses feelings or is affected by an activity or possesses certain things or qualities is either in the nominative or dative or genitive case (Subbarao et al., 2009) as highlighted through the GEN subjects *moi* and *mur* in example 44:

- (44) a. moi b^hoi k^ha-is-u
 I.GEN fear eat-PERF-1
 ‘I am suddenly scared.’ [Lit: ‘I am eating fear.’]
- b. mur b^hoi lag-is-e
 I.GEN fear get-PERF-1
 ‘I am getting scared.’

As pointed out by Woolford (2008), Assamese has only one verb that marks its experiencer subject with a DATIVE, or rather perhaps with an ACCUSATIVE. Furthermore, as mentioned earlier this subject do not show person agreement and takes the default 3rd PERSON agreement as shown in 45:

- (45) a. ram=ok b^hat lag-e
 Ram=DAT rice want-3
 ‘Ram wants rice.’ (Woolford, 2008, p. 30)
- b. muk b^hat lag-e
 I.DAT rice want-3
 ‘I want rice.’

She also points out that since Assamese does have ‘frighten class verbs’ and uses constructions like ‘feed fear to’ or ‘cause to eat fear’, the ERGATIVE subject of such constructions are actually the argument of the verbs ‘feed, and ‘cause to eat’ instead of ‘fear’ as demonstrated in 46:

- (46) a. gan-tu=e xap-tu=k bhoi khuale
 song-CLF=ERG snake-CLF=DAT fear fed/ate
 ‘The song frightened the snake.’ (Woolford, 2008, p. 28)
- b. xap=e tak bhoi khuai
 snake=ERG him.DAT fear feed/eat
 ‘Snakes scare him.’ (Woolford, 2008, p. 28)

Moreover, this is also true for other causal themes like ‘anger’ or ‘calm’ expressed through light verb constructions such as ‘do anger’ or ‘do/make calm’ where the ERGATIVE subject is actually the argument of the verb ‘do/make’ and not ‘anger/calm’. We must also note here that DATIVES that are typically connected with psych predicates and can be syncretic with the accusative as seen in the DATIVE subjects *ram* ‘Ram’ and *muk* ‘I’ in (45) and the objects *xap* ‘snake’ and *tak* ‘him’ in (46). This syncretism is also highlighted for convenience in (47) where the beneficiary *lusi* ‘Lucy’ in (47a) takes the DATIVE case while the DO *lusi* ‘Lucy’ in (47b) triggers the accusative case:

- (47) a. *lusi=k kitap-k^hon diya*
 lucy-DAT book-CLF give-2
 ‘(You) give the book to Lucy.’
- b. *lusi=k mar-il-e*
 lucy=ACC beat-PST-3
 ‘(He/She/It) beat/killed Lucy.’

This may raise the question whether the ACCUSATIVE case marker *-(o)k* in Assamese is in fact syncretic with the DATIVE or are these instances of double object constructions. The main reason behind treating such argument as DATIVE is because these are psyche predicates that perfectly map with the thematic role of beneficiary instead of theme/patient. Even in Hindi, the DATIVE and ACCUSATIVE case markers are syncretic as highlighted through the DOs *ravan* ‘Ravan’ in example 48:

- (48) a. *ram=ne ravan=ko mara*
 ram=ERG ravan=ACC kill.M.SG
 ‘Ram killed Ravan.’
- b. *ravan=ko t^hand lagi he*
 ravan=DAT cold get-INF.M.SG bePRES.3.SG
 ‘Ravan is cold.’

2.6 Summary and Conclusion

We began by looking at the contrasting views about the case marking system in Assamese which shows that its alignment has not been clearly understood so far. We tried to eliminate this confusion by providing a definitive account of the ergative splits observed in the language. We also presented an overview of the animacy and specificity-based hierarchy that triggers a differential object marking in Assamese. Further, we also looked at the non-canonical subjects in the language.

In conclusion, contrary to the popular interpretation that Assamese is a NOMINATIVE-ACCUSATIVE language, the analysis presented in this paper

strongly supports the characterization of Assamese as a split ergative language based on the splits in intransitivity and PERSON and NUMBER. The intransitivity based split in Assamese implies that the unergative subjects in the language exhibit a pattern similar to that of (di) transitive subjects (A), while the unaccusative subjects exhibit a pattern identical to inanimate direct objects of transitive verbs (O). Additionally, the split in the pronominals that is based on a PERSON and NUMBER distinction, ensures that only the 2nd and 3rd PERSON plural pronominal subjects trigger an ERGATIVE subject, while the rest remain unmarked in NOMINATIVE form.

Bibliography

- Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory*, 21(3):435–483.
- Amritavalli, R. and Sarma, P. P. (2002). A case distinction between unaccusative and unergative subjects in Assamese. *Snippets*, 5:6–7.
- Andersen, T. (1988). Ergativity in pari, a nilotic ovs language. *Lingua*, 75(4):289 – 324.
- Anderson, S. R. (1976). On the notion of subject in ergative languages. *Subject and topic*, 1:23.
- Beames, J. (1872). *A Comparative Grammar of the Modern Aryan Languages of India: On sounds*, volume 1. London: Trübner 1872-79.
- Bhatt, R. (2007). Ergativity in indo-aryan languages. In *Talk given at the MIT Ergativity Seminar*.
- Bhattacharjya, M. (n.d.). Case in Assamese: Morpho-Syntactic & Semantic Analysis.
- Bhattacharyya, S. (2011). Experts question roots of 'modern' assamese - reprint of nathan brown's original 1848 grammar creates flutter in the state's literary circles.
- Bickel, B. (1995). In the vestibule of meaning: Transitivity inversion as a morphological phenomenon. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"*, 19(1):73–127.
- Bjorkman, B. M. (2018). Ergative as perfective oblique. *Syntax*, 21(4):321–361.
- Brown, N. (1848). *Grammatical notices of the Asamese language*. American Baptist Mission Press, Sibsagor.
- Butt, M. (2005). Differential Case Marking Theory (DCT).
- Butt, M. (2006). The dative-ergative connection. *Empirical Issues in Syntax and Semantics*, 6:69–92.

- Butt, M. and Deo, A. (2001). Ergativity in Indo-Aryan. *Kurdica*, 5(3).
- Camilleri, M. and Sadler, L. (2012). On the analysis of non-selected datives in Maltese. In Butt, M. and King, T. H., editors, *Proceedings of the LFG12 Conference*, pages 118–138, Stanford. CSLI Publications (online publications).
- Chowdhary, R. (2014). Differential marking of cases in Asamiya. *North East Indian Linguistics*, 6:107–124.
- Comrie, B. (1978). Ergativity. In Lehmann, W. P., editor, *Syntactic Typology: Studies in the Phenomenology of Language*, pages 329–394. University of Texas Press, Austin.
- Comrie, B. (2013). Alignment of case marking of full noun phrases. In Dryer, M. S. and Haspelmath, M., editors, *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.
- Coon, J. (2012). Split ergativity and transitivity in chol. *Lingua*, 122(3):241–256.
- Coon, J. (2013). Tam split ergativity, part i. *Language and Linguistics Compass*, 7(3):171–190.
- Croft, W. (2003). *Typology and universals*. Cambridge University Press, Cambridge.
- Deb, D. (2012). On case marking in assamese bengali and oriya. *International Journal of Applied Linguistics & English Literature*, 1(2):102.
- DeLancey, S. (1981). An interpretation of split ergativity and related patterns. *Language*, pages 626–657.
- Devi, J. (1986). *Ergativity: A Historical Analysis in Assamese*. PhD thesis, University of Delhi, New Delhi.
- Dixon, R. M. (1972). *The Dyirbal language of north Queensland*, volume 9. CUP Archive.
- Dixon, R. M. W. (1979). Ergativity. *Language*, 55:59–138.
- Dixon, R. M. W. (1994). *Ergativity*. Cambridge University Press, Cambridge.
- Dryer, M. S. (1997). On the six-way word order typology. *Studies in Language. International Journal sponsored by the Foundation “Foundations of Language”*, 21(1):69–103.
- Enç, M. (1991). The semantics of specificity. *Linguistic Inquiry*, 22:1–25.
- Garrett, A. (1990). The origin of np split ergativity. *Language*, 66(2):261–96.

- Goswami, G. C. and Tamuli, J. (2003). Asamiya. In Cardona, G. and Jain, D., editors, *The Indo-Aryan Languages*, pages 391–443. Routledge, London.
- Gundel, J. K., Hedberg, N., and Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69:274–307.
- Haddad, Y. A. (2007). *Adjunct Control in Telugu and Assamese*. PhD thesis, UNIVERSITY OF FLORIDA.
- Handschuh, C. (2008). Split Marked-S Case Systems. *Scales*, 86:73–82.
- Harris, A. C. (1982). Georgian and the unaccusative hypothesis. *Language*, 58(2):290–306.
- Harris, A. C. (2008). On the explanation of typologically unusual structures. *Linguistic universals and language change*.
- Johns, A. (1992). Deriving ergativity. *Linguistic Inquiry*, 23(1):57–87.
- Kakati, B. (1941). *Assamese: Its Formantion and Development*. Department of Historical and Antiquarian Studies, Gauhati, Assam.
- Kellogg, S. H. (1972). *A grammar of the Hindi language*. .
- Kornfilt, J. (2009). DOM and two types of DSM in Turkish. In de Hoop H., d. S. P., editor, *Differential Subject Marking*, pages 79–111. Springer, Dordrecht.
- Legate, J. A. (2006). Split absolutive. In *Ergativity*, pages 143–171. Springer.
- Legate, J. A. (2014). Split ergativity based on nominal type. *Lingua*, 148:183–212.
- Li, C. (2007a). Split ergativity and split intransitivity in nepali. *Lingua*, 117(8):1462–1482.
- Li, C. (2007b). Split ergativity and split intransitivity in Nepali. *Lingua*, 117(8):1462 – 1482.
- Masica, C. P. (1982). Ergativity in south asia. *South Asian Review*, 6(3):1–11.
- McGregor, W. B. (2009). Typology of ergativity. *Language and Linguistics Compass*, 3(1):480–508.
- Mithun, M. (1991). Active/agentive case marking and its motivations. *Language*, 67(3):510–546.
- Mohanan, T. (1994). *Argument Structure in Hindi*. CSLI Publications, Stanford.

- Montaut, A. (2018). The rise of differential object marking in Hindi and related languages. In Seržant, I. A. and Witzlack-Makarevich, A., editors, *Diachrony of differential argument marking*, volume 19, pages 281–313. Language Science Press, Berlin.
- Morey, S. (2013). The marking of noun phrases: Some observations on the languages of north east india. *North East Indian Linguistics*, 5:169–191.
- Nath, D. (2003). *Case in Assamese*. PhD thesis, Central Institute of English and Foreign Languages, Hyderabad.
- Onishi, M. (2001). Non-canonically marked s/a in bengali. In Aikhenvald, A. Y., Dixon, R. M., and Onishi, M., editors, *Non-canonical marking of subjects and objects*, volume 46. John Benjamins Publishing.
- Plank, F. (1979). *Ergativity: Towards a theory of grammatical relations*. Academic Press.
- Prince, E. F. (1990). Syntax and discourse: A look at resumptive pronouns. In *Annual meeting of the Berkeley Linguistics Society*, volume 16, pages 482–497.
- Saha, A. and Patgiri, B. (2013). Ergativity in Axomiya. *Language in India*, 13(12):35–46.
- Sarma, K. (2015). Some observations on the characteristics of assamese. *Editorial Board*.
- Sen, S. K. (1973). *Proto-New Indo-Aryan that is Avahattha*. PhD thesis, Calcutta University, Calcutta: Eastern Publishers.
- Silverstein, M. (1976). Hierarchy of features and ergativity. *Grammatical categories in Australian languages*.
- Spencer, A. (2008). Does Hungarian have a case system? In Corbett, G. G. and Noonan, M., editors, *Case and grammatical relations: Studies in honor of Bernard Comrie*, pages 35–57. John Benjamins Publishing Company, Amsterdam/Philadelphia.
- Subbarao, K., V. K., Rabha Hakacham, U., and Thokcham, S. (2009). Case-marked pro: Evidence from rabha, manipuri, hindi and telugu. *R. Biemeier and F. Haller (eds.), Linguistics of the Himalayas and Beyond (Berlin: Mouton de Gruyter, 2007), 291–322*.
- Takeuchi, T., Takahashi, Y., et al. (1995). Split ergative patterns in transitive and intransitive sentences in tibetan: a reconsideration. *Senri ethnological studies*, 41:277–288.
- Thampoe, H. D. (2017). *Sinhala and Tamil: A case of contact-induced restructuring*. PhD thesis, Newcastle University, Newcastle.

- Von Heusinger, K. and Kornfilt, J. (2005). The case of the direct object in Turkish: Semantics, syntax and morphology. *Turkic Languages*, 9:3–44.
- Woolford, E. (2008). Differential subject marking. In *Springer*, pages 17–40. Helen de Hoop and Peter de Swart.
- Zakharyin, B. (2015). Indo-Aryan ergativity and its analogues in languages of central and western eurasia. *Lingua Posnaniensis*, 57(2):63–76.
- Zwart, J.-W. and Lindenbergh, C. (2015). Rethinking alignment typology. *Manuscript. University of Groningen*.

Chapter 3

Acquisition of Assamese Split Ergativity

Abstract

Assamese represents an insightful example of an Indo-Aryan language relying on split ergativity based on PERSON and NUMBER and/or semantics of the verbs, unlike most other Indo-Aryan languages that have an aspect-based split. The current paper provides an analysis of how these ergative splits are acquired by children based on two sets of data elicited through an experiment and a semi-structured language game. Data from forty children aged between 2;6 to 5;11 years that participated in the production experiment show that children produce adult-like structures with appropriate use of ERGATIVE morphology as early as 2;6 years and display a growing proficiency as they progress in terms of age. Furthermore, data from thirty-five children that participated in the semi-structured language game show that children seem to acquire the DOM split in Assamese with little difficulty, typically achieving competence by the age of 3.

Keywords: Child Language Acquisition, Split Ergative, Differential Object Marking

3.1 Introduction

Assamese, an Indo-Aryan language, serves as the first language for around 14 million people, as reported by the Linguistic Survey of India (2001). Since 1960, it has held the position of the official language of the state, with the exception of a few districts in Assam where Bengali and Bodo are also recognized as (other) official languages due to their dominance in those regions. Besides being a significant lingua franca in the area, Assamese is also spoken as creoles by the southern Nagas and Bodo-Kacharis, further emphasizing its linguistic importance and influence.

Assamese is a morphologically split-ergative language with a NOMINATIVE-

ACCUSATIVE agreement system. The subjects of (di)transitive verbs (henceforth A) are obligatorily required to have an ERGATIVE case enclitic. Meanwhile, the direct objects (O) can take an ACCUSATIVE/DATIVE case ending or remain null-marked, depending on their position in the animacy hierarchy. However, in intransitive sentences, marking varies based on the semantics of the verb and/or the PERSON and NUMBER split, particularly when dealing with pronominal subjects. In sentences with one-place predicates, the agent-like subjects (S_a) of unergative verbs, such as ‘dance’ and ‘jump’ trigger ERGATIVE marking. On the other hand, the patient-like subjects of unaccusative verbs (S_o), such as ‘sleep’ and ‘fall’ remains zero-marked in NOMINATIVE case¹

This paper delves into the early development of children’s understanding and production of ergative splits, examining their ability to generate adult-like structures from a very young age. In the subsequent section, I will discuss the specific grammatical phenomenon under investigation, vis-à-vis split ergativity in Assamese. Section 3 will provide an overview of the existing literature on the acquisition of case markers, with a particular emphasis on studies conducted in ergative languages. Section 4 will outline the methods employed in this study. Sections 5 and 6 will introduce, elaborate, and present the results from the two studies conducted as part of this investigation. Finally, in Section 7, I will conclude with my final remarks.

3.2 Split Ergativity in Assamese

Fully ergative languages are a rarity, with the majority of ergative languages exhibiting some form of split (Dixon, 1979; Comrie, 2013; Mithun, 1991; Handschuh, 2008). The most common types of splits observed cross-linguistically revolve around the pronominal system, verb valency, and aspect.

In many languages, including Dyirbal, Punjabi, Thargari, Aranda, and Gumbaynggir, pronominals exhibit distinct morphological case markings (Legate, 2014; Bjorkman, 2018; Butt and Deo, 2001; Garrett, 1990; Handschuh, 2008). In other languages like Cocho (Popolacan), Ikan (Chibchan), proto-Arawak (Arawakan), Ket (a language isolate), and Dakota (Siouan-Catawban), transitive and intransitive verbs employ different case encodings. Even within intransitive verbs, S_a and S_o can trigger distinct case markings depending on the semantics of the verbs.

Additionally, in languages with an aspect-based split, such as Hindi/Urdu, Marathi, Gujarati, and Punjabi, the subject of perfective verbs is marked with the ERGATIVE case, while the subject of progressive verbs remains unmarked (Butt and Deo, 2001). In contrast to most Indo-Aryan languages that adopt an aspect-based split, Assamese relies on a split ergative system based on intransitivity and/or PERSON and NUMBER, which will be further explored in

¹We follow Mohanan (1994)’s terminologies and gloss the inflected A and S SUBJS as ERG, and the uninflected S SUBJS as NOM

the following subsections.

3.2.1 Intransitivity Based Split

In Assamese, the subjects (A) of sentences with two or three place predicates obligatorily takes an ergative case enclitic. However, the language exhibits a split in terms of one place predicates, whereby the (S_a) subjects of unergative verbs behave like an A, and take obligatory ERGATIVE case ending, as demonstrated through Table 3.1.

SUBJ	GF	NP	ANIM	OBJ	GF	INANIM	OBJ	GF
A		-ERG	-ACC/DAT			-∅		
S _a		-ERG	NA			NA		
S _o		-∅	NA			NA		

Table 3.1: Intransitivity Based Split in Assamese

The A arguments of (di)transitive verbs always take an ERGATIVE case ending, as highlighted by the subject NPs *xikk^hjoetri* ‘teacher’ of the ditransitive verb *de* ‘give’, transitive verb *tan* ‘pull’ in (49b). Similarly, the (S_a) subject for the unergative verb *nas* ‘dance’ in example (49c) also gets obligatorily marked, adhering to the verb semantics-based split observed in the language.

- (49) a. *xikk^hjoetri-goraki=e randhoni-zon=ok kolom e-dal d-i*
 teacher.F.HON=ERG cook.CLF=ACC pen one-CLF give-PROG
as-e
 be.PRES-3
 ‘The (female) teacher is giving a pen to the (male)cook.’
- b. *xikk^hjoetri-goraki=e randhoni-zon=ok tan-i as-e*
 teacher.F.HON=ERG cook.CLF=ACC pull-PROG be.PRES-3
 ‘The (female) teacher is pulling the (male)cook.’
- c. *xikk^hjoetri-goraki=e nas-i as-e*
 teacher.F.HON=ERG dance-PROG be.PRES-3
 ‘The (female) teacher is dancing.’

On the other hand, the patient-like (S_o) subjects of unaccusative verbs remain unmarked, as illustrated by the subject NP *xikk^hjoetri* ‘teacher’ of the unaccusative verb *boh* ‘sit’ in example (50)

- (50) *xikk^hjoetri-goraki.∅ boh-i as-e*
 teacher.F.HON.NOM sit-PROG be.PRES-3
 ‘The teacher is sitting.’

3.2.2 Person and Number Based Split

Along with the verb semantics based split, Assamese also displays a PERSON and NUMBER based split in its case marking system. Within this split, the 1.SG, 2.SG, 3.SG and 1.PL pronominal subjects of both di(transitive) and unergative verbs remain unmarked for case. However, the 2.PL and 3.PL pronouns require obligatory ERGATIVE marking. This PERSON and NUMBER based split is illustrated in Table 3.2, highlighting the distinctions in case marking for different pronouns in the language.

PERSON	SG.NUMBER	PL.NUMBER
1	-∅	-∅
2	-∅	-e
3	-∅	-e

Table 3.2: PERSON and NUMBER Based Split in Assamese

As we can see in examples (51), the 2.PL and 3.PL pronominal subjects for the transitive verb *thel* ‘push’ take an ERGATIVE ending while the 1.SG, 2.SG and 3.SG and 1.PL subjects of the same verbs remain unmarked. These examples also highlight the phenomenon of the Differential Object Marking (DOM) in Assamese, as seen through the contrasts in marking for the personal pronouns *tumaluk* ‘2.PL’ and *muk* ‘1.SG’ in the object position and the inanimate object *soki* ‘chair’.

- (51) a. *moi.∅ tumaluk=ok thel-i as-u*
 1.SG.NOM 2.PL=ACC push-PROG be.PRES-1
 ‘I pushing you (PL).’
- b. *ami.∅ tumaluk=ok thel-i as-u*
 1.PL.NOM 2.PL=ACC push-PROG be.PRES-1
 ‘We are pushing you (PL).’
- c. *tumi.∅ muk thel-i as-a*
 2.SG.NOM 1.SG.ACC push-PROG be.PRES-2
 ‘you are pushing me.’
- d. *tumaluk=e muk thel-i as-a*
 2.PL=ERG 1.SG.ACC push-PROG be.PRES-2
 ‘you (PL) ARE PUSHING ME.’
- e. *tai.∅ soki-khon thel-i as-a*
 3.SG.NOM chair-CLF push-PROG be.PRES-3
 ‘She is pushing the chair.’

- f. xihot=e soki-khon thel-i as-a
 3.PL=ERG chair-CLF push-PROG be.PRES-3

‘They (PL) are pushing the chair.’

It must be noted here that the intransitivity-based split takes supremacy over the pronominal split for di(transitive) and unergative verbs. Pronominal subjects for both A and S_a take zero-marking, unless they refer to 2.PL and 3.PL pronouns, as highlighted through examples (52a-d). In contrast to this, pronominal subjects of unaccusative verbs consistently remain unmarked, even for 2.PL and 3.PL pronouns, as shown in examples (52e) and (52f).

- (52) a. xik_kh_jok-goraki=e satro-satri-bur=ok uttor-bohi-bur di
 teacher.hon=erg student-pl=acc answer-scripts give-prog
 as-e
 be.pres-3

‘The teacher is giving the answer-scripts to the students.’

- b. moi.∅ tumaluk=ok uttor-bohi-bur di as-u
 1.SG.NOM 2.PL-ACC answer-scripts give-PROG be.PRES-1

‘I am giving the answer scripts to you (PL).’

- c. suwali-zoni=e kãh-is-e
 girl-CLF=ERG cough-PERF-3

‘The girl has coughed.’

- d. tai.∅ kãh-is-e
 3.F.NOM cough-PERF-3

‘She has caughed.’

- e. tumaluk.∅ boh-is-a
 2.PL.NOM sit-PERF-2

‘You (PL) HAVE SAT.’

- f. xihot.∅ boh-is-e
 3.PL.NOM sit-PERF-3

‘They have sat.’

From the preceding discussion, it becomes evident that a language can display multiple types of splits, as is notably exemplified by Assamese. Figure 3.1, as introduced in Chapter 2 of this thesis, provides an effective model of the ergative splits in Assamese, namely the intransitivity and PERSON and NUMBER-based split. The primary objective of this chapter is to put this model to test and investigate how the child and young adult participants have acquired the intricate ergative splits in Assamese.

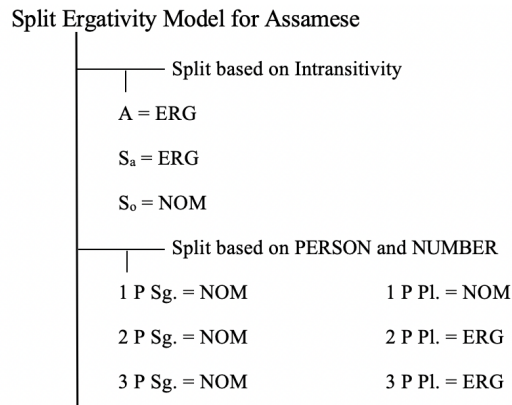


Figure 3.1: Split Ergative Case Model for Assamese

3.3 Earlier Studies on Case Acquisition by Children

The majority of previous research on the acquisition of case has primarily focused on NOMINATIVE-ACCUSATIVE languages, in which both transitive and intransitive subjects are in nominative case, while direct objects exhibit accusative case (van Valin Jr, 1990; Bavin and Stoll, 2013). Consequently, most generalizations and findings about case acquisition have been based on these NOMINATIVE-ACCUSATIVE systems, and challenges emerge when attempting to apply these findings to other languages.

In particular, there has been considerably less research on ergative languages, where transitive subjects are marked with the ERGATIVE case, while intransitive subjects and direct objects carry the absolutive case (Bavin and Stoll, 2013). Furthermore, some languages feature split systems, encompassing both NOMINATIVE-ACCUSATIVE and ERGATIVE-ABSOLUTIVE marking. Despite the added complexities associated with ERGATIVE-ABSOLUTIVE and split ergative languages, recent studies have demonstrated that children display evidence of rule comprehension and produce adult-like utterances from an early age (Allen and Crago, 1996; Narasimhan et al., 2005; Narasimhan, 2013; Paareek, 2018).

Narasimhan et al. (2005); Narasimhan (2013) investigates the acquisition of Hindi ergativity in a longitudinal case study involving three children aged between 1;7 to 3;9 years. Hindi has an aspectual-based split ergativity, which means that ERGATIVE marking is triggered only within the perfective aspect. However, there are other complexities involved with it, such as ‘psych-predicates’ denoting mental states get DAT-marked even in the perfective aspect, argument ellipses, and so on. Despite these syntactic-semantic restrictions, Narasimhan finds that children acquire the case system at an early age. Moreover, they do not overextend the ergative enclitic *-ne* on non-ergative arguments.

A similar pattern is observed in Pareek (2018), (also see Pareek et al. (2016)), which studies the acquisition of case and agreement in Hindi. Her findings concur with Narasimhan (2005, 2013) showing that none of the twenty-one child participants (age 1;11 to 5;11) violates agreement when the subject is in *ERGATIVE* case.

While relatively few in number, some of these studies have a specific focus on the phenomenon of children-led linguistic change in the domain of *ERGATIVE* case-marking. One notable study conducted by Allen and Crago (1996) finds evidence of an ongoing shift from ergative to accusative alignment in Inuktitut, reflected on the speech of the children. Using naturalistic data, Allen examined the use of *ERGATIVE* case morphology in four children (aged 2;0 to 3;6). In Inuktitut, *ERGATIVE* case marking is primarily applied to 3rd *PERSON* nominals and demonstratives of A-arguments for transitive verbs, resulting in only a few instances of obligatory marking in the data from the two-year-olds. Despite this inherent challenge, her overall findings align with what has been previously observed in other languages, namely that children do not overextend *ERGATIVE* affixes in *ABSOLUTIVE* contexts, and vice versa.

Mahalingappa (2013) investigates the acquisition of split ergativity in Kuramanji Kurdish, another language that is undergoing shifts in its case morphology. Kuramanji Kurdish has a tense-based *ERGATIVE* split where the arguments get *ERGATIVE-ABSOLUTIVE* marking in past tense and *NOMINATIVE-ACCUSATIVE* case marking in the present tense. As this language is changing from an *ERGATIVE-ABSOLUTIVE* language to a *NOMINATIVE-ACCUSATIVE* language, the children receive inconsistent inputs from the adult speakers. The research consists of twelve children (age 1;6 to 3;6) and suggests that in spite of the varied and inconsistent inputs from caregivers, children seem to acquire the *ERGATIVE* case marker as early as the age of 2;0.

Acquiring a language that is ongoing a change in progress is not the same as acquiring a language with a stable grammatical system as children receive mixed inputs from the adult speakers (Sankoff, 1990; Cheshire et al., 2011; Kerswill and Williams, 2000; O’Shannessy, 2013). Nevertheless, research has demonstrated that children are adept at processing the surface cues in their language input and could push the changes even further (Sankoff, 1990).

Research on the acquisition of split ergative systems within the Indo-Aryan language family has been relatively limited, except for studies conducted on Hindi (Narasimhan et al., 2005; Narasimhan, 2013; Pareek et al., 2016; Pareek, 2018) and Urdu Saleemi (1995). Assamese, on the other hand, employs a unique split ergative system based on *PERSON*, *NUMBER*, and verb semantics, distinguishing it from most other Indo-Aryan languages that mainly utilize an aspect-based split. Notably, the *ERGATIVE* case marking in Assamese does not block subject agreement, which is in contrast to Hindi and Urdu. This

thesis will explore how children navigate the complexities of such splits and to what extent they are capable of producing adult-like utterances at an early age.

3.4 Methods

Language acquisition researchers primarily resort to naturalistic samples, semi-structured elicitation techniques, and elicited production experiments for collecting language production data (Blom and Unsworth, 2010; Podesva and Sharma, 2014; Hoff, 2013; Lust and Blume, 2016). Spontaneous or naturalistic data can be collected from any category of learners, independent of their age, or ability. Although such data is considered to show fewer observer effects and methodological artefacts than other methods, this methodology is not without issues. One of the major drawbacks of naturalistic data is the considerable amount of incomparable data. In such situations, production experiments and semi-structured tasks/games where researchers specifically make use of techniques aimed at collecting the targeted data in a somewhat controlled environment is deemed more useful. Semi-structured techniques, although controlled to a certain extent, allows flexibility, thereby keeping the recording environment natural to a great extent. Production experiments, on the other hand, make use of ‘standardised procedures and stimuli’ which minimises the elicitation of irrelevant utterances and no responses (Eisenbeiss, 2010).

Depending on the kind of linguistic data required for the given study, the age and/or cognitive ability of the participants, researchers might have to prioritise the use of one production method over another, or incorporate more than one technique in a given study. In order to investigate the use of overt case morphology in Assamese for the current study, I made use of two tasks: The Contrastive elicitation Task and Final Destination Board Game.

The Contrastive Elicitation Task for Testing Case Marking² is an elicited-production experiment, which targets case marking (Ruigendijk, 2015). In this experiment, participants have to describe a set of minimally contrastive pictures that target the elicitation of case marking, mainly NOM-ACC, ERG-ABS, DAT, and GEN.

The Final Destination Board Game, on the other hand, was developed to elicit subject-verb agreement and NP-internal agreement among Kuwaiti Arabic children (Al-Houti, 2013). It makes use of stimulus pictures of the participants themselves conducting a variety of actions to target the use of utterances consisting of verb+noun+adjective(s) constructions.

²This experiment was developed as part of the cross-linguistic COST Action IS0804 “Language Impairment in a Multilingual Society: Linguistic Patterns and the road to Assessment” (<http://bi-sli.org/>)

While the Contrastive Case Elicitation Task allows us to compare case marking on different argument types (transitive and intransitive subjects, direct and indirect objects), an additional task was required to capture the complexity of the Assamese case system, especially the PERSON and NUMBER-based split. Hence, the Final Destination Board Game was included in the study to elicit case marking for 1st and 2nd PERSON contexts.

3.4.1 Participants

For the Contrastive Case Elicitation Task, forty child participants were recruited from the Tinsukia and Dibrugarh Districts of eastern Assam. The participants, who are all native speakers of Assamese, were aged between 2;6 and 5;9 (mean age - 4.50, SD - 0.79) at the time of the study. The primary rationale for selecting this age group for both studies is rooted in the recognition that the initial years of a child's life are pivotal for language acquisition, a concept underscored in prior research (Bates et al., 1992; Allen and Crago, 1996; Narasimhan et al., 2005; Narasimhan, 2013; Friedmann and Rusou, 2015). Additionally, considering Assam's status as a highly multilingual society, the choice of a young age group serves to minimize exposure to other languages and maintain linguistic focus.

The study included a total of forty child participants, consisting of eighteen males and twenty-two females. Among these children, twenty-three lived in rural areas, while the remaining seventeen came from urban settings. Given India's status as a highly multilingual country with a three-language policy implemented in the education system, it was not possible to recruit monolingual children for this study. As a result, only eight out of the forty child participants were monolingual, with the rest being either bilingual or multilingual, speaking up to four languages in some cases.

Interestingly, despite the majority of caregivers primarily using Assamese for interaction with the children, most of the children appeared to have acquired other languages, such as Hindi, entirely through watching cartoons on TV. In addition to Hindi, children also spoke English, Bengali, and Sadri to certain extent (refer to Appendix D for a detailed overview).

The control group consisted of twenty-two adults, consisting of ten males and twelve females, ranging in age from 16 to 25 years. They were either studying at a college or university level program at the time of the study. Thirteen of them lived in a rural area, while the remaining nine lived in an urban location. It's worth noting that all of the adult participants were multilingual.

Additionally, thirty-five children (mean age - 4.55, SD - 0.86), who had previously participated in the production experiment, also took part in the semi-structured elicitation task (The Final Destination Board Game). This

task necessitated a collection of pictures depicting children and other individuals enacting various scenarios to facilitate the elicitation of person and number-based case markings. Therefore, scheduling this task after the contrastive case task provided the opportunity to capture the required images during the initial data collection visit.

All the participants in this study were recruited through personal contacts and from a school in a village in Tinsukia District. The table 3.3 below provides an overview of the study’s participants. In the Appendix, you can find a pseudonymized list of both child and adult participants, along with additional background information such as the number of languages spoken and understood, the language used with immediate family, friends, and neighbours, the medium of instruction in their educational institute for those already in schools or pursuing higher education, and more.

Production Method	Group	No. of Participants	Mean Age	Gender	Place of Residence
Experiment	Child	40	4.46	Male = 18 Female = 22	Rural = 23 Urban = 17
Experiment	Adult	22	19.04	Male = 10 Female = 12	Rural = 13 Urban = 9
Semi-structured Elicitation Task	Child	35	4.55	Male = 16 Female = 19	Rural = 18 Urban = 17

Table 3.3: Summary of Participants

3.4.2 Data Collection

The required cross-sectional data was collected during three field trips undertaken in Dibrugarh and Tinsukia districts of eastern Assam. The first field trip was carried out in May 2016 during which a pilot study of the methods was conducted, and based on which the stimuli were further modified. The second field trip was undertaken from February 2017 to May 2017, when the majority of the data was collected for this research. The final phase of the data was collected in December 2017 and January 2018.

Approval from the Ethics Committee of the researcher’s university was obtained before the start of the fieldwork. A signed consent was obtained from the adult participants of the experiment, while in the case of the child participants, the form was signed by their primary caregiver when recordings were done at home, or by the school headmaster when the recordings were conducted at a school.

Before each session, the child participants were encouraged to engage in the setting up of the recording area. This helped the children to familiarise themselves with the recording equipment, as well as the environment, even when the recordings took place at their own homes. At the end of the final session, each child was given a gift that consisted of drawing books, colour pencils, a pencil, an eraser, a pencil sharpener, and a ruler. All the adult participants of the experiment were given a pen to thank them for their time.

3.4.3 Equipment Used

All the sessions were video recorded using a Panasonic HDC-SD900 camcorder. For the audio recordings a Zoom H1 recorder connected to a Sennheiser microphone via a 3-pin XLR cable was used. Both the camcorder and the microphone were mounted on a tripod and a stand throughout the recordings. For the pictures for the Final Destination Game, a OnePlus 3T 16 mega-pixels mobile phone camera was used. These pictures were then printed out from a Canon Pixma printer at home. In some situations, where the data was collected from areas where electricity was not available to print out the pictures, a Fujifilm Instax Mini 8 Instant camera was used.

3.4.4 Procedure of Data Transcription and Analysis

The recordings were transcribed using the ELAN Linguistic Annotator (version 5.0.0)³. The CHAT transcription format of CHILDES MacWhinney (2000) was used for annotating the data on ELAN. Given the time-constraints, only the utterances of the child-participant were transcribed for both the experiment and the semi-structured task. Moreover, when the participants described the same picture more than once, only their final utterance was considered even if the earlier ones were more target-like.

The target utterances of each participant were then exported from ELAN to Excel and were coded focusing on the verbs, subjects and objects. The CASE marking of both the subjects and objects (direct and indirect) were coded on the basis of five parameters: obligatory use, omissions in case of obligatory overt contexts, null realisations as is the case with the nominative marking on subjects, and inanimate nouns in objects, and whether the target case marker was replaced by a different one. For the semi-structured board game, variation used for the pronouns were also coded separately.

Finally, the data from both these tasks were analysed using different statistical methods, such as chi-square and general linear model in either in SPSS

³Developed by Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands (Wittenburg et al. 2006, url:<http://tla.mpi.nl/tools/tla-tools/elan/>)

or RStudio.

3.5 Study 1: Contrastive Case Elicitation Task

3.5.1 Materials

While designing the stimuli for this task, several conditions, such as verb types, PERSON and NUMBER values, animacy and GENDER⁴ were taken into account. Table 3.4 provides an overview of these conditions.

Item-Type	Condition
Verb Type	Unaccusative Unergative Highly transitive Ditransitive
Person Subject	3P
Person Direct Object	3P
Person Indirect Object	3P
Number Subject	Sg
Number Direct Object	Sg
Number Indirect Object	Sg
Animacy Subject	Animate 2 animate and 2 inanimate for unaccusative verbs
Animacy Direct Object	Inanimate
Animacy Indirect Object	Animate
Gender Subject	Male and Female
Gender Indirect Object	Male and Female
Gender Direct Object	Male and Female

Table 3.4: Stimuli Design for Contrastive Case Elicitation Task

A total of eleven pairs of pictures depicting actions for ditransitive *give*, highly transitive *pull* and *push*, unergative *jump*, *dance*, *swim*, and *run* and unaccusative *fall*, *sit*, *burn*, and *sink* were included in the revised set of pictures. The stimuli consisted of one pair of pictures for the ditransitive verb and two pairs of pictures for the transitive verbs. Further, four pictures pairs each for unergative and unaccusative verbs were included in the design in order to understand how children acquire this complex verb semantics based ergative split in Assamese.

⁴It's important to note that Assamese lacks grammatical gender and does not exhibit gender agreement on any other parts of the sentence. The inclusion of stimuli indicating both masculine and feminine gender in this experiment was primarily done for the purpose of enabling future comparisons with a language possessing grammatical gender, such as Hindi.

Item Number	Condition	Item	Target Sentences in English
1	SOOV	Give	The male-teacher is giving a book to the male-cook. The female-teacher is giving a pen to the female-cook.
2	SOV	Pull	The washerman is pulling a chair. The washerwoman is pulling a bed.
3	SOV	Push	The male-doctor is pushing an old lady. The female-doctor is pushing the hunter.
4	S _a	Jump	The farmer is jumping. The reaper is jumping.
5	S _a	Dance	The male-tailor is dancing. The female-tailor is dancing.
6	S _a	Swim	The king is swimming. The queen is swimming.
7	S _a	Run	The milkman is running. The milkwoman is running.
8	S _o	Fall	The male-pilot is falling. The female-pilot is falling.
9	S _o	Sit	The cricket-player is sitting here. The female-nurse is sitting here.
10	S _o	Burn	The paper is burning. The candle is burning.
11	S _o	Sink	The box is sinking. The boat is sinking.

Table 3.5: Non-randomised Item List of The Contrastive Case Elicitation Task

After finalizing the item list, highlighted in Table 3.5, an artist in India was commissioned to create the pictures. The artist received specific instructions to adhere to the same simplistic cartoon pattern used in the stimuli of Ruigendijk (2015) while making sure that the resulting images would be culturally appropriate for Indian children. A sample of these images is provided in Figure 3.2. The final set of pictures was then printed and laminated to prevent any damage during the experiment.

3.5.2 Procedure

At the start of each session, the pictures were arranged in a random order that was strictly maintained for each participant. While placing the contrasting picture pair in front of the participant, the researcher always prompted the verb as *ei photo-khon diyaa bixoye* ‘this picture is about giving’ while being mindful not to use prompts that contained any case inflections. The researcher then encouraged the participants to identify the characters in the images. In some situations, when the children were not aware of the lexical term used for a particular occupation represented through a character, the researcher prompted the term without using the case marker as *ei-zon randhoni* ‘this is a

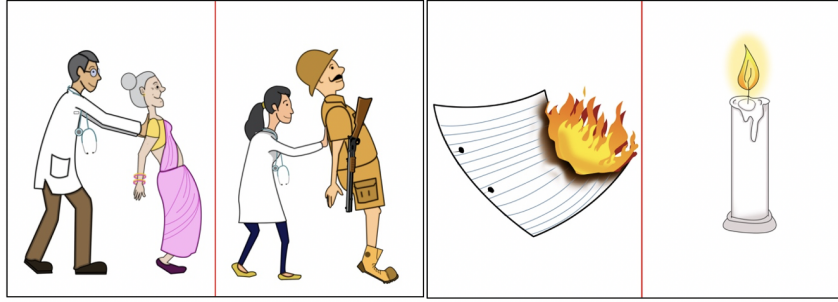


Figure 3.2: Contrastive pictures for transitive ‘push’ and unaccusative ‘burn’

cook’. Once the child was aware of the lexical item for all the characters and objects in that picture, the researcher would then encourage the children to describe what was going on in the picture by saying *etiya kuwasun eyat ki hoi ase* ‘now tell me what is happening here’. If the children did not answer in complete sentences, especially for the ditransitive and the transitive verbs, they were requested to repeat it again by pointing out that they left out a character or and object. Once the participant described the event in the pictures, the researcher would then take those away and place the next pair of pictures in front of the child and follow the exact same narration and procedure until all the pictures were described. Each child participant took approximately 30-40 minutes to complete the task, while adult participants completed the same task in approximately 10 minutes.

3.5.3 Data Analysis

A total of 880 utterances were elicited from forty child participants, out of which 869 were full target-like utterances. There were 9 fragment sentences and 2 irrelevant utterances, all of which were produced by two two-year old participants. Also, as the recommended age for using this elicited production experiment is from 3 years onwards, it was decided that the 44 utterances collected from the two children aged 2;6 and 2;9 will be excluded from future analysis and will be discussed separately in a descriptive manner. The remaining data which consisted of 836 utterances from thirty-eight child participants had all analysable utterances as shown in Table 3.6. This table also gives us an overview of the data collected from the twenty-two adult participants.

Group	No. of Utterances	Irrelevant Utterances		Fragment Utterances		Target-like Complete Utterances		Non-target-like Utterances	
		n	%	n	%	n	%	n	%
Child	836	0	0	0	0	706	84.4	130	15.6
Adult	484	0	0	0	0	387	79.9	97	20.1

Table 3.6: Utterance Types Produced in the Contrastive Case Elicitation Task

3.5.4 Overt SUBJ NP Realisation in Child-data

Table 3.7 below shows the trends of SUBJ realisations for each age category, while a table with the individual trends of SUBJ realisations for each child participants is included in Appendix F. As depicted in Table 3.7, the children correctly produced the targeted SUBJ NP in 765 utterances with an accuracy rate of 89.3%.

Age Group	Target SUBJ NP in Assamese	Similar SUBJ NP in Assamese	SUBJ in English	SUBJ Other	Total
3-3;11	201	16	2	1	220
4-4;11	314	32	4	2	352
5-5;11	250	10	4	0	264
Total	765	58	10	3	836

Table 3.7: Age-wise Distribution of Overt SUBJ Realisation in Child-data

The English SUBJ NPs were used for two lexical items: *xikkhojetri*⁵ ‘female teacher’ and *bakos* ‘box’. In spite of using the English lexical terms, the children used the ERGATIVE case enclitic with *miss* and *ma’am* while describing the image for the ditransitive verb *de* ‘give’ as shown in (53). The English SUBJ NP ‘box’ in (54) produced for the unaccusative verb *dub* ‘sink’ was accurately NOM-marked:

(53) mis=e randhoni=k kolom d-i as-e
 teacher.F=ERG cook=ACC pen give-PROG be.PRES-3
 ‘The (female) teacher is giving a pen to the cook.’

(54) bōks-tu.Ø-tu dub-i as-e
 box-CLF.NOM sink-PROG be.PRES-3
 ‘The box is sinking.’

Furthermore, there were three SUBJ NPs that are coded as *Other* in the dataset. Two instances were when the participants used *k^heti* ‘farm’ instead the target Assamese NP *k^hetijok* ‘farmer’. Further, there is one use of the Hindi lexical term *raza* ‘king’ instead of the Assamese term *roza*. For both these test items, however, the children correctly produced the target-like utterances as shown in (55) and (56).

(55) *k^heti-tu=e zopia-i as-e
 farm-CLF=ERG jump-PROG be.PRES-3
 ‘The farm (intended: farmer) is jumping.’

⁵The more frequent colloquial terms ‘mastorni’ and ‘baidew’ were also coded as target-like usage of the SUBJ NP in Assamese.

- (56) raza=e xatur-i as-e
king=ERG swim-PROG be.PRES-3
‘The king is swimming.’

The children faced some difficulties while producing the lexical terms for some of the characters selected for this stimuli. For instance, there was an issue with the “over-adaptation” effect with respect to the occupation of the farmer in the rural areas. Children were not familiar with the term *k^hetiyok* ‘farmer’ as farming is very common in the rural areas of Assam, and the children are used to identifying or referring to these people with a kinship term instead of their profession. Similarly, the children recognised some of the professions depicted in the pictures with what they do, for example someone who sews clothes, but were not aware of the lexical term for that profession.

3.5.5 Overt OBJ NP Realisation in Child-data

The Table 3.8 below shows how all the direct objects (DO)s in the ditransitive and intransitives sentences were produced by the children.

Age Group	Target SUBJ NP in Assamese	Similar SUBJ NP in Assamese	SUBJ in English	SUBJ Other	Total
3-3;11	47	11	2	0	60
4-4;11	79	15	2	0	96
5-5;11	56	12	4	0	72
Total	182	38	8	0	228

Table 3.8: Age-wise Distribution of Overt OBJ Realisation in Child-data

There were 38 instances where the participants used a kinship term to refer to one of the DO characters in the experiment. Except one utterance, which is presented in (57), all other participants used the kinship term *aita* ‘grandmother’ instead of the targeted lexical item *burhi-manuh* ‘old woman’ to describe the image for the transitive verb *thel* ‘push’. Similar observations were made in the adult data, where only one participant used the targeted lexical item.

- (57) doktor=e buri=k thel-i as-e
doctor=ERG old-woman=ACC push-PROG be.PRES-3
‘The doctor is pushing the old woman.’ (KD; age 4;7)

The English lexical terms were mostly used for the DO *kolom* ‘pen’ of ditransitive *de* ‘give’. Apart from this, there was one utterance with the term ‘cook’ instead of its targeted Assamese item *randhoni*.

- (58) a. sir=e cook=ok book di ase
 teacher.M=ERG cook=ACC book give-PROG be.PRES-3
 ‘The teacher is giving a book to the cook.’ (AB, 5;8)
- b. mis=e pen di ase randhoni=k
 teacher.F=ERG pen give-PROG be.PRES-3 cook=ACC
 ‘The teacher is giving a pen to the cook.’

3.5.6 Overt IOBJ NP Realisation in Child-data

There were 4 omissions of the indirect object in the data, which occurred for the verb *de* ‘give’. However, all these omissions happened in the utterances of the two-year old children whose data is excluded from the final statistical analysis. As shown in the data of the remaining participants in Table 3.9, out of the 76 contexts for indirect objects, there were 3 instances where these children used an english lexical item instead of the Assamese targeted NPs.

Age Group	Target SUBJ NP in Assamese	Similar SUBJ NP in Assamese	SUBJ in English	SUBJ Other	Total
3-3;11	18	0	2	0	20
4-4;11	32	0	0	0	32
5-5;11	23	0	1	0	24
Total	73	0	3	0	76

Table 3.9: Age-wise Distribution of Overt IOBJ Realisation in Child-data

3.5.7 Results and Discussion

The children produced 706 target-like utterances and 130 non-target-like utterances in total. As highlighted in Table 3.10, there were only 9 under-extension/undersupply errors on the SUBJ NP in unergative verbs as opposed to 121 over-extension/oversupply errors in the domain of unaccusative verbs. However, when we look at the adult data from the same task, summarised in Table 3.11, we find that there are 57 errors in terms of under-extension of the ERG case marker in the unergative verbs, and 40 over-extension of the ERG in the unaccusative verbs. Both the child and adult participants did not produce any errors for the SUBJ NPs of the (di)transitive verbs.

3.5.8 Results for Subjects

The 121 non-target-like utterances in unaccusative verbs were due to the over-extension of the ergative case enclitic on S_o SUBJs as seen in example (59). Given the intransitivity-based split in Assamese, these S_o SUBJs *pilot* ‘pilot’ and *baido* ‘elder sister’ should have remain unmarked.

Group	Data	SOOV	SOV	S _a	S _o	Total
Child	Target-like Utterances	76	152	295	183	706
Child	Non-target-like utterances	0	0	9	121	130
	Total	76	152	304	304	836

Table 3.10: Summary of the Full Utterances of the Child Participants

Group	Data	SOOV	SOV	S _a	S _o	Total
Adult	Target-like Utterances	44	88	119	136	387
Adult	Non-target-like utterances	0	0	57	40	97
	Total	44	88	176	176	484

Table 3.11: Summary of the Full Utterances of the Adult Participants

- (59) a. *pilot-tu=e por-i as-e
pilot-CLF=ERG fall-PROG be.PRES-3
‘The pilot is falling.’
- b. *baido-zoni=e boh-i as-e
sister.elder-CLF=ERG fall-PROG be.PRES-3
‘Elder sister is sitting.’ [Intended: The nurse is sitting] (NST, 4;5)

Moreover, out of the 532 obligatory contexts for ERG-marking, as highlighted in Table 3.12, the children produced 523 of them (refer to target-like utterances in Table 3.10), which has an accuracy of 98.3%. However, out of the 304 obligatory contexts for null case realizations, only 183 (60.1%) were in nominative case (refer to non-target-like utterances in Table 3.10). Adults, on the other hand, under-extended the ERG case enclitic in 57 obligatory contexts. Interestingly, the case-marking patterns in the child participants seems to be relatively consistent as opposed to that of the young adults that participated in this study.

Group	Obligatory Context for ERG-marking	ERG Omissions in Obligatory Context	Obligatory Context for NOM-marking	NOM Omissions in Obligatory Context
Child	532	9	304	121
Adult	308	57	176	40

Table 3.12: Summary of ERG and NOM Omissions in Obligatory Contexts

In unaccusative verb category, there are 304 obligatory contexts for NOM-marking or zero-marked subjects, however we find that the children used an ergative case ending 121 times, when it should have been left unmarked. As discussed in the methods section earlier, the S_o SUBJs in this experiment were equally divided into animates and inanimates. Interestingly, the highest

number of non-target-like utterances (n=119) were produced for the animate unaccusative SUBJs. There were only two instances where the participant used the ergative case enclitic for an inanimate S_o SUBJ. The utterances are provided below:

- (60) a. *bakos=e dub-i as-e
 box=ERG sink-PROG be.PRES-3
 ‘The box is sinking.’
- b. *nao=e dub-i as-e
 boat=ERG sink-PROG be.PRES-3
 ‘The boat is sinking.’ (CM, 4;1)

There was a total of 152 animate S_o SUBJ in the full target-like utterances of the children, out of which 121 were marked with an ergative = 76.1%. The same phenomenon was observed in only 22.15% of the adult data. All the ergative omissions in the obligatory contexts for the adult participants were in the unergative subjects. The adults were consistent with overtly marking the SUBJs of the ditransitive and transitive verbs and had no omissions. However, they had 57 instances of zero-marked S_a SUBJs, as shown in example (61) below.

- (61) a. *k^hetijok-zon.∅ zopia-i as-e
 farmer-CLF.NOM jumpPROG be.PRES-3
 ‘The farmer is jumping.’
- b. *dawoni-zoni.∅ zopia-i as-e
 reaper-CLF.NOM jumpPROG be.PRES-3
 ‘The reaper is jumping.’

Table 3.13 displays the observed and expected counts, revealing that child participants exhibited improved proficiency in producing target-like utterances as they advanced in age. Additionally, it’s noteworthy that all the children under-extended the ERGATIVE case marker to a much lesser extent than expected, in contrast to the adult participants.

The data from this table is also visually represented in Figure 3.3, which displays the error types across different age groups, clearly highlighting the stark contrast between adult and child data.

3.5.9 Results for Direct Objects

Assamese exhibits a differential object marking (DOM), whereby the (personal) pronouns and proper names are obligatorily marked with an accusative case enclitic, while [+ human] common nouns and [- human] animate common nouns are optionally marked. The inanimate nouns, on the other hand, remain

Age Group		3-4	4-5	5-6	16-25	Total
Case-marking Accurate	Observed Count	185	288	233	387	1093
	Expected Count	182.2	291.5	218.6	400.8	1093
Case-marking Underextended	Observed Count	5	4	0	57	66
	Expected Count	11.0	17.6	13.2	24.2	66
Case-marking Overextended	Observed Count	30	60	31	40	161
	Expected Count	26.8	42.9	32.2	59	161
Total	Observed Count	220	352	264	484	1320
	Expected Count	220	352	264	484	1320

Table 3.13: Observed and Expected Counts of Case Elicitation Task Data

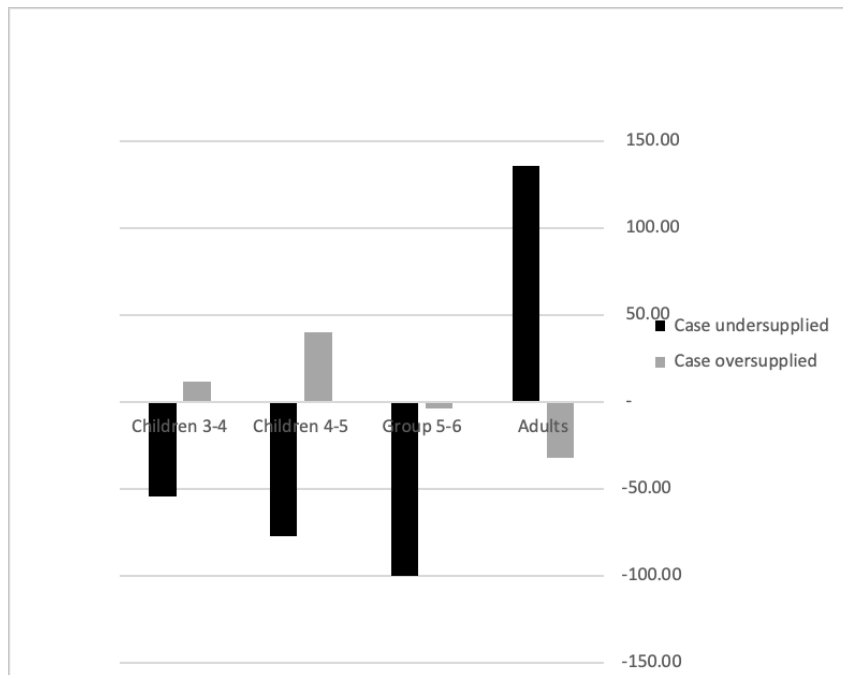


Figure 3.3: Case Elicitation Task: Error Types by Age Group

unmarked. Owing to this hierarchy, only the contrastive pictures for the transitive verb *thel* ‘push’ triggered obligatory ACC marking. The other contrastive pictures for the transitive verb *tan* ‘pull’ had an inanimate objects *soki* ‘chair’ and *bisona* ‘bed’ in its direct object positions. Similarly, the ditransitive verb *de* ‘give’ had two inanimate items as direct objects, namely *kitab* ‘book’ and

kolom ‘pen’. Given the animacy hierarchy, these verbs did not trigger an ACC-marking. As we will see in the discussion below, there were no variation in the data for both direct and indirect objects, due to which we could not conduct any statistical analysis.

There were 152 instances for obligatorily ABS-marking for the inanimate direct objects in the child data, as shown in example (62). The children did not overextend the ACC-marking to the inanimate objects, and produced adult-like utterances with null-marked objects in all the obligatory contexts. Similarly, the adult participants did not over-extend any case markers for the 88 obligatory ABS-marked contexts.

- (62) a. *sir=e randhoni=k kitap di as-e*
 teacher.M=ERG cook=ACC book give-PROG be.PRES-3
 ‘The (male) teacher is giving a book to the cook.’ (KD; age 4;7)
- b. *dhuba=e soki tan-i as-e*
 washerman=ERG chair pull-PROG be.PRES-3
 ‘The washerman is pulling a chair.’ (AB, 5;8)
- c. *dhubuni-e bisona tani ase*
 washerwoman=ERG bed pull-PROG be.PRES-3
 ‘The washerwoman is pulling a bed.’ (MP, 5;9)
- d. *dhubuni-e bisona tani ase*
 washerwoman=ERG bed pull-PROG be.PRES-3
 ‘The washerwoman is pulling a bed.’ (MP, 5;9)

Furthermore, the children as well as the adult participants did not omit the obligatory ACC-marker for the 76 instances in child data and 44 instances in adult data, respectively. The examples in (63) highlights the obligatory ACC-marking for the transitive verb *thel* ‘push’ and its correct usage in the child data.

- (63) a. *doktor=e sikari=k thel-i as-e*
 doctor=ERG hunter=ACC push-PROG be.PRES-3
 ‘The doctor is pushing the hunter.’ (AB, 5;8)
- b. *doctor=e aita=k thel-i ase*
 doctor=ERG grandmother=ACC push-PROG be.PRES-3
 ‘The doctor is pushing grandmother.’ (MP, 5;9)

3.5.10 Results for Indirect Objects

There were 76 contexts for obligatory *-(o)k*-marked object thetas in the ditransitive test item *de* ‘give’ in the child data as opposed to 44 obligatory contexts in the adult data. The contrastive images of this verb had two [+animate] [+human] NPs, vis a vis, a ‘male cook’ and a ‘female cook’, in the indirect object position. Since, the object thetas in Assamese also adhere to the DOM principles in the language, both these NPs triggered the ACC/DAT enclitic as highlighted in examples (64)

- (64) randhoni=k mastorni=e kolom d-i as-e
 cook=ACC teacher.f=ERG pen give-PROG be.PRES-3
 ‘The (female) teacher is giving a pen to the cook.’ (KD; age 4;7)

Although there were no omissions in the obligatory context, there were 4 instances where the children ended up using the ERG case marker instead of the ACC/DAT marker in their utterances as shown in example (65). However, it is likely that the children could not comprehend who was performing the action of giving and ended up assigning the ERG case marker to both the animate NPs. The adult participants, on the other hand, made no errors in their production of the ACC/DAT case marker in obligatory contexts.

- (65) *sir=e randhoni=e kitap di ase
 teacher.m=ERG cook=ERG book give-PROG be.PRES-3
 ‘The (male) teacher/cook is giving a book.’ (PC, age 3;8)

The oblique objects of the ditransitive verbs take the GEN-marking in Assamese. However, the stimuli for this task did not account for oblique objects, hence we do not have any data to show the usage of this particular case marker.

3.5.10.1 Development of Split Ergative Case Marking Among Two-year Olds

It was noted earlier that the elicited production experiment is typically recommended for use from the age of 3 years onwards. Nevertheless, I managed to collect some data from two children aged 2;6 and 2;9 using this task. However, the data collection process with these two-year-olds presented several challenges. The participants were easily distracted and prone to boredom, leading to frequent breaks during the task. Additionally, the two-year-olds were not familiar with the specific lexical terms for most of the professions depicted in the images. Consequently, they resorted to using more generic kinship terms, such as *koka* ‘grandfather’, *aita* ‘grandmother’, *dada* ‘elder brother’, *baa* ‘elder sister’ and so on to describe all the images instead of the target SUBJ NP in Assamese. Despite these limitations, I continued with the task to obtain some data from two-year-olds, as I was unable to collect any naturalistic data from younger child participants to examine the use of the ergative case marker.

Out of the 44 utterances produced by the two two-year olds, only 6 contained the target SUBJ NP, while 26 of the utterances consisted of kinship terms. There were 24 obligatory contexts for the use of the ERG case marker, out of which 8 were marked. On the other hand, out of 16 obligatory contexts for NOM marked subjects, there was one utterance where the ERG case marker was oversupplied as shown in example 66.

- (66) ba=e bohi-i as-e
 2.SG=ERG sit-PROG be.PRES-3
 ‘(Elder) sister is sitting.’

Furthermore, there was one instance where the participant produced the 2.SG pronoun TUMI instead of the targeted NP *gwwaloni* ‘milk-woman’ for the transitive verb run. Since *dour* ‘run’ is an unergative verb, the subject would have otherwise triggered the ergative enclitic *-e*. However, given the PERSON and NUMBER-based split in Assamese, the 2nd PERSON singular pronouns remain unmarked. The participant maintained this split, as demonstrated in the example 67a.

- (67) a. tumi.∅ dour-i as-a
 2.SG.NOM run-PROG be.PRES-2
 ‘You are running.’
 b. guwaloni-zoni=e dour-i as-e
 milkwoman-CLF-ERG run-PROG be.PRES-3
 ‘The milkwoman is running.’

3.6 Study 2: Final Destination Board Game

Four verbs from the Contrastive Elicitation Task were chosen for the second study, namely The Final Destination Board Game (Al-Houti, 2013). The original board game had objects in different shapes and sizes to test adjective inflection and noun-phrase internal agreement (concord) in Kuwaiti Arabic. We, however, eliminated this parameter for the current study, and only focused on PERSON and NUMBER for both subjects and objects. This was necessary as the Contrastive Case Elicitation Task targeted the elicitation of the intransitivity-based split only in the 3rd PERSON.

While designing the stimuli for this board game, we decided to use the same verbs tested in the elicited production experiment, as this allowed for comparison of the same verb across tasks. Moreover, it was difficult to find a broad range of ditransitive and unaccusative verbs that can be acted out by people, particularly with the child participants. This was necessary as later, pictures had to be taken of them performing those actions to describe the scenario during the task. Hence, we decided to use the ditransitive *de* ‘to give’, the highly



Figure 3.4: The handmade felt castle for the Final Destination Board Game

transitive *lo* ‘to pull’, the unergative *nās* ‘to dance’, and the unaccusative *boh* ‘to sit’, to specifically elicit the PERSON and NUMBER-based ergative split in Assamese. All the scenarios were then acted out with the child, the researcher and other known/unknown people to elicit 1st, 2nd, and 3rd PERSON in both singular and plural.

We created the required material for this board game out of felt so that it is lightweight, and yet durable, and can be easily carried around in the rural areas where transportation was an issue (refer to Figure 3.4). The felt castle had velcro hoops and loop tapes sewn into them which meant it could be disassembled into flat pieces for transportation convenience. The child participants were always invited to engage in building the castle which helped to put them at ease for the upcoming recording sessions.

This castle is then placed at the end of a road with several stoppages, all of which are assembled on a felt board. Three finger puppets were used for characters in the narrative to make the game more captivating. The pictures needed for description were taken while enacting four scenarios for different PERSON and NUMBER values as highlighted through the stimuli design in Table 3.14.

3.6.1 Procedure

For this semi-structured board game, the researcher first constructed the castle with the child and then introduced the character of a prince. The researcher then explained to the child that the prince needed help to rescue the princess who was locked up in the castle by the villain. However, to rescue the princess, the child must stop at four different stoppages along the way to the castle, and describe six pictures at each stoppage. The first picture is always of the child giving flowers to the researcher and another person. While showing this picture

Item Number	Verb Condition	Item	SUBJ PERS & NUM	OBJ PERS & NUM	IOBJ PERS & NUM
1	SOOV	de ‘to give’	1 SG	3 SG	2 PL
2	SOOV	de ‘to give’	1 PL	3 SG	3 SG
3	SOOV	de ‘to give’	2 SG	3 SG	3 PL
4	SOOV	de ‘to give’	2 PL	3 SG	1 SG
5	SOOV	de ‘to give’	3 SG	3 SG	1 PL
6	SOOV	de ‘to give’	3pl	3 SG	2 SG
7	SOV	tan ‘to pull’	1 SG	2 PL	NA
8	SOV	tan ‘to pull’	1 PL	3 SG	NA
9	SOV	tan ‘to pull’	2 SG	3 PL	NA
10	SOV	tan ‘to pull’	2 PL	1 SG	NA
11	SOV	tan ‘to pull’	3 SG	1 PL	NA
12	SOV	tan ‘to pull’	3 PL	2 SG	NA
13	S _a	nãs ‘to dance’	1 SG	NA	NA
14	S _a	nãs ‘to dance’	1 PL	NA	NA
15	S _a	nãs ‘to dance’	2 SG	NA	NA
16	S _a	nãs ‘to dance’	2 PL	NA	NA
17	S _a	nãs ‘to dance’	3 SG	NA	NA
18	S _a	nãs ‘to dance’	3 PL	NA	NA
19	S _o	boh ‘to sit’	1 SG	NA	NA
20	S _o	boh ‘to sit’	1 PL	NA	NA
21	S _o	boh ‘to sit’	2 SG	NA	NA
22	S _o	boh ‘to sit’	2 PL	NA	NA
23	S _o	boh ‘to sit’	3 SG	NA	NA
24	S _o	boh ‘to sit’	3 PL	NA	NA

Table 3.14: Stimuli Design for Final Destination Board Game

Obligatory Context for ERG-marking	Under-extension of ERG in Obligatory Context	Obligatory Context for NOM-marking	Over-extension of NOM in Obligatory Context	Total
106	17	429	34	535

Table 3.15: Summary of the Results from The Final Destination Game

to the child, the researcher asked *sini paisane eiya kun hoi?* ‘Can you recognise who this is?’. Following their identification of the people, the researcher would then ask the child to describe what was happening in the picture by prompting them with: *etiya kuwasun eyat ki hoi ase?* ‘Now tell (me) what is happening here?’ After the first set of six pictures visualising the ditransitive verb was described, the researcher would then move on to the next set of pictures. This game always ended with some free play once the final set of pictures were described, whereby the child was then allowed to have full control of the narrative and defeat the villain to rescue the princess according to their own plot.

3.6.2 Data Analysis

A total of 840 utterances were collected from the child participants in the age group of 3 to 6 years (mean age 4.55), using this semi-structured language game. After filtering out utterances that didn’t contain the target structure (n=41) or used a full noun phrase or proper name instead of a pronoun (n=264), we retained 535 utterances for the analysis presented in this chapter. The unusable data were primarily from the 3rd PERSON scenarios where the child used kinship terms instead of the targeted pronouns.

3.6.3 Results and Discussion

Out of the 535 utterances that were included for further analysis, 484 followed the canonical patterns of case marking usage found in adult language, while 51 were non-target-like utterances. In the analyzed data, there were 106 instances where the ERG case marker was required in an obligatory context, out of which 89 utterances were overtly marked. On the other hand, there were 429 instances in the overall data where the pronominal required obligatory zero-marking, but 34 utterances were over-extended with an ERG-marked subject as highlighted in Table 3.15:

Overall, there were 17 instances where the children under-extended the ERG case marker for the pronominal subjects of an unergative verb as shown in examples (68) where the 2nd PERSON plural pronoun *tumaluk* ‘you’ should have been ERG-marked given the PERSON and NUMBER-based split:

- (68) a. *tumaluk.∅ ze muk ze tan-is-a
 2.PL.NOM EM 2.SG.ACC EM pull-PERF-1
 ‘You have pulled me.’ [Lit: It is me that you have pulled.] (SD, 3;5)
- b. *tumaluk.∅ amak phul di asa
 2.PL.NOM 1.PL.ACC flower give-PROG be.PRES-2
 ‘You (pl) have given flowers to us.’ (P.5, 45)
- c. *tumahot.∅ nas-is-e
 2.PL.NOM dance-PERF-2
 ‘You have danced.’ (PC, 3;8)

In contradiction, as highlighted through the examples in (69), there were 34 instances where they had over-extended the ERG case marker for pronominal subjects of unaccusative verbs that should otherwise remain unmarked following the rules of the ergative split in the language. The unaccusative So verbs will always trigger a nom-marked subject irrespective of the fact whether the SUBJ is a pronoun or an NP. Hence, utterances such as (69) for the unaccusative verb *boh* ‘sit’ where the 2nd PERSON plural pronoun *tumaluk* ‘you’ and the 2nd PERSON (honorific) plural pronoun *apunaluk* ‘you’, which are overtly marked with the ergative case ending are not found in adult use of the language.

- (69) a. *tumaluk=e boh-i as-a
 2.PL=ERG sit-PROG be.PRES-2
 ‘You (PL) are sitting.’ (AAS, 4;2)
- b. *eya.t apunaluk=e boh-i as-e
 here.LOC 2.PL.HON=erg sit-PROG be.PRES-2
 ‘You (PL) are sitting.’ (PC, 3;8)

Example (70) highlights one of the rare instances where the child over-extended the use of the ergative case ending as well as the plural marker *-luk*. The 1st PERSON plural pronoun *ami* ‘we’ is a free morpheme, and it does not require a plural, however, in this example, the participant used the plural marker with it. Furthermore, the participant also used the ergative case ending with both the pronominal as well as the numeral modifying it.

- (70) *amiluk=e duta=i boh-i as-u
 1.PL=ERG two=ERG sit-PROG be.PRES-1
 ‘We both are sitting.’ (IKC 3;3)

As can be seen in Table 3.16, the children in the age group of 4-5 years old performed the best with an accuracy rate of 94%. On the other hand, the youngest group (3-4 years old) had the lowest performance with an accuracy rate of 86%.

Utterance Type	Age Group	Age Group	Age Group	Total
	3 - 4	4 - 5	5 - 6	
Target Like Utterances	131 86%	203 94%	150 90%	484 90%
Non-Target Like Utterances	21 14%	14 6%	16 10%	51 10%
Total	152 100%	217 100%	166 100%	535 100%

Table 3.16: Age-wise distribution of the Results from The Final Destination Game

Group	Obligatory Context for ERG-marking	Under-extension of ERG in Obligatory Context	Obligatory Context for NOM-marking	Over-extension of NOM in Obligatory Context	Total
3 - 4	24	10	107	11	152
4 - 5	41	2	162	12	217
5 - 6	24	5	126	11	166
Total	89	17	395	34	535

Table 3.17: Age-wise distribution of the Production Errors in The Final Destination Game

The age-wise distribution of the errors produced in this task are presented in Table 3.17. As highlighted in the table, the youngest group made similar number of errors in terms of over-extending (10 instances) or under-extending (11 instances) the ERG case marker in obligatory contexts. The 4-6 year old children dropped the obligatory ERG case marker in 2 utterances and over-extended it in 12 utterances in a zero-marked NOM context. The 5-6 year old children under-extended the ERG case marker in 5 utterances and over-extended the ERG case marker in a NOM context in 11 instances.

The crosstabulation in Table 3.18 displays the within group developmental patterns in terms of the accuracy of ERG usage in the 535 analysed utterances from 35 child participants. These results are also visually represented in Figure 3.5.

The observed and expected values in Table 3.18 show that the youngest age group (age 3-4) has fewer target-like utterances in terms of the ERG and NOM marked SUBJs than would have been expected. Further, this group produced more non-target-like utterances with higher omission rates of obligatory case markers and overextensions of case markers than expected. This pattern is reversed for the intermediate age (4-5), which has more accurate and fewer non-target-like utterances than expected, and in particular has strikingly few cases of omission of obligatory marking (2 against an expected count of

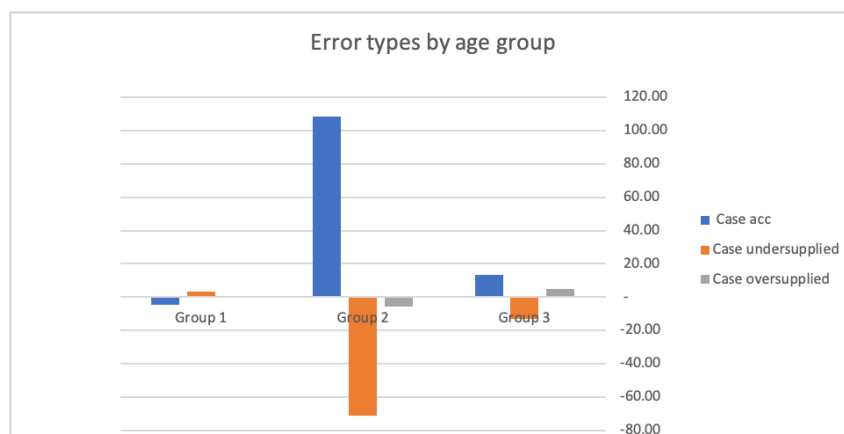


Figure 3.5: Summary of the Error Types by Age Group

Age Group		3-4	4-5	5-6	Total
case-marking accurate	observed count	131	203	150	484
	expected count	137.5	196.3	150.2	484
case-marking underextended	observed count	10	2	5	17
	expected count	4.8	6.9	5.3	17
case-marking overextended	observed count	11	12	11	34
	expected count	9.7	13.8	10.5	34
Total	observed count	152	217	166	535
	expected count	152	217	166	535

Table 3.18: Age-wise distribution of the Production Errors in The Final Destination Game

6.9). The oldest age group (5-6), occupies the intermediate position between the two younger groups, with rates of both accurate and inaccurate usage very close to the expected distribution. The group differences are significant ($\chi^2(1, 4) = 9.998, p < .05$).

3.6.4 Results for Objects

As noted earlier, Assamese has differential object marking based on animacy hierarchy for objects which also extends to the pronominals. In order to test whether young children are sensitive to this phenomenon, we used the ditransitive verb *de* ‘give’ with an inanimate direct object (DO) *phul* ‘flower’ and the transitive verb *tan* ‘pull’ with a [+human] [+animate] DO for the elicitation of data in this task.

Given this animacy hierarchy, the children were expected to use the ACC/DAT marker *-ok* for the [+human] [+animate] DO while omitting the inanimate DO zero-marked for case. There were a total of 260 sentences with direct objects in the data, out of which there were 128 obligatory contexts of ACC/DAT

marking given the [+animate] [+human] DO for the transitive verb *tan* ‘pull’. However, there was one omission of the ACC/DAT-marker in the obligatory context for the transitive verb and one utterance where the participant replaced the ACC/DAT marker with an ERG case marker. In contrast, there were 132 obligatory contexts for an ABS or zero-marked DO for the ditransitive verb *de* ‘give’ and there was not a single instance where the children used any case marker with it.

There is not enough variance in the analysed elicited utterances for DOs (and IOs discussed below) to permit statistical modelling. Out of the 260 utterances that consisted of a DO, there are only two non-target-like utterances, both of which were elicited from the same participant. As shown in example (71), there is a high probability that the participant failed to comprehend the directionality of the action in the images and was confused with who is performing the action of ‘pulling’ halfway through the utterance and ended up using the ERG marker for DO. However, the child correctly coreferences the verb correctly with the 1P marker *u*.

- (71) *moi.∅ tumaluk=e tan-isu
 1.SG.NOM 2.PL=ERG pull-PERF-1
 ‘I, you have pulled.’
 Intended: ‘I have pulled you.’ (PC, 4;6)

In example (72) we see an instance where the ACC case marker for the [+animate] [+human] 3P was been dropped in an obligatory context. Given the animacy hierarchy, the 3P plural pronominal *teuluk* should have been marked with the ACC case marker *-ok*:

- (72) *tumi.∅ teuluk ∅ tan-is-a
 2.SG.NOM 3P.PL pull-PERF-2
 ‘You have pulled them.’ (PC, 4;6)

The following table (3.19) highlights the age-wise distribution of the analysed utterances for the DOs. With the exception of the participant discussed above, all other children produced adult-like utterances while maintaining the accurate animacy hierarchy.

Apart from the DOs, there were 129 obligatory contexts for overtly marked indirect objects (IO) in the data. The ditransitive verb *de* ‘give’, triggers the obligatory ACC/DAT marker for the recipient of the DO. However, out of the 129 utterances that required an obligatory ACC/DAT case marker, there were 2 instances of omission and 1 instance of a different case marker being used. The two omissions or under-extensions of the ACC/DAT marker occurred on the 2nd PERSON plural pronoun *tumaluk* and the 3rd PERSON plural pronoun *teuluk* in example (73) below:

- (73) a. *moi∅ tuma-luk∅ phul d-is-u
 1.SG.NOM 2.PL flower-PERF-1

Utterance Type	Age Group	Age Group	Age Group	Total
	3 - 4	4 - 5	5 - 6	
Target Like Utterances	75 98%	102 100%	81 100%	258 99%
Non-Target Like Utterances	2 2%	0 0%	0 0%	2 1%
Total	77 100%	102 100%	81 100%	260 100%

Table 3.19: Age-wise distribution of the Results for OBJs from The Final Destination Game

Utterance Type	Age Group	Age Group	Age Group	Total
	3 - 4	4 - 5	5 - 6	
Target Like Utterances	32 91%	53 100%	41 100%	126 98%
Non-Target Like Utterances	3 9%	0 0%	0 0%	3 2%
Total	35 100%	102 100%	81 100%	129 100%

Table 3.20: Age-wise distribution of the Results for IOBJs from The Final Destination Game

‘I have given flowers to you.’ (KC, 3;7)

b. *tumiØ teu.lukØ phul d-is-a
 2.SG.NOM 3.PL flower-PERF-2

‘You have given flowers to them.’ (PC, 3;8)

In addition to the two omission errors mentioned above, there is one instance where the children marked the IO with the ERGATIVE case enclitic. Similar to the use of the ERG case marker for the DO in example (74), this non-target-like utterance could have been also produced because the child got confused with who is performing the action of ‘giving’ instead of this being a case of any developmental errors in acquisition of the case markers.

(74) *moiØ tuma-luk=e phul d-i as-u
 1.SG.NOM 2.PL=ERG flower give-PROG be.PRES-1

‘I, you am giving flowers’

Intended: ‘I am giving flowers to you.’ (RPS, 3;7)

The age-wise distribution of the analysed data in Table 3.20 shows that all 3 non-target-like utterances were produced in the age group of 3-4, while the slightly older participants had a 100% accuracy rate.

It should be noted here that there were no ditransitive constructions with oblique objects in the task, so we did not elicit any utterances with an obligatory GEN marking, however there were two non-target-like usage of the marker in the analysed data. As shown in example (75), the participants incorrectly used the GEN case marker with the 2nd PERSON singular pronoun *tumi*.

- (75) *moi∅ tumar du-ta=k phul d-i as-u
 1.SG.NOM 2.GEN two-CLF=ACC flower give-PROG be.PRES-1
 ‘I am giving flowers to you’ ((RPS, 3;7), (KC, 3;7))

In both of these utterances, the participants correctly used the NOM case for the 1st PERSON singular SUBJ *moi*, and the ACC/DAT marker for the numeral *du* ‘two’, which is the head of the NP. However, they used a GEN marker *-(o)r* with the 2nd PERSON in place of the plural marker *luk*.

3.7 Summary and Conclusion

The data presented through the two studies indicates that Assamese children begin producing target-like utterances from a young age, consistent with findings in earlier studies. These children appear to have an understanding of the complex rules related to both the intransitivity and PERSON and NUMBER-based as early as 2;6 years of age. However, the results from both studies reveal that children are more inclined to over-extend the ERGATIVE case enclitics rather than omit them in obligatory contexts. Interestingly, this pattern contrasts with the observations in our data from 22 young adults who also participated in the production experiment.

Furthermore, our analysis of both tasks suggests that children perform better in the PERSON and NUMBER-based split than in the intransitivity-based split. In the first study, which focused on the intransitivity based-split, we noted that both children and young adults appeared to deviate from the verb semantics-based split as outlined in the split ergative model presented in section 3.2 of this chapter. This ongoing shift in alignment appears to be moving towards an animacy-based split, a topic that will be discussed in Chapter 4 of this thesis.

Bibliography

- Al-Houti, S. K. (2013). *The acquisitions of finiteness, nominal and verbal agreement by Kuwaiti children acquiring Arabic*. PhD thesis, University of Essex.
- Allen, S. E. and Crago, M. B. (1996). Early passive acquisition in inuktitut. *Journal of Child Language*, 23(1):129–155.
- Bates, E., Thal, D., and Janowsky, J. S. (1992). Early language development and its neural correlates. *Handbook of Neuropsychology*, 8:109–176.
- Bavin, E. L. and Stoll, S. (2013). *The acquisition of ergativity*, volume 9. John Benjamins Publishing.

- Bjorkman, B. M. (2018). Ergative as perfective oblique. *Syntax*, 21(4):321–361.
- Blom, E. and Unsworth, S. (2010). *Experimental methods in language acquisition research*, volume 27. John Benjamins Publishing.
- Butt, M. and Deo, A. (2001). Ergativity in Indo-Aryan. *Kurdica*, 5(3).
- Cheshire, J., Kerswill, P., Fox, S., and Torgersen, E. (2011). Contact, the feature pool and the speech community: The emergence of multicultural london english. *Journal of Sociolinguistics*, 15(2):151–196.
- Comrie, B. (2013). Alignment of case marking of full noun phrases. In Dryer, M. S. and Haspelmath, M., editors, *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.
- Dixon, R. M. W. (1979). Ergativity. *Language*, 55:59–138.
- Eisenbeiss, S. (2010). Production methods in language acquisition research. *Experimental Methods in Language Acquisition Research*, pages 11–34.
- Friedmann, N. and Rusou, D. (2015). Critical period for first language: the crucial role of language input during the first year of life. *Current opinion in neurobiology*, 35:27–34.
- Garrett, A. (1990). The origin of np split ergativity. *Language*, 66(2):261–96.
- Handschuh, C. (2008). Split Marked-S Case Systems. *Scales*, 86:73–82.
- Hoff, E. (2013). Interpreting the early language trajectories of children from low-ses and language minority homes: implications for closing achievement gaps. *Developmental Psychology*, 49(1):4–14.
- Kerswill, P. and Williams, A. (2000). Creating a new town koine: Children and language change in milton keynes. *Language in Society*, 29(1):65–115.
- Legate, J. A. (2014). Split ergativity based on nominal type. *Lingua*, 148:183–212.
- Lust, B. and Blume, M. (2016). *Research methods in language acquisition: Principles, procedures, and practices*. Walter de Gruyter GmbH & Co KG.
- MacWhinney, B. (2000). *The CHILDES project: The database*, volume 2. Psychology Press.
- Mahalingappa, L. (2013). The acquisition of split-ergative case marking in kurmanji kurdish. *The Acquisition of Ergativity (239-270)*. Amsterdam/Philadelphia: John Benjamins.
- Mithun, M. (1991). Active/agentive case marking and its motivations. *Language*, 67(3):510–546.

- Mohanan, T. (1994). *Argument Structure in Hindi*. CSLI Publications, Stanford.
- Narasimhan, B. (2013). Ergative case-marking in hindi child-caregiver speech. *The acquisition of ergativity*, pages 209–238.
- Narasimhan, B., Budwig, N., and Murty, L. (2005). Argument realization in hindi caregiver–child discourse. *Journal of Pragmatics*, 37(4):461–495.
- O’Shannessy, C. (2013). The role of multiple sources in the formation of an innovative auxiliary category in light warlpiri, a new australian mixed language. *Language*, pages 328–353.
- Pareek, B. (2018). *Nominal Features in Hindi Language Acquisition: A Study of Agreement & Modification*. PhD thesis, Jawaharlal Nehru University, New Delhi.
- Pareek, B., Kidwai, A., and Eisenbeiss, S. (2016). Verb agreement in hindi and its acquisition. *Formal Approaches to South Asian Languages*.
- Podesva, R. J. and Sharma, D. (2014). *Research methods in linguistics*. Cambridge University Press.
- Ruigendijk, E. (2015). Contrastive elicitation task for testing case marking. In Armon-Lotem, S., de Jong, J., and Meir, N., editors, *Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment*, pages 38–54. Multilingual Matters, Bristol.
- Saleemi, A. P. (1995). On the acquisition of split ergativity: Some evidence from urdu. In Clark, E., editor, *Proceedings of the Stanford Child Language Research Forum*, volume 26. CSLI Publications, Menlo Park, CA.
- Sankoff, G. (1990). The grammaticalization of tense and aspect in tok pisin and sranan. *Language Variation and Change*, 2(3):295–312.
- van Valin Jr, R. D. (1990). Semantic Parameters of Split Intransitivity. *Language*, 66:221–260.

Chapter 4

Assamese Case Alignment Shifts in Progress

Abstract

This paper looks at case alignment in Assamese from both a synchronic and diachronic point of view. We take the task of tracing the development of the ergative case marker from the language's proto-period, and see how it evolved. This study, for the first time, provides a comparison of adult and child language data. Beyond the account of Assamese as a split ergative language, our study's results show that the semantic factor guiding this split is changing. From an ergative system based on split intransitivity determined by agentivity, thus realizing a split between unergative and unaccusative SUBJS, ERG marking is emerging on the ANIMATE subjects of unaccusatives. Interestingly, we also find that there are already traces of evidence of Assamese having had possibly the onset of an ANIMACY-based subject marking distinction in its proto-period.

4.1 Introduction

Language change is a constant phenomenon. It has been shown to affect all aspects of language, i.e. the lexicon, semantics, syntax, and phonetics, in a gradual progression over time:(Bauer, 2014; Ohala and Jones, 1993; Heggarty, 2000; Mair and Leech, 2020; Stein, 2011). It is the process through which the grammar of a given language gradually evolves into a new stable system that is distinct from the original language. Such changes have mainly been studied in terms of languages that develop because of contact, such as koines and creoles . In these instances, children and young adults regularise the current contact-induced variations observed in their languages, and maintain these as they grow up (see (Shnukal and Marchese, 1983; Amery, 1993; Cheshire et al., 2011; O'Shannessy, 2013; O'Shannessy et al., 2019)).

In this paper we determine that Assamese, an Indo-Aryan language spoken by 14 million native speakers in the northeastern state of Assam in India,

is showing signs of change in its morphosyntax. Specifically we argue that the language is reanalysing its differential subject marking system guided by semantically-motivated case alternations that are changing the nature of the current status of the language which is one of split ergativity based on agentivity. In effect, we argue that what is taking place in Assamese very much parallels the situation in other New Indo-Aryan (NIA) languages, as argued in, for e.g. Ahmed (2010) and Butt and Ahmed (2011), where the language is recycling its current SUBJ case system to express distinct semantic factors.

Supporting the thrust of this study which is a discussion of a hypothesis that change is in progress, we incorporate a child language data-based study demonstrate that the direction in which change is progressing is magnified by what can be observed through child language data. The employment of child language data as a means which can guide our assessments on, and of, variation and change, is key to the views upheld in Lightfoot (2010). We also argue that the newly evolving split is conditioned by the semantic nature of the NPs, and which is in fact a reflection of the differential marking one finds with respect to objects in the language. Consequently we hypothesise that what is emerging can also be referred to, in parallel, as differential subject marking.

This paper is organised as follows. In section 4.2 we provide an overview of the previous literature that focuses on children-led language change. In section 4.3 we present a characterisation of case marking in Assamese. In section 4.4 we integrate the child language data and its corresponding adult data studies to our overall assessment of the language's grammar and pinpoint the change in progress. In section 4.5 we then provide a summary of our conclusions.

4.2 Previous Studies in Children-Led language Change

The phenomenon of children-led language change is a captivating area of study that sheds light on how linguistic variation and innovation are driven by younger generations. This literature review explores how children, through linguistic proximity and parental influences, play a pivotal role in language change. These studies reveal how children's linguistic practices can lead to structural and functional changes in language, extending beyond traditional language domains.

Kerswill and Williams (2000) studied a single generation of children and young adults, who, having originally spoken different varieties of English, developed a koine after having come to live together in Milton Keynes, England. Kerswill and Williams (2000) concluded that 'the New Town Koine' developed due to the frequent mixing of the children's and young adults' varied linguistic varieties, in their interactions. They argue that fewer linguistic differences between the two varieties, and the less prescriptivist approach of the children's

4.2. PREVIOUS STUDIES IN CHILDREN-LED LANGUAGE CHANGE⁷⁷

parents, or caregivers, have also played vital roles in the development of this koine.

Amery (1993) reports a similar occurrence of language change in Arnhem Land, Australia. This study is based on a baby talk register that has been developed when speakers of different languages within the Yolngu Matha linguistic family frequently socialised together at school and in the community.

O'Shannessy (2013); O'Shannessy et al. (2019) looks at the emergence of Light Warlpiri through a contact-induced language change led by children. Children received input in a baby talk register that contained grammatical elements from Warlpiri (Pama-Nyungu), English and Kriol (a creole spoken in North Australia). The children eventually went on to add their own structural innovations to this emerging mixed system, and gradually, they stabilised it into a new language, now known as Light Warlpiri.

Other than structural change, there is evidence that argues in favour of children-led language change in other linguistic domains. Cheshire et al. (2011) study children in London, England, who have acquired specific vowel positions and morphosyntactic patterns of the local variety while they were exposed to it in schools at 4-5 years of age, even when their home language input was different.

(Queen, 1996, 2001) studies bilingual Turkish-German children who developed new intonation patterns that draw on both German and Turkish, yet with the resulting pattern employed by these children being very different from both.

Acquiring a language in which the grammar has undergone change is not the same as acquiring a language with a stable grammatical system (Cheshire et al., 2011; Kerswill and Williams, 2000; O'Shannessy, 2013; O'Shannessy et al., 2019). Children receive mixed inputs with variations from different languages and have to work through surface cues to process the language. There is evidence that suggests that children develop changes that go further than whatever innovations they may have received in their inputs. For example, Tok Pisin speaking children have been shown to make use of a more novel aspectual use of a tense-mood-aspect marker than their adults, and have as a consequence pushed the change further (Sankoff, 1990).

Assamese is currently undergoing a shift, as highlighted in the data of the previous chapter, wherein the intransitivity-based split is changing to an animacy-based split. In this paper we will look at ergativity in Assamese in its current form, and investigate how native child speakers at the age of 2-6 years are furthering a change that parallels emerging innovations in the language of young adult speakers.

4.3 Distribution of Case in Assamese

4.3.1 Assamese as a Split Ergative Language

Assamese is a head-final SOV dominant language that is syntactically accusative, that is, its different subjects, as we will exemplify later, align in one pattern together with respect to control phenomenon, anaphoras, relativisation, and in particular agreement patterns. When it comes to the morphological characterisation of the case system, several terminologies to code case alignments are provided in the literature (cf. (Dixon, 1979, 1994; Comrie, 1978)), however, we specifically choose the terminologies used in Mohanan (1994), and refer to Assamese as a split ergative language, i.e. a language with two distinct cases associated with subjects, where one is inflected and the other remains uninflected or unmarked. The marked subject is referred to as being ergative, while the unmarked subjects are referred as nominative.

Assamese is often described as a NOMINATIVE-ACCUSATIVE (Kakati, 1941; Goswami and Tamuli, 2003; Nath, 2003; Haddad, 2011) or a (fully) ERGATIVE system (Devi, 1986; Butt and Deo, 2001; Zakharyin, 2015). However, what Assamese really demonstrates is a **split ergative system** with splits conditioned by **intransitivity**, i.e. based on whether the intransitive verb is unergative or unaccusative, which, synchronically, without yet considering the direction of the change in progress, is based purely on **agentivity** (Amritavalli and Sarma, 2002). Within the pronominal system, however differential case marking (DCM) (Aissen, 1999, 2003) is conditioned by PERSON and NUMBER (Saha and Patgiri, 2013).

To understand why we are referring to Assamese as a split ergative language, we provide the data below. NP subjects of (di)transitive verbs (A), irrespective of animacy, obligatorily take an overt ERG case marker in Assamese, as exemplified through the data in (76).¹

- (76) a. lora-tu=e bol-tu d^hor-i as-e
 boy-CLF=ERG ball-CLF hold-PROG be.PRES-3
 ‘The boy is holding the ball.’
 (ANIM ERG SUBJ **of transitive** PRED)
- b. bas-bur=e baik-k^hɔn k^hundi-a-l-e
 bus-PL=ERG bike-CLF knock down-CAUS-PST-3
 ‘The buses knocked down the motorbike.’
 (INANIM ERG SUBJ **of transitive** PRED)

Among intransitive verbs, agent-like subjects of unergatives (S_a), irrespective of animacy, trigger an overt marker on the subject, while the patient-like subjects of unaccusative verbs (S_o) remain unmarked. For example, the S_a NP

¹The Assamese data, unless provided with a citation, is the native speaker author’s own.

referent of unergative verbs like *jump*, *dance*, and *swim* control an activity, as opposed to the S_o NP referents of unaccusative verbs like *fall*, *sink*, and *burn* that have no control over the activity. Further, similar to referents of an O function, the referents of S_o could be affected by the event. Although certain intransitive verbs can be easily categorised as either S_a or S_o, the categorisation of some might vary across languages (see, for instance, Dixon (1979); van Valin Jr (1990); Handschuh (2008)). The contrast between S_a and S_o, in the context of unergative and unaccusative verbs respectively, is illustrated through (77) and (78).

- (77) a. roza-zon=e xātur-i as-e / xātur-is-e
king-CLF=ERG swim-PROG be.PRES-3 / swim-PERF-3

‘The king is / has been swimming.’

(ANIM ERG **NP** SUBJ of unergative PRED)

- b. botah-zak=e huhurija-is-e
wind-CLF=ERG whistle-PERF-3

‘The wind has been whistling.’

(INANIM ERG **NP** SUBJ of unergative PRED)

- (78) a. roza-zon.∅ boh-i as-e / boh-is-e
king-CLF.NOM sit-PROG be.PRES-3 / sit-PERF-3

‘The king is / has been sitting.’

(ANIM NOM **NP** SUBJ of unaccusative PRED)

- b. kat^h-sota.∅ upoŋ-i as-e / upoŋ-is-e
wood-CLF.NOM float-PROG be.PRES-3 / float-PERF-3

‘The piece of wood is / has been floating.’

(INANIM NOM **NP** SUBJ of unaccusative PRED)

Case marking is more complex in the pronominal system. On the basis of the discussion by Saha and Patgiri (2013), specifically in Assamese, only the 2nd and 3rd PERSON **plural pronouns** trigger ERG case marking in the form of an enclitic.²

- (79) a. tumaluk=e xātur-i as-a / xātur-is-a
2.PL=ERG swim-PROG be.PRES-2 / swim-PERF-2

‘You (PL) are / have been swimming.’

(2.PL ERG **pronoun** SUBJ of unergative PRED)

²In another Indo-Aryan language, Punjabi, the pronominal system appears to be sensitive *just* to 1st/2nd vs. 3rd PERSON based split, whereby only the latter set of pronouns (and NPs) take an ERG marking (Butt and Deo, 2001). This, thus differs from the seemingly more complex interaction between PERSON and NUMBER in Assamese.

- b. $x\tilde{i}hot=e$ $x\tilde{a}tur-i$ $as-e$ / $x\tilde{a}tur-is-e$
 3.PL=ERG swim-PROG be.PRES-3 / swim-PERF-3
 ‘They are / have been swimming.’
 (3.PL ERG **pronoun** SUBJ of **unergative** PRED)

Supposedly, the rest of the pronominal SUBJ paradigm remains unmarked, i.e. it expresses NOM case, (as the \emptyset marking is meant to illustrate) demonstrated via the 1.SG and 3.SG.M pronominal subject forms in (80). In this respect, therefore, the split on the basis of PERSON and NUMBER within the pronominal system, in contrast to the neater nominal system, takes supremacy over the requirement of (A) SUBJs of transitive predicates to be ERG-marked as illustrated in (76).³

- (80) a. $moi.\emptyset$ $sur-tu=k$ $d^hor-il-u$
 1.SG.NOM thief-CLF=ACC hold-PST-1
 ‘I caught the thief.’
 (1.SG NOM **pronominal** SUBJ of **transitive** PRED)
- b. $xi.\emptyset$ $sur-tu=k$ $d^hor-il-e$
 3.SG.M.NOM thief-CLF=ACC hold-PST-3
 ‘He caught the thief.’
 (3.SG.M NOM SUBJ of **transitive** PRED)

An internal reviewer suggests that there is a probability that the observation of unmarked pronominal forms may look so only on the basis of their surface morphology, i.e. in the absence of an $-e$ marking. For this reason, an alternative analysis would be to assume that these pronouns are in fact ‘old and have come down (for some reason) in an originally oblique form’. In support of this alternative analysis, we could argue, following Kakati (1941), that the PERSON and NUMBER based split in the pronominal system is itself a remnant from Middle Indo-Aryan (MIA). He observes how for instance, the 1.SG pronoun *moi* or the inferior 2.SG *toi*, and so on maintain the MIA proto-instrumental forms $-\tilde{e}$, $-i$ (synchronically interpreted as ERG) in their extended oblique pronominal bases. On the other hand, (Saha and Patgiri, 2013, pp. 39-40) argue that the split that results is a reflex of a morphophonological constraint, such that since the 1.SG/PL, 2.SG and 3.SG pronominal forms end with a high vowel $/i/$, ERG $-e$ marking is blocked.

Given the above characterisation for the SUBJ case marking system for nominals, inclusive of a split intransitivity governed by the subject’s agentivity, along with an incorporation of the assumption that the pronominal system is actually characterised by DCM based on PERSON and NUMBER, the following table summarises the facts.⁴

³The same parallel behaviour follows for S_a subjects.

⁴It is worth mentioning that pronouns in Assamese have always been discussed with respect to animate reference; the distribution of which, in terms of case marking, is pre-

SUBJ	NP	1.SG/PL	2/3.SG	2/3.PL
A	ERG	NOM	NOM	ERG
S _a	ERG	NOM	NOM	ERG
S _o	NOM	NOM	NOM	NOM

Table 4.1: Distribution of case-marking on SUBJ GFs

Another instance where the ergative split discussed above for NPs as well as pronouns is overridden, is in contexts where a homophonous *-e* marker is present on SUBJs to express what Butt and Holloway King (1991), Butt (2006), Ahmed (2010), and Butt and Ahmed (2011) refer to, with respect to cognate *-ne* marker in Hindi/Urdu, as a marker of volitionality/intentionality. If we consider the contrast in the pair below, *lora* ‘boy’ is unmarked in the context of the intransitive unaccusative verb *por* ‘fall’, in line with our discussion above. Nevertheless, when the semantic interpretation expressed is such that the SUBJ deliberately/purposefully initiates the falling event, then an *-e* marker, which we presume to be the ERG marker, but which here is of a distinct semantic function is marked on the SUBJ. The end result is such that as illustrated in (81b), we have the presence of an ERG marker in the context of an unaccusative verb.

- (81) a. *lora-tu.∅ por-il(-e)*
 boy-CLF.NOM fall-PST(-3)
 ‘The boy fell down.’
- b. *lora-tu=e por-i di-l-e*
 boy-CLF=ERG fall-NF give-PST-3
 ‘*The boy* (deliberately/purposefully) fell down.’ (Chowdhary, 2014, p. 111)

Saha and Patgiri (2013, p. 40) argue that the same follows for pronouns, including the 1st PERSON.SG/PL and 2nd/3rd.SG pronouns, where NOM marking is overridden, and *-e* marking is present, as illustrated through the 1.SG pronominal *moi* ‘I’ in (82).⁵

- (82) *moi=e za-m tumar log-ot*
 1.SG=ERG go-FUT 2.SG.GEN company-LOC
 ‘It is *I*, who will accompany you’ (S.C. Chiring Phukan, p.c.)

sented in Table (1). Reference to inanimate entities, on the other hand, involves a distinct pronominal device; a resort to the use of the demonstrative pronominal paradigm, such as *ei/hei* ‘this/that’ along with the attachment of the default classifier *-tu*, or any of the shape classifiers, such as *-dal* and *-k^hɔn*.

⁵Note that this particular use of the ERG form on the 1.SG pronominal can also be realised as *-ei* or *-(e)he*. The availability of alternations of this sort is also true for the 1.PL and 2/3.SG.

Before proceeding further, a note on agreement behaviour vis-à-vis case is in order. In Assamese, NOM vs. ERG DCM on SUBJs does not block agreement with the subject on verbs, unlike what goes on in other Indo-Aryan ergative languages, such as Hindi. This is exemplified, for instance, through data such as (79a) and (80a), where irrespective of ERG vs. NOM marking, respectively, the verb displays the relevant PERSON agreement with it. Notwithstanding this pattern, DCM *does* matter for SUBJ-verb agreement purposes beyond ERG/NOM-marked SUBJ contexts. Non-canonical subjects, which are expressed via non-ERG/NOM morphology, on the other hand trigger default 3rd PERSON agreement. In (83), we have an illustration of *lag* lit. ‘want’ functioning as a psych predicate, with the meanings ‘feel’ and ‘get (fear)’, respectively. This consequently requires a GEN-marked SUBJ, such as the 1.SG pronominal *mur*. The presence of such a subject triggers 3rd PERSON agreement on the verb, which is the form employed for default agreement contexts. A similar default agreement pattern also follows in the case of the predicate *lag* when used in a desiderative sense, meaning ‘want’, as in (84). In this case, the SUBJ is ACC/DAT-marked via the phonologically-conditioned allomorphs *-(o)k*.⁶

(83) a. *mur ijat niz=ok asohua zen lag-e*
 1.SG.GEN here self=ACC stranger as if get.PRES-3
 ‘I feel as if I am a stranger here.’ (Chowdhary, 2014, p. 115)

b. *mur b^hoi lag-is-e*
 1.SG.GEN fear get-PERF-3
 ‘I am scared’ (Lit: ‘I have got fear.’)
 (GEN SUBJ of psych verb)

(84) *muk b^hat lag-e*
 1.SG.DAT/ACC rice want.PRES-3
 ‘I want rice.’
 (DAT/ACC SUBJ of desiderative ‘want’)

4.3.2 The Historical Development of the Ergative in Assamese

Ergativity was not an inherent grammatical feature of Sanskrit, which is the ancestor of all Indo-Aryan languages. Sanskrit, which is a NOM-ACC lan-

⁶Although not discussed in the literature, one could argue that what we have in the case of (84), is an instance of the verb’s agreement with the OBJ GF, rather than an instance of default agreement. This would in principle parallel Hindi in the sense that when ERG subjects are present in perfect contexts, the agreement which results on the verb is that with the object. While we won’t engage in this discussion here, although such an analysis is a possibility, we refer the reader to (83a), at least if it can be said to constitute a like with like instantiation, and argue that if what we have in (84) were an instance of OBJ-verb agreement, rather than default 3rd PERSON agreement, then we would have expected to see 1st PERSON agreement on the verb in (83a), given the *-ok*-marked reflexive OBJ.

guage, as highlighted in (85a), used the instrumental marker *-ena* on the semantic agent within the passive construction, yet where the subject remains NOM-marked, as in (85b) (Kakati, 1941; Butt and Deo, 2001; Verbeke and De Cuypere, 2009).

- (85) a. devadatta-ḥ kaṭa-m ca-kār-a
 Devadatta-NOM mat-ACC PERF-make-3SG
 ‘Devadatta made a mat.’ (Verbeke & De Cuypere, 2009, p. 2)
- b. devadatt-ena kaṭa-ḥ kṛ-taḥ
 Devadatta-INS mat-NOM make-PST.PASS.PTCP
 ‘The mat is made by Devadatta.’ (Verbeke & De Cuypere, 2009, p. 3)

Although there are several accounts of how ergativity developed in Indo-Aryan languages, the reanalysis of a passive as an ergative construction is the most common hypothesis among scholars. Moreover, the evolution from the Sanskrit instrumental *-ena* or *-ī*, to the Assamese ergative marker *-e* seems highly probable (Kakati, 1941; Coghill, 2016; Kulikov, 2017), *contra*. (Butt and Deo, 2001; Beames, 1872; Kellogg, 1972). If we look at the timeline of this development, we find that Ashokan inscriptions from the Early Indo-Aryan (EIA) period show that the Sanskrit NOM marker *-ah* was being replaced by *-e*, as in *devānāmpiy-e* (*devānāmpri-yah*) ‘the one who is loved by God’ (Bloch, 1965; Devi, 1986).⁷ Tagore (1948), as cited in Devi (1986, p. 68), proposed that the Sanskrit term *putrah* ‘son’ changed to *putte* in the Middle Indo-Aryan (MIA) period, until it eventually became *putti*, due to vowel weakening during the Apabhramśa period.

The Caryā texts composed by the Buddhists between the 8th and 12th century are claimed to bear the earliest evidence of literature stemming from the eastern group of Indo-Aryan languages. Devi (1986) notes several similarities between the Assamese ERG *-e*, and the *-e* and *i* subject markers found in these texts. Since the use of the *-e* marker had not become stable until the New Indo-Aryan (NIA) period, these texts bear the expected inconsistencies of the transition stage. For example, both *-e* and *i* were used, at this stage of the language, with the agent of transitive verbs, as in *sur-e* ‘thief’, *kānhi* ‘Kānhā’.⁸ However, towards the end of the texts, the use of *-e* gets stabilised as the sole subject marker. Devi (1986) points out that there is only one exclusive instance of an unmarked subject of a transitive verb in Caryā 6 of these texts. The example is represented in (86) below, where we observe an instance involving the subject NP *harina* ‘deer’ not taking an ERG marking, in spite of being the subject of two coordinated transitive clauses.

⁷Alternatively, as suggested by an internal reviewer this could potentially be a case of phonological change rather than (direct) replacement.

⁸Name of a Hindu God.

- (86) *tina na echupai hariṇa pibai na pāni*
 grass not touch deer drink no water

‘The deer does not notch any grass nor does (the deer) drink any water.’
 (Devi, 1986, p. 70)

Apart from optionally-marked subjects of transitive verbs, there are also examples of optional *-e* marking on the subjects of intransitive verbs in these texts. For example, in Caryā 48, the reflexive pronoun *apan*⁹ ‘self’, which is the subject of the unaccusative verb *bah* ‘sit’ is marked with the *-e* marker, while the subject of the clause in the first conjunct: *grāhaka* ‘customer’ is unmarked, in line with the synchronic facts when in context of unaccusative intransitive verbs, such as *ai* ‘come’.

- (87) *āilā grāhaka.∅ apan-e bahiā*
 come customer.NOM self-ERG sit.PST

‘Customer came and (himself) sat down.’ (Devi, 1986, p. 71)

This sporadic use of the *-e* marker on the subject of transitive verbs, and some intransitive verbs, can be taken as the stage where a split ergative system started emerging in Assamese. Assamese developed simultaneously with other eastern Indo-Aryan languages, such as Odia (Oriya) and Bengali from the common ancestor: Eastern Magadhi, which branched out of Māgadhi Prakrit in the MIA period (Chatterji, 1926). There is evidence that such NIA languages from the eastern branch, including Maithili also once used *-e* markers on their subjects (Chatterji, 1926; Kakati, 1941).

However, synchronically, Assamese differs from other eastern Indo-Aryan languages, including Bengali, Oriya, Maithili, and Bhojpuri, which have now lost their erstwhile ergative case system, and have become reanalysed as NOM-ACC systems. In contrast, Assamese, together with Sylheti, and Nepali, are the only eastern Indo-Aryan languages that have retained the ERG alignment of their parent language, but which is not based upon an ASPECT-based ergative split system¹⁰. Rather, they collectively display an intransitive-based split, which one could argue to be an influence akin to contact with neighbouring Tibeto-Burman languages.

Devi (1986, p. 63) argues that the consistent use of *-e* that we see on agents in Assamese might be an influence from the Ahom (Tai) and Naga (Sino-Tibetan) languages, which mark their agents with a distinct marker. She further notes that texts from the 13th century that were composed just

⁹Note that Devi (1986) glosses the *-e* in this example as NOM. Here we gloss this morph as ERG. We additionally glossed the unmarked subject *grāhaka* ‘customer’ as a NOM and marked it with a ∅.

¹⁰Nepali *does* maintain an ASPECT-based split. However, this is only internal to the transitive sub-system (Li, 2007).

after the Ahoms conquered Assam show the use of an optional *-ko* marker with both NPs and pronouns. Moreover, the Nagas that were given place in the Ahom court used distinct agentive markers for their NPs and pronouns. The same can also be said for the Tangsa group of languages spoken to the east of Assam. Devi argues that the presence of such language systems in contact with Assamese must have accelerated the use, and later the consolidation of an agent marker in Assamese. A parallel can be drawn to Dakkhini (Stroński, 2010), which has lost its ERG case marking due to isolation from other Indo-Aryan languages, along with its long lasting influence from its neighbouring NOM-ACC Dravidian languages.

Kakati (1941, p. 286), as mentioned earlier, on the other hand, argues that the ERG *-e* in Assamese is a reanalysis of the instrumental *-(er)e*¹¹ that is obligatorily present on the subjects of passive constructions built out of transitive verbs, as in: *hat-(er)e buwa kapur* ‘cloth woven by hand’. He further argues that it is this constant use of the INS *-(er)e* that has led to the habitual use of *-e* in the expression/realisation of agent subjects.

However, the synchronic analysis of the language shows that there is a distinction between the INS *-(er)e* and the ERG *-e*, even if the literature suggests that these were once the same *-(er)e* form in the past. Irrespective of the interchangeable use of the INS *-(er)e* and the ERG *-e*, it is the subjects with *-e* that render an agentive reading, and not the ones marked with *-(er)e*. The data in (88) is meant to demonstrate that although *kotari* ‘knife’ can be marked with *-e*, we still are glossing the morph as INS, as we cannot possibly assume two ERG-marked NPs in the clause. It is clear that in this active sentence, the ERG-marked 3.PL pronoun functions as the SUBJ.

- (88) xīhot=e tak kotari=re kat-il-e
 3.PL=ERG 2.SG.ACC knife=INS cut-PST-3
 ‘They cut him with a knife.’

Moreover, in sentences such as (89), *dak* ‘post’ can only be marked through the INS *-ere* marker. This suggests to us that an NP like *dak* ‘post’ can never be ascribed any agentive role, in contrast to the possibility with respect to *kotari* ‘knife’, which could be what is allowing us an *-e* morph to express the INS case in (88).

- (89) sithi-k^hon dak=ere/*=e ah-il
 letter-CLF post=INS/=ERG come-PST
 ‘The letter came by the post.’

¹¹In the early Assamese period, the INS *-(k)ere* was also used to express accompaniment, as in *jāmāi-ere* ‘with my son-in-law’ (Kakati, 1941, p. 287). The Chittogong dialect takes both genitive *-ar* and instrumental *-di* on the same noun to express accompaniment as in *put-ar-di* ‘with the son’ (Kakati, 1941, pp. 286-287).

4.3.3 Non-SUBJ Case Marking in Assamese

If we are to argue that DCM results in differential subject marking in Assamese, then we here present a context, where elsewhere in the grammar of the language we also observe distinct markings associated with the same GF. Here we consider the distribution of case in the context of non-SUBJ GFs. Just as it has been shown in the literature that the Animacy Hierarchy accounts for a good deal of the cross-linguistic variation in split ergative systems, with differences observed on the basis of the nature of the noun type (McGregor, 2009), the same premise can be applied to behaviours associated with OBJ GFs in Assamese, which come to be marked as ACC with the *-(o)k* marker.

The applicability of the Animacy Hierarchy scale may differ from one language to another. It has, however been shown to have wider impact on a number of distinct grammatical phenomena ranging from agreement to syntactic marking, and the like. Croft's (2003, p. 112) Animacy Hierarchy represented in (90), is indicative of the fact that, for instance, referents higher on the scale, such as 1st/2nd PERSON pronouns are more likely to receive overt case marking than inanimate common nouns lower on the hierarchy.

- (90) **first, second-person pronoun < third-person pronoun < proper names < human common noun < nonhuman animate common noun < inanimate common noun**

While (animate-referring) pronouns in Assamese are always ACC-marked, as illustrated in (91), NPs do not display a uniform behaviour. For example, in (92a) the animate object *Rita* of the transitive verb *d^hor* 'hold' takes the ACC case marker *-(o)k*, while in (92b), the inanimate object *bol* 'ball', associated with the same transitive verb, remains unmarked. Leaving *Rita* unmarked in (92a), results in ungrammaticality. Such behaviours have been referred to as differential object marking (DOM) in the literature. DOM also exists in a number of typologically different languages, such as Turkish (Kornfilt, 2009), Maltese (Camilleri and Sadler, 2012), and Spanish (Comrie, 2013).

- (91) a. xi.∅ muk dek^h-il-e
 3.SG.M.NOM 1.SG.ACC see-PST-3
 'He saw me.'
- b. moi.∅ xihot=ok dek^h-il-u
 1.SG.NOM 3.PL=ACC see-PST-1
 'I saw them.'
- (92) a. nitu=**e** rita=k/*rita d^hor-il-e
 nitu=ERG rita=ACC/rita.∅ hold-PST-3
 (Lit. 'Nitu held Rita.')
- 'Nitu caught Rita.'

- b. nitu=e bol-tu/*bol-tu=k d^hor-il-e
 nitu=ERG ball-CLF/*ball-CLF=ACC hold-PST-3
 (Lit. ‘Nitu held the ball.’)
 ‘Nitu caught the ball.’

Although non-human animates are higher on the Animacy Hierarchy than inanimates, in Assamese no distinction appears to be made between animate categories such as animals, birds, or trees and inanimates. This is illustrated through the data in (92b) and (93) that take no *-(o)k* ACC marking.

- (93) a. goru-zoni.∅/*=k band^h-il-i-ne
 cow-CLF.∅/=ACC tie-PST-1-Q
 ‘Did you tie the cow?’
 b. tamul-zupa.∅/*=k ne-kat-ib-i
 areca-nut-CLF.∅/=ACC NEG-cut-FUT-2
 ‘Do not cut the arica nut tree.’

As things stand, it seems therefore that DOM in Assamese is conditioned by a HUMAN feature. However, there is added intricacy to when and in which contexts does case marking appear even on non-HUMANS. For instance, if the *goru* ‘cow’ in (93a) is given a Proper Name, this will be *-(o)k* marked. If on the other hand, the Proper Name of an inanimate is in OBJ position, such as the *Taj Mahal* (a heritage monument), this will not get *-(o)k* marked (Chowdhary, 2014, p. 117). Beyond (ANIM) Proper Names (and pronouns), DOM on HUMAN NPs is interrelated with concerns that pertain to SPECIFICITY. This is in fact something that has been discussed quite amply for Hindi (see e.g. Butt (1993) and Montaut (2018)).

Assamese is a numeral classifier language. This implies that once a classifier attaches onto the right-edge of an NP¹² OBJ (be it HUMAN or non-HUMAN) in the absence of a numeral, which would otherwise take the classifier, that NP becomes DEFINITE. While the behaviour of DEF/INDEF cuts across the board irrespective of whether a HUMAN or non-HUMAN OBJ is involved, as illustrated through (94) and (95 a-b) below, the addition of case becomes obligatory in the context of a [+SPEC] reading in association with HUMAN NPs, as illustrated in (95 c).

- (94) a. moi.∅ kitap.∅/*=ok porh-i b^hal pa-o
 1.SG.NOM book.∅/=ACC read-PROG good get.1
 (Lit: ‘I feel good reading book.’)
 ‘I love reading books.’ ([+/- DEF] [- SPEC])
 b. moi.∅ kitap-k^hon/*=ok porh-i b^hal pa-o
 1.SG.NOM book-CLF/=ACC read-PROG good get.1
 ‘I love reading the book.’ ([+ DEF] [+/- SPEC])

¹²Note that when there is a classifier as well as a case marker attached onto an NP, the classifier always precedes the case marker.

- (95) a. *pulis=e sur.∅ d^hor-e*
 police=ERG thief.∅ hold.PRES-3
 (Lit: ‘Police hold thief.’)
 ‘Police catches thieves.’ ([+/- DEF] [- SPEC])
- b. *pulis=e sur-tu.∅ d^hor-il-e*
 police=ERG thief-CLF.∅ hold-PST-3
 ‘Police caught the thief.’ ([+ DEF] [+/- SPEC])
- c. *pulis=e sur-tu=k d^hor-il-e*
 police=ERG thief-CLF=ACC hold-PST-3
 ‘Police caught the thief.’ ([+ DEF] [+ SPEC])

What this implies therefore is that HUMAN NPs are made SPECIFIC via the very presence of *-(o)k* marking; a strategy which is not morphosyntactically available for non-HUMAN NPs. In contrast, SPECIFICITY in non-HUMAN INANIM NPs such as *kitab* ‘book’ comes solely from the context. Furthermore, although the majority of the literature (e.g. Gundel et al. (1993), Eng (1991)) suggests that DEFINITENESS also implies SPECIFICITY, this does not hold true for Assamese, given that the numeral classifier *-tu* attached to the HUMAN NP *sur* ‘thief’ implies that it already takes a DEFINITE reference, but which is not yet made SPECIFIC, necessarily, until a case marker is present on OBJ. While DOM brings out SPECIFICITY effects in both Hindi and Assamese, with Assamese allowing this only in the context of HUMAN NP OBJs, there are other Indo-Aryan languages like Sinhala/Sinhalese, where an ANIMACY is all that matters in the determination of whether objects can be optionally ACC-marked or not (Thampoe, 2017).

It should finally be noted here that DOM in Assamese only applies to objects in neutral contexts. If the object is placed in a non-neutral context, such as in a topicalised position, typically left-adjacent to the verb (along with additional intonation cues) the inanimate indefinite NP must be ACC-marked. Such a behaviour is highlighted in (96) below through the inanimate, indefinite NPs *zibon* ‘life’ and *d^hopat* ‘tobacco’.

- (96) *zibon-ok adɔr-ok d^hopat-ok nɔ-hɔi*
 life-ACC welcome-IMP tobacco-ACC NEG-be.PRES
 ‘Welcome life, not tobacco’ (Chowdhary, 2014, p. 118)

The morphosyntax and the structure of (96) would imply that a structure such as (97), although displaying a parallel string, cannot be understood as a topicalisation structure. What we have in (97) is an instance where the INANIM *pani* ‘water’ and *mod* ‘alcohol’ are unmarked, unlike the ACC marking on the topicalised *zibon* ‘life’ and *d^hopat* ‘tobacco’. The non-marking of the NPs in (97) is in line with them being INANIM NPs sitting low on the Animacy

Hierarchy. For this reason therefore, as also suggested by an internal reviewer, what we have here is a case of an SOV structure with the SUBJ dropped by virtue of the imperative mood of the structure.

- (97) pani.∅ kha-ok mod.∅ nə-hɔi
 water.NOM drink-IMP alcohol-NOM NEG-be.PRES
 ‘Drink water not alcohol.’

So far we have only considered what goes on with primary/direct objects, or OBJ GFs in LFG terms. When we turn our attention to indirect objects, i.e. those GFs that function as recipients in ditransitive constructions, we have evidence, although not given any attention in the literature on Assamese that morphologically, the OBJ_θ takes a distinct case distribution, even if, the marker which we here, for expository convenience refer to as DAT, takes a homophonous *-(o)k* form just as the ACC, (as is also the case in Hindi/Urdu). What is key for us, in the light of the data paradigm in (98), is that the recipient, which can be a Proper Name, as in (98a), a HUMAN NP as in (98b), a non-HUMAN ANIM NP as in (98c), and an INANIM NP as in (98d) is that of a double object construction. Evidence that the recipient in Assamese maps onto an OBJ_θ, rather than onto an OBJ, is clear from the distribution of *-(o)k* ACC marking on the theme, which patterns exactly what we have just discussed above. On the other hand, *-(o)k* as a DAT marker on the OBJ_θ does not display a similar behaviour. Rather, such marking is present throughout. Note that in Assamese a clear constituent order preference holds whereby it is more likely to have the recipient argument preceding the theme.

- (98) a. tai.∅ pinki=k/*∅ puna=k/*∅ hop-il-e
 3.SG.NOM Pinki=DAT/∅ Puna=ACC/∅ entrust-PST-3
 ‘She entrusted (the custody of) Puna to Pinki.’ (Adapted from
 (Chowdhary, 2014, p. 119))
- b. tai.∅ mastor=ok/*∅ lora-tu(=k) hop-il-e
 3.SG.NOM teacher=DAT/∅ boy-CLF(=ACC) entrust-PST-3
 ‘She entrusted (the custody of) the boy to the teacher.’ (Adapted
 from (Chowdhary, 2014, p. 119))
- c. tɛɔluk=e kukur-tu=k//∅ b^hat/*∅ d-il-e
 3.PL=ERG dog-CLF=ACC/∅ rice/*∅ give-PST-3
 ‘They gave rice to the dog.’
- d. tɛɔluk=e xoŋgram-tu=k/*∅ notun ort^ho/*∅ d-il-e
 3.PL=ERG revolution-CLF=ACC/∅ new meaning/∅ give-PST-3
 ‘They gave new meaning to the revolution.’

We here, finally, consider prepositional objects that are GEN-marked via the phonologically-conditioned allomorphs *-(o)r*, as shown by the oblique objects *Pinki* in (99a), *deutak* ‘father’ in (99b), and *duwar* ‘door’ in (99c).

- (99) a. razu=e pinki=r karone p^hul kin-is-e
 razu=ERG pinki=GEN for flower buy-PERF-3
 ‘Razu has bought flowers for Pinki.’
- b. tai.∅ deutak=or karone sit^hi lik^h-is-e
 3.F.NOM father=GEN for letter write-PERF-3
 ‘She has written a letter for her father.’
- c. tai.∅ duwar-k^hɔn=or karone tola e-ta kin-il-e
 3.F.NOM door-CLF=GEN for lock one-CLF kin-PST-3
 ‘She bought a lock for the door.’

What this means for us is that OBL OBJs, i.e. the OBJ GFs which Ps sub-categorise for, are special. Beyond the fact that they get GEN-marked, such GEN marking appears to ‘block’ a distribution that parallels what we have described above in the context of ACC-marked OBJs.

Table (4.2) below now provides a characterisation of the distribution of case marking across the non-SUBJ GFs.

Value	OBJs	OBJ _θ	OBL OBJs
Pronoun	ACC	DAT	GEN
Proper Names	ACC	DAT	GEN
HUMAN NP	(ACC)	DAT	GEN
ANIM/INANIM NPs	∅	DAT	GEN

Table 4.2: Distribution of case-marking on non-SUBJ GFs

From the above discussion it transpires that the observations associated with the OBJ GF, in relation to the distribution of case correlates with SPECIFICITY. One could say that the behaviours attributed to the Animacy Hierarchy fall out in an expected manner, since for instance personal pronouns and Proper Names are inherently specific, and hence precisely illustrate the contexts where we get to observe obligatory ACC marking. In contrast, OBJ_θ and OBL OBJs display a uniform behaviour and do not pertain to any Animacy Hierarchy-based observations.

4.3.4 Current Predictions of Change in Progress

Focusing specifically on the SUBJ GF of intransitive verbs in Assamese, and the distribution of ERG case, it seems to us that change is in progress. The hypothesis of the change we envisage can be summarised as follows. The ERG-based split in intransitives does no longer seem to be *solely* motivated by agentivity, but rather, it has started infiltrating within the unaccusative domain, and wherein, it is being guided by a distinct semantic factor, namely ANIMACY. A discussion of the study and its results follows below.

4.4 The study

4.4.1 Methodology

All the data for this study was collected through two¹³ primary methods: a production experiment and a grammaticality judgement test. For the first part of this study, 40 children (2-6 years) whose primary language input is Assamese took part in a Contrastive Elicitation Task for Testing Case Marking (Ruigendijk, 2015). Further, to develop a set of comparable data, 22 Assamese speaking adults were asked to take part in the same elicitation task experiment. The participants in this control group were young adults (16-25 years) studying at a higher educational institute. All the participants of both the studies were from Tinsukia and Dibrugarh district in eastern Assam. A summary of both child and adult participants that took part in this study is presented in Table 4.3.

Group	No. of Participants	Mean Age	Gender	Place of Residence
Child	40	4.46	Male = 18 Female = 22	Rural = 23 Urban = 17
Adult	22	19.04	Male = 10 Female = 12	Rural = 13 Urban = 9

Table 4.3: Summary of Participants

All the participants of this production task were asked to describe 11 pairs of minimally contrastive images, some of which are shown in Figure 4.1¹⁴.

Contrastive Elicitation Task (Ruigendijk, 2015) was originally designed for two verb conditions: ditransitive and transitive. However, since what this PHD study is after, which encompasses the whole case alignment in Assamese, and with the knowledge that Assamese has an intransitivity-based split, intransitives, specially four unergative, and four unaccusative verbs were included in the stimuli, and a new set of pictures were designed to suit any Indian language and culture. The stimuli were controlled for conditions such as verb type, PERSON, NUMBER and ANIMACY to elicit target utterances for the specific case markers. The intransitive verb types, in particular, included the unergative *nas* ‘dance’, *xãtur* ‘swim’, *zopia* ‘jump’, and ‘dour’ run, and the unaccusative *por* ‘fall’, *zol* ‘burn’, *boh* ‘sit’, and *dub* ‘sink’. Out of the set of these

¹³We excluded the data from The Final Destination Board Game in the chapter because we did not observe any changes occurring in the domain of the person and number-based split.

¹⁴The entire task involved describing the illustrations of eleven different verbs in both progressive and perfect structures. However, notwithstanding the incorporation of this grammatical ASPECTual distinction, no correlation was observed with respect to SUBJ case marking, and hence we do not discuss it further.

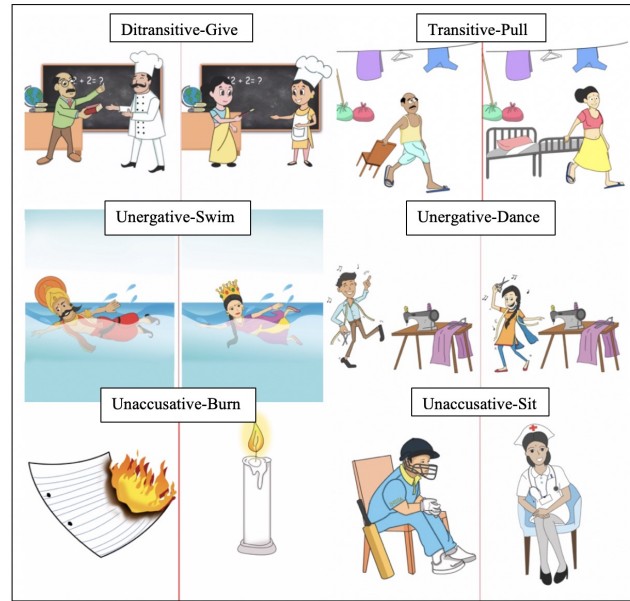


Figure 4.1: Contrastive Elicitation Task for Testing Case Marking (based on Ruigendijk (2015))

eight verbs, only the subjects of ‘fall’ and ‘sit’ were HUMAN. ANIM/HUMAN subjects for ‘burn’ and ‘sink’ were avoided given the projected violent nature, as majority of our participants were small children. All the unergative verbs in the study involved HUMAN subjects. We here deem important to reiterate why the task did not include any stimuli that involved unergative verbs with INANIM subjects in the elicitation task. This is because, as illustrated clearly in §2.1, through the pair in (77), ANIMACY plays no role in the assignment of ERG case marking. Moreover, as the hypothesis posited in §2.4 already mentions, the observed change is exclusively taking place in the domain of unaccusative verbs.

In total, we elicited 880 utterances from the forty child participants. However, the data of the two two-year old participants were excluded from statistical analysis¹⁵. The remaining data which consisted of 836 utterances from thirty-eight child participants and 484 responses from the 22 young adult participants were then coded in excel and statistically analysed using SPSS and R.

Furthermore, we also conducted a Grammaticality Judgement Test among 88 adults (15-60) to have a better understanding of the intransitivity based split in Assamese. This test was designed at a much later course of the study, when we realised that the young adults in the Contrastive Case Elicitation Task did not produce the expected target-like utterances that were based on our research of previous literature. We used a 4 point Likert scale to avoid participants choosing a mid-scale. There were 48 questions in total, out of which 11 were

¹⁵See Section 3.5.10 of Chapter 3 for the qualitative description of the two two-year old participant’s data

test items and the remaining 37 were fillers. The test items included all the target utterances from the Contrastive Case Elicitation Task. These questions were then randomised using a Latin Square design to ensure that there is no carryover effects. The stimuli was then uploaded on the online survey software Qualtrics and the links were then distributed among the participants to be completed on their mobile phones. The responses of 18 participants were excluded from the final analysis as they had either not answered all the questions or had chose only one scale for all their questions. Finally, the responses of 70 participants were exported to excel and then were analysed using SPSS.

4.4.2 Results

4.4.2.1 Results of Contrastive Case Task

All the participants of the elicitation task were observed to be adhering to the description of the transitive and ditransitive structures as provided in §4.3. However, in the context of intransitive verbs, our findings yielded a more complex and nuanced response. Since the stimuli were developed following the in-depth description in Chapter 2, which is also summarised in §4.3 of this chapter, we were expecting that the subjects of unaccusative verbs will be \emptyset , that is, NOM, while the subjects of unergative verbs will maintain their *-e* ERG marking.

Contrary to our expectations, the elicited data did not reflect such a clear agentivity-based intransitive split. In fact, we found that both children and adult participants alike were ignoring the intransitivity split discussed earlier, and were rather embracing a new case marking pattern emerging among children, which appears to be conditioned by a distinct semantic factor of the nominal, i.e ANIMACY. On the other hand, the young adult participants did not show a leaning towards this ANIMACY based split and were instead more inclined towards leaving all the SUBJ GFs of the intransitive verbs zero-marked, thereby simplifying the split on the intransitive verbs.

This divergence in behaviour is visually presented in Figure 4.2, which highlights the contrasting trends in case marking observed among children and young adults. As is evident in Figure 4.2, our young adult participants demonstrated a consistent tendency to under-extend or undersupply the ERGATIVE case marker in contexts necessitating obligatory marking, such as in the case of unergative verbs. In stark contrast, the children predominantly exhibited an opposing trend, tending to oversupply the ERGATIVE case marker even in contexts where its presence was not mandatory, as notably observed for the S_0 subjects of unaccusative verbs.

We performed a general mixed effects (binomial) analysis to assess if the use of ERGATIVE enclitics in elicited utterances were dependent on ANIMACY and/or age. The results from the child data shows that the effects of all the

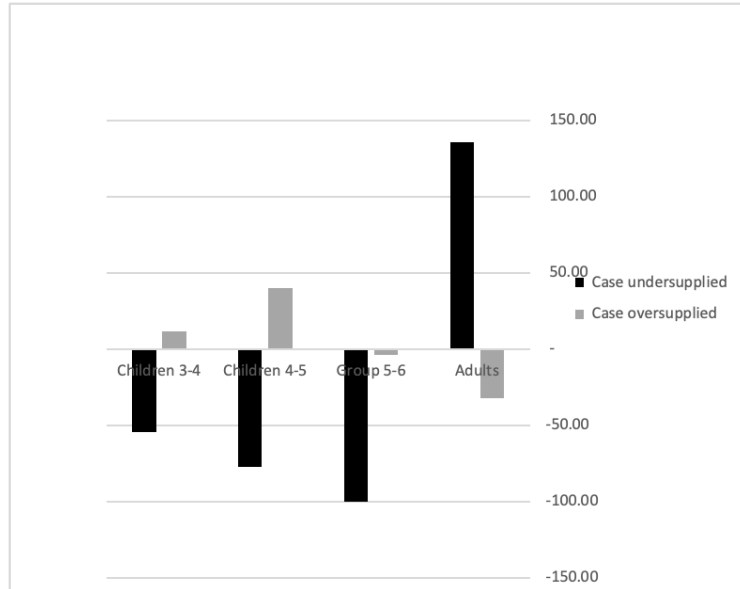


Figure 4.2: Case Elicitation Task: Error Types by Age Group

predictor variables are significant ($p = .0001$). As illustrated in Figure 4.3, children are less likely to mark the SUBJ GFs when they are inanimates¹⁶ (OR: (-2.62), $z = (12.73)$, $p = .0001$).

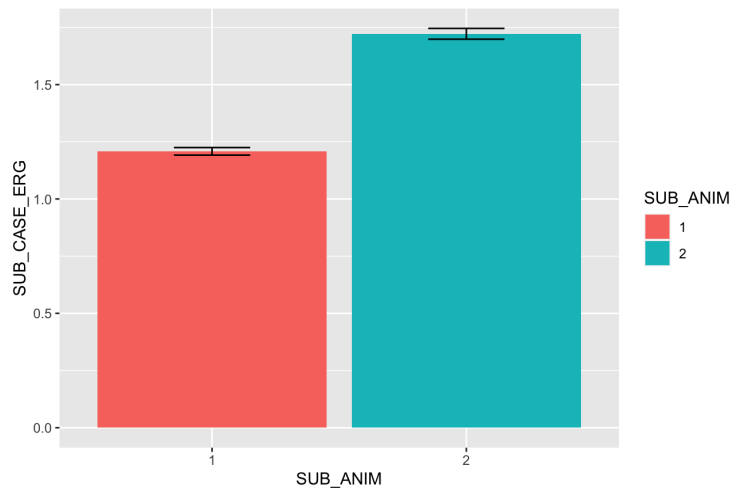


Figure 4.3: GLM Output for Adult Data

Furthermore, a significant age effect was observed within the child data ($p = .0001$), which is visually represented in Figure 4.4. In contrast, no significant age effects were measured among the adult participants ($p = .529$).

The empirical support for this observation in child data, where they consistently over-extended the ERGATIVE case enclitic in the context of unaccusative verbs with animate subjects, is highlighted through the contrast in examples

¹⁶Inanimates are coded as 2

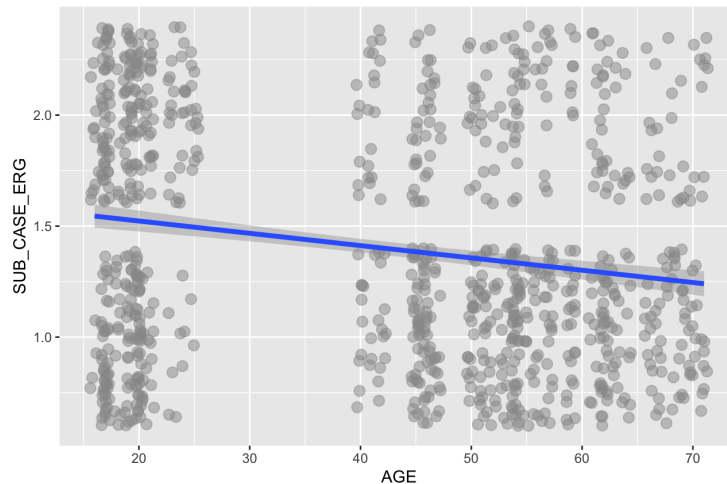


Figure 4.4: Significance of Age in Child Data

(100) and (101). Here we have the unaccusative predicates *boh* ‘sit’ and *por* ‘fall’ taking ANIM subjects where we observe the emergent *-e* ERG marking as opposed to the predicates *zol* ‘burn’ and *dub* ‘sink’ with INANIM subjects which in turn remain \emptyset -marked. In fact, all the participants consistently maintained a \emptyset marking in utterances consisting of inanimate subjects.

(100) a. k^h eluwoi-zon=e/narse-goraki=e boh-i as-e
 sportsman-CLF=ERG/nurse-CLF=ERG sit-PROG be.PRES-3
 ‘The sportsman/nurse is sitting.’

b. bimansalok-zon=e/bimansalika-goraki=e por-i as-e
 pilot.M-CLF=ERG/pilot.F-CLF=ERG fall-PROG be.PRES-3
 ‘The pilot (M/F) IS FALLING.’
 (**Emergent ANIM ERG SUBJ of unaccusative PRED**)

(101) a. k agos- k^h ɔn.∅/mom-dal.∅ zol-i as-e
 paper-CLF.NOM/candle-CLF.NOM burn-PROG be.PRES-3
 ‘The paper/candle is burning.’

b. nao- k^h ɔn.∅/bakos-tu dub-i as-e
 boat-CLF.NOM/box-CLF.NOM sink-PROG be.PRES-3
 ‘The boat/box is sinking.’
 (**INANIM NOM SUBJ of unaccusative PRED**)

We hypothesise this new emerging situation to have arisen as a result of a reanalysis of what the morphological form that is responsible for the unaccusative- unergative split, i.e. the *-e* that exists in the intransitive domain, comes to express. The ERG’s erstwhile agentive marking has, within the unaccusative domain of intransitives seemingly come to express an ANIMACY

distinction. Consequently, the split is being overhauled, in the sense that it is now being conditioned by a semantic feature in the lexical entry, rather than by a theta-role - GF association at the argument-structure.

The emerging system highlighting the trend observed in current child data is represented in Table (4.4).

Value	A	S _a	S _o
Animate	-e	-e	-e
Inanimate	-e	-∅	-∅

Table 4.4: **The emergent ANIMACY-based split**

4.4.2.2 Results of Grammaticality Judgement Test

From the data presented in Table 4.5 and Table 4.6, it becomes evident that participants aged 35-60 years were more consistent in maintaining the traditional intransitivity-based split compared to the 15-34 year olds. The 34-65 year olds performed better than the 35-60 year olds for all test items, except for the unaccusative verb *burn* where the younger group had more accurate responses. A Chi Square Test was conducted to assess the relationship between the two variables *Age Group* and *Completely Grammatical Responses*. The results of the test did not indicate a significant relationship ($p = .087$), suggesting that the differences in response accuracy across age groups may not be statistically significant.

Age Group	Test Items	Completely Grammatical	Somewhat Grammatical	Somewhat Ungrammatical	Completely Ungrammatical
15-34	Dance	19.35	16.13	22.58	41.94
	Jump	25.81	3.23	19.35	51.61
	Run	22.58	12.90	29.03	35.48
	Swim	54.84	6.45	19.35	19.35
35-60	Dance	64.10	7.69	5.13	23.08
	Jump	48.72	7.69	15.38	28.21
	Run	53.85	10.26	10.26	25.64
	Swim	43.59	25.64	20.51	10.26

Table 4.5: Summary of Grammaticality Judgement Test of Unergative Verbs

For a visual representation of the collected responses from the Grammaticality Judgement Test, please refer to Figure 4.5, which further enhances our understanding of the study's outcomes.

Age Group	Test Items	Completely Grammatical	Somewhat Grammatical	Somewhat Ungrammatical	Completely Ungrammatical
15-34	Burn	74.19	6.45	19.35	0
	Fall	54.84	16.13	12.90	16.13
	Sink	54.84	9.68	22.58	12.90
	Sit	58.06	22.58	12.90	6.45
35-60	Burn	66.67	12.82	20.51	0
	Fall	71.79	7.69	10.26	10.26
	Sink	66.67	10.26	12.82	10.26
	Sit	66.67	12.82	15.38	5.13

Table 4.6: Summary of Grammaticality Judgement Test of Unaccusative Verbs

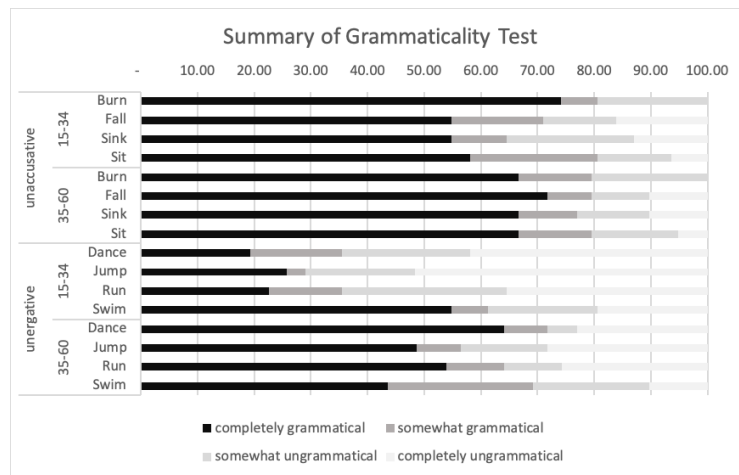


Figure 4.5: Summary of Grammaticality Judgement Test for Intransitive Verbs

4.5 Conclusion

The case alignment system in Assamese is currently undergoing a change, and there seems to be a shift in the marking of SUBJs. However, it is impossible to definitively point out when the changes in the intransitive verb based on animacy started to emerge as we do not have any clear data to represent it. As previously noted in our data from the Grammaticality Judgement Test, the older participants seemed to maintain the difference in the intransitivity based split, although the analysis did not reveal any statistically significant results. This could be assigned to the fact that there was not much difference between the two age groups in the Grammaticality Judgement Test as opposed to the age groups for the Contrastive Elicitation Task where we see a significant correlation.

Bibliography

- Ahmed, T. (2010). The unaccusativity/unergativity distinction in Urdu. *Journal of South Asian Linguistics*, 3(1).
- Aissen, J. (1999). Markedness and subject choice in optimality theory. *Natural Language & Linguistic Theory*, 17(4):673–711.
- Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory*, 21(3):435–483.
- Amery, R. (1993). An australian koine: Dhuwaya, a variety of yolu matha spoken at yirrkala in north east arnhemland. *International Journal of the Sociology of Language*, 1993(99):45–64.
- Amritavalli, R. and Sarma, P. P. (2002). A case distinction between unaccusative and unergative subjects in Assamese. *Snippets*, 5:6–7.
- Bauer, L. (2014). *Watching English change: An introduction to the study of linguistic change in standard Englishes in the 20th century*. Routledge.
- Bloch, J. (1965). *Indo-Aryan: From the Vedas to modern times*. Librairie d'Amérique et d'Orient Adrien-Maisonneuve, Paris.
- Butt, M. (1993). Object specificity and agreement in Hindi/Urdu. In *29th Regional Meeting of the Chicago Linguistic Society*, volume 1, pages 80–103.
- Butt, M. (2006). The dative-ergative connection. *Empirical Issues in Syntax and Semantics*, 6:69–92.
- Butt, M. and Ahmed, T. (2011). The redevelopment of Indo-Aryan case systems from a lexical semantic perspective. *Morphology*, 21(3-4):545–572.
- Butt, M. and Deo, A. (2001). Ergativity in Indo-Aryan. *Kurdica*, 5(3).
- Butt, M. and Holloway King, T. (1991). Semantic Case in Urdu. In *27th Regional Meeting of the Chicago Linguistic Society*, volume 27, pages 31–45.
- Camilleri, M. and Sadler, L. (2012). On the analysis of non-selected datives in Maltese. In Butt, M. and King, T. H., editors, *Proceedings of the LFG12 Conference*, pages 118–138, Stanford. CSLI Publications (online publications).
- Chatterji, S. K. (1926). *The origin and development of the Bengali language*. Calcutta University Press, Calcutta.
- Cheshire, J., Kerswill, P., Fox, S., and Torgersen, E. (2011). Contact, the feature pool and the speech community: The emergence of multicultural London English. *Journal of Sociolinguistics*, 15(2):151–196.

- Chowdhary, R. (2014). Differential marking of cases in Asamiya. *North East Indian Linguistics*, 6:107–124.
- Coghill, E. (2016). *The rise and fall of ergativity in Aramaic: Cycles of alignment change*, volume 21. Oxford University Press, Oxford.
- Comrie, B. (1978). Ergativity. In Lehmann, W. P., editor, *Syntactic Typology: Studies in the Phenomenology of Language*, pages 329–394. University of Texas Press, Austin.
- Comrie, B. (2013). Alignment of case marking of full noun phrases. In Dryer, M. S. and Haspelmath, M., editors, *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.
- Croft, W. (2003). *Typology and universals*. Cambridge University Press, Cambridge.
- Devi, J. (1986). *Ergativity: A Historical Analysis in Assamese*. PhD thesis, University of Delhi, New Delhi.
- Dixon, R. M. W. (1979). Ergativity. *Language*, 55:59–138.
- Dixon, R. M. W. (1994). *Ergativity*. Cambridge University Press, Cambridge.
- Enç, M. (1991). The semantics of specificity. *Linguistic Inquiry*, 22:1–25.
- Goswami, G. C. and Tamuli, J. (2003). Asamiya. In Cardona, G. and Jain, D., editors, *The Indo-Aryan Languages*, pages 391–443. Routledge, London.
- Gundel, J. K., Hedberg, N., and Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69:274–307.
- Haddad, Y. A. (2011). *Control into conjunctive Participle Clauses: The case of Assamese*. De Gruyter Moutone, Berlin.
- Handschuh, C. (2008). Split Marked-S Case Systems. *Scales*, 86:73–82.
- Heggarty, P. (2000). Quantifying change over time in phonetics. *Time Depth in Historical Linguistics*, 2:531–62.
- Kakati, B. (1941). *Assamese: Its Formantion and Development*. Department of Historical and Antiquarian Studies, Gauhati, Assam.
- Kerswill, P. and Williams, A. (2000). Creating a new town koine: Children and language change in milton keynes. *Language in Society*, 29(1):65–115.
- Kornfilt, J. (2009). DOM and two types of DSM in Turkish. In de Hoop H., d. S. P., editor, *Differential Subject Marking*, pages 79–111. Springer, Dordrecht.
- Kulikov, L. (2017). Indo-aryan. In Kapović, M., editor, *The Indo-European Languages*, pages 214–262. Routledge, London.

- Li, C. (2007). Split ergativity and split intransitivity in Nepali. *Lingua*, 117(8):1462 – 1482.
- Lightfoot, D. (2010). Language acquisition and language change. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(5):677–684.
- Mair, C. and Leech, G. N. (2020). Current changes in english syntax. *The Handbook of English Linguistics*, pages 249–276.
- McGregor, W. B. (2009). Typology of ergativity. *Language and Linguistics Compass*, 3(1):480–508.
- Mohanan, T. (1994). *Argument Structure in Hindi*. CSLI Publications, Stanford.
- Montaut, A. (2018). The rise of differential object marking in Hindi and related languages. In Seržant, I. A. and Witzlack-Makarevich, A., editors, *Diachrony of differential argument marking*, volume 19, pages 281–313. Language Science Press, Berlin.
- Nath, D. (2003). *Case in Assamese*. PhD thesis, Central Institute of English and Foreign Languages, Hyderabad.
- Ohala, J. J. and Jones, C. (1993). The phonetics of sound change. *Historical linguistics: Problems and perspectives*, 237:278.
- O’Shannessy, C. (2013). The role of multiple sources in the formation of an innovative auxiliary category in light warlpiri, a new australian mixed language. *Language*, pages 328–353.
- O’Shannessy, C., Disbray, S., Martin, B., and Macdonald, G. (2019). (re) turning research into pedagogical practice: A case study of translational language research in warlpiri. In L. Barwick, J. G. and Vaarzon-Morel, P., editors, *In Archival Returns: Central Australia and Beyond: LDC Special Publication*, volume 18, pages 139–51. University of Hawai’i Press, Honolulu/Sydney: University of Hawaii Press Sydney University Press.
- Queen, R. M. (1996). *Intonation in contact: A study of Turkish-German bilingual intonation patterns*. The University of Texas at Austin.
- Queen, R. M. (2001). Bilingual intonation patterns: Evidence of language change from turkish-german bilingual children. *Language in Society*, 30(1):55–80.
- Ruigendijk, E. (2015). Contrastive elicitation task for testing case marking. In Armon-Lotem, S., de Jong, J., and Meir, N., editors, *Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment*, pages 38–54. Multilingual Matters, Bristol.
- Saha, A. and Patgiri, B. (2013). Ergativity in Axomiya. *Language in India*, 13(12):35–46.

- Saikia, P. (in prep.). *Children's Acquisition of Assamese Split Ergativity*. PhD thesis, University of Essex, Colchester.
- Sankoff, G. (1990). The grammaticalization of tense and aspect in tok pisin and sranan. *Language Variation and Change*, 2(3):295–312.
- Shnukal, A. and Marchese, L. (1983). Creolization of nigerian pidgin english: a progress report. *English World-Wide*, 4(1):17–26.
- Stein, D. (2011). *The semantics of syntactic change: Aspects of the evolution of 'do' in English*, volume 47. Walter de Gruyter.
- Stroński, K. (2010). Variation of ergativity patterns in Indo-Aryan. *Poznań Studies in Contemporary Linguistics*, 46(2):237–253.
- Thampoe, H. D. (2017). *Sinhala and Tamil: A case of contact-induced restructuring*. PhD thesis, Newcastle University, Newcastle.
- van Valin Jr, R. D. (1990). Semantic Parameters of Split Intransitivity. *Language*, 66:221–260.
- Verbeke, S. and De Cuypere, L. (2009). The rise of ergativity in Hindi: Assessing the role of grammaticalization. *Folia Linguistica Historica*, 30:1–24.
- Zakharyin, B. (2015). Indo-Aryan ergativity and its analogues in languages of central and western eurasia. *Lingua Posnaniensis*, 57(2):63–76.

Chapter 5

Conclusion

This thesis was undertaken to investigate three core research questions concerning the case alignment system in Assamese and its acquisition by native speakers at an early age. These questions are outlined below:

- 1) Does Assamese conform to the nominative-accusative language pattern, as commonly believed among researchers?
- 2) To what extent do children between the ages of 2 and 6 demonstrate adult-like case marking skills?
- 3) Are there any observable changes in the current case marking system?

To address our research questions, we undertook a thorough cross-sectional study involving both children and adult participants. Our analysis in this thesis was primarily founded on the empirical data obtained through three key methods: a production experiment, a semi-structured language game, and a grammaticality judgment test. Our research findings offer insights into various aspects of the Assamese case marking system, the developmental progression of language acquisition in children, and the emergence of a new *ERGATIVE* split based on animacy. These key insights are summarized in the following section.

5.1 Summary of the Main Findings

The first research question aimed to clarify the case alignment pattern of Assamese, specifically whether it conforms to the *NOMINATIVE-ACCUSATIVE* or *ERGATIVE/split ergative* model. While there has been an ongoing debate regarding the case alignment of Assamese, our thorough investigation of the language, as well as that of the available literature, confirms that Assamese is not a *NOMINATIVE-ACCUSATIVE* language in spite of this being the most popular consensus among linguists. Instead, it exhibits a split ergative system, which is determined by intransitivity and distinctions related to *PERSON* and

NUMBER.

To put it simply, the subjects of (di) transitive and unergative verbs obligatorily receive an ERGATIVE case ending in Assamese, while the subjects of unaccusative verbs remain unmarked. On the other hand, in the case of pronominals, only the 2nd, and 3rd PERSON plural subjects trigger an ERGATIVE marking while the rest of the pronominal subjects remain unmarked. However, complexity arises when the two splits are combined. As a result of this combination, the 2nd and 3rd PERSON plural pronominal subjects of unaccusative verbs remain null-marked, even though they typically trigger obligatory ERGATIVE case marking in the contexts of (di) transitive and unergative verbs. This complexity highlights the nuanced nature of the case marking system in Assamese and shows that the unaccusative split takes precedence over the PERSON and NUMBER-based split.

The second research question aimed to investigate whether children between the ages of 2 to 6 years demonstrate adult-like case marking realizations. Our study has revealed that, despite the complexities of Assamese's split ergative system, children begin to acquire this linguistically intricate structure as early as 2;6 years, demonstrating increasing proficiency as they grow older. This finding adheres to a systematic acquisition process, mirroring patterns identified in other (split) ergative languages. Our results also suggest that children seem to grasp the Differential Object Marking (DOM) split in Assamese with relative ease, typically achieving competence by the age of 3. Furthermore, the results from this chapter shed light on an emerging split in language based on animacy that has not been explored previously.

The third research question delved into the direction of change observed in the case marking system in Assamese that was highlighted in the previous chapter. The data from the production experiment unveiled a new semantic shift underpinning this split, which is based on animacy as opposed to the previously documented intransitivity split based on agentivity. Early on in this thesis, we had introduced a split ergative model for Assamese that was based on our understanding of the language, drawing from native language intuition and existing research. However, as we delved deeper into this study, we realised that although the children still maintain a split in the intransitives, there was a stark contrast as to the factors that determined this split. Instead of adhering to the traditional agentivity-based split proposed in our model, the children seemed to maintain a more simplified split that is based on animacy. Nevertheless, despite the transition from an agentivity-based split to an animacy-based one, Assamese still retains its classification as a split ergative language.

5.2 Limitations and scope of future research

Considering the time constraints inherent to a PhD research project and my inability to spend an extended period in Assam, I opted for a cross-sectional study involving a production experiment and a semi-structured elicitation task, rather than pursuing a longitudinal study.

Additionally, due to limitations in the age range suitable for the Contrastive Case experiment, I couldn't gather data from younger children to explore the onset of case realizations. In the future, if the opportunity arises, I would like to collect naturalistic data from younger participants to comprehensively examine the entire acquisition process.

Furthermore, considering the current state of the language, which appears to be undergoing a change in terms of its split ergativity, it would be interesting to investigate the nature of adult input received by the children. Hence, for future research endeavors, I am interested in transcribing and analyzing the child-directed speech data I collected through two further semi-structured games that were played with the caregivers. Unfortunately, due to time constraints associated with transcribing such vast amount of data, this data could not be included in the present thesis.

Appendix A

Ethical Approval Application

Application for Ethical Approval of Research Involving Human Participants

This application form should be completed for any research involving human participants conducted in or by the University by masters and undergraduate students. 'Human participants' are defined as including living human beings, human beings who have recently died (cadavers, human remains and body parts), embryos and fetuses, human tissue and bodily fluids, and human data and records (such as, but not restricted to medical, genetic, financial, personnel, criminal or administrative records and test results including scholastic achievements). Research should not commence until written approval has been received (from Departmental Research Director / Ethics Officer, Faculty Ethics Committee (FEC) or the University's Ethics Committee). This should be borne in mind when setting a start date for the project.

Applications should be made on this form, and submitted electronically, to your Departmental Research Director / Ethics Officer. A signed copy of the form should also be submitted. Applications will be assessed by the Research Director / Ethics Officer in the first instance, and may then passed to the FEC, and then to the University's Ethics Committee. A copy of your research proposal and any necessary supporting documentation (e.g. consent form, recruiting materials, etc.) should also be attached to this form.

A full copy of the signed application will be retained by the department/school for 6 years following completion of the project. The signed application form cover sheet (two pages) will be sent to the Research Governance and Planning Manager in the REO as Secretary of the University's Ethics Committee.

1.

Title of project: Children's Acquisition of Assamese Case Markers

2. Principal Investigator (i.e. name of student)

Name: Ms. Pori Saikia	Department: Language and Linguistics
--------------------------	---

3. Name of supervisor(s):

Name: Prof. Louisa Sadler	Department: Language and Linguistics

4.

Proposed start date of research (note ethical approval cannot be granted retrospectively): 05.01.2017
--
5.

Probable duration: 3 years

6. Will this project be externally funded? Yes / No
If Yes,
7.

What is the source of the funding?

Appendix B

Consent Form for Child Participants

Department of Language and Linguistics, University of Essex, Consent and Ethical Approval

University of Essex

Form of Consent to take part in a Research Project (CONFIDENTIAL)

Title of Project:

Children's Acquisition of Assamese Case Markers.

Name of Researcher: Ms. Pori Saikia

Brief Outline of Project: This project is part of my PhD program at the University of Essex that aims to study how the Assamese speaking children acquire the structures of the language. For this purpose, about 60 children from Dibrugarh and Tinsukia district of Assam in the age-group of 2 to 6 years will be asked to participate in four language games.

What does participating involve?

The children will be asked to take part in four language games. Two games will involve describing pictures and two other games will involve playing with their parents/caregivers using two sets of child-friendly toys in a bag that the researcher will bring. All these sessions will be recorded using a video and audio recorder.

The identity of the children will be completely confidential, and pseudonyms will be used while transcribing and later describing their use of the language. Anonymised transcriptions of these recordings will be submitted to a data archive like CHILDES if permitted by the legal guardian of the child. If consent is given, these transcriptions will also be made available to genuine researchers for research purpose only. Further, all the recordings will be shown to the PhD supervisor and external examiner of the researcher for analysis.

If the parents/guardians are willing to let their child participate, they will be asked to read through this information sheet which will also be available in Assamese, sign the consent form and return it. At that point, their child's participation can be confirmed, and arrangements can be made for the sessions.

The sessions will take about 2 hours in total for playing all the games and filling in the consent form and a brief questionnaire about the child's language and family background. The children, however, will not be asked to play more than two games in one day, and they will be free to withdraw from any session at any point of time. Further, they will not be paid any money for participating in the study but small gifts like colouring books, toys, stationary items, etc. will be given to them at the end.

Appendix C

Consent Form for Adult Participants

Department of Language and Linguistics, University of Essex, Consent and Ethical Approval

University of Essex

Form of Consent to take part in a Research Project (CONFIDENTIAL)

Title of Project:

Children's Acquisition of Assamese Case Markers.

Name of Researcher: Ms. Pori Saikia

Brief Outline of Project: This project is part of a PhD research at the University of Essex that aims to study how the Assamese speaking children in Assam (Dibrugarh and Tinsukia district) acquire the structures of the language. As a part of this project, I will collect data from 20 adults studying at a college or university in Dibrugarh or Tinsukia district of Assam to compare their use of the Assamese language with that of the children.

What does participating involve?

This study comprises taking part in one language game where each participant will have to describe a set of not more than 15 minimally contrastive picture pairs. The session will be recorded using a video and audio recorder after written consent of the participants. Further, they will be free to withdraw from the session at any point of time.

The identity of the participants will be completely confidential, and pseudonyms will be used while transcribing and later describing their use of the language. Anonymised transcriptions of these recordings will be submitted to a data archive like Childes if permitted. These transcriptions will also be available for genuine researchers for research purpose only. The recordings, however, will be made available to the PhD supervisor and external examiner of the researcher for analysis.

The sessions will take about 30 minutes in total for playing the game and filling in the consent form. Although the participants will not be paid any money, they will be given small gifts like stationary items at the end of the study.

Appendix D

Language Background Information: Child Participants

Participant Number	Initials	Age (in months)	Gender	Location	No. of Languages Spoken	Spoken Languages	Language Used with Caregivers	Language Used with Siblings	Language Used with Friends	Language Used with Neighbours	Number of Months in School	Medium of Instruction in School
1.	IKC	40	M	Urban	1	Assamese	Assamese	NA	Assamese	Assamese	19	Assamese
2.	RPS	45	M	Rural	4	Assamese, Hindi, Sadri	Assamese, Hindi, Assamese	Assamese, Hindi	Assamese, Hindi, Sadri	Assamese, Hindi, Bengali	3	Assamese
3.	KC	45	F	Rural	2	Assamese, Hindi	Assamese	NA	Assamese, Hindi	Assamese, Hindi	16	English
4.	PC	46	F	Rural	3	Assamese, Hindi, English	Assamese, Hindi, English	NA	Assamese	Assamese	16	English
5.	AM	45	M	Urban	3	Assamese, Hindi, English	Assamese	Assamese, Hindi	Assamese, Hindi	Assamese	18	Assamese, Hindi, English
6.	DC	46	M	Rural	3	Assamese, Hindi, English	Assamese	NA	Assamese, Hindi	Assamese, Hindi	14	Assamese, Hindi, English
7.	DG	46	M	Rural	1	Assamese	Assamese	NA	Assamese	Assamese	3	English
8.	SD	42	M	Urban	2	Assamese, Hindi	Assamese, Hindi	NA	Assamese, Hindi	Assamese	15	English
9.	PK	47	F	Urban	2	Assamese, Hindi	Assamese	NA	Assamese, Hindi	Assamese, Hindi	0	Hindi, English
10.	DB	41	F	Rural	2	Assamese	Assamese, Hindi	NA	Assamese	Assamese	12	Assamese
11.	CM	50	F	Urban	3	Assamese, Hindi, Sadri	Assamese	NA	Assamese	Assamese	20	Hindi
12.	NST	54	F	Rural	3	Assamese, Hindi	Assamese	Assamese, Hindi	Assamese, Hindi	Assamese	24	English
13.	AAS	51	M	Urban	3	Assamese, Hindi	Assamese	NA	Assamese, Hindi	Assamese, Hindi	24	English
14.	RB	55	F	Rural	1	Assamese	Assamese	NA	Assamese	Assamese	12	Assamese

15.	HNB	50	F	Rural	2	Assamese, Hindi	Assamese	Assamese	Assamese, Hindi	14	English
16.	BS	54	M	Rural	2	Assamese, Hindi	Assamese	Assamese	Assamese	4	Assamese
17.	MK	56	M	Rural	2	Assamese, Hindi	Assamese	Assamese	Assamese	4	Assamese
18.	DH	59	F	Rural	1	Assamese	Assamese	Assamese	Assamese	14	Assamese
19.	PRB	51	M	Urban	2	Assamese, Hindi	Assamese	Assamese, Hindi	Assamese	27	English
20.	NG	52	F	Urban	2	Assamese, Hindi	Assamese	Assamese	Assamese	14	English
21.	AC	54	F	Urban	3	Assamese, Hindi	Assamese, Hindi, English	Assamese	Assamese, Hindi	14	English
22.	ND	53	F	Urban	2	Assamese, Hindi	Assamese	Assamese	Assamese	14	English
23.	KD	57	F	Rural	2	Assamese, Hindi	Assamese, Hindi	Assamese, Hindi	Assamese, Hindi	15	English
24.	AP	57	M	Urban	2	Assamese, Hindi	Assamese	Assamese, Hindi	Assamese, Hindi	15	English
25.	NK	62	M	Urban	2	Assamese, Hindi	Assamese	Assamese, Hindi	Assamese, Hindi	14	English
26.	SA	59	F	Urban	2	Assamese, Hindi	Assamese, Hindi	Assamese, Hindi	Assamese, Hindi	24	English
27.	MSB	68	M	Rural	1	Assamese	Assamese	Assamese	Assamese	14	Assamese
28.	NB	67	F	Rural	3	Assamese, Hindi, Bengali	Assamese	Assamese, Hindi	Assamese, Hindi, Bengali	16	English
29.	NC	64	F	Rural	1	Assamese	Assamese	Assamese	Assamese	16	Assamese
30.	PB	62	F	Rural	2	Assamese, Hindi	Assamese	Assamese	Assamese	16	Assamese
31.	SC	62	F	Rural	1	Assamese	Assamese	Assamese	Assamese	16	Assamese

32.	GJN	69	M	Rural	3	Assamese, Hindi, English	Assamese, Hindi, English	NA	Assamese, Hindi	Assamese, Hindi	16	English
33.	SI	66	F	Rural	4	Assamese, Hindi, English, Bengali	Assamese, Hindi, English, Bengali	NA	Assamese, Hindi, English, Bengali	Assamese, Hindi, English, Bengali	24	English
34.	KG	61	F	Rural	1	Assamese	Assamese	NA	Assamese	Assamese	24	Assamese
35.	AB	70	M	Urban	4	Assamese, Hindi, English, Bengali	Assamese, Hindi, English, Bengali	Assamese	Assamese, Hindi	Assamese, Bengali	28	English
36.	NG	65	M	Urban	3	Assamese, Hindi, English	Assamese, Hindi, English	NA	Hindi	Assamese	12	English
37.	BS	63	M	Urban	4	Assamese, Hindi, Bengali, Sadri	Assamese, Hindi, Bengali, Sadri	Assamese	Bengali, Sadri	Assamese, Sadri	14	Assamese
38.	KG	60	F	Urban	3	Assamese, Hindi, English	Assamese, Hindi, English	NA	Assamese, Hindi	Assamese	24	English
39.	RS	54	M	Urban	2	Assamese, Hindi	Assamese, Hindi	NA	Assamese, Hindi	Assamese	14	English
40.	MP	71	M	Urban	4	Assamese, Hindi, Bengali, Sadri	Assamese, Hindi, Bengali, Sadri	Assamese	Assamese, Hindi, Bengali, Sadri	Assamese, Hindi, Bengali	3	Assamese, Hindi, English
41.	AB	65	F	Urban	2	Assamese, Hindi	Assamese, Hindi	NA	Hindi	Assamese	24	NA
42.	AS	32	M	Rural	2	Assamese, Hindi	Assamese, Hindi	NA	Assamese, Hindi	Assamese	0	NA
43.	AB	32	M	Rural	2	Assamese, Hindi	Assamese, Hindi	NA	Assamese	Assamese	0	NA
44.	JB	35	F	Urban	2	Assamese, Hindi	Assamese, Hindi	NA	Assamese	Assamese	2	Assamese

Appendix E

Language Background Information: Adult Participants

Rural	2	Assamese, Hindi	Assamese	Assamese	Assamese	Assamese	Assamese	Assamese
Rural	2	Assamese, Hindi	Assamese	Assamese	Assamese	Assamese	Assamese	Assamese
Rural	2	Assamese, Hindi	Assamese	Assamese	Assamese	Assamese	Assamese	Assamese
Rural	3	Assamese, Hindi, Sadri	Assamese	Assamese	Assamese, Sadri	Assamese	Assamese	Assamese
Urban	2	Assamese, Hindi	Assamese	Assamese	Assamese	Assamese	Assamese	Assamese
Rural	3	Assamese, Hindi, English	Assamese	Assamese, Hindi	Assamese, Hindi	Assamese, English	English	English
Rural	2	Assamese, Hindi	Assamese	Assamese	Assamese	Assamese, Hindi	Assamese	Assamese
Rural	2	Assamese, Hindi	Assamese	Assamese, Hindi	Assamese, Hindi	Assamese, Hindi	Assamese	Assamese
Rural	3	Assamese, Hindi, English	Assamese, Hindi	Assamese, Hindi	Assamese, Hindi	Assamese	Assamese	English
Rural	3	Assamese, Hindi, English	Assamese	Assamese, Hindi	Assamese, Hindi	Assamese	Assamese	English
Urban	3	Assamese, Hindi, English	Assamese, English	Assamese, Hindi, English	Assamese, Hindi	Assamese, Hindi, English	English	English

Appendix F

SUBJ Realisations for Child Participants

No.	Initials	Age in Months	Target SUBJ NP in Assamese		Similar SUBJ NP in Assamese		SUBJ Pronoun		SUBJ in English		SUBJ Other	
			n	%	n	%	n	%	n	%	n	%
1	IKC	40	23	96	0	0.0	0	0.0	0	0.0	1	4
2	RPS	45	23	96	0	0.0	0	0.0	1	4.2	0	0.0
3	KC	45	24	100	0	0.0	0	0.0	0	0.0	0	0.0
4	PC	46	19	83	4	17	0	0.0	0	0.0	0	0.0
5	AM	45	24	100	0	0.0	0	0.0	0	0.0	0	0.0
6	DC	46	21	91	2	9	0	0.0	0	0.0	0	0.0
7	DG	46	23	100	0	0.0	0	0.0	0	0.0	0	0.0
8	SD	42	24	100	0	0.0	0	0.0	0	0.0	0	0.0
9	PK	47	19	83	4	17	0	0.0	0	0.0	0	0.0
10	DB	41	16	70	6	26	0	0.0	1	4	0	0.0
11	CM	50	16	70	6	26	0	0.0	1	4	0	0.0
12	NST	54	23	96	0	0.0	0	0.0	0	0.0	1	4
13	AAS	51	22	92	0	0.0	0	0.0	1	4	1	4
14	RB	55	24	100	0	0.0	0	0.0	0	0.0	0	0.0
15	HNB	50	19	83	4	17	0	0.0	0	0.0	0	0.0
16	BS	54	24	100	0	0.0	0	0.0	0	0.0	0	0.0
17	MK	56	19	83	4	17	0	0.0	0	0.0	0	0.0
18	DH	59	21	91	2	8	0	0.0	0	0.0	0	0.0
19	PRB	51	24	100	0	0.0	0	0.0	0	0.0	0	0.0
20	NG	52	19	83	4	17	0	0.0	0	0.0	0	0.0
21	AC	54	17	77	4	18	0	0.0	1	4	0	0.0
22	ND	53	19	83	4	17	0	0.0	0	0.0	0	0.0
23	KD	57	24	100	0	0.0	0	0.0	0	0.0	0	0.0
24	AP	57	23	96	0	0.0	0	0.0	1	4	0	0.0
25	NK	62	23	100	0	0.0	0	0.0	0	0.0	0	0.0
26	SA	62	21	84	4	16	0	0.0	0	0.0	0	0.0
27	MSB	68	24	100	0	0.0	0	0.0	0	0.0	0	0.0
28	NB	67	22	92	2	8	0	0.0	0	0.0	0	0.0
29	NC	64	24	100	0	0.0	0	0.0	0	0.0	0	0.0
30	PB	62	24	100	0	0.0	0	0.0	0	0.0	0	0.0
31	SC	62	24	100	0	0.0	0	0.0	0	0.0	0	0.0
32	GJN	69	22	96	0	0.0	0	0.0	1	4	0	0.0
33	SI	66	23	100	0	0.0	0	0.0	0	0.0	0	0.0
34	KG	61	20	99	1	4	0	0.0	1	4	0	0.0
35	AB	70	22	92	0	0.0	0	0.0	2	8	0	0.0
36	NG	63	19	87	4	17	0	0.0	0	0.0	0	0.0
37	BS	54	24	100	0	0.0	0	0.0	0	0.0	0	0.0
38	KG	71	20	84	4	17	0	0.0	0	0.0	0	0.0
39	RS	32	3	20	11	73	1	6.7	0	0.0	0	0.0
40	MP	35	3	17	15	83	0	0.0	0	0.0	0	0.0
Total			828	89.3	85	9.2	1	0.1	10	1.1	3	0.3