

Niladri Sekhar Dash
S. Arulmozi
N. Ramesh *Editors*

Handbook on Endangered South Asian and Southeast Asian Languages

Handbook on Endangered South Asian and Southeast Asian Languages

Niladri Sekhar Dash · S. Arulmozi · N. Ramesh
Editors

Handbook on Endangered South Asian and Southeast Asian Languages

Editors

Niladri Sekhar Dash
Linguistic Research Unit
Indian Statistical Institute
Kolkata, India

S. Arulmozi
Centre for Applied Linguistics
and Translation Studies
University of Hyderabad
Hyderabad, India

N. Ramesh
Centre for Applied Linguistics
and Translation Studies
University of Hyderabad
Hyderabad, India

ISBN 978-3-031-80751-0 ISBN 978-3-031-80752-7 (eBook)
<https://doi.org/10.1007/978-3-031-80752-7>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2025

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

Dedicated to



*This little Sabar girl who can smile against
all odds!*

Preface

The standard praxis of language documentation faces different kinds of challenges in the endangered languages of South Asia and Southeast Asia. These differences are caused due to differences in perspectives, purposes, methods, goals, and implementation of processes. The challenges that we normally find among South Asian and Southeast Asian languages are unique in nature due to the presence of different kinds of linguistic and extralinguistic factors and elements relating to the languages and communities. Moreover, ecological, geoclimatic, and sociocultural environments pose challenges to data collectors and data analysts. Many a time, we find that people who are trying to work on minority and endangered languages of South Asia face some typical and unique language- and culture-specific challenges that are not found in other parts of the world.

Keeping these in mind, we focus on unique aspects and issues that are related to language documentation of South Asian and Southeast Asian languages. Although our primary goal is to highlight the issues and challenges faced in the documentation and digitization of minority languages of South Asia and Southeast Asia, we also want to address some more challenging issues like writing grammar, compiling word-books and dictionaries, making communities aware about their languages, addressing language policies and planning. This book not only presents a general description of minority languages and analysis of some linguistic features of endangered South Asian and Southeast Asian languages, but also refers to theoretical, methodological, and applicational issues and challenges linked with works of collection, management, processing, analysis, and utilization of language data and information from some endangered languages of South Asia and Southeast Asia.

UNESCO's 'Atlas of the World's Languages in Danger' raises awareness about the issue of 'language endangerment'. It also invites people across the globe to make realistic attempts to safeguard the world's linguistic and cultural diversities. The open call is extended not only to policy-makers, linguists, and speakers of the communities but also to the general public, private agencies, government departments, and non-governmental organizations, who can join the campaign to protect and preserve endangered languages and help in the preservation and promotion of linguistic diversities at the global level. In several studies by UNESCO, it is reported

that more than 50% of the world's endangered ethnic languages are actually located in eight countries: India, Brazil, Mexico, Australia, Indonesia, Nigeria, Papua New Guinea, and Cameroon. These countries and around them are the areas that are the most linguistically diverse in the world.

Keeping this in view, this volume reports on some of the endangered languages of the region to show how these languages need attention and collective effort for documentation, protection, preservation, and promotion so that these languages, and their speakers, can overcome threats of death and extinction in new global ecosystem. To achieve this mission, this volume includes unpublished, original research works, and field studies that present before us many unknown events, facts, and data about the indigenous endangered languages of South Asia and Southeast Asia. It gives us a great opportunity to know how the speakers are struggling with their language and life, how we need to rethink to improve the present condition of the languages, and how we should design new methods and policies for their sustenance and revitalization.

This handbook has some specific goals. It includes a not-so-elaborate description of some of the endangered languages of South Asia and Southeast Asia. The descriptions primarily include analysis and presentation of language data of these endangered languages. It records and highlights the linguistic features and properties as well as extralinguistic aspects and issues that are linked with the protection, preservation, and promotion of endangered South Asian and Southeast Asian languages. It includes an ethnolinguistic profile of an endangered language with reference to its name, official status, state of endangerment, demography, usage variation, and other issues of the language and speakers. It also discusses the purposes of the study, goals and objectives, unique linguistic features, and visibility of a language within a larger geo-cultural context, and refers to the unique linguistic identity with respect to other languages. It addresses the issues of planning for field surveys, questionnaires to be used, hurdles faced during data collection, and challenges involved in the documentation of linguistic and non-linguistic data from communities. The description and analysis of language data include phonology, morphophonemics, morphology, word class analysis and description, morphosyntax, sociolinguistic cues, major findings from analysis, and new findings from studies. It shows how linguistic data and information gathered from surveys are utilized to understand the language and the society to reflect on communities that carry many new linguistic features and social practices. The databases that are presented here may be used for making knowledge-based resources like dictionaries, grammar, and language planning with benefits for target communities.

The volume makes some theoretical contributions to this particular field of study to show how new methods and approaches are used to study indigenous languages and how these methods can bring change in theoretical understanding of methods and ways used to document endangered indigenous languages, analysis of indigenous linguistic data and information, and generation of usable linguistic resources and knowledge texts for empowering the linguistically backward and economically weaker communities.

This is the first book of its kind to present studies on endangered languages of South Asia and Southeast Asia. There has never been any effort of this kind for

compiling a handbook on South Asian and Southeast Asian endangered languages although scholars have made several efforts to represent one or two languages in their individual chapters. This volume becomes a significant contribution in the ten years (2022–2032) declaration by UNESCO as the ‘decade of endangered languages’, where the goal is to collect language and culture-related data from indigenous languages and record their ethnic identities and knowledge resources. This volume directly serves this mission with a representation of linguistic features of several endangered indigenous languages. This volume will partly serve the effort for documentation of endangered South Asian and Southeast Asian languages, which has been long due in language documentation. The scholars who are working on various endangered languages will find this book useful for their reference, research, and academic activities.

Kolkata, India
Hyderabad, India
Hyderabad, India
August 2024

Niladri Sekhar Dash
S. Arulmozi
N. Ramesh

Contents

1	Preface: Documentation of Endangered South Asian and Southeast Asian Languages: A Saga of Endless Struggle and Marginal Success	1
	Niladri Sekhar Dash, S. Arulmozi, and N. Ramesh	
2	Noun Morphology in Hajong	17
	Albina Narzary and S. Arulmozi	
3	The Constraints of Linguistic Fieldwork and Documentation Concerning the Languages of Nagaland	55
	Dinkur Borah and Amalesh Gope	
4	Language Shift from Balochi to a Regional Language: Language Ideologies and the Symbolic Power of the Sindhi Language in Sindh, Pakistan	73
	Ameer Ali and Maya Khemlani David	
5	The Punchi Language: A Typological Sketch	91
	Amitabh Vikram Dwivedi and Ayushi	
6	Language Shift: Intergenerational Language Choice in Family Discourse Among the Plain Tiwas of Assam	119
	Ananya Baruah and Ajit Kumar Baishya	
7	Language Contact and Borrowing: A Case of Nominal Compounds in Kumaonī	131
	Anshikha Adhikari	
8	Preserving Indigenous Languages in Indonesia to Maintain the Plurality of Indonesian Linguistic Canvas	155
	Gatut Susanto, Hero Patrianto, Awaludin Rusiandi, and Suparmi	
9	Juray: An Endangered Variety of the Sora Group of Lects	173
	Mendem Bapuji, Opino Gamango, and P. Phani Krishna	

10	Ethnolinguistic Revitalization Among Nepali Communities in Darjeeling: In Search of Forgotten Identity	211
	Birendra Bhujel and Sweta Sinha	
11	Kom: An Endangered Language of Manipur, India	229
	Nilanjana Roy Chowdhury and Awadesh Kumar Mishra	
12	Dawoodi: A Highly Endangered Language in Northern Pakistan	247
	Chris Donlay and Zafeer Hussain Kiani	
13	Bugun: An Endangered Language and Community of Arunachal Pradesh, India	275
	Madhumita Barbora	
14	Developing a Dictionary for Kharia Sabar: An Indigenous and Endangered Tribal Speech Community of Eastern India	309
	Niladri Sekhar Dash, Meghna Majumder, and Srija Deb	
15	Kadar—An Endangered Dravidian Tribal Language of India	331
	P. Chandramohan	
16	The Bidayuh Language of Sarawak, Malaysia: Language Use and Proficiency	365
	Su-Hie Ting, Teresa Wai See Ong, and Florence Gilliam Kayad	
17	Assessing Language Endangerment in Northeast India Through a Four Language Prism	383
	Suranjana Barua and Dimple Choudhury	

Editors and Contributors

About the Editors

Niladri Sekhar Dash Ph.D. is a Professor and Head of the Linguistic Research Unit, the Indian Statistical Institute, Kolkata. Since 1992, he has been working in the areas of Corpus Linguistics and Language Technology, Language Documentation and Digitization, Digital Lexicography, Computer Assisted Language Teaching, Digital Humanities and Ethnography, and Dementia of Bilingual Aphasic and Alzheimer's patients. Prof. Dash is an International Visiting Fellow of the British Academy, UK (2018), Visiting Scholar of the Language and Brain Laboratory, University of Oxford, UK (2019), and a Visiting Scholar at the Institute of Education, University College London, UK (2023). Since 2018, he has been a Visiting Scientist at the School of Psychology and Clinical Language Sciences, University of Reading, UK (2018–2027). He acts as the Principal Investigator for more than 30 research-based projects funded by the Government of India, international funding agencies, and corporate houses. To his credit, Prof. Dash has published 20 research monographs/books and 290+ research papers in international and national journals, edited anthologies, and conference proceedings. He has delivered academic lectures and taught courses in more than 50 universities and institutions in India and abroad. Details: <https://lru.isical.ac.in/lru-faculty>

S. Arulmozi Ph.D. is a Professor and Head at the Centre for Applied Linguistics and Translation Studies, University of Hyderabad, Hyderabad, India. He obtained his Ph.D. from the University of Hyderabad in 1999. Before joining his current position, he worked at Dravidian University where he carried out some mission projects on Telugu WordNet and Indian Languages Corpora Initiative (Telugu) funded by the Department of Information Technology, Ministry of Electronics and Information Technology, Government of India. To his credit, he has several books and research papers published in reputed national and international journals. His areas of research include Corpus Studies, Language Endangerment Studies, and Sociolinguistics. <https://centres.uohyd.ac.in/calts/people/>

N. Ramesh Ph.D. is an Associate Professor at the Centre for Applied Linguistics and Translation Studies, University of Hyderabad, Hyderabad, India. He obtained his Ph.D. from the Bharathiar University, Tamil Nadu, India, in 1997. Before joining the current position, he worked at the Bharathiar University where he carried out several projects related to tribal languages and cultures of the Nilgiri regions. His areas of research include Corpus Studies, Language Endangerment Studies, and Sociolinguistics. <https://centres.uohyd.ac.in/calts/people/>

Contributors

Anshikha Adhikari Department of General Studies, University of Engineering and Management, Kolkata, India

Ameer Ali University of Sindh, Jamshoro, Pakistan

S. Arulmozi Centre for Applied Linguistics and Translation Studies, University of Hyderabad, Hyderabad, India

Ayushi School of Languages & Literature, Shri Mata Vaishno Devi University, Katra, JKUT, India

Ajit Kumar Baishya Department of Linguistics, Assam University, Silchar, India

Mendem Bapuji Linguistics Research Unit, Indian Statistical Institute, Kolkata, West Bengal, India

Madhumita Barbora Department of Linguistic and Language Technology, Tezpur University, Tezpur, Assam, India

Suranjana Barua Department of Humanities and Social Sciences, Indian Institute of Information Technology Guwahati, Assam, India

Ananya Baruah Department of Linguistics, Assam University, Silchar, India

Birendra Bhujel Department of Humanities and Social Sciences, Indian Institute of Technology Patna, Patna, Bihar, India

Dinkur Borah Department of Linguistics and Language Technology, Tezpur University, Tezpur, Assam, India

P. Chandramohan CAS in Linguistics, Annamalai University, Chidambaram, Tamilnadu, India

Dimple Choudhury Faculty of Humanities and Social Sciences, Assam down town University, Assam, India

Nilanjana Roy Chowdhury English and Foreign Languages University, Shillong Campus, Meghalaya, India

Niladri Sekhar Dash Linguistic Research Unit, Indian Statistical Institute, Kolkata, Kolkata, India

Maya Khemlani David University of Malaya, Kuala Lumpur, Malaysia

Srija Deb Linguistic Research Unit, Indian Statistical Institute, Kolkata, Kolkata, India

Chris Donlay Department of Linguistics and Language Development, San Jose State University, San Jose, USA

Amitabh Vikram Dwivedi School of Languages & Literature, Shri Mata Vaishno Devi University, Katra, JKUT, India

Opino Gamango Centre for ALTS, University of Hyderabad, Hyderabad, India

Amalesh Gope Department of Linguistics and Language Technology, Tezpur University, Tezpur, Assam, India

Florence Gilliam Kayad Faculty of Education, Language and Communication, Universiti Malaysia Sarawak, Kota Samarahan, Sarawak, Malaysia

Zafeer Hussain Kiani Department of English, University of Azad Jammu & Kashmir, Muzaffarabad, Pakistan

Meghna Majumder Linguistic Research Unit, Indian Statistical Institute, Kolkata, Kolkata, India

Awadesh Kumar Mishra English and Foreign Languages University, Shillong Campus, Meghalaya, India

Albina Narzary Centre for Applied Linguistics and Translation Studies, University of Hyderabad, Hyderabad, India

Teresa Wai See Ong Teaching & Learning Centre, Singapore University of Social Sciences, Singapore, Singapore

Hero Patrianto Balai Bahasa Provinsi Jawa Timur (Language Agency of East Java Province), Pasuruan, Indonesia

P. Phani Krishna Department of Linguistics, University of Delhi, Delhi, India

N. Ramesh Hyderabad Central University, Hyderabad, India

Awaludin Rusiandi Balai Bahasa Provinsi Jawa Timur (Language Agency of East Java Province), Pasuruan, Indonesia

Sweta Sinha Department of Humanities and Social Sciences, Indian Institute of Technology Patna, Patna, Bihar, India

Suparmi Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia (State Islamic University of Maulana Malik Ibrahim Malang), Malang, Indonesia

Gatut Susanto Universitas Negeri Malang (State University of Malang), Malang, Indonesia

Su-Hie Ting Faculty of Education, Language and Communication, Universiti Malaysia Sarawak, Kota Samarahan, Sarawak, Malaysia

Chapter 1

Preface: Documentation of Endangered South Asian and Southeast Asian Languages: A Saga of Endless Struggle and Marginal Success



Niladri Sekhar Dash, S. Arulmozi, and N. Ramesh

The Mission

The standard processes of language documentation that are followed in advanced countries encounter different kinds of challenges in the documentation of endangered languages of South Asia and Southeast Asia. These differences are caused due to differences in perspectives, purposes, methodologies, goals, and implementation of documentation techniques. In simple terms, the challenges that are normally found among the South Asian and Southeast Asian languages are unique in nature due to the presence of several different kinds of linguistic and extralinguistic factors and elements related to these languages and communities. Moreover, the ecological, geoclimatic, and sociocultural environments pose serious problems to data collectors who are engaged in linguistic field surveys for language data and information collection, data analysis, and information interpretation. Many a time, it is noted that the people who are trying to work on minority and endangered languages of South Asia often face some typical and unique language and culture-specific challenges that are not found in other parts of the world.

Keeping many of these events in mind, the present anthology desires to focus on the unique aspects and issues that are related to language documentation activities of the South Asian and Southeast Asian languages. Although the primary goal of this anthology is to highlight the issues and challenges faced in the documentation and digitization of minority languages of South Asia and Southeast Asia, it also

N. S. Dash (✉)
Indian Statistical Institute, Kolkata, India
e-mail: ns_dash@yahoo.com

S. Arulmozi · N. Ramesh
Hyderabad Central University, Hyderabad, India
e-mail: arulmozi@uohyd.ac.in

desires to address some more challenging issues like the development of descriptive grammars, compilation of wordbooks and dictionaries, making people aware about their languages and issues involving language policies and planning for the endangered languages. This volume not only desires to present a general description of a minority language and description and analysis of some linguistic features of endangered South Asian and Southeast Asian languages, but also desires to highlight the theoretical, methodological, and applicational issues, aspects, and challenges that are directly linked with the activities of collection, management, processing, analysis, and utilization of language data and information from the endangered languages of South Asia and Southeast Asia.

The Background

The UNESCO's 'Atlas of the World's Languages in Danger' is intended to raise awareness about the issue of 'language endangerment'. It also desires to invite people across the globe to make realistic attempts to safeguard the world's linguistic and cultural diversities. The open call is extended not only to policy-makers, linguists, and speakers of the communities but also to the general public, private agencies, government departments, and non-governmental organizations, who can actively join in the campaign to protect and preserve endangered languages and extend all possible help in preservation and promotion of linguistic diversities at the global level. In several studies by UNESCO, it is reported that more than 50% of the world's endangered ethnic languages are actually located in eight countries: India, Brazil, Mexico, Australia, Indonesia, Nigeria, Papua New Guinea, and Cameroon. These countries and around them are the areas that are the most linguistically diverse in the world.

Keeping these reports in view, the present volume (*Handbook on Endangered South and Southeast Asian Languages*) reports on some of the endangered languages of these regions to show how these languages need attention and collective effort for their documentation, protection, preservation, and promotion so that these languages, as well as their speakers, can defy the threats of death and extinction to survive and sustain in the new global ecosystem. To achieve this mission, this handbook includes several unpublished, original research works and field studies that present before us many unknown events, facts, and data about the indigenous endangered languages of South and Southeast Asia. It gives us a great opportunity to know how the speakers are struggling with their language and life, how we need to rethink to improve the present condition of their language, and how we should design new methods and policies for their sustenance and revitalization.

The Goals

The proposed handbook has the following specific goals:

- It includes a not-so-elaborate description of some of the endangered languages of South and Southeast Asia. The descriptions primarily include analysis and presentation of language data of these endangered languages.
- It records and highlights the linguistic features and properties as well as extralinguistic aspects and issues that are linked up with the protection, preservation, and promotion of endangered South and Southeast Asian languages.
- It refers to the typical and unique language and culture-specific issues as well as some ecological, geoclimatic, and sociocultural factors that pose challenges in language data and information collection, data analysis, and information interpretation by scholars working on endangered South and Southeast Asian languages.
- It focuses on some of the basic linguistic and extralinguistic aspects and issues relating to the documentation, analysis, and processing of endangered South and Southeast Asian languages and presents the same for global access.
- It addresses theoretical, methodological, and applicational issues, aspects, and challenges that are linked with the collection, management, processing, annotation, analysis, and utilization of data and information from endangered South and Southeast Asian languages for different sociopolitical works like language planning and education.
- Most of the chapters include the ethnolinguistic profile of an endangered language with reference to its name, official status, state of endangerment, demography, usage variation, and other issues of the language and speakers.
- Some of the chapters also discuss the purposes of the study, goals, and objectives, unique linguistic features, visibility of the language within a larger geo-cultural context, and refer to the unique linguistic identity with respect to other languages.
- Some chapters also address the issues for planning for a linguistic field survey, a questionnaire to be used, major issues faced during data collection, and the major challenges involved in the documentation of linguistic and non-linguistic data from the communities.
- Description and analysis of language data include phonology, morphophonemics (phonological processes), morphology, word class analysis and description, morphosyntax, sociolinguistic issues, major findings from analysis, and new findings from studies.
- Some chapters show how linguistic data and information gathered from surveys may be utilized to understand the language and the society to reflect on the community which carries many new linguistic features and social practices. The databases that are presented in these chapters may also be used for making knowledge-based resources like dictionaries and grammar as well as for language planning, which have direct benefits for the target indigenous communities.

- Finally, it makes some theoretical contributions to this field of study to show how new methods and approaches are being used to study Indigenous languages and how these methods bring a notable change in theoretical understanding of methods and strategies used in the documentation of endangered indigenous languages, analysis of indigenous linguistic data and information, and generation of usable linguistic resources and knowledge texts for empowering the linguistically backward and economically weaker speech communities.

Chapter Summary

In Chap. 2 (*Noun Morphology in Hajong*), Albina Narzary and S. Arulmozi present a descriptive analysis of the Hajong noun morphology, one of the small minority language groups living in Assam, Meghalaya, Arunachal Pradesh, West Bengal in India, and Bangladesh. Morphologically, nouns are marked by the following categories: gender, number, and case. In Hajong, a noun is a word that is generally inflected for a number or case and is followed by noun attributes such as adjectives, numerals, quantifiers, and so on. Although earlier works cover a wide range of case marking in Hajong, in most cases, they fail to reflect on various aspects linked with case marking. The present study seeks to examine this particular aspect in Hajong with reference to the case marking systems of the language. The nominal case system in Hajong records seven major cases, namely, nominative, objective, genitive, allative, ablative, instrumental, and locative. The case formatives have possible cognates from the Indo-Aryan and Tibeto-Burman languages. The pronouns of this language are differentiated by three persons and two numbers, although the second and third person lacks any formality distinction (e.g., *honorific* and *non-honorific*) and, therefore, are evenly inflected for the cases. The language possesses the natural gender system while its noun phrases can be modified by modifiers. The pre-nominal modifiers include determiners, adjectives, and adjective phrases. In the case of usage, the determiners and modifiers precede the head in Hajong. The data that are presented in this paper is actually elicited from the informants of the younger generation of the targeted community living in various areas in the Goalpara district of Assam, India.

In Chap. 3 (*The Constraints of Linguistic Fieldwork and Documentation Concerning the Languages of Nagaland*), Amalesh Gope and Dinkur Borah briefly discuss the scenario of linguistic diversity and endangerment and various challenges (including linguistic and non-linguistic factors) that are faced by field linguists while conducting fieldworks in the state of Nagaland, India. Nagaland is the 26th largest state, covering 16,579 square kilometres of land in the eastern part of India. The state is considered a region of 25 major Indian languages (Census Report of India, 2011), including several undocumented languages. Even though the state shares a small part of the land, the state has a high number of indigenous languages, many of which may be tagged ‘endangered’. Besides, the majority of the indigenous languages are not yet recorded following the standard methods applied in language documentation.

The state is quite unique in linguistic diversities and cultural varieties due to its location in mountainous geographical contours. In fact, the diverse geographical location of the state and the mountainous roadway communication make the entire process of language documentation a task with several constraints and logistic challenges. The state flaunts a unique cultural-cum-economic identity due to its indigenous cultural practices such as head-hunting, self-dependency on agriculture, foraging, and dependency on traditional medicinal herbs. Moreover, customized household necessities also play a crucial role in minimizing the scope for intervillage communication. Keeping these aspects in view, this chapter elaborates on some personal data collection experiences when the data collection team worked closely with several indigenous language communities of Nagaland and other northeastern states of India.

In Chap. 4 (*Exploring the Reasons for Language Shift in Larkana's Baloch Community in the Sindh Province of Pakistan*), Ameer Ali and Maya Khemlani David focus on the reasons why the Baloch community living in district Larkana in Pakistan's Sindh province have shifted away from the habitual use of their heritage language. There is no documented evidence to trace the beginning of the migration of the Baloch community into the Sindh Province. It has been assumed that many of them have fled from their one-time home, Baluchistan, a province in Pakistan, due to various ethnological and geoclimatic factors like tribal conflicts, extreme weather conditions, and natural disasters. At various points in time these factors, individually or collectively, have played crucial roles in their migration to Larkana and other parts of Sindh. Today, the Baloch community considers Sindh as their homeland, and they live in different villages and cities in Sindh. The authors in this paper focus their study on the Baloch community in Larkana to understand the event of language shift occurring from one neglected language (Balochi) to another neglected language (Sindhi). They want to investigate why in a region dominated by English and Urdu, a language shift takes place towards the Sindhi language which is not a dominant language of the region. The findings of the issues and reasons may help language revitalizers understand that dominance and economic reasons are not always the factors causing language shifts within a community. Interviews with members of the Baloch community also help them to trace some other reasons for the language shift and they argue that there are cultural, solidarity, and psychological factors behind language shift.

In Chap. 5 (*The Punchi Language: A Typological Sketch*), Amitabh Vikram Dwivedi and Ayushi report on the typological features of the Punchi language, which is at present considered as one of the lesser-known and vulnerable (at risk of becoming endangered in the near future) languages in India. This language belongs to the Indo-Aryan family of languages and is spoken primarily in the Poonch district in the union territories of Jammu and Kashmir in the Republic of India. Unlike the majority of the Indo-Aryan languages, the Punchi language does not find any mention as a language in the official records and documents (e.g., in the 8th Schedule of the Constitution of India). The authors present a synchronic account of the distinctive grammatical properties of the Punchi language such as its diverse pronominal system, subject-auxiliary agreement, optional ergative marking, elaborate system of verb morphology, and notable syntactic aspects. The majority of the data reported in

this paper are elicited from the speakers through linguistic field surveys. Linguistically, the Punci is a verb-final language (SOV type) with a relatively flexible word order. It exhibits a complete set of nasalized vowels that produce phonemic contrast with oral vowels. It also shows that all aspirated voiced and voiceless consonants are in contrastive distribution. At the morphology level, a high amount of inflection is observed in nominal and verbal morphology as the Punci nouns and pronouns inflect for number, gender, and person. This language employs postpositions while nouns and pronouns are found to be used in the oblique case. An empirical record of this kind extends a positive help in the development of descriptive grammar for the language and works as a constructive step towards the protection and preservation of this vulnerable language.

In Chap. 6 (*Language Shift: Intergenerational Language choice in Family Discourse Among the Plain Tiwas of Assam*), Ananya Baruah and Ajit Kumar Baishya inform that in a multilingual setting, there is a chance for minority languages to fall into the clutches of very common linguistic phenomena such as language shift, language variation, and language identity loss. They inform that the Tiwa language community, which is living in the plains of Assam (also known as the *Datiyalis*), provides a serious example in this regard. Most of the plain Tiwas, which is considered an endangered language by UNESCO, have stopped using their mother language and have shifted to the dominant language of the state (i.e., Assamese) owing to various political, economic, and social reasons. Since intergenerational language transmission is considered to be the ‘golden standard’ for the survival and sustenance of an endangered language (Fishman 1991), the event of language transmission from one generation to the next generation becomes an area of serious investigation to understand how it takes place and how it helps in sustenance of an endangered language. Moreover, since it has been argued that an endangered language can survive through intergenerational language transmission, the processes of transmission need to be critically investigated to observe whether it leads to language convergence, language divergence, language shift, or language death. This paper sheds some light upon the intergenerational language choice in families as well as intra-community discourse such as peer interactions. Moreover, it delves into some noted sociolinguistic phenomena (e.g., *language shift, variation*) to explore how these are taking place in the Tiwas community. The standard research methods such as *conversational recordings* and *participant observation* are applied to examine if processes such as code-switching and code-mixing are taking place in the language of the community. The paper also informs us how questionnaire-based data elicitation and analysis help in studying the attitudes of the speakers towards their own language.

In Chap. 7 (*Language Contact and Borrowing: A Case of Nominal Compounds in Kumaoni*), Anshikha Adhikari discusses the compounds that are borrowed in Kumaoni due to language contact. It also looks into the creation of mixed nominal compounds in Kumaoni, one of the less-known Indo-Aryan languages spoken in the Kumaon Region of the state of Uttarakhand in northern India and parts of the Doti Region in Western Nepal. The Kumaon region has remained a place of dynamic contact among Kumaoni and non-Kumaoni speakers. The language is under contact influence of dominant languages due to various socioeconomic and sociocultural

reasons and as a result of this, the phenomenon of bilingualism is a characteristic feature that is common among the Kumaoni speakers. It is noted that although Kumaoni and Hindi have their own specialized domains of application, there is a significant overlap between the two. The newly attested data reveal the recurrent use of Kumaoni-English and Persian-Kumaoni compounds. Keeping these issues in view, this chapter deals with two main issues, namely, the description of nominal compounds borrowed into Kumaoni from various languages and the description of the creation of mixed compounds in Kumaoni. In an organized manner, the author provides an ethnolinguistic profile of Kumaoni including an introduction to the language, number of speakers, demography, and linguistic situation of Kumaoni; looks into literature review of borrowing as a result of contact and compound formation in Kumaoni; refers to the data collection methods and the secondary sources used in for this purpose; gives a brief introduction to the types of compounds found in Kumaoni; analyses and classifies compounds borrowed from Sanskrit, Perso-Arabic, Arabic, English, and Hindi; and discusses the patterns of formation of mixed compounds. In essence, this chapter contributes to the studies in contact-induced unidirectional lexical borrowings and compound-internal language mixing in different language pairs.

In Chap. 8 (*Preserving Indigenous Languages in Indonesia to Maintain the Plurality of Indonesian Linguistic Canvas*), Gatut Susanto, Hero Patrianto, Awaludin Rusiandi, and Suparmi inform us that Indonesia is globally known as a plurilingual country. As the fourth largest population in the world, Indonesia is home to many ethnic groups and hundreds of indigenous languages. Based on the Indonesian Center of Statistics Bureau (*Biro Pusat Statistik* or BPS) in 2010, it is reported that there are 1340 ethnic groups in Indonesia. With regard to indigenous languages in Indonesia, the Summer Institute of Linguistics (SIL) informs that there were at least 735 indigenous languages in Indonesia in 2018. Nonetheless, according to the National Language Development and Advising Board (NLDAB), there are at least 139 indigenous languages that are nearly non-existent at present. The existence of these indigenous languages acts as valuable inputs for cultural heritage and a mirror of linguistic pluralism in Indonesia. In order to maintain the linguistic diversity of Indonesia, they argue that the indigenous languages need to be protected and preserved. This paper reports some endeavours undertaken by the academics and government of Indonesia in the preservation of the indigenous languages of the country. This paper also reports on the results conducted through a documentary study. The results indicate that there are several activities as well as endeavours to preserve the indigenous languages in Indonesia, which can be divided into three important activities, namely language mapping, language vitalization, and language registration. The paper ends with practical and essential recommendations related to the results of this study.

In Chap. 9 (*Juray: An Endangered Variety of the Sora Group of Lects*), Mendem Bapuji, Opino Gamango, and P. Phani Krishna inform that the Juray language is one of the endangered ethnic languages which is rich with indigenous knowledge about ecology, environment, and ethnomedicine which need immediate recording and documentation. The Juray is one of the six varieties of the Sora group of lects

used in the central regions of the state of Odisha in India. According to their findings, in comparison to its sister varieties, the Juray language has not been studied in an extensive manner due to which many of its unique linguistic features are not yet known to the people. The language is so rich with resources that it can be studied either from a linguistic point of view or the perspective of anthropolinguistics. The variety, which is referred to in this paper is found mainly in the Gajapati District in the southern part of the state of Odisha. The language is a severely endangered language among the South Munda languages. Since the language has various interesting grammatical features, which are not been documented so far, the authors give a comprehensive description of the ethnolinguistic profile of the language and a few aspects of its noun morphology. Their main aim is to discuss the linguistic as well as non-linguistic and cultural behaviour of the people who use Juray as their mother language. Apart from the ethnolinguistic profile, they also highlight the major challenges encountered during the field studies. Finally, they discuss the issues of language endangerment and propose some steps to be taken for the sustainable development of the language.

In Chap. 10 (*Ethnolinguistic Revitalization among Nepali Communities in Darjeeling: In Search of Forgotten Identity*), Birendra Bhujel and Sweta Sinha elaborate on the ethnolinguistic revitalization movement among the Nepali ethnic groups that have been noticed in the last few decades in the district of Darjeeling in the state of West Bengal, India. The existing studies claim the movement to be an ethnic politics for scheduled tribe (ST) status. Historically, the multilingual texture of the place turned into a linguistically homogenized territory as various ethnic groups of Nepali origin adopted Nepali as their lingua franca. The Nepali language hegemonized all other ethnic languages and the Nepali ethnic groups hampered their ethnolinguistic identities for the umbrella Nepali/Gorkha identity leading ethnic languages to the verge of endangerment. The realization of the fact among them could become possible after they failed to recognize their adopted Nepali/Gorkha identity in the Indian nation-state during the 1980s. To satisfy the realized identity crisis, the ethnic groups found another venture for recognition in India in terms of ethnolinguistic revitalization. Ethnolinguistic revitalization as a social movement is significant for reviving or revitalizing endangered languages. It is also important for reviving ethnic identity. The ethnolinguistic revitalization movement in Darjeeling in this sense needs insight, especially from the language endangerment vantage point. Contrary to the previous studies the result of the present study reveals that the revitalization movement is an identity negotiation process in Darjeeling. The method of integrative literature review has been employed in the study along with the unstructured interview and non-participant observation methods.

In Chap. 11 (*Kom: An Endangered Language of Manipur, India*), Nilanjana Roy Chowdhury and Awadesh Kumar Mishra inform us that Kom is one of the scheduled languages of the state of Manipur in India. It is recognized as a minority language with less than fifteen thousand speakers. The Koms are believed to have entered Manipur in 1600AD from the hills lying on the southern side of the Manipur valley. Although they do not associate themselves with any of the major tribal groups of the state like Nagas, Kukis, or Mizos, they share genetic similarities in their structure and forms

with the Kukis and cultural similarities with the Nagas. The typological features of this language appear to be of the agglutinating type with an SOV (subject-object-verb) word order. The lack of their script forces the community to use the Roman script with a slight modification of thirty-three alphabets soothing their language. The intelligibility of the Kom dialects, according to the concept of a dialect continuum, is not always bidirectional. Since it does not share the status of a modern Indian language, it is not taught at any school, college, or university in the state. Field-based investigation into the Kom community shows that there is a tendency among the members to shift towards the dominant Manipuri language as it provides them with some economic advantages. In fact, a tendency of ‘clan shift’ is also noticed due to societal, political, and economic dominance.

In Chap. 12 (*Dawoodi: A Highly Endangered Language in Northern Pakistan*), Chris Donlay and Zafeer Hussain Kiani inform us that Dawoodi, which was previously known as Domaaki in the earlier literature, is at present considered a severely endangered language spoken in two villages in the northern-most Gilgit-Baltistan Province of Pakistan. Each village is home to a separate dialect, one in Hunza Valley and the other in Nagar Valley, although the two are mutually intelligible. The Dawoodi language belongs to the central zone of the Indic branch of the Indo-Iranian language family and is believed to be a close relative of Romani and Domari. Due to negative attitudes and pressure from the regional languages, a language shift is well underway in the language. It is estimated today that there are only four or five dozen speakers of the Hunza dialect and roughly a dozen speakers of the Nagar dialect. This article provides an overview of the language, its history, sociolinguistic situation, and grammar. Based on the analysis of natural discourse data from dozens of speakers it sheds greater light on the lexicon, phonology, morphology, and syntax of this little-known language.

In Chap. 13 (*Bugun: An Endangered Language and Community of Arunachal Pradesh, India*), Madhumita Barbora informs that Bugun, one of the endangered languages of ancient origin, is spoken in the West Kameng District in the state of Arunachal Pradesh, India. This language belongs to the *ko-bwa* cluster of languages. Bugun has a low population count of 1434 as per the Census Report of India 2011. The Census Report of India places Bugun under the ‘other’ mother tongue category while Ethnologue labels Bugun as a ‘definitely’ endangered language. The Bugun language, like most of the languages of Arunachal Pradesh, is an *oral* language. Their folklore, cultural, agricultural and religious practices, and rituals have been passed orally from one generation to the other. With the advent of modern education and the shift to *Arunachalli Hindi*, a lingua franca of the state, has adversely impacted the language as well as the age-old traditional, religious, and cultural practices of the community. The change in the topography has made the community adapt to new agricultural practices instead of their traditional ones. The Bugun people practice ‘animism’ (believing that like human beings, every animal, plant, rock, river, weather, object, place, or creature is animated and alive). In recent years, however, most of the native Bugun people have converted to the Mahayana Sect of Buddhism and Christianity, yet they have their festivals and rituals venerating nature. The author provides a brief account of the Bugun community as well as the language and the ethnographic

status of the community. The paper presents a language sketch on the phonemic inventory, tone, and syllabic structure; followed by an account of morphological strategies used in the language for word formation and analysis of grammatical or inflectional markers, and lastly, an account of age-old Bugun phrases and clauses. The primary data are collected from Sinchung, Wanghoo, and Kaspi villages in the state of Arunachal Pradesh India from 2009 to 2014 through several field surveys. The language shows some Tibeto-Burman features; it is homophonous, syncretic and has split-ergativity.

In Chap. 14 (*Developing a Dictionary for Kheria Sabar: An Indigenous and Endangered Tribal Speech Community of Eastern India*) Niladri Sekhar Dash, Meghna Majumder, and Srija Deb address the challenges and problems that they face while they develop a dictionary for the Kheria Sabar speech community—an indigenous and endangered tribal community living in the district of Purulia in the state of West Bengal, India. The challenges that they face are classified into two broad types, namely, linguistic challenges, and extralinguistic challenges. The linguistic challenges are primarily linked with a collection of lexical data from the community members; sufficiency, diversity, and variety of the lexical data types; the paucity of lexicographic details for entry words; citation of example sentences for determining usage-based sense variations of polysemous entries; inadequacy of linguistic description of lexical items for addressing referential and pedagogical requirements; utilization of pictures, images, and diagrams for a visual representation of complex concepts and ideas of the communities and some other issues. The extralinguistic challenges, on the other hand, are primarily linked with awareness about the importance of such a knowledge resource among the members of the community; careful investigation of the attitude of the community members relating to the procurement of data and information from their life, living, culture, history, heritage, and ecology; logistic issues in data collection from *Urheimat* (i.e., the *primaeval habitation*) of community through on-spot interviews; ethical/humanistic issues in the selection of appropriate respondents; availability of funds for conducting elaborate linguistic field surveys; a collection of lexical data covering all major aspects of the community life; availability of the trained human resource for lexical processing, analysis and dictionary compilation; and availability of agencies willing to publish dictionary as a commercial product. These issues are also linked with several theoretical and ethical issues—all of which combine together to make the process of dictionary compilation for the Kheria Sabar community an upheaval task fretted with many caveats and shortcomings particularly in those contexts when their folk texts, verbal narratives, written materials or historical records become non-available for reference and utilization. Keeping all the challenges in view, in this paper, they discuss the strategies they adopt to overcome the hurdles they face while compiling a dictionary that may help in the process of preserving the indigenous language against a backdrop of camouflaged aggression of more powerful neighbouring languages. The proposed dictionary that they develop may be used by the Kheria Sabar speakers for general reference and pedagogic purposes as well as by external agencies for academic, commercial, and localization purposes.

In Chap. 15 (*Kadar—An Endangered Dravidian Tribal Language of India*), P. Chandramohan informs that there are twenty-six languages that belong to the Dravidian language family and they are broadly classified into four groups, viz. South Dravidian, South Central Dravidian, Central Dravidian, and North Dravidian. Four of them are literary languages, viz. *Tamil*, *Malayalam*, *Kannada*, and *Telugu*, and they have recorded history from the early centuries of the Christian era and are the major Dravidian languages of India. Among the non-major languages, the Kadar language is an endangered Dravidian tribal language spoken by the Kadar tribe, who inhabit the mountain regions of the Western Ghats in the state of Tamil Nadu, India. At present, the Kadar language is at risk of disappearing, along with the cultural traditions and knowledge systems that it represents. The language is currently listed by UNESCO as one of the critically endangered languages with only a few hundred speakers left. The Kadar language is characteristically different from others in grammar and vocabulary, with a complex system of noun classes and verb conjugations. However, due to the influence of Tamil, the regional as well as the state language of Tamil Nadu, the Kadar language is rapidly declining in usage and is in danger of being lost forever. If Kadar is allowed to go extinct, it will represent a significant loss not only to the Kadar community but also to the broader field of linguistics. The author makes an effort to document and preserve the language through linguistic surveys for digital documentation. This paper reports on the phonology, morphology, and syntax of this language with data elicited from linguistic field surveys.

In Chap. 16 (*The Bidayuh Language of Sarawak, Malaysia: Language Use and Proficiency*), Su-Hie Ting, Teresa Wai See Ong, and Florence Gilliam Kayad report that Bidayuh is an indigenous community living in Sarawak, Malaysia. The Bidayuh language is currently not taught in schools but experimented as a medium of instruction in some kindergartens. Past studies have shown that factors such as rural–urban migration, intermarriage, and higher education are causing a shift away from the Bidayuh language. However, they are conducted on a small scale and little is known about the language maintenance situation among its community. The authors present the findings of Bidayuh language use and proficiency in the community. Analysis of the questionnaire-based data from the respondents shows that not a large number of people spoke Bidayuh as their first language. Based on studies, a good percentage considered themselves as pure Bidayuh, many as half-Bidayuh, and a few feel that they had lost most of their Bidayuh characteristics. The findings showed that the respondents have a better ability to understand a conversation in the Bidayuh language compared to speaking it. They have lower literacy in the Bidayuh language because it is a spoken language and thus, not many learn it through formal instruction. Additionally, the large regional variations in the Bidayuh varieties accentuate problems of standardizing the Bidayuh language for instruction in schools. A comparison of the results for the youngest and oldest groups of respondents indicates that the Bidayuh language maintenance is difficult in the face of the importance of English for a career, and perceptions of Sarawak Malay dialect, English, and Iban as strong languages in Sarawak.

In Chap. 17 (*Assessing Language Endangerment in Northeast India Through a Four-Language Prism*), Suranjana Barua and Dimple Choudhury define the purpose

of their study which includes conducting a comparative analysis of the vulnerability of languages by examining actual language usage in order to assess the sociocognitive impact that the *status* of a language can have on its speakers. For this purpose, they select four Tibeto-Burman languages, each with a different status with regard to language endangerment, spoken in two Northeastern states of India. Of the four languages, two are spoken in Assam, viz. Bodo (a scheduled language and an associate official language of Assam) and Garo (an associate official language in Meghalaya and with speakers also in Assam). The other two languages chosen for this study are Miju and Digaru: both endangered languages, primarily spoken in the Lohit and Anjaw districts of Arunachal Pradesh, India. Neither of these two languages has any official recognition in any part of the state. The vulnerability of these languages is investigated in relation to their status from two vantage points: the functionality of language and the political question of representation. This study emphasizes the critical nature of mother tongue transmission from one generation to another and the decline of the use of these languages with a significant impact on language ecology, identity, heritage, and associated knowledge system. The current study elaborates on the endangered status of indigenous languages from multiple perspectives in order to raise multilevel concerns spanning from sociocultural to cognitive while addressing the political question of the linguistic identity of a community in two Northeastern states of the nation of India. Thus, they study Bodo and Garo in the context of their status and functionality in contrast to Assamese (perceived as the dominant language of Assam spoken by the majority) while the endangerment threat posed to Miju and Digaru, despite being indigenous languages. Their study is based in the context of the nearly irreversible rise of Hindi as the lingua franca of the state.

In Chap. 18 (*Indonesian Language Policy and Perspectives on Its Implementation in Promoting Bahasa Indonesia as an International Language*), Gatut Susanto, David Pickus, DeAndree Espree-Conaway, Suparmi, and Awaludin Rusiandi examine the language policy and its implementation to campaign *Bahasa Indonesia* to be an international language. They use a qualitative design and collect data by conducting a literature review and semi-structured interviews with academicians and activists of *Bahasa Indonesia bagi Penutur Asing* (aka, BIPA). The results show that the Indonesian language policy is enshrined by the law of the land, Government policies and regulations of the ministry. Although the language policy implementation in law and government legislation is successful, the ministerial regulations fail to implement them. They conclude that a language policy should be issued by the authorities to campaign *Bahasa Indonesia* to be an international language and efforts should be made to measure the success of teaching and learning BIPA both nationally and internationally. This indicates that the language policy to campaign *Bahasa Indonesia* to be an international language is inseparable from the development of BIPA teaching and learning.

The Lacunae

As editors of an anthology that desires to cover almost all the language varieties available in South Asian and Southeast Asian countries, we made all possible efforts to reach potential authors and researchers who could contribute chapters to enrich diversity in the number of languages and content of the volume. Moreover, as the editors of the volume, we had in mind that this volume is particularly designed keeping in mind the ‘ten-year window of 2022 to 2032’ declared by UNESCO as the ‘decade of endangered languages’; desired to represent language and culture-related data from indigenous languages and their unique ethnicity and knowledge system. Although the present volume is capable of addressing a small portion of it, a major portion is still untapped, which we believe and hope, can be done by some more young and enthusiastic scholars who are committed to the preservation and promotion of endangered languages and people.

Another major lacuna of this volume is the lack of any study that relates to an effort to develop a new critical perspective towards the prevailing matrices and frameworks designed to measure the level of endangerment and/or language vitality of the endangered indigenous languages keeping in view the South Asian and Southeast Asian languages. All the authors have more or less accepted the ‘language vitality index’ identified by UNESCO and tried to assess the vitality of the South Asian and Southeast Asian languages based on the existing parameters. The rich empirical data and information that are represented in this volume could have given us a great opportunity to make new observations and formulate new parameters to critique the prevailing views and framework(s) from the perspective of South Asian and Southeast Asian languages. This would have given us a scope to frame region-specific indices for language vitality assessment as well as generate avenues for reorienting the direction of prevalent discourse on language endangerment. However, we keep this issue open for more intensive analysis of language-specific data and information before we venture out to propose something new to challenge the existing (and working) framework.

Some Recommendations

While addressing various aspects and issues of endangered indigenous languages, it is necessary to develop some guidelines that can be localized and appropriated keeping in view the problems faced by different indigenous language communities. Moreover, there arises a need for defining policies for measuring and assessing the status of the endangered indigenous languages spread across the world.

- We should prepare some policy documents for language planning of the endangered indigenous languages by which we shall be able to assess their linguistic status, linguistic vitality and the possibility for revitalization. Moreover, we should define their revitalization strategies through corpus and education planning.

- The existing policies for endangered language revitalization should be localized to the ground level based on the status, vulnerability, and requirement of a particular endangered indigenous language and its speakers.
- Advanced systems, methods, and technology of corpus linguistics and language technology should be applied in an organized manner in documentation, digitization, archiving, resource generation, and revitalization of the endangered indigenous languages.
- Educational planning is also required for the endangered indigenous languages. This may be realized through resource development, the use of indigenous languages as mediums of instruction at primary and secondary levels of education, development of course syllabi, study materials, and reference materials in the indigenous languages.
- The involvement of government and non-governmental organizations, academicians, corporate sectors, and individuals may be encouraged in the protection, preservation, and promotion of endangered and minority languages, culture, history, and heritage.
- It is necessary to carry out elaborate ethnographic studies to know the factors responsible for the migration of speakers of indigenous languages to dominant languages as well as propose appropriate methods and strategies for controlling this migration.
- The members of endangered indigenous language communities should be involved in language documentation and revitalization tasks to empower them as well as make them aware of their own linguistic rights and responsibilities. This will not only develop a new generation of people who are groomed as trained human resources but also generate many job opportunities for the community members for their economic growth and advancement.
- We should reach out to endangered indigenous language communities by conducting on-the-spot meetings, community-learning sessions, seminars, conferences, and workshops on a regular basis with the members of the language communities. We should invite people engaged in language documentation and revitalization activities to look into the status of progress and the hurdles thereof.
- All organizations and departments engaged in these activities should be encouraged to develop a digital community profile (DCP) which will contain details about ethnographic and ecological information of the endangered indigenous communities so that the DCP becomes a window for their global exposure and accessibility.

Conclusion

This is the first book of its kind to present studies on endangered languages of South Asia and Southeast Asia. There has never been any effort of this kind for compiling a handbook on South Asian and Southeast Asian endangered languages although

scholars have made several efforts to represent one or two languages in their individual chapters. This volume becomes a significant contribution in the ten years (2022–2032) declaration by UNESCO as the ‘decade of endangered languages’, where the goal is to collect language and culture-related data from indigenous languages as well as record their ethnic identities and knowledge resources. This volume directly serves this mission with a representation of linguistic features of several endangered indigenous languages. This volume will partly serve the effort for documentation of endangered South Asian and Southeast Asian languages, which has been long due in the scheme of language documentation at the world level. A large number of scholars who are working on various endangered languages will invariably find this book useful for their reference, research, and academic activities.

Chapter 2

Noun Morphology in Hajong



Albina Narzary and S. Arulmozi

Abstract This study is a descriptive analysis of the Hajong Noun Morphology, one of the small minority groups living in Assam, Meghalaya, Arunachal Pradesh, West Bengal, and Bangladesh. Applying a descriptive approach and documenting the inflectional noun morphology of Hajong is highly relevant to the preservation of the endangered Hajong language. The study focuses on documenting Hajong noun morphology to contribute to the linguistic record and ensure data accuracy for future research. The result of the study shows that noun morphology in Hajong is mostly aligned with Indo-Aryan (IA) languages of the eastern zone, Bengali, Assamese while retaining a few morphological features from Tibeto-Burman languages. Morphologically, nouns in Hajong are inflected for number, case and can be followed by adjectives, numerals, and quantifiers. Phillips (Northeast Indian linguistics. vol 3, Cambridge, Delhi, pp 2011) covers the wide range of Case Marking in Hajong. However, the present study seeks to examine the various aspects of nominal morphology in Hajong including case marking. The Hajong case formatives have possible cognates from both IA (Assamese, Bangla) and TB languages (Koch, Garo, Bodo) (Phillips in Northeast Indian linguistics. vol 3, Cambridge, Delhi, pp 2011). Pronouns are differentiated by three persons and two numbers. However, the second and third person lacks any formality distinction (honorific and non-honorific) but are inflected for the case. Hajong possesses the natural gender system. A noun phrase in Hajong can be modified by the modifiers. The pre-nominal modifiers include determiners, adjectives, and adjective phrases. Determiners and modifiers precede the head in Hajong. The present data were collected from a total of 12 informants: six informants were aged 24 to 49 from the targeted community, while the other six were 50 years and above from the Goalpara district of Assam.

Keywords Assam · Hajong language · Noun morphology · Gender · Number · Case

A. Narzary (✉) · S. Arulmozi
Centre for Applied Linguistics and Translation Studies, University of Hyderabad, Hyderabad,
India
e-mail: narzary.albi@gmail.com

S. Arulmozi
e-mail: arulmozi@uohyd.ac.in