

THE LANGUAGES OF MAINLAND SOUTHEAST ASIA

Mainland Southeast Asia is one of the most complex cultural and linguistic areas in the world. This book provides a rich and comprehensive survey of the history and core systems and subsystems of the languages of the region. Drawing on his depth of expertise in mainland Southeast Asia, N. J. Enfield includes more than a thousand examples from over a hundred of the languages of Cambodia, southern China, Laos, Malaysia, Myanmar, Thailand, and Vietnam, and brings together data and analyses that have not previously been available in one place. Chapters cover the many ways in which these languages both resemble and differ from each other, and the diversity of the area's languages is highlighted by an emphasis on minority languages, which outnumber the national languages by nearly a hundred to one. The result is an authoritative treatment of an important and fascinating linguistic area.

N. J. ENFIELD is Professor of Linguistics at the University of Sydney. His publications include *Linguistic Epidemiology* (2003), 'Areal linguistics and mainland Southeast Asia' (*Annual Review of Anthropology* 2005), *A Grammar of Lao* (2007), *Dynamics of Human Diversity* (2011) and *Languages of Mainland Southeast Asia* (2015, co-authored with Bernard Comrie). He is working on a reference grammar of Kri, a Vietic language spoken in Laos.

CAMBRIDGE LANGUAGE SURVEYS

General editors

P. Austin (School of Oriental and African Studies, London)
B. Comrie (University of California, Santa Barbara)
R. Lass (University of Cape Town)
D. Lightfoot (Georgetown University)
K. Rice (University of Toronto)
I. Roberts (University of Cambridge)
S. Romaine (University of Oxford)

This series offers general accounts of the major language families of the world, with volumes organized either on a purely genetic basis or on a geographical basis, whichever yields the most convenient and intelligible grouping in each case. Each volume compares and contrasts the typological features of the languages it deals with. It also treats the relevant genetic relationships, historical development and sociolinguistic issues arising from their role and use in the world today. The books are intended for linguists from undergraduate level upwards, but no special knowledge of the languages under consideration is assumed. Volumes such as those on Australia and the Amazon Basin are also of wider relevance, as the future of the languages and their speakers raises important social and political issues.

Volumes already published include

Australian Languages R. M. W. Dixon
The Dravidian Languages Bhadriraju Krishnamurti
The Languages of the Andes Willem Adelaar with Pieter Muysken
The Slavic Languages Roland Sussex and Paul Cubberley
The Germanic Languages Wayne Harbert
Sign Languages Diane Brentari
The Afroasiatic Languages Zygmunt Frajzyngier and Erin Shay
The Languages of Mainland Southeast Asia N. J. Enfield
Turkic Lars Johanson

Forthcoming titles

The Balkan Languages Victor Friedman and Brian Joseph

THE LANGUAGES OF MAINLAND SOUTHEAST ASIA

N. J. ENFIELD
University of Sydney



CAMBRIDGE
UNIVERSITY PRESS

Cambridge University Press & Assessment
 978-1-108-70021-4 — The Languages of Mainland Southeast Asia
 N. J. Enfield
 Frontmatter
[More Information](#)



CAMBRIDGE
UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom
 One Liberty Plaza, 20th Floor, New York, NY 10006, USA
 477 Williamstown Road, Port Melbourne, VIC 3207, Australia
 314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India
 103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment, a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
 Information on this title: www.cambridge.org/9781108700214

DOI: 10.1017/9781108605618

© N. J. Enfield 2021

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press & Assessment.

First published 2021
 First paperback edition 2023

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

Library of Congress Cataloging-in-Publication data

Names: Enfield, N. J., 1966– author.

Title: The languages of mainland Southeast Asia / N. J. Enfield.

Description: New York : Cambridge University Press, 2021. | Series: Cambridge language surveys | Includes bibliographical references and index.

Identifiers: LCCN 2020027669 (print) | LCCN 2020027670 (ebook) | ISBN 9781108476331 (hardback) | ISBN 9781108605618 (ebook)

Subjects: LCSH: Southeast Asia – Languages.

Classification: LCC PL3501 .E54 2021 (print) | LCC PL3501 (ebook) | DDC 409.59–dc23

LC record available at <https://lcn.loc.gov/2020027669>

LC ebook record available at <https://lcn.loc.gov/2020027670>

ISBN 978-1-108-47633-1 Hardback

ISBN 978-1-108-70021-4 Paperback

Cambridge University Press & Assessment has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

CONTENTS

List of Figures page ix
List of Maps xi
List of Tables xii
Preface xix
Acknowledgements xxi
Note on the Nature and Organization of the Book xxii
List of Abbreviations xxiv

1 Context 1
1.1 The Mainland Southeast Asia Region 1
1.2 Mainland Southeast Asian Languages 3
1.3 Nomenclature, System Ontology, and Language Data Selection 4
1.4 History of MSEA 10
1.4.1 Prehistory 10
1.4.2 Hoabinhian Neolithic Tradition 12
1.4.3 Earliest Agriculture 13
1.4.4 Bronze and Iron Age States 15
1.4.5 Summary: Fall of First Millennium States, Rise of Second Millennium States 24
1.4.6 Away from States: Zomia 26
1.4.7 Dynamics of Ethnolinguistic Diversity in MSEA 28
1.5 On Modern Dominant Languages 32
1.6 Resources and Developments in MSEA Linguistics 35
1.6.1 Conferences and Publications 35
1.6.2 New Descriptive Work 36
1.6.3 New Methods 37

2 Language Histories and Classifications 38
2.1 Current Developments 40
2.1.1 Unprecedented Access to Linguistic Data 40

2.1.2	Unprecedented Access to Other Sources of Relevant Information	40
2.1.3	New Methods of Historical Linguistic Analysis	43
2.2	Austroasiatic	43
2.2.1	Austroasiatic Subgrouping	46
2.2.2	Munda Languages	47
2.2.3	Non-Munda ('Mon-Khmer') Subgrouping	49
2.3	Tai-Kadai	51
2.3.1	Subgrouping of Tai-Kadai Languages	53
2.3.2	The Tai Branch	59
2.4	The Chamic and Moklenic Branches of Austronesian	66
2.4.1	Chamic	69
2.4.2	Moklenic	70
2.4.3	Structural Change in Chamic and Moklenic	71
2.5	Hmong-Mien	72
2.5.1	Comparative-Historical Hmong-Mien	73
2.5.2	Hmongic versus Mienic	74
2.6	Sino-Tibetan	79
2.6.1	Sinitic	81
2.6.2	Tibeto-Burman Subgroupings	85
2.6.3	Affixes in Historical Tibeto-Burman	89
2.7	Proposed Connections among the Families	92
2.8	Problems and Challenges for Comparative-Historical Research	94
2.9	Signed Languages	96
2.9.1	Home Sign Systems	97
2.9.2	Village Sign Systems	97
2.9.3	Alternate Sign Languages	99
2.9.4	National-Level Sign Languages	100
2.9.5	Prospects	105
2.10	Classical Languages	106
2.11	Scripts and Writing Systems	107
3	Typological Overview	118
3.1	Preliminaries	118
3.2	Typological Overview	122
3.3	Morphological Profile	123
3.4	Syntax	128
3.5	Basic Phrasal Constituent Order Patterns	129
3.5.1	Relative Position of Verb and Object	131
3.5.2	Relative Position of Modifier and Head in Noun Phrases	132
3.5.3	Relative Position of Noun and Relative Clause	135

3.5.4	Relative Position of Possessor and Possessed	137
3.5.5	Relative Position of Adposition and Noun	138
3.5.6	Relative Position of Adjective and Standard of Comparison	139
3.6	Sentence-Type Distinctions	141
3.7	Sentence-Final Particles	144
3.8	Expressive Language	146
3.8.1	Ideophones/Expressives	146
3.8.2	Poetic Idioms (Four-Syllable Expressions)	151
4	Phonological Systems	154
4.1	Consonants	154
4.2	Vowels	161
4.3	Phonotactics	166
4.4	Register	168
4.5	Tone	174
4.6	Tonogenesis	183
4.7	Tone Sandhi	195
4.8	Intonation	203
4.9	Prosodic Hierarchy and Phonological Words	204
5	Word Formation	209
5.1	Form Classes	210
5.1.1	Noun–Verb Distinction Is Usually Clear	210
5.1.2	‘Adjectives’ Are Verbs, or Are Verb-Like	211
5.1.3	‘Adpositions’ Are Often Nouns or Verbs	214
5.1.4	Verbs Can Act as ‘Adverbs’ without Derivation	216
5.2	Compounding	216
5.3	Psycho-collocations	222
5.4	Reduplication	227
5.5	Affixation	242
5.5.1	Existential/Locative Marking and Impersonal Predicate	243
5.5.2	Lexical Category Derivation	243
5.5.3	Reciprocal/Collective Marking	246
5.5.4	Causative	247
5.5.5	Further Morphological Derivations	250
5.6	Tone in Word Formation	252
5.6.1	Tone Sandhi in Compound Formation	252
5.6.2	Regular Grammatical Distinctions	253
5.6.3	Forming Ideophones	255

6	Reference and Nominal Syntax	257
6.1	Nominal Modification	259
6.1.1	Simple Head-Modifier Relations	259
6.1.2	Relative Clause Constructions	266
6.1.3	Possessive and Part–Whole Constructions	277
6.2	Nominal Classification	282
6.2.1	Numeral Classifier Constructions	283
6.2.2	Class Terms	300
6.3	Personal Pronoun Systems	308
6.4	Demonstrative Systems	318
6.4.1	Two-Term Systems	318
6.4.2	Three-Term Systems	321
6.4.3	Four-Term Systems	322
6.4.4	Five-Term Systems	323
6.4.5	Larger Systems	324
6.4.6	Comment on Demonstrative Systems	325
7	Predication and Clausal Syntax	326
7.1	Marking Syntactic Functions	326
7.2	Verbal Marking	332
7.2.1	Negation	332
7.2.2	Aspect and Related Marking	335
7.2.3	Modality and Related Marking	340
7.3	Multi-verb Constructions	342
7.3.1	Possible Number of Verbs in Series	345
7.3.2	Events Occurring in Series	346
7.3.3	Events or Event-Facets Occurring as Elements of a Single Event	354
7.3.4	Complementation Strategies	362
7.4	Valency-Changing Strategies	368
7.4.1	Participant-Adding: Syntactic Causatives and Related Constructions	368
7.4.2	Single Participant in a Transitive Event: Reflexives and Related Constructions	375
7.4.3	Two-Way or Distributed Action: Reciprocals and Related Constructions	384
	<i>Postface</i>	393
	<i>References</i>	396
	<i>Author Index</i>	429
	<i>Language Index</i>	431
	<i>Subject Index</i>	435

FIGURES

- 1.1 Chronology, I: Mainland Southeast Asia during the last two Marine Isotope Stages (MIS 1, 2) *page* 14
- 1.2 ‘Schematic depiction of succession of dominant states’ 25
- 1.3 Chronology: Mainland Southeast Asia over the last 4,000 years 29
- 2.1 Mon-Khmer classification according to Thomas and Headley 50
- 2.2 Austroasiatic (after Diffloth 2005: 79) 51
- 2.3 Austroasiatic (after Sidwell 2015: 179) 52
- 2.4 Tai-Kadai languages 58
- 2.5 ‘Tentative tree diagram of Austronesian languages’ 68
- 2.6 Proposed Malayo-Chamic grouping 69
- 2.7 ‘Malayo-Sumbawan’ 70
- 2.8 Mienic subgrouping with four main subgroups 77
- 2.9 ‘Classification of Mjuenic dialects and sub-dialects’ 78
- 2.10 Mienic, with three subgroups 78
- 2.11 Subclassification of Mienic languages 79
- 2.12 The ‘standard view’ of Sino-Tibetan as consisting of two top-level branches: Tibeto-Burman and Sinitic 82
- 2.13 ‘Chinese “dialects” (languages)’ 84
- 2.14 ‘Schematic chart of Sino-Tibetan groups’ 86
- 2.15 ‘The Tibeto-Burman language family tree, reflecting the Sino-Bodic and Western Tibeto-Burman hypotheses’ 89
- 2.16 ‘Sign languages in Thailand classified by traditional language family trees’ 102
- 2.17 ‘Sign languages in Vietnam classified by a traditional language family tree’ 103
- 2.18 ‘Linguistic relations between sign languages in Thailand and Vietnam’ 104
- 2.19 Modern Burmese writing 110
- 2.20 Modern Lao writing 110
- 2.21 Modern Lao Song (Black Tai) writing 111

x *List of Figures*

2.22	‘Historical development of Indic scripts in Southeast Asia’	113
4.1	Segmental structure of the Kri (AA) word	166
4.2	Hakha Lai (TB) tones	175
4.3	Modern Standard Chinese (SN) tones	176
4.4	Cantonese (SN) tones	177
4.5	Northern Vietnamese (AA) tones	178
4.6	White Hmong (HM) tones	178
4.7	Dong (TK) tones	179
4.8	Lao (TK) tones	180
4.9	The Xiamen (SN) tone circle	196
6.1	Accessibility hierarchy for the construction of relative clauses in languages of the world	268

MAPS

- 1.1 Mainland Southeast Asia *page 2*
- 1.2 ‘Sundaland’ 12
- 2.1 Approximate distribution of languages of the Austroasiatic family 44
- 2.2 Approximate distribution of Austroasiatic languages in Core MSEA 45
- 2.3 Approximate distribution of Tai-Kadai languages 54
- 2.4 Approximate distribution of Austronesian languages 67
- 2.5 Approximate distribution of languages of the Chamic and Moklenic branches of Austronesian 68
- 2.6 Approximate distribution of Hmong-Mien languages in MSEA 73
- 2.7 Approximate distribution of Sino-Tibetan languages in MSEA 80
- 3.1 Order of adjective and noun in Tibeto-Burman languages of northwestern MSEA and beyond 135

TABLES

- 1.1 Languages of MSEA: (a) numbers of languages in Core MSEA, by family; (b) numbers of languages in Greater MSEA, by family *page 4*
- 2.1 Lexical correspondences in Santali, Mon, and Muong 44
- 2.2 Proto-Austroasiatic inventory of syllable-initial consonants 46
- 2.3 Different patterns of merger of the Proto-Austroasiatic three-way initial bilabial plosive split, showing retention of implosive bilabial *ɓ- survives in Old Mon (Monic), Bahnar (Bahnaric), and Katu (Katuic) 46
- 2.4 Different patterns of merger of the Proto-Austroasiatic three-way initial apical plosive split, showing retention of implosive apical *ɗ- in Old Mon (Monic), Bahnar (Bahnaric), and Katu (Katuic) 47
- 2.5 Summary of Austroasiatic proto-branch level initial-consonant correspondences 48
- 2.6 Thirteen branches of Austroasiatic 53
- 2.7 Some basic vocabulary in Tai-Kadai languages from distant branches 54
- 2.8 Lexical correspondences supporting the grouping of Kra languages (Gelao, Lachi, Laha, Paha Buyang, and Pubiao) as distinct from languages of other branches of Tai-Kadai 56
- 2.9 Rice farming vocabulary, establishing a split between Gelao-Hlai and the rest of Tai-Kadai 57
- 2.10 Early Tai tone categories for ‘live’ syllables resulting from a historical split of three proto tones (A, B, and C) 61
- 2.11 Early Tai tone categories for ‘dead’ syllables, with long versus short vowels 62
- 2.12 Gedney’s complete system of twenty pan-Tai tone boxes 63
- 2.13 Modern reflexes in Tai languages of syllables in boxes 1–3 under the A category in Proto-Tai 65
- 2.14 Lexical evidence for a historical distinction between Hmongic and Mienic languages 75

- 2.15 Modern reflexes of Proto-Hmong-Mien tight (*NC-) versus loose (*N-C-) prenasalized obstruents 75
- 2.16 Evidence for a correspondence between historical prefixes (also found in Tibeto-Burman) and modern voiceless initials in Hmongic languages 76
- 2.17 ‘Sino-Tibetan comparisons’ 81
- 2.18 Selected core vocabulary of Chinese varieties 83
- 2.19 Proto Lolo-Burmese initial simple plosives 90
- 2.20 Modern reflexes of Proto Lolo-Burmese initial bilabial stops 90
- 2.21 Lexical correspondences illustrating modern reflexes of Proto Lolo-Burmese initial bilabial stops 91
- 2.22 Modern reflexes of Proto Lolo-Burmese voiced and voiceless initial bilabial stops, with and without the *s- and *m- prefixes 91
- 2.23 Lexical correspondences illustrating modern reflexes of Proto Lolo-Burmese voiced and voiceless initial bilabial stops, with and without the *s- and *m- prefixes 92
- 2.24 Data on selected sign languages of core MSEA 100
- 2.25 ‘Neologisms and lexical specifications’ based on Pali/Sanskrit in Thai 107
- 2.26 ‘Cham alphabet letters: Consonants’ 109
- 2.27 Features common to cases of the independent development of orthography in ‘Zomian’ minority languages 115
- 3.1 Overview of sound system features in MSEA languages 122
- 3.2 Overview of morphosyntax-semantics system features 123
- 3.3 Features of Cantonese sentence-final particles (SFPs) 145
- 3.4 A selection of Sgaw Karen (TB) ideophones 147
- 3.5 Some ideophones in Chrau (Austroasiatic) 150
- 4.1 Simple syllable-initial consonants in Khmu (AA) 156
- 4.2 Khmu (AA) initial clusters 156
- 4.3 Syllable-final consonants in Khmu (AA) 157
- 4.4 Set of consonants possible in the initial position of main syllables in Kri (AA) 157
- 4.5 Final segments of stressed syllables in Kri (AA) 157
- 4.6 Syllable-initial consonants in Lao (TK) 158
- 4.7 Syllable-final consonants in Lao (TK) 158
- 4.8 Syllable-initial consonants in Sgaw Karen (TB) 159
- 4.9 Syllable-initial consonant clusters in Sgaw Karen (TB) 159
- 4.10 Syllable-final consonants in Sgaw Karen (TB) 160
- 4.11 Set of consonants possible in the initial position of main syllables in Moklen 160
- 4.12 Set of consonants possible in the final position of main syllables in Moklen 160

- 4.13 Initials of Dànánshān Miao, Guizhou (HM) 161
- 4.14 Cambodian (AA) vowel phonemes 162
- 4.15 Dànánshān Miao (HM) vowel phonemes 162
- 4.16 Cantonese (SN) vowel phonemes 162
- 4.17 Shanghainese (SN) vowel phonemes 163
- 4.18 Sgaw Karen (TB) vowel phonemes 163
- 4.19 (Eastern) Khmu (AA) vowel phonemes 164
- 4.20 Lao (TK) vowel phonemes 164
- 4.21 Modern Western Cham (AN) vowel phonemes 164
- 4.22 Moken (AN) vowels 165
- 4.23 Bahasa Malaysia (AN) vowels 165
- 4.24 Moklen (AN) vowel phonemes 166
- 4.25 Some Kri (AA) (near) minimal pairs contrasting in register alone 169
- 4.26 Long vowel phonemes of Kri (AA) showing vowels paired by register 170
- 4.27 Long vowel phonemes of Kri (AA) laid out in a standardized nine-place system 170
- 4.28 Selection of Rengao (AA) long vowel phonemes and their realization in tense versus lax register 171
- 4.29 Hypothetical scenario based on the Rengao situation shown in Table 4.28 172
- 4.30 Vowel phonemes of Cambodian (AA) 172
- 4.31 The relation between modern Khmer vowels (shown in Table 4.30) and the earlier two-register system in Old Khmer 173
- 4.32 Takhian Thong Chong registers 174
- 4.33 Illustration of falling, rising and low tones on smooth syllables as realized after *ka* ‘my’ in Hakha Lai (TB) 176
- 4.34 The three classes of syllable-initial consonant in Lao 182
- 4.35 Modern Lao (TK) examples of reflexes of traditional Tai tones 183
- 4.36 Vietnamese (AA) and Chinese (SN) lexicon 184
- 4.37 Eight traditional Chinese tone categories 185
- 4.38 Traditional Vietnamese tone categories 186
- 4.39 Stage 1 of tonogenesis, pre-tonal 187
- 4.40 Stage 2 of tonogenesis, pitch contours being phonologized 187
- 4.41 Stage 2 of tonogenesis, showing syllable types distinguished by voicing of syllable-initial segment 187
- 4.42 Stage 3 of tonogenesis, initial contrasts lost (here, all Cs are voiceless), incidental pitch differences arising from previous \pm voicing now make lexical contrast 188
- 4.43 Comparison of Vietnamese tone category C words with Mon, Mnong, and Kri cognates (all AA) 188

- 4.44 Comparison of Vietnamese tone category C words with Khmer cognates (both AA) 189
- 4.45 Comparison of Vietnamese tone category B words with Khmu, Riang, and Kri cognates (all AA) 189
- 4.46 Comparison of Vietnamese tone category A words with Khmu cognates (both AA) 189
- 4.47 The six Vietnamese tones, plotted as a function of two non-tonal parameters of syllables in the proto-language 190
- 4.48 Haudricourt's model of tonogenesis applied to White Hmong 190
- 4.49 'Karen tone box' 191
- 4.50 Three values for terminance in Kri (AA) 192
- 4.51 The six Kri (AA) syllable types defined by intersections of terminance and register distinctions 192
- 4.52 The six Vietnamese tones plotted as a function of two non-tonal parameters of syllables in the proto-language 192
- 4.53 Associations between syllable type and pitch in Moklen (AN) 193
- 4.54 Moklen (AN) pitch-based minimal pairs 194
- 4.55 Examples of Xiamen (SN) tone sandhi following the tone circle pattern 196
- 4.56 Pattern of tone replacement in progressive tone sandhi in White Hmong 197
- 4.57 Biao Min Yao (HM) tones 198
- 4.58 Hakha Lai (TB) tones, with values represented by start and end points 199
- 4.59 Five combinations of tone in Hakha Lai (TB) that do not undergo tone sandhi change 200
- 4.60 The four combinations of tone in Hakha Lai (TB) that undergo tone sandhi change 200
- 4.61 Wuming Zhuang (TK) tones 201
- 4.62 Two patterns of tone alternation in Wuming Zhuang (TK) 201
- 4.63 Wuming Zhuang (TK) tones re-described in terms of register (first letter) and contour within that register 202
- 4.64 Two patterns of tone alternation in Wuming Zhuang (TK) re-stated using a register/contour description of tone 202
- 4.65 Elements of the phonological/prosodic hierarchy 205
- 4.66 Kinds of unstressed syllables in Lao (TK) distinguished by their identity on the implicational scale minor > light > unstressed 206
- 4.67 Comparison of phonological and other grammatical features of Munda and Mon-Khmer (= non-Munda) Austroasiatic languages 208
- 5.1 Kratochvíl's five-way distinction in 'basic types of morphological construction' in Mandarin Chinese (SN) 209

- 5.2 Example psycho-collocations in Bangladesh Khumi (Kuki-Chin, TB) 223
- 5.3 Example psycho-collocations in Pwo Karen (TB) 223
- 5.4 Example psycho-collocations in Yongning Na (Mosuo) (TB) 224
- 5.5 Example psycho-collocations in Southern Min (SN) 224
- 5.6 Example psycho-collocations in Wa (Paraok) (AA) 224
- 5.7 Example psycho-collocations in Vietnamese (AA) featuring reference to *lòng* ‘entrails, bosom’ 225
- 5.8 Example psycho-collocations in Thai (TK) featuring reference to *cay* ‘heart’ 225
- 5.9 Example psycho-collocations in Hmong (Mong Leng, Hmong-Mien) mostly featuring reference to the liver 226
- 5.10 Example psycho-collocations in Colloquial Eastern Cham (AN) featuring reference to the liver, body, head, and belly 226
- 5.11 Example psycho-collocations in Mon (AA) 227
- 5.12 Example psycho-collocations in Khmer (AA) 227
- 5.13 Example psycho-collocations in Burmese (TB) 227
- 5.14 Four types of chameleon affix in Vietnamese (AA) 237
- 5.15 Expressive reduplication in Lao (TK) with stem alternation, back vowel changed to front vowel 241
- 5.16 Expressive reduplication in Lao (TK), front vowel changed to schwa 242
- 5.17 Doublets in White Hmong (HM) arising from Sandhi effects 252
- 6.1 Dominant properties of relativization in non-Sinitic MSEA languages 267
- 6.2 Type I classifier constructions in MSEA and beyond 285
- 6.3 Type II classifier constructions in MSEA and beyond 285
- 6.4 Khmu (AA) classifiers 286
- 6.5 Mandarin classifiers: Categories and inventory sizes 291
- 6.6 Class nouns in White Hmong (Hmong-Mien) 301
- 6.7 Some variable features of personal pronouns across a sample of MSEA languages 307
- 6.8 Personal pronouns in Cantonese (SN) 308
- 6.9 Personal pronouns in Mandarin (SN) 309
- 6.10 Personal pronouns in Hakka (SN) 309
- 6.11 Personal pronouns in Lahu (TB) 310
- 6.12 Personal pronouns in Xuyong Hmong (HM) 310
- 6.13 Personal pronouns in Green Miao (HM) 310
- 6.14 Personal pronouns in Mulao (TK) 311
- 6.15 Personal pronouns in Chadong (TK) 311
- 6.16 Personal pronouns in Chrau (AA) 312
- 6.17 Personal pronouns in Phan Rang Cham (AN) 312

- 6.18 Personal pronouns in Zhuang (TK) 312
- 6.19 Personal pronouns in Jahai (AA) 313
- 6.20 Personal pronouns in Kri (AA) 314
- 6.21 Some kin categories that are consequential for selection of person
reference form in Kri (AA) 314
- 6.22 Kri (AA) basic kin terms 315
- 6.23 Khmer (AA) personal reference paradigm 316
- 6.24 Burmese (TB) personal reference paradigm 317
- 6.25 Cantonese (SN) demonstrative pronouns 318
- 6.26 Mulao (TK) demonstrative pronouns 319
- 6.27 Zhuang (TK) demonstrative pronouns 319
- 6.28 Hainan Cham (AN) demonstrative pronouns 319
- 6.29 Phan Rang Cham (AN) demonstrative pronouns 319
- 6.30 White Hmong (HM) demonstrative pronouns 320
- 6.31 Lao (TK) demonstrative determiners 320
- 6.32 Lao (TK) demonstrative adverbs 320
- 6.33 Semelai (AA) basic demonstratives 321
- 6.34 Sedang (AA) demonstratives 321
- 6.35 Hakka (SN) demonstratives 321
- 6.36 Aizhai Miao (HM) demonstratives 322
- 6.37 Green Miao (HM) demonstratives 322
- 6.38 Chrau (AA) demonstratives 322
- 6.39 Akha (TB) demonstratives 322
- 6.40 Lai (TB) demonstratives 323
- 6.41 Lahu (TB) demonstratives 323
- 6.42 Kri (AA) demonstratives 323
- 6.43 Jahai (AA) demonstratives 324
- 6.44 The semantics and pragmatics of Jahai (AA) demonstrative
categories 324
- 7.1 Some idealized grammatical categories of verbal marking that ground
a proposition by linking (1) the ‘narrated event’ to (2) the coordinates of
the speech event 331
- 7.2 Cantonese (SN) path-of-motion verbs 355
- 7.3 Cantonese (SN) directional SVCs 355
- 7.4 Elements of the Lao (TK) Manner–Path–Direction construction 356
- 7.5 Semelai (AA) associated motion multi-verb constructions 358

PREFACE

To understand the nature of human language, we need to know the properties that languages can have, the distribution of those properties in the world, and the reasons for that distribution. This book brings together information relevant to these questions in relation to one geographical area: mainland Southeast Asia (MSEA). In line with the remit of the Cambridge Language Surveys series, we present a technical survey of the languages of this area, with two main points of thematic focus: first, the histories of the languages and their speakers, and second, the structural properties of the languages from sound systems to the make-up of phrases and clauses.

There is of course much more to language: verbal art and oratory, ritual speech, multilingualism, dialectology, politeness and impoliteness, discourse structure, toponymy, personal names, kinship terminologies, ethnobiological classification, narrative style and practice, language ideology, et cetera. These areas await a more sociologically and anthropologically oriented volume, tapping into a wealth of past and current research in these domains. For example, on the sociolinguistics of language endangerment, and associated issues including language protection and revitalization, see Phattharathanit (2012) on identity maintenance in Lanna (cf. Bradley 2007, Premsrirat 2007). Research on linguistic politeness is being done, mostly in relation to national languages, and with reference to the languages' elaborated systems of social deixis, for example in their systems of personal pronouns, and the pragmatic alternatives that effectively create open class systems for person reference (see Cooke 1968, Haas 1969, Luong 1990). The more complex documented systems of person reference belong to the major literate languages of the area, including Thai, Cambodian, Vietnamese, and Burmese (Cooke 1968). There has been recent work in this domain on languages including Lao (Enfield 2007: ch. 5; 2015: ch. 5). On Vietnamese, see Srichampa (2008) on politeness strategies, and Sidnell and Shohet (2013) on avoidance strategies (see also Luong 1988). Linking social life to central concerns of historical linguistics and typology, there has been recent work on sociolinguistic conditions for borrowing (Alves 2009). For similar work see Thurgood (2010) comparing two varieties of Cham with the Tibeto-Burman language Anong. And a new line of work in MSEA is in conversation analysis: Enfield (2013) presents several case studies of

Lao language in conversation; Hà (2010, 2013) presents studies of Vietnamese conversation with a focus on the role of prosody, for example in repair and backchannelling (see also Umaporn 2007 on backchannelling in Mon); and a series of in-progress studies by Jack Sidnell and colleagues examine strategies of repair in Vietnamese. Constraints of feasibility and the remit of the series mean that the coverage of linguistic topics in this book is necessarily incomplete. I look forward to the day when a full-length volume on the sociocultural anthropology of language in MSEA can complement the present book's focus on historical and structural linguistic matters.

ACKNOWLEDGEMENTS

I thank all those who have given me comments on material that has found its way into this book, improving it immeasurably: especially Peter Bellwood, Marc Brunelle, Tony Diller, Bob Dixon, Bill Foley, Gwen Hyslop, Nerida Jarkey, Weijian Meng, Pittayawat Pittayaporn, Mark Post, Martha Ratliff, Jack Sidnell, Paul Sidwell, David Solnit, and Martin Stuart-Fox. I thank my research assistants for indispensable help: Weijian Meng, Gus Wheeler, Marlena Lutz-Hughes, and Naomie Nguyen. The maps and figures were created by Gus Wheeler. The Chinese language sources were accessed by Weijian Meng. I am grateful to Angela Terrill, of Punctilious Editing Sweden, for her genuinely punctilious work in compiling the three indexes. I am grateful to the anonymous reviewers for their generous advice, to the editorial board of the Cambridge Language Surveys series, and to Kate Brett and Helen Barton at Cambridge University Press for their guidance and support. This work was supported in part by the Australian Research Council (Discovery Project DP170104607).

NOTE ON THE NATURE AND ORGANIZATION OF THE BOOK

This book incorporates most of the material included in *Mainland Southeast Asian Languages: A Concise Typological Survey* (Cambridge University Press 2019), which focused solely on typological features of the spoken languages of core MSEA. At nearly twice the length of that book, the present book adds two main things: (1) it expands the data coverage and analysis in the descriptive/typological chapters (3–7), and (2) it places the linguistic diversity of the area in historical context, both in terms of the social (pre)history of the mainland Southeast Asia area (in Chapter 1) and in the historical-comparative linguistics of the area's languages (in Chapter 2).

Readers interested in the historical and social context of the languages of mainland Southeast Asia will get most out of Chapter 1, which gives an overview of the prehistory and history of the area. Chapter 2 focuses on historical-comparative research, outlining current and evolving views on the relatedness of modern MSEA languages, with many references for those who wish to follow up on the details. The rest of the book covers grammatical topics, moving through broad categories of structural organization. A general outline of the typology of the area's languages is provided in Chapter 3, and subsequent chapters delve into details under four rubrics: Phonological systems (Chapter 4), Word formation (Chapter 5), Reference and nominal syntax (Chapter 6), and Predication and clausal syntax (Chapter 7). The organization of chapters 6 and 7 is based loosely on a functional distinction between reference and predication, the two basic elements of any proposition. Of equally crucial functional importance are devices for textual cohesion and for managing interpersonal relations in language use. These are covered under the appropriate functional domain in which they occur (for example, politeness-based choice occurring in pronoun systems, discussed in Chapter 6) or are included in Chapter 4 (e.g., sentence-final particles for the expression of speech act and other interpersonally grounded pragmatic distinctions).

This book cites language examples from a wide range of sources. We have endeavoured to represent data exactly as given in the original sources, though we note an occasional

exception made out of convenience: we have sometimes used ‘*a*’ (italic a) for ‘a’ (roman a) and vice versa, as long as this has not introduced ambiguity internal to the example cited (i.e., by collapsing the two characters). For this reason (and also as a matter of general principle), readers are advised to consult the primary sources before citing the examples.

ABBREVIATIONS

I	first person	3SGA	third person singular
2	second person		A-argument
3	third person	3SS	third person singular
IDU	first person dual		subject
IP	first person	A	A-argument
IPL	first person plural	ACC	accusative
IS	first person singular	ACHV	achievement
ISG	first person singular	ACOP	attributive copula
ISS	first person singular	ADJT	adjunct
	subject	ADVRS	adversative
2DU	second person dual	AFF	affirmative
2HON	second person honor	AGR	agreement
2P	second person	ALL	allative
2PL	second person plural	ALP	attributive linking
2S	second person singular		particle
2SG	second person singular	AMB	ambient noun with
3A	third person A-argument		weather predication
3D	third person dual	AMSL	ambifunctional selector
3DU	third person dual	ANIM	animate
3P	third person plural	ANTICAUS	anticausative
3PL	third person plural	AO	agent orientating
3PLA	third person plural	ASP	aspect
	A-argument	ASS	associative
3PLS	third person plural	ASSOC	associative
	S-argument	AT	locative preposition
3POSS	third person possessor	ATT	attainment
3S	third person S-argument	AUG	augmented
3SG	third person singular	AUX	auxiliary

List of Abbreviations

xxv

B	bare	EXPR	expressive
BRO	brother	EXIST	existential
CAUS	causative	EXST	existential
CAUSE	causative	EXT	external
CL	classifier	F	female
CLASS	classifier	FA	familiar
CLF	classifier	FUT	future
CMPL	completive	FW	functional word
CMPR	comparative	GEN	genitive
COLL	collective	GRP	group marker
COM	comitative verb particle	GSL	group selector
COMP	complementizer	HIRESP	high respect
COND	conditional	IDEO	ideophone
CONT	continuous	IMM	imminent
COP	copula	IMP	imperative
CT	class term	IMPEF	imperfective
DAT	dative	IMPERF	imperfective
DECL	declarative	IMPF	imperfective
DEM	demonstrative	IMPFV	imperfective
DEP	dependent	IMPORT	importance
DET	determiner	INCL	inclusive
DID	verbal prefix (<i>d</i>) <i>id-</i>	INDEF	indefinite
DIM	diminutive	INSTR	instrumental
DIR	directional	INTER	interrogative
DISC	discourse particle	INTJ	interjection
DIST	distal	INTNS	intensive
DO	direct object	INTR	intransitive
DST	distal	INTRG	interrogative
DU	dual	IO	indirect object
DUR	durative	IRR	irrealis
E	elder	LOC	locative
EBR	elder brother	LORESP	low respect
EMP	emphatic	LP	linking particle
ERG	ergative	M	male
EUPH	euphemism	MC	modifier classifier
EVID	evidential	MID	middle voice
EX	exclusive	MOD	modality
EXCL	exclusive	N	noun
EXP	experiential	NEG	negation

xxvi *List of Abbreviations*

NEG1	first position negator	QM	question marker
NEG2	second position negator	QPLR	polar question marker
NFIN	non-finite	QUE	question marker
NFUT	non-future	QW	quantity word
NMLZ	nominalizer	R	realis
NMZ	nominalizer	RCNT	recent
NOM	nominalizer	RCP	reciprocal
NON	non	RCPL	reciprocal
NVOLT	non-volitional	RDP	reduplication
OBJ	object	RDUP	reduplication
OBL	oblique	REC	reciprocal
OBLIG	obligatory	RECIP	reciprocal
OCOMP	object of comparison	RECP	reciprocal
P	polite	RED	reduplication
PART	particle	REF	reflexive
PASS	passive	REFL	reflexive
PAST	past tense	REL	relativizer
PATIENT	patient marker	RESP	respect
PCL	particle	RT	relational tense
PERF	perfective	S	S-argument
PFV	perfective	SAT	quantifier/intensifier
PL	plural		(s)at
PN	proper name	SC	speaker conclusion
POL	polite	SEQ	sequential
POSS	possessive	SFP	sentence-final particle
PP	pragmatic particle	SFP1	first-position sentence-
PREF	prefix		final particle
PRF	perfect	SFP2	second-position
PROG	progressive		sentence-final particle
PROH	prohibitive	SG	singular
PROX	proximal	SIB	sibling
PRT	particle	SP	species
PRX	proximal	SPMY	socially-conditioned
PTC	particle		pronoun, male, younger
PTCL	particle	SPNY	socially-conditioned
PV	verbal particle		pronoun, neutral,
PVF	final verbal particle		younger
PVP	postverbal particle	STAT	static aspect
Q	question marker	SUB	subject