

WESTERN AUSTRONESIAN APPLICATIVE CONSTRUCTIONS:
TYPOLOGICAL AND FUNCTIONAL APPROACHES

A dissertation submitted to the Graduate Division of the University of Hawai'i at Mānoa
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

IN LINGUISTICS

May 2024

By

Christina L. Truong

Dissertation Committee:

Bradley McDonnell, Chairperson

Shelece Easterday

Gary Holton

William O'Grady

Joni Sasaki

*Keywords: applicatives, Austronesian, constructionist approaches, Indonesia, Malaysia,
Sundanese language, linguistic typology*

Copyright © 2024 Christina L. Truong

In loving memory

Liza M. Lai

Acknowledgments

First, I wish to acknowledge the Hawaiian people as the original and present inhabitants and caretakers of the lands on which I lived and conducted this research at the University of Hawai'i at Mānoa, and the Kizh Nation, on whose lands I have continued this work after leaving Hawai'i. I wish to express and offer my *mahalo* for the opportunity to live in Hawai'i and to learn from Hawaiian people, past and present, which has been a formative part of my education. *Mahalo nui loa* to my friends, classmates, and students of Native Hawaiian ancestry who extended aloha to me and taught me about Hawai'i, especially Noah Ha'a Solomon, and Nalani Cook.

Second, I thank my friend and collaborator, Dewi Setiani, for sharing her language with me and lending her voice and expertise in support of the documentation and description of the Sundanese language as part of this project. I also thank Eti for her contribution as a speaker of Sundanese. I express my appreciation to Dr. Eri Kurniawan, Dr. Chye Retty Isnendes, Dr. R. Dian Dia-an Muniroh, and Dr. Hernawan Sunda for sharing with me their expertise about the Sundanese language, and to Adinda, Taasya, and Yanti who answered many questions from me during our work together transcribing Baduy language recordings for my graduate assistantship.

I also express my thanks to the many people at the University of Hawai'i at Mānoa who have supported me throughout the doctoral program. Thank you to Brad McDonnell: without your steady support, collaboration, and guidance, this research would not have taken shape. Thank you to my committee members Shelece Easterday, Gary Holton, William O'Grady, and Joni Sasaki, for your encouragement, patience, practical help, and guidance. Thank you to my professors, Kamil Dean, Andrea Berez-Kroeker, Victoria Anderson, Rory Turnbull, Amy Schafer, and the late Bob Blust, along with my committee members: you have shaped me as a scholar and I am grateful for your candor and kindness. I am especially grateful to Dr. Blust for all that he taught me about the Austronesian language family and for his support of me as a scholar. Thank you to Jen Kanda and John Kawahara, our administrative team: you make good things happen everyday and I am indebted to you for your help on so many occasions. Thank you to the Bilinski Educational Foundation for funding the collection of Sundanese data used in this project, and the completion of this dissertation through the generous Bilinski Dissertation Fellowship. Thank you to Hiroko Sato and Victoria Chen for welcoming me to UH and cheering me on in my research. Thank you to my classmates for the camaraderie, laughter, and tears we have shared, especially to members of my cohort and research buddies including Jenny Sou, Sydney Ludlow, Khairunnisa, Dan Lin, Jacob Hakim, and Greg Vondiziano.

Thank you to my friends on O'ahu, especially Sandy Chong, Nalani Cook, Josie Young, Kerrie Wong, Linda Yamamoto and the Bluewater Ohana for the many meals we shared, for being great listeners, and for helping me through the pandemic years and my move off-island.

I express my gratitude and love for my parents, who sacrificed for us to have opportunities

they did not. Thank you also to my Auntie Mooi Hiang, my late Auntie Mee Kin, and my mother-in-law Alice, who supported me in my studies, and to my Uncle Kok Chew, who hosted me on an earlier fieldwork trip to Malaysia. It is not lost on me that from our humble beginnings, through your example and support, I have been able to complete my doctoral studies, and I will always be grateful for our family.

To Eldwin, thank you for believing in me all along the way, for humoring my half-baked ideas, for helping me to pick up my literal and figurative messes, and for caring for us in your actions daily. To my daughter and son, thank you for making me laugh, for helping me to stop work and get out to play, and for teaching me new words, ideas, and ways to do things. I am proud of you three every day and I love you.

Finally, thanks be to God, the Creator and Sustainer of all good things.

Abstract

This dissertation investigates applicatives in the western Austronesian languages of Indonesia, Malaysia, Brunei, and Singapore—that is, West Nusantara—and adjacent areas of the Philippines and mainland Southeast Asia. As used in this study, an applicative construction (AC) is a kind of clausal construction in which overt morphology on the verbal complex coincides with the selection of a peripheral semantic role (e.g. beneficiary, goal, instrument) as a core clausal argument. In many of these languages, applicative alternations signalled by such verbal morphology—as well as causative, aspectual, and semantic alternations signalled by the same morphemes—shape and color the use of verbal predicates throughout the entire language.

A primary goal of the study is to understand the applicative systems of West Nusantara in typological context, but also on their own terms, in the context of the diachronic and synchronic systems in which they developed and are used. Special attention is also given to broadening the description and cross-linguistic comparison of West Nusantara ACs and their functions, properties, and usage.

In Part I of the study, I present a case study of applicatives in Sundanese and show how these data and similar examples in other West Nusantara languages present several problems under previous approaches. These include the non-discrete nature of applicative and causative functions marked by applicative morphemes (AMs), and marking of canonical and non-canonical ACs with the same AMs. I argue for a constructional approach—one in which ACs are viewed as associations of form and meaning at different levels of specificity—which allows these data to be better accounted for. I also show how the unique Philippine-type voice systems found in western Austronesian languages may be understood using a *descriptive category* of symmetrical voice and a *comparative concept* of applicative (see Haspelmath 2010), and I propose the terms *pivot-selecting applicatives* for Philippine-type locative voice (LV) and circumstantial voice (CV) constructions as opposed to *pivot-neutral applicatives* found in the so-called ‘Indonesian-type’ languages.

Part II presents a typological survey examining 85 languages from lower-level subgroups indigenous to West Nusantara. Based on distributional patterns and evidence from languages showing transitional states, I argue that the development of pivot-neutral applicatives is associated with the demise of Philippine-type voice, but not the rise of a coherent ‘Indonesian-type’ grammatical profile. I further argue that the pivot-neutral ACs selecting locations and goal roles are derived from earlier LV constructions in Proto Malayo-Polynesian, while the pivot-neutral ACs selecting beneficiaries, instruments, and/or themes are derived from earlier CV constructions. Earlier LV morphology gives rise to pivot-neutral locative/goal AMs, while many benefactive/instrumental AMs are reflexes of the Proto Austronesian CV imperative suffix *-an. However, this *-an has been replaced with newer suffixes like *-kan* and *-akən* in a number of subgroups.

In Part III, I develop a typology of ACs and other AM-marked constructions in West Nusantara languages according to functional and formal properties. Notably, beneficiary-selecting ACs are much more likely to be valency-increasing while most other ACs are more likely to show remapping of the selected peripheral role and patient/theme. The observed patterns underscore that ACs have their own consistent, non-derived properties. I also explore the relationship between lexical semantics and functions of AMs. Across languages, some lexical bases show consistent attraction to constructional meanings of AMs based on compatible semantic properties. Large variance is observed, however, in the productivity of constructions with different functions across the lexicon and across languages. I conclude that functional patterns for applicatives observed in better-known languages like Indonesian and Javanese cannot be generalized to Sulawesi languages and other West Nusantara languages spoken outside of a narrow band of western Indonesia.

Table of Contents

List of figures	xii
List of tables	xiv
Abbreviations	xv
Part I: The problem and current approaches	1
1 Introduction	2
1.1 About this study	2
1.2 Background and sources of data	3
1.3 About “the researcher”	4
1.4 Framework and Assumptions	5
1.4.1 Applicatives as constructions	5
1.4.2 Grammatical relations and syntactic properties of arguments	7
1.4.3 Semantic roles and mapping of roles to argument structure	9
1.5 Scope of the study	10
1.6 Previous research on applicatives in West Nusantara languages	13
1.7 Organization	15
2 Case study: Sundanese applicatives	16
2.1 The Sundanese language and community	17
2.2 Basic morphosyntax	18
2.2.1 Intransitive verbal clauses	19
2.2.2 Monotransitive verbal clauses	21
2.2.3 Ditransitive verbal clauses	23
2.3 Overview of applicative morphology	26
2.4 Locative-selecting constructions marked with <i>-an</i>	29
2.4.1 Location-selecting constructions	29
2.4.2 Goal-selecting constructions	31
2.4.3 Path-selecting constructions	32
2.4.4 Category-changing locative-selecting construction	34
2.5 Theme- and instrument-selecting constructions marked with <i>-keun</i>	34
2.5.1 Theme-selecting constructions	35
2.5.2 Instrument-selecting constructions	38
2.6 Mapping alternations for theme- and locative-selecting applicatives	39

2.7	Beneficiary-selecting constructions marked with <i>pang-</i> <i>-keun</i>	41
2.8	Beneficiary-selecting constructions marked with <i>-keun</i>	46
2.9	Other ACs	47
2.10	Other AM-marked constructions in Sundanese	53
2.10.1	Causative constructions	53
2.10.2	Pluractional aspect	55
2.10.3	Marking with multiple AMs	56
2.10.4	Comparatives and other degree constructions	57
2.11	Summary and discussion	58
3	Approaches to applicatives in the literature and this study	61
3.1	What are applicatives?	61
3.1.1	Origins of the term <i>applicative</i>	61
3.1.2	Commonalities in definitions of applicatives	62
3.1.3	Differences in definitions of applicatives: What do applicatives do?	65
3.2	Problematic constructions and their significance	68
3.2.1	Problematic constructions in previous research on other languages	68
3.2.2	Polyfunctionality and problematic constructions in languages of West Nusantara	72
3.3	Constructional approaches as a lens for West Nusantara applicative systems	75
3.4	On the problem of Philippine-type voice in a study of applicatives	77
3.5	On the relationship between serial verb constructions and applicatives	84
3.6	West Nusantara applicatives in typological perspective	86
	Part II: Typological survey	88
4	Introduction to the typological survey of applicative systems in West Nusantara	89
4.1	Goals of the survey	89
4.2	Selection of languages	90
4.3	General properties of languages considered	93
4.3.1	Geographic location	93
4.3.2	Genetic grouping	93
4.4	Structural features of language considered	93
4.4.1	Word order	93
4.4.2	Morphological and syntactic alignment	94
4.4.3	Symmetrical voice and diathesis	96
4.4.4	Morphological case marking	97
4.4.5	Order of Noun-Possessor	97
4.4.6	Morphological causative constructions	98
4.5	Properties of applicative systems considered	99
4.5.1	Presence of applicatives	99
4.5.2	Co-occurrence with other constructions	99
4.5.3	Indexing of the applied phrase	100

4.5.4	Syntactic properties of the applied phrase	101
4.6	Properties of applicative morphemes considered	101
4.6.1	Applicative morphemes and source morphology	102
4.6.2	Semantic role of the applied phrase	103
4.6.3	Co-occurrent marking with other morphology	103
4.6.4	Other functions and constructional meanings	104
4.7	Summary	104
5	Results: Distribution and properties of languages of West Nusantara with and without applicatives	106
5.1	Overview of results and multivariable analysis	107
5.1.1	Testing of individual features for association with applicatives	107
5.1.2	Multivariable modeling of features: Random forest analysis	108
5.2	Detailed results for structural properties of language	110
5.2.1	Word order	110
5.2.2	Morphological and syntactic alignment	111
5.2.3	Symmetrical voice and diathesis	114
5.2.4	Other structural features	116
5.3	Location and genetic affiliation: Summary of results	116
5.4	Sumatra and the Barrier Islands	120
5.4.1	Outliers in Sumatra	123
5.5	Java and Madura	124
5.6	The Lesser Sundas	126
5.7	Mainland Southeast Asia and Peninsular Malaysia	128
5.7.1	Language contact, restructuring, and the loss of applicatives	131
5.8	Borneo and the Southern Philippines	132
5.8.1	Sabahan Languages	132
5.8.2	Land Dayak languages	136
5.8.3	North Sarawak and Melanau-Kajang languages	138
5.8.4	Greater Barito languages	142
5.8.5	Other languages spoken in Borneo	146
5.8.6	Possible explanations for the lack of applicatives in Borneo south of Sabah	148
5.9	Sulawesi	149
5.9.1	South Sulawesi languages	149
5.9.2	Celebic languages	153
5.9.3	Outliers in Sulawesi	155
5.9.4	The transition from Philippine-type voice to pivot-neutral applicatives in Sulawesi languages	155
5.10	Summary of major findings	164
6	Results: Properties of applicative constructions and their distribution in West Nusantara	166
6.1	Applicative morphemes marking pivot-selecting constructions	167
6.1.1	Voice morphology in Proto Austronesian and Proto Malayo-Polynesian	167
6.1.2	Distribution and functions of LV and CV in the sample	168

6.1.3	Forms of morphemes marking LV in the sample	171
6.1.4	Forms of morphemes marking CV in the sample	171
6.1.5	Implications for diachronic development	173
6.1.6	On the emergence of *-an as the marker of CV in indicative mood	174
6.2	Applicative morphemes marking pivot-neutral constructions	176
6.2.1	Applicative systems by number and function of applicative morphemes	176
6.2.2	Forms and development of pivot-neutral locative/goal applicatives	185
6.2.3	Forms and development of pivot-neutral applicative morphemes selecting beneficiaries, instruments, and themes	189
6.3	Irregular morphological marking of ACs	195
6.4	Polyfunctionality of West Nusantara applicative morphemes	196
6.4.1	Causative functions of applicative morphemes	196
6.4.2	Use in irregular causative constructions	199
6.4.3	Aspectual and intensive meanings of applicative morphemes	200
6.4.4	Comparative and other degree constructions	203
6.5	Syntactic properties of applicative constructions	204
6.5.1	Co-occurrence with other major voice constructions	204
6.5.2	Syntactic properties of the applied phrase in pivot-selecting ACs	206
6.5.3	Syntactic properties of the applied phrase in pivot-neutral ACs	207
6.5.4	Alternations in the realization of the applied phrase	214
6.6	Summary of findings and implications	218

Part III: Form and meaning in West Nusantara applicative constructions 220

7	A functional typology of applicative constructions in languages of West Nusantara	221
7.1	Introduction	221
7.2	Sampling of languages	222
7.3	Basic morphosyntax	224
7.4	Applicative morphology	230
7.5	Beneficiaries and recipients	230
7.5.1	Distribution and general properties of beneficiary-selecting ACs	230
7.5.2	Beneficiary-selecting ACs with special properties	234
7.5.3	Properties of the applied phrase and interactions with voice/valency	236
7.6	Instruments and themes	241
7.6.1	Instrument-selecting ACs	241
7.6.2	Theme-selecting ACs	244
7.7	Goals and locations	247
7.7.1	Distribution and general properties	247
7.7.2	Locative/goal-selecting ACs with special properties	249
7.7.3	Addressees and recipients marked like goals	250
7.8	Circumstantial and comitative roles	252
7.9	Other applied phrases	254
7.9.1	Content-selecting ACs	254
7.9.2	Stimulus-selecting ACs	254

7.9.3	ACs with verbs of bodily function verbs	256
7.10	Aspect, intensity and other semantic effects	256
7.11	Causative AM-marked constructions	259
7.12	Applicatives and voice	263
7.13	Summary of findings	267
8	Exploring West Nusantara applicative constructions through the lexicon	269
8.1	Introduction and rationale	269
8.2	Language sample and sources of data	271
8.3	Methods used for coding and sampling lexical meanings	272
8.4	Overview of results	273
8.5	Lexical meanings showing consistent patterns of association	278
8.5.1	Beneficiary-selecting constructions	278
8.5.2	Theme-selecting constructions	280
8.5.3	Instrument-selecting ACs	282
8.5.4	Goal-selecting ACs	283
8.5.5	Locative-selecting ACs	285
8.5.6	Causative constructions	286
8.5.7	Pluractional meanings	286
8.6	Summary of findings and implications	288
9	Conclusion	292
Appendix A: Language sample for typological survey		298
Appendix B: Questionnaires used in the typological survey		302
B.1	Questionnaire A	302
B.2	Questionnaire B	307
Appendix C: Database for typological survey		310
Appendix D: Sampling of meanings for the lexical study		316
Appendix E: Database for lexical study		319
Appendix F: Glossary of terms		338
References		343

List of Figures

1.1	Map of Southeast Asia	12
4.1	Languages included in the sample for the typological survey	92
5.1	Importance of features for classification by presence of applicatives	109
5.2	Importance of features for classification by presence of pivot-neutral applicatives	110
5.3	Map of languages of the sample by presence of applicatives and type	119
5.4	Overview map of languages of Sumatra and the Barrier Islands	121
5.5	Overview map of languages of Java and the Lesser Sundas	125
5.6	Map showing location for Suriname Javanese in South America	125
5.7	Overview map of languages of mainland Southeast Asia and northern Peninsular Malaysia	129
5.8	Overview map of languages of Borneo and the Southern Philippines	133
5.9	Map showing location for Merina Malagasy	143
5.10	Overview map of languages of Sulawesi	150
6.1	Pivot-selecting and mixed applicative systems of West Nusantara	169
6.2	Pivot-neutral and mixed applicative systems of Sulawesi	178
6.3	Pivot-neutral and mixed applicative systems of West Nusantara	179
6.4	Causative functions of applicative morphemes in West Nusantara	197
6.5	Causative functions of applicative morphemes in Sulawesi	198
6.6	Aspectual, intensive, and comparative functions of applicative morphemes in Su- lawesi	201
6.7	Aspectual, intensive, and comparative functions of applicative morphemes in West Nusantara	202
8.1	Productivity of functions across all languages and AMs	276
8.2	Productivity of functions across AMs by type and language	277
8.3	Lexical meanings associated with beneficiary-selecting ACs	279
8.4	Lexical meanings associated with theme-selecting ACs	281
8.5	Lexical meanings associated with inst-selecting ACs	282
8.6	Lexical meanings associated with goal-selecting ACs	284
8.7	Lexical meanings associated with locative-selecting ACs	285
8.8	Lexical meanings associated with causative AM-marked constructions	287
8.9	Lexical meanings associated with pluractional AM-marked constructions	289

List of Tables

1.1	Semantic role categories used in this study	10
2.1	Monotransitive Sundanese verbs	22
2.2	Sundanese applicative morphemes	26
2.3	Sundanese goal-selecting applicative constructions compared	58
3.1	Commonly-accepted definitions of applicatives	64
3.2	Constructions which pose challenges to definitions of applicatives	70
3.3	Constructions in West Nusantara languages that pose challenges to definitions of applicatives	73
3.4	Overview of Balinese voice system	79
3.5	Partial paradigm for Balinese verbs	79
3.6	Overview of Kimaragang voice system	81
3.7	Typological dimensions relating applicatives to similar clausal constructions	86
5.1	Results of Fisher’s exact tests by feature	108
5.2	Verb-P word order and presence of applicatives	111
5.3	Clausal word order and presence of applicatives	112
5.4	Morphological alignment and presence of applicatives	112
5.5	Voice system and presence of pivot-neutral applicatives	114
5.6	Location and presence of applicatives	117
5.7	Genetic affiliation and presence of applicatives	118
5.8	Typological survey results for Sumatra and the Barrier Islands	120
5.9	Typological survey results for Java and Madura	126
5.10	Typological survey results for the Lesser Sundas	127
5.11	Typological survey results for mainland SE Asia and Peninsular Malaysia	129
5.12	Typological survey results for Sabahan languages	134
5.13	Partial paradigm for Tatana verbs	135
5.14	Typological survey results for Land Dayak	137
5.15	Typological survey results for North Sarawak and Melanau-Kajang	139
5.16	Typological survey results for Greater Barito languages	142
5.17	Applicatives in Sama-Bajaw languages	144
5.18	Typological survey results for South Sulawesi languages	151
5.19	Typological survey results for other Sulawesi languages	154
5.20	Partial paradigm for voice and pivot-neutral applicatives in Bobongko	161

5.21	Partial paradigm for voice and pivot-neutral applicatives in Pendau	163
6.1	Proto Austronesian voice morphology (Chen 2017: 151)	167
6.2	Morphologically-marked voice distinctions in Philippine-type systems of the sample	168
6.3	Proto Austronesian non-indicative verbal morphology revisited	174
6.4	Morphological marking for pivot-neutral applicatives	177
6.5	Languages with two distinct forms marking pivot-neutral applicatives	180
6.6	Languages with one distinct form marking pivot-neutral applicatives	181
6.7	Languages with more than two distinct forms marking pivot-neutral applicatives	182
6.8	Locative/goal-selecting applicative morphemes	187
6.9	Morphological marking for locative applicatives and TAM in selected languages .	188
6.10	Beneficiary/instrument/theme-selecting applicative morphemes	190
6.11	Pivot-neutral applicative morphemes of the -AN type	191
6.12	Benefactive/instrumental applicatives and TAM in Javanese	192
6.13	Applicative morphemes of the -K and -AK types	193
6.14	Applicative morphemes with a comparative degree function	203
6.15	Bound-indexing of the applied phrase	208
6.16	Mapping of the applied phrase to the privileged syntactic argument	208
6.17	Access to relativization for the applied phrase	209
6.18	Fusion of Mori Bawah <i>-ako</i> with absolutive pronouns (Mead 2005: 704)	210
6.19	Muna pronominal suffixes for direct and indirect objects	211
7.1	Language sample used for the functional typology	223
7.2	Voice and case marking in languages of the sample	225
7.3	Applicative morphology by semantic role of the applied phrase	231
8.1	Language sample used for the lexical study	271
8.2	Coding categories used for the lexical study	274
8.3	Lexical data collected by language	275
8.4	Applicative morphemes by language and functional type	276
A.1	Languages included in sample for typological survey	299
C.1	Survey data for Questionnaire A	311
C.2	Survey data for Questionnaire B	313
E.1	Lexical data by base meaning	320

Abbreviations

Abbreviation	Description
I, II, III	denotes inclusion in sets of pronominal forms
1	first person
2	second person
3	third person
A	most agent-like argument of transitive clause
ABS	absolutive
ABIL	abilitative
AC	applicative constructive
ABS	absolutive
ACC	accidental or unexpected
ADDR	addressee role
ADVZ	adverbializer
AF	actor focus
AGT	agent role
AM	applicative morpheme
APPL	applicative
APASS	antipassive
AppP	applied phrase
AV	voice in which A is the privileged syntactic argument
BC	base construction
BEN	benefactive, beneficiary role
BIV	bivalent
BV	beneficiary voice
CAUS	causative
CIRC	circumstantial role, e.g. reason or purpose
CL	classifier
CMPL	completive
CN	common noun
CONT	content role
COM	comitative role
COMP	comparative
CORE	core argument
CN	common noun

CV	circumstantial voice
DAT	dative
DEF	definitive
DEIC	deictic marker
DEM	demonstrative
DIR	directional
DIST	distal
DISTR	distributive
DV	dative voice
EMPH	emphatic
EXCL	exclusive
FAM	familiar
GEN	genitive
GER	gerund
GOAL	goal role
HON	honorific
HORT	hortative
ICP	incompletive
IMP	imperative
INCL	inclusive
INST	instrumental, instrument role
INTENS	intensive
INTR	intransitive
IRR	irrealis
IV	instrumental voice
LNK	linker
LOC	locative, location role
LV	locative voice
MED	medial
NEG	negative
NM	noun marker
NMLZ	nominalizer
NOM	nominative
NP	noun phrase
NPST	non-past tense
NPIV	non-pivot core argument
NRLS	non-realis
NVOL	non-voluntary
OBJ	object
OBL	oblique
P	patient-like argument of transitive clause
PART	participial
PASS	passive
PAT	patient role
PERF	performance role

PFV	perfective
PFX	prefix
PL	plural
PLUR	pluractional
PIV	pivot argument
PN	personal name marker
POSS	possessive
POT	potentive
PP	prepositional phrase
PRED	predicative
PREP	preposition
PROG	progressive aspect
PRON	pronominal marker
PROX	proximal
PRT	particle
PST	past tense
PUNCT	punctiliar aspect
PV	voice in which P is the privileged syntactic argument
R	recipient-like argument of a ditransitive clause
REAS	reason role
RV	referent voice
RDP	reduplication
REL	relativizer
RLS	realis
SEQ	sequential
SF	stem former
SG	singular
STIM	stimulus role
STR	semi-transitive
SUBJ	subject
T	theme-like argument of a ditransitive clause
THM	theme role
TR	transitive
TV	theme voice
VERB	verb marker

Part I: The Problem and Current Approaches

Chapter 1

Introduction

1.1 About this study

The term applicative, as used in this study, refers to a clausal construction in which overt morphology on the verb or verbal complex coincides with the selection of a peripheral semantic role, such as a beneficiary, goal, or instrument, as a core clausal argument. In this study, I seek to understand the nature of such applicative constructions (ACs) in the Western Austronesian languages of western Indonesia, Malaysia, Brunei, and Singapore—a region that I refer to as West Nusantara—and adjacent areas of the Philippines and mainland Southeast Asia. In many of these languages, the applicative alternations signalled by such verbal morphology—that is, applicative morphemes, or AMs—as well as causative, aspectual, and semantic alternations signalled by the same AMs, are a foundational and distinctive property that shape and color the use of verbs throughout the entire language.

In this study, I seek to understand the applicative systems of West Nusantara in cross-linguistic typological context, but also especially, on their own terms, in the context of the diachronic and synchronic systems of language in which they developed and are used. To this end, I have adopted a conceptual framework and approach that allows me to capture two important properties of such applicatives, which have proved problematic for previous approaches. First, the morphology that marks ACs in these languages is highly polyfunctional, yet the associated functions or meanings—applicativization, causativization, aspectual meanings, and so on—are far from discrete, but instead interconnected. Second, the distribution of these functions or meanings, and the forms of argument structure by which they are expressed, show a complex and varied distribution—across lexical stems, across clausal types (e.g. voice, transitivity, pragmatic meaning), and across AM forms, when multiple are available in a given language. This distribution is difficult or nigh impossible to predict or describe by rule, yet competent speakers of these languages clearly have internalized the patterns—both those which are generalizable and those which are highly specific—that give rise to this observed distribution.

In the remainder of this introductory chapter, I describe the background of the study and sources of data used (§1.2), my positionality as “the researcher” (§1.3), the framework and assumptions I have adopted, including the definition of applicatives used (§1.4), the scope of the study (§1.5), previous research on applicatives in West Nusantara languages (§1.6), and the organization of the dissertation (§1.7).

1.2 Background and sources of data

This study began as a project for a two-semester series of courses on Field Methods in Linguistics at UH Mānoa in Fall 2018 and Spring 2019. At that time, I met Dewi Setiani, who is a native speaker of Sundanese and was in residence as a graduate student in the Education Department of UH Mānoa. I began working with Dewi to describe the verbal system of Sundanese, starting with spoken narrative texts that my classmates and I recorded with Dewi, and with a particular interest in applicative alternations.

Following this, Bradley McDonnell and I began to compile a database of languages of western Indonesia which show causative and applicative alternations in their verbal systems. This collaboration led to a number of small scale studies of applicatives in western Indonesian languages that I conducted together with him and independently, including some which were presented in panel discussions and published in collections addressing typological variation in ACs in the world's languages (Truong & McDonnell 2022, Truong & McDonnell 2021; McDonnell & Truong 2024; Truong 2021). These discussions and collaborations revealed ways in which western Austronesian applicatives have been inadequately represented in the typological literature, and that constructions marked with AMs in the languages show a number of properties that are not commonly found for AMs in other language families, as well as some others which are found, but are often not considered core properties of applicatives (see Pacchiarotti & Zuñiga 2022, and further discussion in Chapter 3).

Around the same time, I put on hold some prior plans to conduct fieldwork in Southeast Asia due to practical difficulties with travel and personal concerns. In the summer of 2021, Dewi Setiani and I began working together again, by which point she had returned to Bandung, Indonesia following the completion of her M.A. degree. We met remotely via video call for elicitation sessions and annotation of original and published Sundanese language materials. She also facilitated the making of new recordings in Bandung with herself as a speaker, and when safe, with a few other speakers. The initial period of data collection in 2021 was funded by a Pre-dissertation Research Award from the Bilinski Educational Foundation. I continued working with Dewi into Fall of 2022, using personal funds and proceeds of a generous merit-based award from the Graduate Student Organization of UH Mānoa for excellence in teaching. I also began work on an expanded typological database of languages of West Nusantara and their attested AM-marked constructions based on existing corpora, descriptive, and pedagogical resources. Data compilation, analysis, and the writing of this dissertation were funded by a generous Bilinski Dissertation Fellowship. Some parts of the database were compiled based on personal fieldnotes from Bradley McDonnell including data provided by Khairunnisa for Ampenan Sasak, Hendi Feriza for Besemah (South Barisan Malay), and Johan Safri and Wawan Sahrozi for Nasal. Together, these sources of data constitute the basis for this study.

1.3 About “the researcher”

*and all of my visions and interpretations / depend on what I see, /
and between my eyes is always / the rain, the migrant rain.*¹

Li-Young Lee, Chinese-American poet born in Jakarta, Indonesia

In order to provide context about the perspective represented in this study, in this section I discuss my positionality as “the researcher” behind it. I am a non-indigenous scholar in the context of West Nusantara, and I identify as a member of a minoritized community in my own national context—the Asian American community—and a member of a larger diaspora community—the ethnic Chinese diaspora. I was born to immigrant parents of Teochew and Hakka ethnic descent who emigrated from Malaysia to the United States and raised me in a working-class neighborhood of the East Bay Area in northern California. I was born into a Cantonese-speaking household, but after the time I entered school in the Bay Area, I retained only passive language ability in my mother language. In this environment, English was the dominant language and white American identity was the dominant cultural paradigm.

My sense of cultural identity and the wider context for it was influenced by educational experiences first at UC Berkeley, and then at UCLA in the Asian American Studies Department, and the study of world literatures and languages at UCLA. I later earned an M.A. in Applied Linguistics and my advisor, Dr. Brian O’Herin, helped shaped my view of the diversity of language through his familiarity with languages of the Caucasus and Aboriginal languages of Australia, and his interest in linguistic theories in which such languages might be analyzed according to their observed natural structures.²

As an adult, I had the opportunity to spend several years in Indonesia and later, Malaysia. My studies in the BIPA program (Indonesian for Foreign Speakers) at Universitas Indonesia, Depok, first peaked my interest in applicatives. At the time, I did not know to call these *applicatives*, only that I needed to painstakingly memorize the meaning of verbs bearing the suffixes *-i* and *-kan*—and learn to interpret verbs marked with *-in* from Betawi—to have any hope of communicating with Indonesian people around me. Afterwards, I worked with an international NGO to support multilingual education in Indonesia through research, training, and materials development, in partnership with local governmental and non-governmental organizations. This work, my relationships with my Indonesian colleagues, and our field visits around the country impressed on me the importance of small and underresourced languages. In 2009, I had the opportunity to share about some of our work in community engagement at the First International Conference for Language Documentation and Conservation, where I became interested in language documentation and the University of Hawai‘i. I later lived in Malaysia for a few years to connect with my extended family, while pursuing opportunities for involvement in research and language documentation with speakers of Austronesian languages and speakers of Austroasiatic languages from *Orang Asli* communities.

In the course of my studies at the University of Hawai‘i at Mānoa, I have had the opportunity to make a number of short exploratory field trips to Central and West Sulawesi in Indonesia

¹From the poem “Visions and Interpretations” by Li-Young Lee (1986: 68–69).

²Though I was not aware of it at the time, O’Herin’s (2001) work on applicatives in Abaza, a language of the Caucasus, pushed the boundaries of the typological understanding of applicatives (see §3.2.1).

and Perak State in Malaysia, funded by the Bilinski Educational Foundation. From 2020–2022, I worked from Honolulu as a graduate assistant on an NSF-funded project to document the Baduy Dalam language of West Java, Indonesia. During this time, I interacted regularly via video call with professors on the research team from Universitas Pendidikan Indonesia (UPI), and I supervised three student research assistants from UPI to transcribe and annotate audio/visual documentary resources in Baduy. As described in the previous section, I have worked in person and remotely with my friend and collaborator Dewi Setiani, to collect data for this study and produce documentary resources in Sundanese.

In the context of my research and interactions with people from various community identities in West Nusantara, I view myself as occupying an intermediate—sometimes paradoxical—space between a cultural insider and outsider. I am at times treated with hesitancy or undue deference as a foreigner, a Westerner, and an outsider. At other times, I am perceived and treated as a person of Chinese racial identity in the local context, which may be colored with distrust based on historical acts of exploitation by people of my ethnic background against others, or alternately, marked with solidarity on the basis of more favorable experiences. Sometimes, I am regarded as “one of us,” or close enough to it, on the basis of my family background, appearance, mannerisms, and shared goals and experiences with community members.

In settings where I am perceived as sharing local or insider identity, I am mindful of my responsibilities to act according to local roles as appropriate—guest, sister, co-worker, and friend—and to participate in communal work, play, talk, and sharing of resources. In settings where I am perceived as holding Western or outsider identity, I am mindful of my responsibility to respect the experiences and stated interests of others, and I seek to share access to recognition, resources, and relationships in ways that others value. Often, these two realities overlap, in ways that are both confusing and familiar to me as an Asian American and the child of immigrants. To this day, I am still learning to navigate intercultural relationships well, and to follow the words of the Indonesian proverb: *Di mana bumi dipijak, di situ langit dijunjung*.³

1.4 Framework and Assumptions

The methods and lines of inquiry adopted in this study have been influenced by a number of typological, functional, and constructional approaches in linguistics. In this section, I lay out the assumptions that underlie this study and approach that I use in discussing applicatives.

1.4.1 Applicatives as constructions

The definition of applicative that I will use is given below in (1).

- (1) An applicative construction (AC) is a kind of clausal construction in which overt morphological marking on the verbal complex coincides with the selection of a non-agent, non-patient semantic role to map to a core argument in the clause.

³This saying may be translated, “Wherever one walks on earth, there one upholds the sky.”

An example of an AC from Sundanese is given in (2).⁴ I use the term *applicative morpheme* (AM) to refer to the overt morphological marking that identifies an AC together with its argument structure, as described in (1). These markers themselves have also been referred to as applicatives elsewhere in the literature. In example (2), the AM is the circumfix *pang-* *-keun*.

(2) Sundanese, Beneficiary-selecting applicative

- a. *Asep m-(b)awa sangu keur Lilis.*
 A. AV-bring cooked.rice for L.
 ‘Asep brought food for Lilis.’ (BC) (SLD-025)
- b. *Asep m-(p)ang-mawa-keun Lilis sangu.*
 A. AV-BEN.APPL-bring-BEN.APPL L. cooked.rice
 ‘Asep brought food for Lilis.’ (AC) (SLD-043)

One reason that ACs have been noticed, both by analysts and speakers alike, is that pairs of clausal constructions are observable in language, for which the lexical verb used is the same in both, but the presence of the AM co-varies with selection of a non-agent, non-patient role as a core argument. Such a pair of constructions is found in the relationship between (2a) and (2b).

Note that in (2b), the verb marked with the circumfix *pang-* *-keun* takes a semantic beneficiary, *Lilis* ‘personal name’ as a core argument, as evident from its coding without any prepositional marking, while in (2a), the sole non-agentive core argument is a theme *sangu*, ‘rice’, and the beneficiary is marked with the preposition *keur* ‘for’. Following convention, in such a pair of clausal examples, I refer to the example which lacks marking with the AM as a *base construction* (BC) example. When a pair of constructions like that shown in (2) exists, I treat this like a minimal pair in phonological analysis; the contrast between the two example clauses constitutes one type of evidence that helps an analyst to identify the morphologically marked construction as a likely example of an applicative. I do not, however, assume that the clause constituting an AC and the corresponding clause constituting a BC are related by means of a derivational process, or that the correspondence seen between two such clauses is definitional of applicatives (cf., for example Aikhenvald & Dixon 2011; Dixon & Aikhenvald 2000, see also discussion in §3).

Though I am certainly not the first to define an applicative as a type of construction (e.g. Peterson 2007), the adoption of a constructionist approach is a distinctive foundation for this study. Throughout this study, I argue that adopting a constructional approach to applicatives allows us to develop an insightful typology of western Austronesian applicative systems based on meaningful characteristics that influence the distribution, usage and development of ACs and other AM-marked constructions. While this approach is reflected in the definition given above, I consider it important to further qualify a number of the components of an AC.

I use the term *construction* to refer to a conventionalized pairing of form and meaning (see Goldberg 1995; 2006). I understand an applicative to be a type of clausal construction, meaning that, in examples of clauses found in usage which constitute ACs, aspects of a *fixed form* and *consistent meaning* must coincide across variable non-fixed lexical content that fills open slots in the construction.

⁴For the sake of clarity, in glosses examples, initial stem consonants which are opaque due to morphophonological processes have been included in parenthesis, as in *m-(b)awa* ‘to bring, to carry, AV’ in (2), which is pronounced [mawa], and is formed on the stem *bawa*. Further details are given in §2.2.1.

The fixed form of an AC includes a number of key components. As mentioned above, ACs are characterized by overt morphological marking on the verbal complex, i.e. the presence of an AM. An AM is part of the fixed form of an AC, while compatible verbal bases constitute part of the variable, non-fixed lexical content that may fill an open slot in the construction. Furthermore, an AC is specified to be clausal; its form is consistent with that of a verbal clause. Based on Perek (2015), I understand that the fixed form of a clausal verbal construction, of which ACs are one type, includes the number and structural position of arguments of the verb that occur across observed examples in usage that have a consistent meaning. In the example of an AC in (2b), the fixed form includes two post-verbal NP slots, with the first mapping to a beneficiary role, and the second mapping to a theme role. In some cases, the fixed form of an AC also includes other fixed words or morphemes, such as a particular preposition or modifier, for example, or slots filled by one of a small, fixed set of words or morphemes.

In addition to a fixed form, as stated above, an AC has a consistent meaning. I understand this consistent meaning to include the semantic roles selected for clausal arguments and the mapping of these to slots in argument structure (see Perek 2015). In addition, the meaning of a AC can include particular semantic components ascribed to the event or state of affairs described by clauses that exemplify the construction. This is admittedly a broad statement, however, it must stand that such components are consistent in clauses with the same AM-marking, same (general) argument structure, and same mapping of semantic participants. Part of the undertaking in this study is to develop a typology of such components of consistent meaning observed across ACs in languages of West Nusantara, and to show that these meanings are integral to the development and use of these constructions. For example, the example in (2b) includes a semantic component of substitution, whereby the fact that the agent brings the expressed theme for the beneficiary relieves her of the need to do so herself.

Based on the discussion above, we may augment the definition given in (1), to form the summary statement below:

- (3) An applicative is a clausal construction in which overt morphological marking on the verbal complex coincides with the selection of a non-agent, non-patient semantic role to map to a core argument in the clause. Across verbal bases, an AC may be identified by the specific number and/or structural position of its arguments, the semantic roles that map to these positions, and other consistent semantic properties of the event or state of affairs described by the clause.

As evident in the components of ACs that I have identified thus far, I consider both syntactic and semantic properties to be identificational of applicatives. In light of this, I also lay out here in brief my assumptions about some important related topics: (i) grammatical relations and syntactic properties of arguments, and (ii) the nature of semantic roles and the relationship between semantic roles and argument structure.

1.4.2 Grammatical relations and syntactic properties of arguments

A *grammatical relation* is a syntactic relation that exists between a clausal argument and a clausal construction. I understand these relations to be construction-specific (see Bickel 2010), though

they may be compared across constructions to form meaningful generalizations. Syntactic properties are used to identify grammatical relations, and it is helpful here to distinguish two types of syntactic properties: (i) *syntactic coding* of arguments and (ii) *syntactic behavior* of arguments.

Syntactic coding typically refers to morphological case marking of the argument, agreement marking or indexing of the argument on the verbal complex, and linear word order of the argument vis-a-vis other syntactic constituents in the clause. In many languages of West Nusantara, one important type of syntactic coding is the selection of pronominal forms used to express or index an argument from a particular set, which is commonly observed to co-vary with argument role, and thus grammatical relations.

An example from Pendau, a language of Central Sulawesi, is shown in (4) below.⁵ In Pendau, pronominal core arguments are in most cases coded with forms drawn from the set labelled ‘absolute’ by Quick (2007), which includes first plural exclusive *ami* and third plural *jimo*, as shown in (4a). In certain verbal clauses, however, a pronominal argument expressing an agentive core argument must be drawn from the set labelled ‘genitive’, which includes third plural *nijimo* as shown in (4b) (for more on alignment patterns in West Nusantara languages, see §4.4.2). Oblique pronominal arguments are coded by means of the absolutive set marked with a preposition, which is not shown in the example, e.g. *sono jimo* ‘with them’.

(4) Pendau, Use of pronominal sets

- a. *Ami non-(t)uju jimo.*
 1PL.EXCL.ABS AV.RLS-send 3PL.ABS
 ‘We sent them.’ (AV)
- b. *Ami ni-tuju nijimo.*
 1PL.EXCL.ABS PV.RLS-send 3PL.GEN
 ‘They sent us.’ (PV)

(Quick 2007: 141)

Syntactic behavior refers to compatibility of a clausal argument with syntactic operations, such as relativization, raising, and control, among others. While syntactic behavior may be an useful indicator of grammatical relations, a number of important studies have shown that the syntactic behavior of arguments may differ between ACs and corresponding BCs, or across various types of ACs in particular languages (see e.g. Baker 1988b; Alsina & Mchombo 1990; Bresnan & Moshi 1990; Peterson 2007; among others).

Among the languages of West Nusantara we observe important differences across systems of voice and diathesis that drive alternations in the mapping of grammatical relations to argument structures (for more on this see §3.4, and §5.2.3), as well as differences in the categories and terms used to label grammatical relations in particular studies. Thus to facilitate meaningful comparison across languages and across constructions within languages, in describing and classifying ACs and other AM-marked constructions, I make frequent use of grammatical *macro-role labels* for arguments: S for intransitive clauses, A & P for monotransitive clauses, and A, T, & R for ditransitive clauses (see Comrie 1989; Haspelmath 2015).

For example, in Pendau, Quick (2007: 127–132) identifies a grammatical relation that may be called *pivot*, subject, or privileged syntactic argument. This relation is identified by relatively free

⁵In examples from cited sources, orthographic conventions have generally been kept. Glosses and glossing conventions have been adjusted, however, for the sake of consistency.

word order, eligibility to head a relative clause, quantifier float, and control patterns. The pivot relation is exhibited by the A argument in clauses like (4a), where the verb is marked with the prefix *noN-* in realis mode, while the pivot relation is exhibited by the P argument in clauses like (4b), where the verb is marked with the prefix *ni-* in realis mode. The former is called AV (for A-Voice, sometimes also called actor voice), while the latter is called PV (for P-Voice, sometimes also called patient voice or undergoer voice).

As part of the typological survey in this study, in order to include broad representation of the languages of West Nusantara, I have reviewed a large body of available literature on the morphology and syntax of individual languages and small clusters of languages (see §4.4 and §4.5). Unfortunately, in the majority of cases, information about the syntactic behavior of arguments in ACs is not addressed in any level of detail, and more often than not, the same is true of BCs. Thus, I have often relied on syntactic coding of arguments as evidence of grammatical relations when comparing and classifying ACs in particular languages. When information about syntactic behavior of arguments is available in source material (or can be determined from data therein), I have tried to include it, and in my own analysis of primary data from Sundanese (see Chapter 2), I have made some observations relevant to syntactic behavior of arguments, though this is not my primary focus.

1.4.3 Semantic roles and mapping of roles to argument structure

Semantic roles are labels that describe types of semantic relationships between a verb or other predicate and its arguments. Semantic roles form part of the semantic representation of a verbal clause, and are also commonly referred to as thematic relations, or participant roles. There are many approaches to semantic roles, and generally no agreement on a definitive list of roles which may have grammatical correlates in the languages of the world (see Comrie 1989: 58–59).

In this study, the set of labels for semantic roles that I have used largely corresponds to the thematic relations used by Van Valin (2005: 53–59); which are more generalized than verb-specific roles and less generalized than what Van Valin refers to as “semantic macro-roles.” However, for some lexical verbs that are commonly found to occur as bases in ACs in the languages of West Nusantara, none of the available labels for thematic relations in that source are appropriately descriptive, e.g. the role of an addressee in events of communication. In such cases, I have augmented the set of semantic roles with labels for semantically similar roles (‘frame elements’) used in Frame Semantics (Fillmore & Baker 2012) and the Framenet lexical database⁶ (see Ruppenhofer et al. 2016; Johnson & Fillmore 2000). The primary semantic role labels used in this study are listed in Table 1.1. Semantic role labels are also listed in the glossary in Appendix F.

A particular semantic role (or set of possible semantic roles) is assumed to be related to the argument structure of a clausal construction by a *mapping*. For the purposes of this study, the formal properties of such a mapping are not considered (for some examples of formalisms that express such relations see e.g. Van Valin 2005; Levin 1993; Perek 2015). My assumption is that a mapping relation constrains the interpretation of a clausal meaning, such that the referent of a particular position in the syntactic structure is interpreted as the type of participant specified by one semantic role (or one of a small set of roles) in the event or state expressed by the clause.

For example, in Pendau, based on the observed meanings of clauses with the lexical verb *tuju*

⁶<https://framenet.icsi.berkeley.edu/>

Table 1.1: Semantic role categories used in this study

Semantic role	Abbr.	Description
agent	AGT	an entity that intentionally causes some entity to be affected
addressee	ADDR	the recipient of a message in an event of communication
beneficiary	BEN	an entity that is affected advantageously by an event
comitative	COM	an entity that accompanies another in an event
content	CONT	the content or topic of an event of communication or cognition
goal	GOAL	the ending location of a entity that changes location
instrument	INST	an inanimate entity manipulated to some effect
location	LOC	the general or static location of an entity or event
reason	REAS	the prior reason to which the occurrence of an event is attributed
stimulus	STIM	the phenomenon that is perceived or brings about a sensation
theme	THM	an entity that is located in space or undergoes a change of location
patient	PAT	an entity that exists in an indicated state, or undergoes a change of state
performance	PERF	the experience generated by performers in a performing arts event
purpose	PURP	the intended purpose for which an event is undertaken

‘send,’ it can be inferred that *tuju* takes two core clausal arguments, one, an agent that initiates the sending movement, and the other, a theme that undergoes the sending movement. When the verb is marked with the AV prefix *noN-* as in (4a), the preverbal argument position maps to A, an agentive argument. Thus, the preverbal NP, *ami* ‘we (excl.)’, must be interpreted as the agent. The immediate postverbal argument position maps to P, a patientive argument. Thus *jimo* ‘them’ must be interpreted as the theme. Throughout this study, I will often refer to the mapping of semantic roles to argument structure, using the word *select*, e.g. in example (4a), a theme argument is selected as P.

Likewise for the AC from Sundanese in (2b), it can be said that a beneficiary and a theme role are both selected as non-A core arguments. In this example, the first postverbal NP slot maps to a beneficiary role, and the second postverbal NP slot maps to a theme role. Or, generalizing across similar clausal structures with two non-A core arguments in Sundanese, we may say that a beneficiary role is selected to map to the R argument, and a theme role is selected to map to the T argument in this type of AC, which is observed when *pang-* *-keun* appears on a lexical verb expressing caused motion or transfer, such as *bawa* ‘to carry, to take’.

1.5 Scope of the study

This study is focused on a subset of Austronesian languages, namely the Austronesian languages of the West Nusantara region and those languages spoken outside of this region that are most closely-related to these by way of exclusive subgrouping. The justification of this focus within the Austronesian family is described in the remainder of this section.

Because Bethwyn Evans (2003) has already conducted a comprehensive cross-linguistic study of applicatives in the Oceanic languages, which make up nearly all of the 450 or so Austronesian languages of the Pacific, it is reasonable to leave these aside in favor of the Austronesian lan-

guages of Asia and Madagascar. These total more than 700, and among them are found many languages with applicatives but not with uniform distribution. In particular applicatives have been associated by some authors with symmetrical voice languages of western Indonesia (Sumatra, Java, Kalimantan, and Bali) and central Sulawesi (Himmelman 2005: 175). Here, I would add that applicative morphology has been found in some Austronesian languages of East Nusantara (Indonesia east of Sulawesi in the north and Lombok in the south, and East Timor). Kambara, spoken in the lesser Sundas, is known to have applicatives (Klamer 1998), as is Taba (Bowden 2001) and perhaps a few other languages of the SHWNG (South Halmahera-West New Guinea) subgroup spoken on Halmahera Island and the Bird's Head of Papua. Still, it appears that applicative morphology may be less prevalent in East Nusantara than in West Nusantara, based on typological descriptions of characteristic features of languages of East Nusantara (see Klamer & Ewing 2010). As for Formosan languages—the 20 or so indigenous languages of Taiwan—and Philippine languages—the nearly 200 indigenous languages of the Philippines, together with the more than 150 Austronesian languages of East Nusantara, I set these aside for practical reasons, given that languages of western Indonesia and Sulawesi already number more than 200. Thus, in this study I primarily focus on a subset of Austronesian languages in the western and southwestern part of their total range.

Geographically, I include in this study the indigenous Austronesian languages spoken in Peninsular Malaysia, Singapore, Brunei, Sumatra, the Barrier Islands, Borneo, Sulawesi, Java, Madura, Bali, Lombok, and Sumbawa (see map of Southeast Asia in Figure 1.1). These have sometimes been referred to in the literature as western Indonesian languages but I find that this term unnecessarily excludes Malaysia, Singapore, and Brunei. Thus, as mentioned earlier, I use the term *West Nusantara* to refer to this geographic area and the *languages of West Nusantara* to refer to the languages belonging to the Austronesian family spoken therein.⁷ The languages of West Nusantara may be considered a subset of the western Austronesian languages. Following Himmelman (2002), the term *western Austronesian*, as used in the title of this study, is “a cover term for all Austronesian languages spoken in Taiwan, the Philippines, mainland Southeast Asia, western Indonesia (Sulawesi and all islands to the west of it), Borneo and Madagascar, and also including Palauan and Chamorro” (7).⁸ Using West Nusantara as a geographic focus has the effect of including a sizeable number of the western Austronesian languages that are most likely to be rich in applicatives, while limiting the scope of the study for practical reasons.

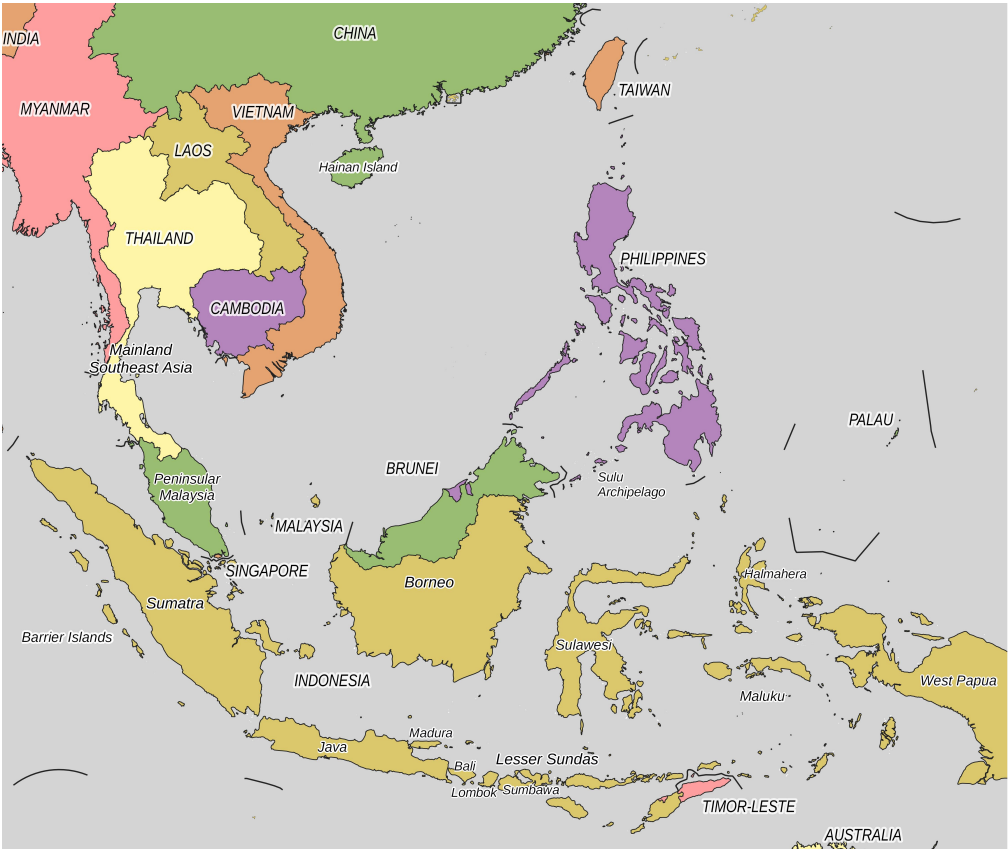
It also makes sense to include languages which reflect an exclusively shared period of common development with some subset of the Austronesian languages of West Nusantara, even if they are presently spoken outside of the area due to historical migration. Thus I include all languages belonging to certain genetic affiliations within Malayo-Polynesian. First, I include three well-defined higher-level subgroups primarily spoken in West Nusantara: Northwest Sumatra-Barrier Islands, including Gayo and Enggano (see Nothofer 1986);⁹ Celebic, including Tolitoli (see Mead 2003); and South Sulawesi (see Mills 1975; Adelaar 1994). Second, I include the Malayo-

⁷Nusantara is a term referring to “the Indo-Malaysian archipelago, generally, without respect to national borders” (Klamer & Ewing 2010: 1). See also Pappas 2022: 4, who defines West Nusantara as “the region of Austronesian languages west of the Wallis line outside of the Philippines.”

⁸This is a geographic designation rather than a genetic subgroup, and also should not be confused with the previously proposed subgroup called Western Malayo-Polynesian.

⁹In a forthcoming paper, McDonnell and Billings propose that Nasal, Enggano, and Gayo should be classified together with the Batak and Barrier Islands languages in an exclusive subgroup called Sumatran.

Figure 1.1: Map of Southeast Asia



Includes geodata from Natural Earth (public domain).

Chamic languages, which are broadly distributed across West Nusantara and parts of mainland Southeast Asia. Third, I include all other lower-level subgroups representing non-Malayic indigenous Austronesian languages of Borneo (e.g. Land Dayak, Greater Barito linkage languages including Malagasy, North Sarawakan, Sabahan, Melanau-Kajang, among others), Sumatra (i.e. Rejang, Lampungic, Nasal) and Java and the lesser Sundas as far east as Lombok (i.e., Javanese, Sundanese, Madurese, Bali-Sasak-Sumbawa). While the organization of these into higher-level groups within Malayo-Polynesian is not entirely clear, all members of these three sets have been proposed to subgroup at a higher level with Malayic and Chamic to the exclusion of Northwest-Sumatra-Barrier-Islands, Celebic, and South Sulawesi in one or more recent proposals (see Adelaar 2005a; R. A. Blust 2010; Smith 2017). Finally, I will exclude Moklen-Moken, Chamorro, and Palauan. These languages have been proposed to originate in Indonesia or the Philippines, but their genetic position within Malayo-Polynesian and migration history remain quite unclear.¹⁰ Altogether the languages defined here include a degree of typological diversity, while being centered on West Nusantara.

1.6 Previous research on applicatives in West Nusantara languages

In this dissertation, I aim to more closely investigate applicatives in languages of West Nusantara than has been done thus far in existing literature. In other language families, studies detailing the functions and properties of ACs and their non-canonical counterparts are found in increasing numbers and an increasingly wide range of languages (e.g. Bantu: Pacchiarotti 2020, Abkhaz-Adyge: O’Herin 2001, Salishan: Kiyosawa 2006, among others, see also §3.2). In contrast, only a few languages of West Nusantara and limited types of data are frequently represented in discussion of applicatives, either in broad cross-linguistic surveys (such as Polinsky 2013 and Peterson 2007) or in article-length investigations of applicatives (see below).

The languages of West Nusantara in which applicatives have received more attention include Indonesian (Dardjowidjojo 1971, 1974; Chung 1976; Kaswanti 1995, 1997; Arka 1993, 2009; Cole & Son 2004; Kroeger 2007), *Tukang Besi* (Donohue 1999, 2001), Javanese (Hemmings 2013; Nurhayani 2014; Vander Klok & Evans 2022), and to a lesser extent Balinese (Austin 2001; Arka 2003) and *Totoli* (Himmelmann & Riesberg 2013; Riesberg 2014b). A few authors have also written on the historical development of certain morphemes with applicative functions in particular languages or subgroups found in West Nusantara (e.g. Beratha 1992 for Balinese, Mead 1998 for *Bungku-Tolaki*, Adelaar 2011 for Javanese). Some have also written on the development of AMs in relation to proposed historical reconstruction or subgrouping at higher levels (e.g. Starosta, Pawley & Reid 1982; Sirk 1996; Ross 2002; Peterson 2007; Kikusawa 2012), though the data from languages of West Nusantara considered in such studies is not particularly detailed. On the whole, review of the literature on applicatives in West Nusantara languages reveals a number of significant gaps.

First, though grammatical descriptions of many individual languages of West Nusantara include verbal morphology with applicative functions, most of these languages have not been represented in any depth in typological studies of applicatives. Some examples include *Toba Batak* (van

¹⁰See Smith (2017), which classifies each of these three as a separate primary branch of Malayo-Polynesian.

der Tuuk 1971 [1864-1867]), Karo Batak (Woollams 1996), Nias (Brown 2001), Besemah (South Barisan Malay) (McDonnell 2016), Salako (Adelaar 2005b), West Coast Bajau (Miller 2007), Sundanese (Hanafi 1997), Betawi (Ikranagara 1975), Madurese (Davies 2010), Pendau (Quick 2007), Wolio (Anceaux 1952) and many more. Some new work on applicatives is becoming available, which is a promising sign (see e.g. Aznar, Döhler & Klok 2023). Still, the extent to which the properties of ACs across West Nusantara languages are similar is not at all clear; in fact it appears that even closely-related languages or dialects may show differences in their inventory of AMs and distribution of functions, e.g. Ampenan Sasak (Truong & McDonnell 2022) as compared to Ngenó-Ngené Sasak (Austin 2001). Thus, data from a broader group of West Nusantara languages should be represented in an adequate typology of ACs.

Second, the range of syntactic properties found for arguments in ACs (see §4.6.4) is not adequately represented in the existing literature. There is reason to believe that within individual languages, variance in the structural properties of ACs formed with the same stems might well be under-reported (Truong & McDonnell 2022). Many studies primarily exemplify constructed and elicited sentence data, and rely on only a limited number of examples. But even limited investigation of naturally-occurring language use reveals that there are different possible syntactic realizations in ACs for a selected peripheral participant, i.e. the *applied phrase*, and multiple treatments of the participant expressed as P in the base clause, i.e. the *companion phrase*. For example, simple searches in available corpora of Indonesian or Sundanese sentences (Goldhahn, Eckart & Quasthoff 2012) show that the applied phrase can be expressed either as a core argument or an apparent oblique, and the companion phrase can be retained as P, encoded as R or T in a ditransitive structure, or encoded as an apparent oblique (i.e. “demoted”). Sometimes argument structure co-varies with participant role of the applied argument. But other times, there appears to be variation in a single language between different possible argument structures for constructions with the same verbal stem, same AM, and same set of participant roles. Thus, even for ACs in languages of West Nusantara that are often-discussed and neatly exemplified in the literature, the true range of syntactic properties and possible structures extends beyond that which is commonly acknowledged.

Third, there has been little investigation of the influence of lexical semantics on applicatives in these languages, even though, there is reason to believe that it plays an important role. In English, it is well-documented that the verbal semantics conditions valency processes, e.g. ‘dative shift’, and other verbal alternations, e.g. passives, causatives, resultatives, and preposition drop (see Thompson 1990; Bresnan et al. 2007; Levin 1993; Rappaport Hovav & Levin 1998). Recent work in Bantu languages has also proposed that lexical semantics conditions selection of participant role for the applied argument, but does not directly determine syntactic structure in ACs (Jerro 2016, 2021). Verbal semantics is also implicated in some analyses of applicatives in West Nusantara languages (Kroeger 2007, Himmelmann & Riesberg 2013). Despite this, in most studies such ACs are exemplified with only a small number of verbal stems. For example, Cole & Son (2004) argue that Indonesian *-kan* functions to license syntactic arguments. To support this analysis, they exemplify ten verbal stems in ACs with *-kan* (plus a number of others in non-applicative constructions). Yet it appears that many fine-grained semantic distinctions which influence verbal alternations are observed only over much larger samples of lexemes (cf. Levin 1993 in which 3000 English verb stems are classified into semantic subclasses that predict verbal alternations). As a result, it is not yet clear what components of semantic meaning influence the function of AMs in particular western Indonesian languages, and to what extent this explains variation in properties

of ACs. Nor is it clear whether semantic effects are consistent across languages.

In light of the gaps in the literature identified here, this study gives special attention to broadening description and cross-linguistic comparison of ACs in languages of West Nusantara, and their functions, properties, and usage.

1.7 Organization

This study is divided into three parts. Part I lays the conceptual groundwork for the dissertation. Following this introductory chapter, in Chapter 2, I present a descriptive account of applicatives in the Sundanese language. I use this account as a case study by which to ground the discussion of applicatives that follows in one clearly exemplified linguistic system. In Chapter 3, I show how applicatives have been defined in the linguistic literature, and some problematic constructions that raise questions about the adequacy of previous approaches. I then present my rationale for adopting a constructionist approach to the ACs and other AM-marked constructions of the languages of West Nusantara, in which such constructions are defined as conventionalized pairings of form and meaning.

Part II of the dissertation describes a large-scale typological survey of West Nusantara languages, which explores the distribution of the applicatives within West Nusantara, and properties of applicative systems therein. In Chapter 4, I describe the goals of the survey, sampling, and methods used. In total 85 languages of West Nusantara are sampled in the survey. In Chapter 5, I report survey results related to location, genetic lineage, and typological features of the languages. The results show that the presence of applicatives that alternate independently of symmetrical voice in these languages is related to the breakdown of the Philippine-type voice system. In addition, the distribution of languages that lack applicatives is highly influenced by geographic patterns and contact effects. In Chapter 6, I report results of the survey related to properties of applicative systems and AMs. I find that the distinct pattern observed for the distribution of applicative functions across forms of AMs are most likely inherited from earlier Philippine-type voice systems, while the distribution of non-applicative functions may be more influenced by geographic subregion within West Nusantara.

Part III of the dissertation is centered on forms and functions of AC and other AM-marked constructions in these languages, and how these relate to other types of semantic meaning. In Chapter 7, I lay out a functional typology of AM-marked constructions based on twenty-six languages of West Nusantara, with descriptive notes on variance in the syntactic and semantic properties of these that is observed across these languages. Chapter 8 presents an exploratory study of the relationship between lexical semantic meaning and AM-marked constructional meanings, based on nine languages of West Nusantara and a set of eighty lexical meanings. I identify some components of semantic meaning that influence attraction of lexemes to particular constructional meanings, and suggest ways that other structural resources in a language shape the patterns between lexical meaning and constructional meaning observed. A short summary of major findings and discussion of implications concludes the dissertation in Chapter 9.

Chapter 2

Case study: Sundanese applicatives

In this chapter, I describe applicative constructions (ACs) in Sundanese, the principal language spoken in West Java, Indonesia, as a case study. In doing so, I aim to provide a careful, detailed description of applicatives in a major language of West Nusantara, which have not yet been studied in depth. Sundanese exhibits some particularly interesting ACs, such as an innovative and productive substitutive benefactive construction and constructions in which the applied phrase is marked with a preposition, yet has some properties characteristic of core arguments. However, Sundanese also clearly demonstrates common characteristics of applicative systems in West Nusantara. Because of this, I will also use this chapter to establish key terminology for the dissertation.

In this chapter, the approach to applicatives that I adopt has two distinctive properties. First, as mentioned in the previous chapter (§1.4), I use a constructional approach, in which an AC is understood as a clausal construction which constitutes a conventionalized pairing of form and meaning. I consider the form of an AC to include the particular morphological marking found on the predicate, i.e., the AM, and elements of its clausal argument structure, including the number and structural position of clausal arguments. I consider the meaning of an AC to include the semantic relationships between clausal arguments and the events or states expressed, which have been called semantic, thematic, or participant roles, as well as other semantic components of verbal meaning. Second, in describing Sundanese applicatives and related constructions, I give careful attention to define and distinguish key components of form and meaning. With respect to form, I am sensitive to the fact that a particular verb stem bearing an AM may be used with more than one possible argument structure, and that the realization of argument structure may be influenced by syntactic characteristics of the verb root (e.g. base valency), syntactic properties of the clause (e.g. grammatical voice), as well as pragmatic context. With respect to meaning, I am concerned with specific and somewhat narrowly defined semantic roles (following Van Valin 2005, see §1.4.3), semantic subclasses of verbs (e.g. emotion verbs, communication verbs), and the pragmatic inferences that clauses give rise to.

An important advantage of the constructional approach taken here, is that it allows clausal example data with diverse characteristics to be classified into types of constructions on the basis of meaningful and well-defined similarities and differences in function and form, with more specific patterns and more general ones both being acknowledged (see Goldberg 2013 on varying levels of specificity in knowledge of language). Furthermore, a constructional approach avoids

the need to strictly rely on correspondence between a base construction (BC) and an applicative construction (AC) taken to be derived or otherwise related to the BC (see Post & Modi 2022 on languages in which there are no BCs expressing the meaning or paraphrased meaning expressed by ACs, and related discussion in §3.2). Instead, for any clausal construction that itself has a form and meaning, we may consider its characteristics to be evidence of its place vis-à-vis other constructions in a typology of applicatives. This allows for better identification of the broad usage-based knowledge that speakers rely on in producing and interpreting constructions marked with AMs, especially the highly polyfunctional AMs found in West Nusantara. In addition, this approach allows for meaningful and well-defined comparison of applicatives and related constructions across languages.

Data in this chapter were either collected using elicitation, taken from original recordings of natural connected speech, or compiled on the basis of published textual material which was glossed by the author with a native speaker of Sundanese. The data presented will primarily include glossed example sentences, with most of these taken from about 80 hours of recorded elicitation and glossing sessions with Dewi Setiani, a female speaker of Sundanese as spoken in Bandung, West Java.

Dewi was raised in West Bandung Regency by parents who are both native Sundanese speakers and who use conversational Indonesian only in limited contexts with non-Sundanese people. She attended primary and secondary school in West Bandung Regency, before attending the Universitas Pendidikan Indonesia in Bandung City; in these contexts both Indonesian and Sundanese were used as languages of instruction, with English later taught as a subject. Dewi subsequently worked as science teacher and earned a Master's degree in Education at UH Mānoa, where she and I were introduced. The period of time we worked together has spanned her early to mid-thirties. Dewi also facilitated the checking of some elicited sentence examples a small number of other Sundanese speakers, including her mother and sister, in West Bandung Regency.

Elicited sentences include examples representing the 80 verbal meanings from the Leipzig Valency Questionnaire (see Appendix D). Original recordings of connected speech include five short oral narrative texts and forty minutes of recorded conversation. A portion of these were collected as part of a Field Methods course at UH Mānoa in Fall 2018 and Spring 2019. The majority of the sessions, however, were conducted by video call during 2021–2022 with participants joining from their place of residence in West Java. Written text used includes three published short stories and sentence data from from the Leipzig Corpora Collection for Sundanese (see Goldhahn, Eckart & Quasthoff 2012).¹

2.1 The Sundanese language and community

Sundanese is the language of the Sundanese ethnic group, and it is spoken over roughly the western third of the island of Java in Indonesia, including the provinces of West Java and Banten, and some parts of Central Java. The Sundanese ethnic community today is made up of more than thirty million people, and almost all of these speak the Sundanese language (Eberhard, Simons & Fennig 2021). Most Sundanese people today are bilingual in Indonesian, the Malay-based national

¹Sentences from the Leipzig Corpora Collection for Sundanese were taken from the 2016 Wikipedia 30K corpus, which is available at <https://wortschatz.uni-leipzig.de/en/download/Sundanese>. In references given for glossed examples, this source is represented as MPI-SUN, followed by the line number of the sentence.

language of Indonesia, and there are some indications that Sundanese is at risk of marginalization in favor of Indonesian, especially in urban centers (see Ravindranath & Cohn 2014 on endangerment of local languages with large speaker populations in Indonesia). Despite this, Sundanese is used in all domains, including local print and broadcast media, and in education at all levels, from primary to university, alongside Standard Indonesian.

Sundanese has had extensive historical contact with Javanese and Malay (see Ewing 1999). In some proposals, Sundanese has been suggested to subgroup with the Malayo-Chamic languages, along with Madurese and Balinese-Sasak-Sumbawa (Adelaar 2005a), or with all of the aforementioned languages and Javanese at a higher level (Smith 2017). Despite this, I find a number of non-negligible differences between the applicative system of Sundanese and those of Javanese (Hemmings 2013; Vander Klok & Evans 2022) and Indonesian (a standardized variety of Malay) with respect to the inventory of AMs, productivity of ACs, and properties of clausal arguments in such constructions. Like Javanese, Sundanese is known for a system of speech levels, the use of which is governed by social rank and etiquette.²

Previous research on Sundanese includes an early grammar by Coolsma (1904), language descriptions by Van Syoc (1959), Robins (1959), Hardjadibrata (1985), Clynes (1995a), Hanafi (1997), and Kurniawan (2013), Nothofer's (1980) dialect geography of West Java, and a number of dictionaries, such as Rigg (1862), Coolsma (1913), Satjadibrata (1948), Tamsyah (1997), Eringa (1984), and Hardjadibrata's (2003) Sundanese-English Dictionary based on Eringa's volume. However, despite the fairly large body of literature on Sundanese and its status as an important language of West Nusantara with a large number of speakers, Sundanese has attracted relatively little attention in research on western Austronesian voice and applicatives, and to date a detailed examination of the Sundanese applicative system has not been undertaken.

Sundanese examples given in italics in this paper are transcribed with the following orthographic conventions: <c> is a voiceless post-alveolar affricate, <j> is a voiced post-alveolar affricate, <y> is a palatal approximant, <ng> is a velar nasal, <ny> is a palatal nasal, <e> is schwa /ə/, <é> is an unrounded mid front vowel /e/ (or less frequently, /ɛ/), <eu> is an unrounded mid back vowel /ɤ/ (this vowel is often described as central and transcribed as /i/ in the literature on Sundanese). Other symbols used are consistent with their expected phonetic values. These orthographic choices are fairly standard in written Sundanese, though the acute accent on the mid front vowel is not consistently used by all community members and outside authors.

2.2 Basic morphosyntax

In this section I will describe properties of Sundanese morphosyntax which are relevant for understanding the syntax and semantics of Sundanese ACs.

Sundanese is a mildly agglunative language. Common morphological processes include reduplication and affixation. Several types of reduplication are attested in addition to morphological prefixes, suffixes, circumfixes, and infixes. Stacking of up to two prefixes and/or two suffixes on a single stem is not uncommon e.g. *di-pika-sieun* 'to be feared' cf. *sieun* 'afraid'; *nga-réndéng-an-*

²Thus at times, Sundanese data used in this dissertation may show more than one lexeme with identical glossing, due to the fact that these items differ in register. For example, the pronoun *abi* or *abdi* '1sg' is used in polite speech in some social contexts, and belongs to the *lemes* ('refined') register, while *kuring* '1sg' is used in more casual or familiar speech and belongs to the *kasar* ('common') register.

keun ‘to sit next to each another’ cf. *réndéng* ‘be next to (s.t. or s.o.)’. Reduplication and (multiple) affixation may generally occur together, however a single stem appears to be limited to up to one reduplication process only.

Sundanese generally shows no morphological case marking on nouns. Prepositions are used to indicate logical, spatial, and temporal relationships, among others. Word order is predominantly subject-initial, though predicate-initial word order is also possible and reflects differences in information structure. In addition, arguments of the verb are often unrealized. This is commonly the case when an argument’s referent is recoverable in the discourse context or left open to interpretation through pragmatic inference (see Ewing 2019 for discussion on pragmatic motivations for use of predicates without arguments in Indonesian).

Grammatical tense is not marked on the verb in Sundanese. Temporal and aspectual information may be indicated by use of auxiliaries or adverbial modifiers but this is not required.³ Marking for number and person agreement on the predicate is very limited and always optional (see Kurniawan 2013: 23–26).

In the remainder of this section, I will describe some basic properties of declarative intransitive, monotransitive, and ditransitive verbal clauses. In discussing types of clauses and clausal arguments, I employ the following terms following Comrie (1989) and Haspelmath (2015):

- *S* refers to the single core argument of an intransitive verb.
- *A* and *P* refer, respectively, to the most agent-like argument and the most patient-like argument of a monotransitive verb. Some authors also use *O* for the latter.
- *A*, *R*, and *T* refer to the arguments of a ditransitive verb. Following Haspelmath (2015), such verbs commonly denote “transfer of an entity (*T*) from an agent (*A*) to a recipient (*R*)”.

2.2.1 Intransitive verbal clauses

In this section, I describe the basic properties of declarative intransitive main clauses. In such clauses, the single core argument (*S*) of an intransitive verb is encoded as an unmarked NP in preverbal position, if it is realized.

In intransitive clauses, some lexical roots appear as the main verb without morphological marking. Examples are shown below in (5) and (6), with the *S* argument bolded. Peripheral participants in intransitive clauses may be coded as oblique and expressed with a prepositional phrase, as with the goal phrase, *ka sakola* ‘to school’ in (6).

(5) Sundanese, Intransitive clause

Abi *luncat*.
1SG jump
‘I jumped.’

³In many declarative examples, I have chosen to use past tense in the English translation, but a correct translation could also use the present tense (or vice versa). For example, the clause in (5) below could be translated either as ‘I jump.’ or ‘I jumped.’

(6) Sundanese, Intransitive clause

Asep lumpat ka sakola.

A. run to school

‘Asep ran to school.’

(FM4-032)

Other verbal roots that appear without any morphological marking in intransitive clauses include *leumpang* ‘to walk’, *indit* ‘to depart’, *dongkap* ‘to come, arrive’, *diuk* ‘to sit’, *saré* ‘to sleep’, *paéh* ‘to die; be dead’, *ulin* ‘to play’, *cicing* ‘to live, reside’, *batuk* ‘to cough’, *ceurik* ‘to cry’, *lapar* ‘to be hungry’, *sedih* ‘to be sad’, and *resep* ‘to be fond (of s.t.)’.

Another set of lexical roots must be marked with a voice prefix when they appear as the main verb in an intransitive clause. Examples (7) and (8) below show intransitive clauses in which the verb root is obligatorily marked with the nasal prefix, here realized as the allomorph *nga-* or *ng-*.⁴ Again, the S argument is realized as an unmarked NP in preverbal position (bolded) and a peripheral participant may be expressed as a PP, e.g. *dina para* ‘in the attic, loft’ in (7).

(7) Sundanese, Intransitive clause

Bal nga-gulutuk.

ball AV-roll

‘The ball rolled.’

(CT1-002)

(8) Sundanese, Intransitive clause

Cai ng-(k)ucur dina para.

water AV-flow in attic

‘Water flowed (down) in the attic (e.g. because the roof was leaking).’

(CT1-024)

Other verbal roots that obligatorily take the nasal prefix in intransitive clauses include *tang-tung* ‘to stand/stand up’, *iring* ‘to follow along (with s.o.), as in an activity’, *sumput* ‘to hide (o.s.)’, *pikir* ‘to think’, and *carios* (or *omong*) ‘to talk’.⁵

More examples of intransitive clauses in Sundanese are given below. The specific role of the participant that maps to S (e.g. agent, experiencer, etc) is selected by the verb. Note that S is typically encoded as an unmarked NP in preverbal position, but may also be unrealized and interpreted from context.

(9) Sundanese, Intransitive posture verb

Abi diuk dina samak.

1SG sit on mat.

‘I sat on a mat.’

(CT1-014)

⁴The nasal prefix has the underlying form *ng-* (IPA /ŋ-/) and its surface forms include many stem-conditioned allomorphs. It triggers the morphophonemic process known as *nasal substitution* in which the nasal segment of the prefix coalesces with initial segment of a stem when that segment is a voiceless obstruent (and sometimes /b/). This results in a single nasal stop segment at the same place of articulation as the underlying initial stem consonant. For the sake of clarity, this initial consonant has been included in parentheses in glossed examples, as in *ng-(k)ucur* ‘to flow, AV’ in (8), which is pronounced [ŋuf̥çur], and is formed on the stem *kucur*.

⁵See also Hanafi (1997: 5-7), which includes discussion of a small number of verbs marked with the voice prefix *di-* that appear to be intransitive.

- (10) Sundanese, Intransitive communication verb
Euis ng-omong ka Udi kamari.
 E. AV-talk to U. yesterday
 ‘Euis talked to Udi yesterday.’ (CT1-015)

- (11) Sundanese, Intransitive emotion verb
Mariam ceurik di kamar.
 M. cry in room
 ‘Mariam cried in (her) room.’ (CT1-006 and Hanafi 1997: 22)

2.2.2 Monotransitive verbal clauses

For transitive verbs, the primary diathetical alternation is between A-oriented clauses marked with *ng-* and P-oriented clauses marked with *di-*. Example (12a) shows that when the verb *te-unggeul* ‘to hit’ is marked with the nasal prefix *ng-*, the A argument, i.e. the agent, or hitter, is encoded as an unmarked NP in preverbal position, and the P argument, i.e. the patient, or entity that is hit, is encoded as an unmarked NP in postverbal position. Example (12b) shows that when the same verb is marked with prefix *di-*, the P argument is realized as an unmarked NP in preverbal position, and the A argument is realized as an NP marked with the preposition *ku* ‘by’ in postverbal position. Example (13) shows a similar alternation with the verb *cukur* ‘to shave (s.o.) / cut (s.o.’s) hair’.

- (12) Sundanese, Transitive voice alternation
- a. *Icih n-(t)eungeul Asep.*
 I. AV-hit A.
 ‘Icih hit Asep.’ (AV)
 - b. *Asep di-teunggeul ku Icih.*
 A PV-hit by S.
 ‘Asep was hit by Icih.’ (PV) (CT1-019)

- (13) Sundanese, Transitive voice alternation
- a. *Saep keur ny-(c)ukur palura.*
 S. PROG AV-shave village.head
 ‘Saep is shaving/cutting the hair of the village leader.’
 - b. *Palura keur di-cukur ku Saep.*
 village.head PROG PV-shave by S.
 ‘The village leader is being shaved/having his hair cut by Saep.’ (CT1-003)

Other syntactically transitive verbs roots which show a simple alternation between AV and PV marked by *ng-* and *di-* (and no other obligatory morphological marking) are listed in Table 2.1.

In the examples above, we have seen agent mapped to A and patient mapped to P. However, other mappings are possible, in accordance with the verbal semantics, e.g. agent and theme with verbs of conveyance and caused motion, or perceiver and stimulus with verbs of perception. Examples of these are given below in (14) and (15). As seen in (15c), the PV prefix *di-* may be replaced with the non-volitional prefix *ka-*, without a change in argument structure. The prefix *ka-* denotes a non-volitional action and can also have an abilitative meaning, i.e. ‘was able to’.

Table 2.1: Monotransitive Sundanese verbs

<i>Processing of materials</i>		
a.	<i>beuleum</i>	‘to roast (s.t.)’
b.	<i>duruk</i>	‘to burn (s.t.)’
c.	<i>kumbah</i>	‘to wash (s.t.)’
d.	<i>pésék</i>	‘to peel (s.t.)’
e.	<i>teukteuk</i>	‘to cut (s.t.) with scissors’
<i>Acts of creation</i>		
f.	<i>jieun</i>	‘to make (s.t.)’
g.	<i>pasak</i>	‘to cook (s.t.)’
<i>Transfer & conveyance</i>		
h.	<i>bawa</i>	‘to carry, bring (s.t.)’
i.	<i>cokot</i>	‘to take (s.t.)’
j.	<i>kirim</i>	‘to send (s.t.)’
k.	<i>pinjeum</i>	‘to borrow (s.t.)’
<i>Striking & application of force</i>		
l.	<i>surung</i>	‘to push (s.t.)’
m.	<i>tabrak</i>	‘to strike, crash into (s.t.)’
<i>Caused motion & placing</i>		
o.	<i>pelak</i>	‘to plant (s.t.), as of crops’
p.	<i>teundeun</i>	‘to put, to place’
<i>Perception</i>		
q.	<i>ambeu</i>	‘to smell/sniff (s.t.)’
r.	<i>déngé</i>	‘to hear (s.t.)’
s.	<i>tinggali (~ tingali)</i>	‘to see (s.t.)’

(14) Sundanese, Transitive caused-motion verb

- a. *Abi n-(t)eundeun cangkir dina méja.*
 1SG AV-put cup on table
 ‘I put the cup on the table.’
- b. *Cangkir di-teundeun dina méja ku abi.*
 cup PV-put on table by 1SG
 ‘The cup was put on the table by me.’

(CT1-022)

(15) Sundanese, Transitive perception verb

- a. *Abi n-(t)inggal i poster film.*
1SG AV-see poster movie
'I looked at the movie poster.'
- b. *Poster film di-tinggali ku abi.*
poster movie PV-see by 1SG
'The movie poster was looked at by me.'
- c. *Poster film ka-tinggali ku abi.*
poster movie NVOL-see by 1SG
'The movie poster was seen by me (by chance).'

(CT1-043)

In the above examples we see that the mapping of semantic roles to argument structure—as evident from NP marking and word order—is accompanied by morphological marking on the predicate (i.e. verbal prefixes *ng-* and *di-*). In the *ng-*marked clauses, the A argument, if realized, is encoded as an unmarked NP in preverbal position, and the P argument, if realized, is encoded as an unmarked NP in postverbal position immediately following the verb. Conversely, in the *di-* marked clauses, the P argument is encoded as an unmarked NP in preverbal position, and the A argument is encoded as an unmarked NP in postverbal position. Because the preverbal argument in such clauses shows evidence of syntactic privilege by virtue of behavioral properties (e.g. access to relativization, and patterns of control, see Kurniawan 2013), I will henceforth refer to the *ng-* marked clauses as A-Voice (AV) and the *di-* marked clauses as P-Voice (PV).

A of a monotransitive AV clause and P of a monotransitive PV clause show the same encoding as S in intransitive clauses. P of an AV clause is likewise encoded as an unmarked NP, but it has a different position in word order. I will consider all of these to be core clausal arguments.

Although A arguments in PV clauses, are marked with the preposition *ku*, there are some reasons to believe that these phrases are in fact arguments of the verb as they exhibit certain properties characteristic of core arguments which distinguish them from other obliques or adjuncts (see Riesberg 2014 and Arka 2009 for more on the intermediary status of non-subject actors in western Austronesian languages).⁶ For this reason, I will consider *ku*-marked agents in Sundanese PV clauses to be verbal arguments, and will consider PV clauses to be transitive rather than intransitive. However, the status of *ku*-marked agents has little bearing on our understanding of ACs, as it is P arguments rather than A arguments that participate in applicative alternations.

2.2.3 Ditransitive verbal clauses

As discussed above, syntactically transitive verb roots in Sundanese show a diathetical alternation between AV clauses marked with *ng-* and PV clauses marked with *di-*. A small subset of such roots that show this type of alternation appear to be able to take three core arguments with no

⁶For example, Kurniawan (2013: 38–41) shows that A arguments in PV clauses can bind a reflexive subject (i.e., P in a PV clause) and control the subject of an embedded clause. Furthermore, optional morphological marking on the verb for plural number with *-al-* (or its allomorph *-ar-*) agrees with the *ku*-marked A argument of a PV clause just like it does with the unmarked A argument of a AV clause, and the unmarked S argument of intransitives. Kurniawan also shows that *ku* may be optionally omitted under certain conditions, resulting in A being encoded as an unmarked postverbal NP in some PV clauses. Similar patterns are also discussed by Kroeger & Riesberg (forthcoming).

additional morphological marking on the verb. Clausal constructions with more than two core arguments some show differences in AV compared to PV, with ditransitive mapping in PV being highly preferred as more natural with most lexical bases, and very few examples of AV ditransitive clauses being found in corpora.

Consider the examples shown below in (16), which illustrates two possible mappings for the verbal base *ajar* ‘to teach (s.t.) (to s.o.)’ in AV. In (16a), the clause is syntactically monotransitive. The semantic agent role, i.e., *Pa Guru* ‘the teacher’, maps to A, and the semantic content role, i.e. *matematika* ‘math’, maps to P, while a third experiencer participant, i.e., the student, is expressed as an oblique PP, *ka Asep* ‘to Asep’. In (16b), the clause is syntactically ditransitive and there are three unmarked NP arguments. The semantic agent maps to A, and the semantic experiencer and content are both realized as unmarked NPs occurring after the verb. The experiencer may be said to map to R, and the content made be said to map to T. Thus, (16b) is a ditransitive clause, and represents one possible argument structure for the verb *ng-ajar*. In the structure shown in (16b), the NP expressing the semantic content role may not appear as the first postverbal argument; it is not possible to reverse the structural position of the content and experiencer roles.

(16) Sundanese, Ditransitive construction in AV

- a. *Pa Guru ng-ajar matematika ka Asep.*
 mister teacher AV-teach math to A.
 ‘The teacher teaches math to Asep.’ (Monotransitive, AV)
- b. *Pa Guru ng-ajar Asep matematika.*
 mister teacher AV-teach A. math
 ‘The teacher teaches Asep math.’ (Ditransitive, AV) (CT1-031)

A PV construction corresponding to (16b) is given in example (17) below. Here we see that the semantic experiencer participant is realized in the preverbal position as an unmarked NP, that is, with encoding characteristic of a core argument. The semantic content participant is realized in the postverbal position as an unmarked NP, also encoding characteristic of a core argument. Finally, the agent is realized as an NP marked with the preposition *ku*, which is characteristic of the A argument in PV clauses. This is a ditransitive clause in PV; in addition to one A argument, there are two non-A core arguments. Of these, R maps to the semantic experiencer participant, which can be said to receive some knowledge, and T maps to the content participant, which can be said to be transferred in the event.

(17) Sundanese, Ditransitive construction in PV

- Asep di-ajar matematika ku Pa Guru.*
 A. pv-teach math by mister teacher
 ‘Asep was taught math by the teacher.’ (Ditransitive, PV) (CT1-031)

The Sundanese verbal root *béré* ‘to give (s.t.) (to s.o.)’ can also form ditransitive clauses in PV without additional morphological marking on the verb. PV clauses marked with *di-* on the verb stem *béré* have two possible mappings as shown in (18) below. The first is syntactically monotransitive, as shown in (18a). Here, the theme role (i.e., *buku* ‘book’, the entity that changes possession) maps to P, while the recipient (i.e., *ka budak* ‘to the child’) is realized as an oblique PP

headed by the allative preposition *ka* ‘to’. The agent role maps to A and is realized as a *ku*-marked agent phrase. The second possible mapping is syntactically ditransitive, as shown in (18b). Now, the recipient role *budak* ‘child’ is realized as an unmarked NP in preverbal position, and thus shows encoding consistent with core argument status. The theme, *buku* ‘book’ is realized as an unmarked NP in postverbal position, which also reflects core argument status. The agent again is realized as a *ku*-marked agent phrase. As in (17) above, there are two non-A core arguments, a recipient (R) and and theme (T).

(18) Sundanese, *Give* construction in PV

- a. *Buku di-béré ka budak ku abi.*
 book PV-give to child by 1SG
 ‘The book was given to the child by me.’ (Monotransitive, PV)
- b. *Budak di-béré buku ku abi.*
 child PV-give book by 1SG
 ‘The child was given a book by me.’ (Ditransitive, PV) (CT1-026)

However, as shown in (19) below, ditransitive constructions are not possible in AV clauses with *béré*. In (19a), a grammatical monotransitive clause is shown. Here, the semantic agent *abi* ‘I’ maps to A and is realized as an unmarked NP in preverbal position, and the semantic theme maps to P, and is realized as an unmarked NP in immediate postverbal position. The semantic recipient is realized as an oblique PP, *ka budak* ‘to the child’. It is not possible for recipient and theme to both appear as unmarked NPs after the verb, as shown in the ungrammatical example in (19b). It is also not possible for the recipient to map to P and be realized as an unmarked NP in postverbal position while the theme is realized as an oblique PP, as intended in the ungrammatical example in (19c).

(19) Sundanese, *Give* construction in AV

- a. *Abi m-(b)éré buku ka budak.*
 1SG AV-give book to child
 ‘I gave the book to the child.’
- b. **Abi m-(b)éré budak buku.*
 1SG AV-give book child
 Intended: ‘I gave the child a book.’
- c. **Abi m-(b)éré budak ku buku.*
 1SG AV-give child with book.
 Intended: ‘I gave (to) the child using a book.’ (CT1-026)

In summary, a small number of roots in Sundanese allow syntactically ditransitive clausal constructions when the verb stem bears only a voice prefix *ng-* or *di-* and no other morphological marking. In such cases, R and T are selected by the verb and may be realized as unmarked NPs. Further restrictions on the syntactic realization of ditransitive constructions may apply. For example, a ditransitive construction may be grammatical in PV but not AV, as shown for the

verb stem *béré* ‘to give’. There may also be restrictions on the linear order of the two non-A core arguments, as the R argument generally precedes the T argument in both AV and PV.

Besides *ajar* ‘to teach’ and *béré* ‘to give’, no other verbal roots in Sundanese appear in ditransitive constructions without suffixation with an AM, even those with cross-linguistically “typical” ditransitive verb meanings (see Haspelmath 2013, 2015). In elicitation sessions, ditransitive constructions of this sort were reported as ungrammatical in both AV and PV with the verbal roots *beuli* ‘to buy’, *jual* ‘to sell’, *pinjam* ‘to borrow’, *bawa* ‘to carry, bring’, and *jieun* ‘to make’. For these and almost all other verbal roots, we will only observe ditransitive constructions when morphological marking in addition to voice, such as a causative or applicative affix, appears on the verb.

2.3 Overview of applicative morphology

Having described the basic morphosyntax of Sundanese verbal clauses, I now turn to the main topic of this chapter, Sundanese applicatives. In this section, I begin with an overview of Sundanese AMs, that is, morphemes which may mark the predicate in a construction in which a peripheral semantic role is selected to map to a core argument. The constituent that expresses this semantic role will hereafter be referred to as the *applied phrase* (following Zuñiga & Creissels 2024). Detailed evidence showing that these affixes should be considered AMs will be exemplified in following sections.

The inventory of AMs in Sundanese is given in Table 2.2. The second column shows possible semantic roles (SR) of the applied phrase in ACs marked by affixation of each AM. These are listed using labels largely taken from Van Valin (2005) with some necessary additions and include: addressee (ADDR), beneficiary (BEN), content (CONT), goal (GOAL), instrument (INST), location (LOC), stimulus (STIM), theme (THM), and performance (PERF). The third column lists other functions of the morpheme when it is found on verbal predicates, including marking causative constructions (CAUS), marking comparative degree constructions (COMP), and indicating pluractional aspect (PLUR), which denotes iterative, habitual or durative aspect and events with multiple actors and/or undergoers. Further description and delineation of Sundanese AMs is given below.

Table 2.2: Sundanese applicative morphemes

Affix	SR of Applied Phrase	Other verbal functions
<i>-an</i>	LOC, GOAL, ADDR, CONT, STIM	CAUS, PLUR, COMP
<i>-keun</i>	THM, INST, CONT, STIM, PERF, BEN	CAUS
<i>pang-</i> <i>-keun</i>	BEN	CAUS*

* This morpheme may select a causer and a peripheral applied phrase in the same clause.

The suffix *-an* has a variety of functions that include marking ACs in which the applied phrase is a location, goal, addressee, content item, or stimulus, marking causative constructions, marking comparative degree constructions, and indicating pluractional aspect.⁷ It also forms verbs from

⁷Hanafi (1997) indicates that Sundanese *-an* also marks ACs in which the applied phrase is a beneficiary. However, he offers a single example with the root *hirup* ‘to live, alive’, and I consider the alternation in question to be more consistent with a causative construction than a benefactive applicative.

non-verbal roots, e.g. *uyah-an* ‘to put salt on (s.t.)’ from *uyah* ‘salt (n.)’.⁸

The suffix *-keun* marks ACs in which the applied phrase is a theme, instrument, content item, stimulus (e.g. in an event of sensation, emotion, or perception), type of performance, or beneficiary. The use of *-keun* in beneficiary-selecting ACs is less productive than the use of the circumfix *pang-* *-keun*. In addition to ACs, *-keun* forms causative constructions, and forms verbs from non-verbal roots, e.g. *tali-keun* ‘to tie (s.t.) to s.t.’, from *tali* ‘rope, string’.

The circumfix *pang-* *-keun* exclusively marks beneficiary-selecting ACs. I use the term *substitutive benefactive* to refer to constructions marked with *pang-* *-keun* because the presence of the circumfix indicates that the agent performs the action on behalf of, and in lieu of, a beneficiary. This type of construction is also called *deputative-benefactive* (see Peterson 2007; Van Valin & LaPolla 1997: 383; Kittilä & Zúñiga 2010: 14). The circumfix *pang-* *-keun* triggers nasalization of the initial consonant of the stem to which it attaches when this segment is a voiceless obstruent (or sometimes /b/). This occurs both in AV and PV verbal forms. This is not the same morphophonemic process as nasal substitution (which is triggered by the AV prefix *ng-*), because the velar nasal of the substitutive benefactive affix and the initial stem consonant do not coalesce into a single surface segment but remain separate segments. The AV prefix *ng-* may be prefixed to stems bearing *pang-* *-keun*, but *ng-* always occurs farther to the left from the root. For example, the word *mangmalingkeun* meaning ‘to steal s.t. for s.o. (AV)’ is made up of the morphemes, *ng-* ‘AV’ + *pang-* *-keun* ‘BEN.APPL’ + *paling* ‘to steal’. The first /p/ of the circumfix undergoes nasal substitution and fuses with the AV prefix into a single word-initial consonant /m/, while the initial /p/ of the root is nasalized to /m/ but remains a distinct segment from the preceding velar nasal /ng/. The corresponding PV verb is *dipangmalingkeun*, which shows the same pattern at the morpheme boundary between the root and circumfix.

Note that I do not consider *pang-* *-keun* to be a co-occurrence of the *-keun* suffix and a separate prefix with the form *pang-* in contemporary usage. The circumfix *pang-* *-keun* occurs with many verbal roots for which affixation of *pang-* and/or *-keun* alone is not attested. In the case of the verbal root *paling* ‘steal’, for example, there is no stem or word with the form **palingkeun*, nor the form **pangmaling*. Some nominalizations do occur in Sundanese with a prefix with the shape *pang-* alone, especially instrument or agent nominalizations, e.g. *pangasuh* ‘nanny’ cf. *asuh* ‘to take care of (as of a child)’. However, unlike affixation with *pang-* *-keun*, when the first consonant of the root is a voiceless obstruent, the *pang-* nominalizing prefix shows nasal substitution with fusion of the final prefix nasal and initial root consonant. For example, for the verbal root *tutu* ‘to pound (as of rice in the husk)’, the substitutive benefactive forms are *mangnutukeun* (AV) or *dipangnutukeun* (PV), each showing two adjacent nasal segments, but the nominalization given by Hardjadibrata (2003: 856) is *panutu*, which refers to the person who pounds rice or instrument used to pound rice, and shows a single fused nasal segment. Even though the nominalizing prefix *paN-* and the substitutive benefactive circumfix *pang-* *-keun* are distinct in modern Sundanese, it is quite plausible that one source morpheme in an earlier form of Sundanese gave rise to both the modern *paN-* prefix and the benefactive meaning of *pang-* *-keun*. While it is not clear how Sundanese *pang-* *-keun* developed, to my knowledge, similar circumfixal forms are not found in other languages of West Nusantara as distinct AMs, though I note that prefixal components with forms like *pa-* and *paN-* are found to indicate locative or instrumental voice constructions in

⁸Sundanese also shows a homophonous nominalizing suffix *-an*, e.g. *pikir-an* ‘(a) thought’ cf. *pikir* ‘to think’, *dagang-an* ‘product, ware’ cf. *dagang* ‘to sell, to trade’.

some Philippine-type languages of West Nusantara (see §6.1.3 and §6.1.4), and in some irregular paradigmatic alternations for benefactive/instrumental applicatives (e.g. *Pendau*, see §5.9.4.4; *Toba Batak*, see §6.5.1).

As is evident from the discussion in this section thus far, Sundanese AMs are highly polyfunctional. Furthermore, the functions listed above are not always discrete; in some constructions, the presence of a single AM on the verb may result in more than one semantic and/or syntactic effect in the clause, and it is not always easy to tease apart these functions. Thus for the sake of clarity, below I note the glossing conventions that I have followed in the remainder of the chapter.

- In constructions that select a peripheral semantic role as a clausal argument, the AM is glossed using an abbreviation for that semantic role followed by *APPL* for applicative.
- Some semantic roles are collapsed under more general labels in glosses:
 - *BEN.APPL* is used when the role selected is beneficiary. In Sundanese, some but not all beneficiaries are also semantic recipients. This is not noted in glosses.
 - *LOC.APPL* is used when the role selected is a locative relation, including location, path and goal.
 - *THM.APPL* is used when the role selected is theme or instrument, as these cannot always be distinguished, though from the free translation an affinity with one or the other meaning is often apparent.
- In constructions where a instigating causer role is introduced to the argument structure and mapped to A, the AM will be glossed as causative (*CAUS*). Note that an AM may be glossed as both causative and applicative (see numerous examples in §2.6).
- In comparative degree constructions where an AM appears on a root denoting a property concept, e.g. ‘high’, ‘strong’, it is glossed *COMP*.
- If the only salient effect of an AM is to indicate pluractional aspect, the AM is glossed *PLUR*. While some applicative clauses appear to include pluractional aspect as part of their semantic meaning, this is difficult to determine definitively. Thus pluractional aspect will not be indicated in glossing of ACs, though the free translation may suggest aspectual effects.

In the remainder of this chapter, I will describe clausal constructions in Sundanese that meet the definition of an AC used in this study in (1) in §1.4.1. In these constructions, a peripheral semantic role is selected to map a core argument position, and this coincides with morphological marking of the verbal predicate. I define a peripheral semantic role, as a non-agentive and non-patientive semantic role. The AMs introduced in this section constitute the attested morphological marking for ACs in Sundanese.

2.4 Locative-selecting constructions marked with *-an*

In this section, I will describe morphologically marked constructions in which a locative expression is selected as a core argument. A locative expression serves as the basis for describing the spatial position of an entity or event. In addition to expressions that describe a static or generalized spatial position (i.e. location), I also include here various types of expressions that describe the spatial position of an entity that undergoes a change of location (e.g. source, path, direction, or goal). Constructions which select a locative expression are marked with the suffix *-an* on the verb in Sundanese. For a number of bases, we observe a location-selecting construction marked with *-an* in alternation with a theme-selecting construction marked with *-keun*. This alternation is discussed in §2.6 below.

2.4.1 Location-selecting constructions

With some verbal bases, the location of an event is selected as a core argument when the verb is marked with *-an*. In examples (20)–(23), when the verb is marked with *-an* in the AC, the location role is selected to map to P, and is expressed as an unmarked NP in immediate postverbal position. In the corresponding unmarked BCs, the location is expressed as a PP headed by the locative preposition *di* ‘in, at, on’ or its near equivalent *dina*.

(20) Sundanese, Location-selecting construction

- a. *Abi cicing di imah.*
1SG stay in house
‘I live in a house.’ (BC)
- b. *Abi ny-(c)icing-an imah.*
1SG AV-stay-LOC.APPL house
‘I live in a house.’ (AC) (CT1-004)

(21) Sundanese, Location-selecting construction

- a. *Abi diuk dina samak.*
1SG sit on mat
‘I sat on a mat.’ (BC)
- b. *Abi nga-diuk-an buku.*
1SG AV-sit-LOC.APPL book
‘I sat on a book.’ (AC) (CT1-014)

(22) Sundanese, Location-selecting construction

- a. *Beurit nga-liang di buruan imah.*
rat AV-hole in yard house
‘The rat made a hole in the yard of the house.’ (BC)

- b. *Bapa nga-liang-an témbok.*
 father AV-hole-LOC.APPL wall
 ‘Father made a hole in the wall.’ (AC) (CT1-027)

(23) Sundanese, Location-selecting construction

- a. *Asep m-(p)elak tangkal cau di kebon.*
 A. AV-plant tree banana in field
 ‘Asep planted banana trees in the field.’ (BC)
- b. *Asep m-(p)elak-an kebon ku tangkal cau.*
 A. AV-plant-LOC.APPL field with tree banana
 ‘Asep planted the field with banana trees.’ (AC) (CT1-004)

Location-selecting constructions marked with *-an* in Sundanese are monotransitive. If the BC is intransitive, as in (20)–(22), we observe an increase in valency when the verb is marked with *-an*. If the BC is transitive, as in (23), there is no increase in valency; the AM-marked construction remains monotransitive. This type of construction may be called a *remapping* (Comrie 1985a) or *valency-preserving AC* (Truong & McDonnell 2022). The semantic role that maps to P in the BC, in this case, the semantic theme, i.e., that which is planted and put into the field, is “remapped.” In the construction marked with *-an*, the semantic theme is no longer realized as an unmarked NP, but as an oblique PP, i.e. *ku tangkal cau* ‘using/with banana trees’. I will use the term *companion phrase* to refer to the constituent in the AC that expresses the semantic role that maps to P in the BC. Because the companion phrase is encoded differently in different constructions, I give attention to its syntactic realization in ACs when applicable. Both the valency-increasing and the valency-preserving location-selecting constructions exemplified here would be considered applicatives in most frameworks, because the presence of the morpheme *-an* can be said to “license” or “result in” the selection of a location as a core clausal argument.

In Sundanese, if a verb may select the location role to map to P in a BC with no AM-marking, when the verb is suffixed with *-an*, no change in argument structure is observed, as in (24). Instead, the construction marked with *-an* has pluractional aspect as part of its meaning. This is illustrated in example (24b).

(24) Sundanese, Pluractional location-selecting construction

- a. *Wadah runtah di-koréh ku abi.*
 container trash PV-dig.in by 1SG
 ‘The trash can was dug in by me’ (e.g. to see if my lost ring was in there). (BC)
- b. *Tanah di-koréh-an ku Lilis n-(t)eang-an huwi.*
 ground PV-dig.in-PLUR by L. AV-look.for-PLUR yam
 ‘The ground was dug in **repeatedly** by Lilis (while she was) looking for yams.’ (AC) (CT1-027)

The pluractional location-selecting construction described here would not be considered an AC in most frameworks, because we do not observe a change in argument structure under affixation with *-an* with verbs like *koréh* ‘dig’. Despite this, it is clearly related to the location-selecting

ACs shown in (20)–(23) in form and meaning. Both show morphological marking of the verb with *-an*, and both select a location role to map to P. Constructions in which *-an* is associated with a pluractional meaning are further discussed in §2.10.2.

2.4.2 Goal-selecting constructions

With some verbal roots, the goal semantic role is selected to map to a core argument position when the verb is marked with *-an*, as in (25). Here, goal refers to an expression that describes the endpoint of an entity that changes location. In (25b), the goal, *lima nagara di Asia* ‘five countries in Asia’ is expressed as the P argument of the verb, *nganjang-an* ‘visit (AV)’, and is realized as an unmarked NP in postverbal position. As shown in (25c), the goal argument can also be expressed as the P argument of a verb in PV that is marked with *-an*, in which case it may be realized as an unmarked NP in preverbal position. In the corresponding BC, shown in (25a) the goal is expressed as an oblique PP, headed by the allative preposition *ka* ‘to’. The clause marked with *-an* in this example would be non-controversially considered an AC.

(25) Sundanese, Goal-selecting construction

- a. *Wakil Presiden Kamala ng-angjang ka Indonesia.*
deputy president K. AV-visit to Indonesia
‘Vice President Kamala (Harris) visited Indonesia.’ (BC)
 - b. *Wakil Presiden Kamala ng-anjang-an lima nagara di Asia.*
deputy president K. AV-visit-LOC.APPL five country in Asia
‘Vice President Kamala (Harris) visited five countries in Asia.’ (AC)
 - c. *Lima nagara di Asia di-anjang-an ku wakil President Kamala.*
five country in Asia PV-visit-LOC.APPL by deputy president K.
‘Five countries in Asia were visited by Vice President Kamala (Harris).’ (AC)
- (CT1-028)

This example shows that when the BC is intransitive, the goal-selecting AC in Sundanese is valency-increasing, resulting in a monotransitive AC.

In some constructions marked with *-an*, in addition to the selection of a goal as a core argument, we also observe a change in the semantic meaning of the verb, as shown in (26) below. In the BC in (26a), the verb *datang* means ‘arrive’ or ‘come’. The clause is intransitive, with the goal expressed as an oblique PP, *ka imah* ‘to the house’. In the AC in (26b), the goal *imah Euis* ‘Euis’ house’ is expressed as P, a core argument. In addition, there is a subtle change in meaning, as the suffixed verb *ngadatangan* means ‘to visit, to call on (s.o.)’ rather than simply ‘to arrive (at), to come (to)’. It is possible that this semantic change is related to the aspectual functions of *-an*; as we have already seen, the suffix can have a durative meaning.

(26) Sundanese, Goal-selecting construction

- a. *Tah ai peuting=na polisi datang ka imah...*
EMPH EMPH night=DEF police arrive to house
‘That night the police came to the house...[telling that the man had already died].’ (BC)
- (FM4-067)

- b. *Abi nga-datang-an imah Euis.*
 1SG AV-arrive-LOC.APPL house E.
 ‘I visited Euis’ house.’ (AC) (CT1-002)

Aspectual changes are also observed in the goal-selecting construction shown in (27) with the verb *teundeun* ‘to put, to place’. In the AC in (27b), the goal *lomari* ‘wardrobe’ is encoded as an unmarked NP reflecting its status as a core argument. In addition, the clause takes on pluractional aspect, as reflected in the translation ‘has many cups put in it’. Again we see both a change in argument structure and a semantic change in meaning when the verb is marked with *-an*.

(27) Sundanese, Goal-selecting construction

- a. *Cangkir di-teundeun dina méja ku abi.*
 cup PV-put on table by 1SG
 ‘The cup was put on the table by me.’ (BC)
- b. *Lomari di-teundeun-an cangkir.*
 wardrobe PV-put-LOC.APPL cup
 ‘The wardrobe has many cups put in it (i.e. the wardrobe is full of cups).’ (AC) (CT1-022)

In example (27), we also see that the goal-selecting AC allows for ditransitive argument structure. In a BC, the verb *teundeun* selects a semantic agent to map to A and a semantic theme to map to P, while the semantic goal maps to an oblique PP. In the AC in (27b), the semantic goal *lomari* ‘wardrobe’ is selected as the R argument and realized as an unmarked NP in preverbal position. The companion phrase, the semantic theme, is selected as the T argument and realized as an unmarked NP in immediate postverbal position.

In this section, we have seen that the goal-selecting AC in Sundanese is consistently *valency-increasing*. When the unmarked verb root is intransitive, the goal-selecting construction is monotransitive, and P maps to the goal participant. When the unmarked verb root is monotransitive, the goal-selecting AC is ditransitive, and R maps to the goal participant, while T maps to the theme, which represents the companion phrase.

2.4.3 Path-selecting constructions

With certain verbal roots, a path is selected as a core argument when the verb is marked with *-an*. Here, path refers to an expression that describes the trajectory of an entity that undergoes a change of location. Path may refer to any such description of a trajectory that is not a source or goal.

In examples (28)–(29), the path is expressed as the P argument of a verb in AV in the AC marked with *-an*. In both ACs, the path is realized as an unmarked NP in postverbal position. In corresponding BCs, the verb selects only one core argument, S, which maps to an agentive mover. Other semantic information about the spatial orientation of the motion described may not be expressed as an unmarked NP, but instead is realized as a PP. Note that the precise semantic relationship of the path argument to the event in ACs marked with *-an* depends on the base verb. In (28b), for which there is no exact monoclausal equivalent, the verb *luncat-an* means ‘to jump

over’. In (28) the verb *lumput-an* means ‘to overtake, surpass (s.o or s.t.)’ and can refer to literally running past someone or simply surpassing another in some measure, e.g., test scores or ranking.

(28) Sundanese, Path-selecting construction

- a. *Abi luncat ti solokan.*
 1SG jump from canal
 ‘I jumped away from the drainage trench.’ (BC)
- b. *Abi nga-luncat-an solokan.*
 1SG AV-jump-LOC.APPL canal
 ‘I jumped over the drainage trench.’ (AC) (CT1-003)

(29) Sundanese, Path-selecting construction

- a. *Ucing di-gebah-keun téh meni langsung lumpat ka jalan.*
 cat PV-shoo-THM.APPL PRT very immediately run to street
 ‘The cat that was shooed off ran to the street right away.’ (BC) (*Lumpat* 2021)
- b. *Ujang nga-lumpat-an Asep.*
 U. AV-run-LOC.APPL A.
 ‘Ujang overtook Asep (lit. Ujang ran past Asep).’ (AC) (CT1-003)

In the above examples, the path-selecting construction marked with *-an* is monotransitive. Thus we see that with intransitive verbal roots, this construction is valency-increasing. Examples of this type are non-controversially considered applicatives in most frameworks.

It appears that some transitive verbs in Sundanese may select a path as a core argument without AM-marking. In these cases, suffixation of the verb with *-an* does not result in a change to argument structure, but an aspectual change with a pluractional meaning. An example is shown below in (30) with the verb root *térékél* ‘to climb up (s.w.)’. In the BC in (30a), the agentive mover is selected to map to A, and is expressed as the *ku*-marked agent phrase following the verb *ditérékél*, which is in PV. The semantic path is selected to map to P, and is expressed as an unmarked NP in preverbal position, i.e. *tangkal* ‘tree’. In the AC in (30b), there is no change to argument structure; the clause remains monotransitive, and the mapping of roles to arguments is also unchanged. However, the clause has a pluractional meaning, as reflected in the translation ‘climbed up and down / all over’, and denotes repeated or distributive climbing action in the event described.

(30) Sundanese, Path-selecting construction

- a. *Tangkal di-térékél ku abi.*
 tree PV-climb.up by 1SG
 ‘The tree was climbed up by me.’ (BC)
- b. *Tangkal di-térékél-an ku abi.*
 tree PV-climb.up-LOC.APPL by 1SG
 ‘The tree was climbed up and down / all over by me.’ (AC) (CT1-004)

In summary, the path-selection construction is found with intransitive and transitive verbal roots in Sundanese. With intransitive verb roots, when the root is affixed with *-an*, we observe a change in argument structure, as the path is now selected to map to a core argument. Such *-an* marked clauses represent valency-increasing ACs. On the other hand, with some transitive verbs, a semantic path may be selected to map to P without any AM-marking. In such cases, when the verb is marked with *-an*, the path continues to be expressed as P, and the clause has a pluractional meaning.

2.4.4 Category-changing locative-selecting construction

With certain nominal bases, denominal verbs marked with the suffix *-an* select a location to map to a core argument. For example, the denominal verb *uyah-an* is composed of the noun *uyah* ‘salt’ and the locative applicative suffix *-an*. This verb takes a goal as a core argument, as shown in (31), where the P argument *sayur* ‘soup’ describes the entity to which salt is to be applied. A similar example with the base *céngék* ‘chili pepper’ is shown in (32). This type of construction is monotransitive.

(31) Sundanese, Category-changing locative construction

Sayur teu di-uyah-an.

soup not PV-salt-LOC.APPL

‘The soup was not salted.’

(CT1-017)

(32) Sundanese, Category-changing locative construction

Sayur di-céngék-an.

soup PV-chili-LOC.APPL

‘To the soup, chili pepper was added.’

(CT1-017)

For category-changing locative-selecting constructions, there is no possible comparison of a BC and an AC because the base cannot function as a verb without affixation with an AM. Despite this, they are clearly related to locative-selecting applicatives. In terms of form, they are marked with *-an* on the verb, and show (minimally) monotransitive clausal argument structure. In terms of meaning, like the goal-selecting applicatives described in §2.4.2, they describe motion events involving a change of location and select a semantic goal to map to a core argument. While these constructions do not show selection of a theme or mover participant as a clausal argument, there is still an inherent semantic theme expressed by the root, i.e. *uyah* ‘salt’ or *céngék* ‘chili pepper’, which undergoes a change of location in the event described. Thus we see a number of similarities between these category-changing constructions and locative-selecting applicatives in Sundanese.

2.5 Theme- and instrument-selecting constructions marked with *-keun*

In this section I will describe morphologically marked constructions in which a participant with the semantic role of theme or instrument is selected as a core argument. In Sundanese, this type of

construction is marked with the verbal suffix *-keun*. I use the label *theme* to refer to an entity that undergoes a change of location or is located in space (see Van Valin 2005: 54). *Instrument* refers to an inanimate entity manipulated to some effect (see Van Valin 2005: 58–59; Fillmore 1968). It is not always possible to clearly distinguish instruments from themes; the employment of material objects that are classified as instruments commonly involves direction of the object into motion by an agent. See also the discussion in Kroeger (2007), where the term “displaced themes” is used to describe certain instruments selected as core arguments in constructions marked with Indonesian *-kan*. In this chapter, I use the term instrument only when there is a clear association between manipulation of an entity and an (intended) effect in the meaning of the clause. Constructions which select a (non-instrument) theme role as the applied phrase are described in §2.5.1. Constructions which select an instrument role as the applied phrase are described in §2.5.2.

2.5.1 Theme-selecting constructions

With certain verbal bases, a theme argument is selected as the P argument and mapped to a core argument position when the verb is marked with the suffix *-keun*. An example is shown in (33) below. The BC in (33a) is monotransitive and the verb *tangkeup* ‘to put one’s arms around (s.t.), to embrace (s.o. or s.t.)’ selects an agent as its A argument, and a patient (or perhaps a target or goal) as its P argument. In the AC in (33b), the verb is marked with the suffix *-keun*. Now the verb selects a different kind of participant role to map to P, neither an agent nor a patient. In this case, it is a theme, i.e. an entity that undergoes a change in position, and expresses the part of the agent’s body, *leungan* ‘arm’, that is directed to wrap around the patient. This example would furthermore be considered a valency-preserving applicative, as the companion phrase, which expresses the patient argument in the AC, is realized as an oblique PP, *ka pamjakan* ‘around (his) wife’.

(33) Sundanese, Theme-selecting construction

- a. *Kuring ngan bisa n-(t)angkeup sirah budak dina dada.*
 1SG only can AV-hug head child on chest
 ‘I could only hug the child’s head on(to) (my) chest.’ (BC) (nar. text by Idris 2021)
- b. *Udi n-(t)angkeup-keun leungeun ka pamajikan.*
 U. AV-hug-THM.APPL arm to wife
 ‘Udi hugged his arm around (his) wife.’ (AC) (CT1-003)

(34) Sundanese, Theme-selecting construction

- a. *Sendal palid.*
 sandal drift.away
 ‘The sandal drifted away (i.e. was carried away by flowing water, as in a rain storm).’
 (BC)
- b. *Runtah di-palid-keun ka walungan ku Udi.*
 trash PV-drift.away-CAUS.THM.APPL to river by U.
 ‘The trash was set adrift into the river by Udi.’ (AC) (CT1-021)

Now consider the example shown in (34). In the BC in (34a), the verb *palid* ‘to drift away (as if carried by flowing water)’, is intransitive, and selects an entity in motion, that is, a theme, as the participant that maps to S. In the AC in (34b), the verb is now marked with the suffix *-keun*. The theme now maps to P, and in this case, it is expressed as the unmarked NP in preverbal position in a PV clause. In addition, an agent maps to A, which is realized as a *ku*-marked agent phrase, as is characteristic of A arguments in PV clauses. This type of construction may be considered as a *causative* construction, because a volitional, instigating participant is added to the suffixed clause. However, there are also clear similarities between the construction in (34b) and that shown above in (33b). Both select a theme participant as the P argument of the clause when the verb is marked with *-keun*, while in neither case may the theme map to P in the unmarked BC. These examples show that there is not necessarily a clear-cut distinction between the causative and applicative functions of AMs like *-keun*.

(35) Sundanese, Theme-selecting construction

a. *Udi tos ny-(s)urung mobil ka imah.*
 U. CMPL AV-push car to house
 ‘Udi pushed the car to/towards the house.’ (BC)

b. *Udi tos ny-(s)urung-keun mobil ka imah.*
 U. CMPL AV-push-THM.APPL car to house
 ‘Udi pushed the car to the house.’ (AC)

(CT1-027)

Another type of theme-selecting construction is shown in (35) above. In the BC in (35a), the verb *surung* ‘to push (s.t.)’ selects an agent as its A argument and a theme as its P argument. In the AC in (35b), the same verb is suffixed with *-keun*, but we see no change in argument structure; the agent again maps to A, and the theme again maps to P.

In this type of construction, the presence of the applicative suffix *-keun* might be considered to have “no effect” (see Kroeger 2007: 15–16 on cases in Indonesian where affixation of *-kan* “has no effect on the semantic or syntactic properties of the verb”). However, there are subtle semantic differences between the two clauses, as they have slightly different truth conditions. The sentence in (35a) is true in the case that the Udi applied force to the cart in the direction of the house and Udi is done pushing, that is, he has stopped applying force. This statement can be true regardless of whether or not the car reached the house. The sentence in (35b) is true in the case that Udi applied force to the cart in the direction of the house, and the cart arrived at the house. In the former, completive aspect as indicated by the aspectual particle *tos*, applies to the event of applying directional force to something only. In the latter, completive aspect as indicated by *tos* applies to the event of applying directional force to something thereby causing it to move to somewhere.

Thus we see that the addition of an AM can reflect a change in the status of a semantic role in argument structure even if that role does not map to an unmarked (or direct) core argument position in the AC but instead is encoded with prepositional marking. This suggests that applicative marking may be associated with changes in clausal argument structure that are not observable only from surface level encoding. One way of interpreting this type of evidence is to adopt the criteria used by Riesberg (2014a) for argumenthood, including the examination of pragmatic inferences bearing on the semantic relationship between a phrase and a verb. Apply-

ing these criteria, it is possible to infer from a reduced version of the AC, i.e. *Udi nyurungkeun mobil*, a version of the same clause in which the goal is replaced with its indefinite form, i.e. ‘*Udi pushed the car to somewhere*’. Under this analysis, the companion phrase expressing a semantic goal in clauses like (35b) has at least some intermediate degree of argumenthood distinguishing it from an adjunct. A reduced version of the BC, i.e. *Udi nyurung mobil*, does not support this same inference, by which evidence the phrase expressing the goal (or direction) in the BC has the status of an adjunct, not a verbal argument.

A similar situation is found with the verbal root *alung*, which means ‘to throw (s.t.)’. It appears that this verb may select a theme as its P argument without AM-marking, as suggested by the phrase *ngalung bal/geulang* defined by Hardjadibrata (2003: 16) as “throw a ball/ring (in the night fair, etc.)”. However the base verb forms *ngalung/dialung* appear to be uncommon in contemporary usage based on lack of occurrence in corpora. Much more common are the AM-marked forms *ngalungkeun/dialungkeun* and *ngalungan/dialungan*.

As shown in (36), when the root *alung* ‘throw’ is marked with *-keun*, the theme role continues to be selected to map to P. In this example, the theme is expressed as an unmarked NP in postverbal position, *hiji batu gedé* ‘one large stone’. In addition, a semantic goal is expressed in the clause as an oblique PP *ka imah Asep* ‘at Asep’s house’. The existence of both an entity in motion as expressed by the theme role and an end-point as expressed by the goal role, is inferrable from the reduced clause *Udi ngalungkeun*. This is reflected in the definition for *ngalunkeun* given by Hardjadibrata (1985: 16): “throw st. to so./somewhere, throw with st.”⁹

(36) Sundanese, Theme-selecting *throw* construction

Udi ng-alung-keun hiji batu gedé ka imah Asep.

U. AV-throw-THM.APPL one stone large to house A.

‘Udi threw a large stone at Asep’s house.’ (AC)

(CT1-006)

If *alung* is suffixed with *-an*, the resulting construction shows no change in argument structure, but instead some semantic differences compared to clauses where *alung* is marked with *-keun*. This is exemplified in (37) below. Here the P argument, *batu* ‘stone’, again expresses a semantic theme, and the oblique PP *ka imah Asep* ‘at Asep’s house’, again expresses a semantic goal. But in addition, the *-an* marked clauses takes a pluractional meaning, as reflected by the translation ‘threw stones / threw stones repeatedly’.

(37) Sundanese, Theme-selecting *throw* construction

Udi ng-alung-an batu ka imah Asep.

U. AV-throw-THM.APPL stone to house A.

‘Udi threw stones (repeatedly) at Asep’s house.’ (AC)

(CT1-006)

The fact that base verb forms of *alung* ‘throw’ without AM-marking appear to be much less common in usage compared to the AM-marked forms suggests one possible stage of development for ACs: the base form of a verb might over time cease to be used as a predicate, while the affixed

⁹The final phrase of this definition, ‘throw with s.t.’, suggests that *alung* may be used in *-keun* marked constructions that select a semantic instrument as a core argument. However, I have not found clear clausal examples illustrating this.

form is lexicalized. Thus, a speaker may be exposed to clauses marked with *-keun* on the root *alung*, but never a corresponding BC. Nevertheless, in Sundanese there is a strong generalized pattern of association between a set of formal elements, i.e. monotransitive clause structure and verbal suffixation with *-keun*, and certain elements of meaning, i.e. selection of a theme to map to P and a goal to map to an oblique PP and the eventive meaning ‘cause s.t. to move s.w.’ That clauses marked with *-keun* on the root *alung* fit into this generalized pattern remains easily evident from observable usage of the theme-selecting construction with other lexical bases filling the verbal slot, even if a corresponding BC with *alung* does not exist as a reference point in the mind of a speaker or the paradigm of an analyst.

Another type of theme-selecting construction marked with *-keun* is shown in (38a) with the verb *béré* ‘to give’. In the BC in (38a), the theme maps to P, and the agent maps to A. This sentence can refer to the act of giving a gift or simply a transfer of physical position that is performed by the agent, e.g. ‘to hand’. In the construction marked with *-keun* in (58b), the theme is also expressed as P, but the meaning of the sentence only refers to the act of physical transfer, i.e. ‘to hand over’, not the giving a gift. This appears to represent lexicalization of the affixed form of the verb, with a shift in meaning to ‘hand over, facilitate physical transfer’ from ‘give’.

(38) Sundanese, Lexicalized change in meaning

- a. *Buku di-béré ka budak ku abi.*
 book PV-give ALL child by 1SG
 ‘The book was given to the child by me / was handed to the child by me.’ (BC)
- b. *Buku ti abi di-béré-keun ka budak ku mang=na.*
 book ABL 1SG PV-give-THM.APPL ALL child by uncle=3SG
 ‘The book from me was handed over to the child by his uncle.’ (AC) (CT1-026)

Finally, as mentioned above, for a number of verbal roots, we observe a location-selecting construction marked with *-an* in alternation with a theme-selecting construction marked with *-keun*. This alternation is discussed in §2.6 below, which includes more examples of theme-selecting constructions.

2.5.2 Instrument-selecting constructions

With some verbal bases, an instrument argument is selected as the P argument and mapped to a core argument position when the verb is marked with the suffix *-keun*.

(39) Sundanese, Instrument-selecting construction

- a. *Bal di-teunggeul maké tongkat kasti (ku Udi).*
 ball PV-hit using bat kasti by U.
 ‘The ball was hit by Udi using a *kasti* bat. (*Kasti* is a game similar to baseball.)’ (BC)
- b. *Tongkat kasti di-teunggeul-keun kana bal ku Udi.*
 bat kasti PV-hit-THM.APPL to ball by U.
 The *kasti* bat was used to hit the ball by Udi. (AC) (CT1-019)

In the BC in (39a), the patient is selected to map to P and is realized as an unmarked core argument, and the agent is selected to map to A and is realized as a *ku*-marked agent phrase. The instrument is realized as an oblique phrase headed by *maké* ‘using’.¹⁰ In the AC in (39b), the instrument, i.e. *tongat kasti* ‘kasti bat’ is now expressed as the P argument, and realized as an unmarked core argument in preverbal position. The companion phrase, the semantic patient, is now realized as an oblique PP, headed by the preposition *kana* ‘to, in, on, at’. Like the theme-selecting constructions we have seen so far, this example of an instrument-selecting AC is a monotransitive construction and, as the BC is also monotransitive, it is valency-preserving.

Example (40) below shows an instrumental AC with the verbal root *coél* ‘to scoop up, take a little of (s.t.) by dipping (s.t.)’. In the BC in (40a), the semantic agent is selected to map to A and is realized as an unmarked core argument in preverbal position, and the semantic theme (the entity that undergoes a change of location) is selected to map to P and is realized as an unmarked core argument in immediate postverbal position. The instrument manipulated as the effector of the scooping action, i.e. *témpé* ‘tempeh (k.o. fermented soybean cake)’, is encoded as an oblique phrase, *maké témpé* ‘using tempeh’. In the AC in (40b), the instrument role is now selected to map to the P argument, which is realized as an unmarked NP in immediate postverbal position. The semantic theme is now encoded as a PP headed by the preposition *kana* ‘to, in, on, at’. Again we see that the instrument-selecting AC is monotransitive, and valency-preserving.

(40) Sundanese, Instrument-selecting construction

- a. *Udi ny-(c)oel sambel maké témpé.*
 U. AV-scoop chili.sauce using tempeh
 ‘Udi scooped up (some) chili sauce using (a piece of) tempeh.’ (BC)
- b. *Udi ny-(c)oel-keun témpé kana sambel.*
 U. AV-scoop-THM.APPL tempeh to chili.sauce
 ‘Udi used (a piece of) tempeh to scoop up (some) chili sauce.’(AC) (CT1-020)

2.6 Mapping alternations for theme- and locative-selecting applicatives

For a sizeable number of bases, we observe a location- (or goal-) selecting construction marked with *-an* in alternation with a theme- or instrument-selecting construction marked with *-keun*. Several examples are given in the remainder of this section. Note that in many of these, the base verb is intransitive, and the AM-marked constructions are both causative and applicative, as an instigating causer is mapped to A.

¹⁰In Sundanese, the word *maké* and its alternate form *paké* derive from the verb *paké* meaning ‘to wear; to use’. However, these words have been grammaticalized and are used much like prepositions. Similar cases are found in other languages of western Indonesia including *Tukang Besi* (Donohue 1999: 333–334) and *Manado Malay*, a Malay-based creole spoken in North Sulawesi (Stoel 2005), though it is not always possible to differentiate clearly between grammaticalized prepositional use and use in serial verb constructions.

- (41) Sundanese, Theme/Location alternation
- a. *Gelas di-kucur-an cai ku Icih.*
 glass PV-flow-CAUS.LOC.APPL water by I.
 ‘The glass has water poured onto it by Icih (e.g. because it needs rinsing).’ (AC)
- b. *Cai di-kucur-keun ku Icih kana gelas.*
 water PV-flow-CAUS.THM.APPL by I. to glass
 ‘Icih poured water into in a drinking glass.’ (AC) (CT1-020)
- (42) Sundanese, Instrument/Location alternation
- a. *Udi keur di-kurubun.*
 U. PROG PV-bundled
 ‘Udi is all bundled up (i.e. wrapped up, as in blankets)’ (BC)
- b. *Mobil di-kurubun-an maké terpal ku Udi.*
 car PV-bundled-CAUS.LOC.APPL using tarp by U.
 ‘The car was covered up with a tarp by Udi.’ (AC)
- c. *Terpal di-kurubun-keun ka budak.*
 tarp PV-bundled-CAUS.THM.APPL LOC child
 ‘The tarp was used to bundle up the child.’ (AC) (CT1-022)
- (43) Sundanese, Theme/Location alternation
- a. *Acan gé anclo m useup=na, geus di-samber mantén ku*
 hardly also touch.water fish.hook=DEF, already PV-pounce.on in.meantime by
gabus.
 snakehead
 ‘The fish hook had hardly touched the water, when already it was swooped upon by a
 snakehead fish.’ (BC) (“anclo m” 2019)
- b. *Sayur kacang di-anclo m-an daun salam.*
 soup bean PV-touch.water-CAUS.LOC.APPL leaf bay.leaf
 ‘The bean soup had a bay leaf submerged into it.’ (AC) (CT1-020)
- c. *Kue di-anclo m-keun kana kopi ku Ujang.*
 cookie PV-touch.water-CAUS.THM.APPL to coffee by U.
 ‘The cookie was submerged into the coffee by Ujang.’ (AC) (CT1-020)

In AM-marked verbal constructions with *rungkup* ‘cover, spread over’, we observe an alternation between a goal-selecting construction with *-an* as in (44a) and a theme- or instrument-selecting construction with *-keun* as in (44b).

- (44) a. *Udi nga-rungkup-an lanca maké mangkok.*
 U. AV-cover-CAUS.LOC.APPL spider using bowl
 ‘Udi enclosed the spider with a bowl.’ (AC)
- b. *Udi nga-rungkup-keun mangkok kana lanca.*
 U. AV-cover-CAUS.THM.APPL bowl to spider
 ‘Udi used a bowl to enclose a spider (lit. enclosed a bowl onto a spider).’ (AC)
 (CT1-022)

In the example below with the same verb taken from the Leipzig corpus, the instrumental meaning—manipulating an inanimate entity as an effector—is quite clear from the context.

- (45) Sundanese, Instrumental use of *rungkup-keun*
Ng-alung-keun atawa nga-rungkup-keun kecrik aya dua
 AV-throw-THM.APPL or AV-cover-CAUS.THM.APPL cast.net exist two
ka-tangtu-an supaya hasil=na mekar jeung ng-arungkup sasar-an anu
 NMLZ-certain-NMLZ so.that result=DEF open.up and AV-cover seek-NMLZ REL
geus di-tangtu-keun.
 already PV-certain-CAUS
 ‘With respect to throwing or casting a small cast-net (lit. using a small cast-net to enclose s.t.), there are two rules (to follow) so that the result (is) (the net) opening up and covering the target which was already determined.’ (CT1-044, based on MPI-SUN 18000)

2.7 Beneficiary-selecting constructions marked with *pang- -keun*

Sundanese has two morphologically marked constructions which select a beneficiary participant as a core argument. A beneficiary represents an entity (typically, a person, though other animate entities are possible) who is affected advantageously by an event without being the agent or primary affected participant of the event (Kittilä & Zúñiga 2010: 2). In this section I describe the beneficiary-selecting construction marked with the circumfix *pang- -keun*. Beneficiary-selecting constructions in Sundanese may also be marked with the suffix *-keun* in a limited number of cases, which is discussed in §2.8.

Benefactive constructions marked with the circumfix *pang- -keun* are frequent in usage and extremely productive across lexical bases in Sundanese. This type of construction has a substitutive benefactive meaning, because it indicates that an agent performs an action on behalf of a beneficiary, who accrues a benefit in that the beneficiary does not have to perform this action her/himself (Kittilä 2005; Van Valin & LaPolla 1997).

- (46) Sundanese, Substitutive benefactive construction
- a. *Bapa nga-betot parahu Asep.*
 father AV-pull boat A.
 ‘Father pulled Asep’s boat.’ (BC)

- b. *Bapa m-(p)ang-metot-keun parahu Asep.*
 father AV-BEN.APPL-pull-BEN.APPL boat A.
 ‘Father pulled Asep’s boat (for him).’ (AC) (CT1-034)

Consider the example shown above, in which the minimal difference between (46a) and (46b) is that in (46b), the verb is marked with the circumfix *pang-* *-keun*. Both clauses describe the action of pulling a boat to move it (e.g. to the edge of a body of water) with the verbal root *betot* ‘to pull, to yank’. In both clauses, the semantic agent is selected as the A argument and is realized as an unmarked core argument in preverbal position, while the semantic theme is selected as the P argument and is realized as an unmarked core argument in immediate postverbal position. However, the two clauses differ in meaning. When the verb is marked with *pang-* *-keun*, the sentence is understood to mean that Father pulled the boat belonging to Asep *on Asep’s behalf*, thereby relieving Asep of the need to do so. When the verb is unmarked, the sentence merely indicates that Father pulled the boat belonging to Asep, and makes no comment on whether Asep (or any other party) benefited from this event, or whether Asep had a need to perform this action.

Note also that in (46b), the semantic beneficiary is not realized with encoding that we expect to see for a core argument (an unmarked NP, or *ku*-marked agent phrase). Instead the beneficiary participant is expressed as the possessor of the P argument (theme), and it is understood that the beneficiary is the same entity as the possessor through pragmatic inference. The expression of a beneficiary as a possessor is very common in Sundanese. Also common is for the beneficiary to be left entirely unmentioned in the clause.

An example of non-realization of the beneficiary is shown in (47) below, where the verb, *pang-neundeun-keun* ‘place (s.t.) for (s.o.)’, is a polite imperative form. In the pragmatic context, the hearer infers that the speaker is the beneficiary of the requested event, because the request implicates that the speaker desires the hearer—and not the speaker herself—to carry out the event.

(47) Sundanese, Substitutive benefactive imperative

- Udi, pang-neundeun-keun balanja-an di dapur.*
 U. BEN.APPL-place-BEN.APPL shopping-NMLZ in kitchen
 ‘Udi, please put the shopping purchases in the kitchen (for me).’ (CT1-022)

Apart from non-realization and expression as a possessor NP, the beneficiary participant can be overtly expressed in a clause marked with *pang-* *-keun* in two other ways.

First, the beneficiary can be selected as the R argument in a ditransitive clause and realized as unmarked NP. An example is given below in (48). In the BC in (48a), the theme is selected to map to P, and is realized as unmarked NP immediately after the verb, which is AV. In the AC shown in (48b), the verb again is in AV and it is marked with *pang-* *-keun*. We see two unmarked NPs immediately following the verb, first the beneficiary *Lilis* ‘female personal name’, which maps to R, and then the theme *sangu* ‘cooked rice (and accompanying dishes)’, which maps to T. In the AC shown in (48c), the verb is in PV and marked with *pang-* *-keun*. The beneficiary again maps to R and is expressed as an unmarked NP in preverbal position. The theme again maps to T, and is expressed as an unmarked NP in immediate postverbal position.

(48) Sundanese, Beneficiary encoded as unmarked NP

- a. *Asep m-(b)awa sangu keur Lilis.*
 A. AV-bring cooked.rice for L.
 ‘Asep brought food for Lilis.’ (BC)
- b. *Asep m-(p)ang-mawa-keun Lilis sangu.*
 A. AV-BEN.APPL-bring-BEN.APPL L. cooked.rice
 ‘Asep brought food for Lilis.’ (AC) (CT1-025)
- c. *Lilis di-pang-mawa-keun sangu ku Asep.*
 L. PV-BEN.APPL-bring-BEN.APPL cooked.rice by A.
 ‘Lilis had food brought for her by Asep.’ (AC) (CT1-025)

Second, the beneficiary can be realized as an oblique PP in a monotransitive clause. In these cases the semantic role that is selected to map to P in the BC continues to map to P in AC. Thus, there is no change in the mapping of the companion phrase in the clause marked with *pang--keun*. This is illustrated in example (49). Here, the verb *bawa* ‘to bring’ is marked with *pang--keun*, but the semantic theme *sangu* continues to be expressed as P and realized as unmarked core argument. This holds both when the marked verb is in AV as shown in (49a) and when it is in PV, as shown in (49b). In both cases, the beneficiary is expressed with the PP *keur Lilis* ‘for Lilis’.

(49) Sundanese, Beneficiary encoded as PP

- a. *Asep m-(p)ang-mawa-keun sangu keur Lilis.*
 A. AV-BEN.APPL-bring-BEN.APPL cooked.rice for L.
 ‘Asep brought food for Lilis.’
- b. *Sangu di-pang-mawa-keun ku Asep keur Lilis.*
 cooked.rice PV-BEN.APPL-bring-BEN.APPL by A. for L.
 ‘Food was brought by Asep for Lilis.’ (CT1-025)

Examples (48a) and (49) show that it is possible to for a beneficiary to be expressed with a PP in both a BC and a AC with the verb *bawa* ‘to bring’. However, the semantic relationship between the verb and a beneficiary PP in the clause shows differences based on whether the verb is marked with *pang--keun* or not. Let’s return to the reduction and inference tests used by Riesberg (2014a) and apply them to examples (2a) and (49a). When the verb does not bear any AM, the reduced clause *Asep mawa sangu* does not support the inference that *Asep brought rice for someone*. This supports the analysis that the beneficiary PP in the BC is an adjunct. On the other hand, when the verb is marked with *pang--keun*, from the reduced clause *Asep mawakeun sangu* it is possible to infer that *Asep brought rice for someone*. This suggests that the beneficiary PP in clauses marked with *pang--keun* has at least an intermediary status because it shows some properties of argumenthood. For this reason, I will treat the beneficiary PP in constructions marked with *pang--keun* as oblique arguments.

The beneficiary-selecting construction marked with *pang--keun* is highly productive across lexical bases in Sundanese. Bases that combine with *pang--keun* in benefactive constructions may be agentive or non-agentive and stative or dynamic. However, verbs marked with *pang--keun*

always describe agentive and volitional actions, in keeping with the meaning of the construction, which denotes that an agent performs some action on behalf of another party.

An example of a benefactive construction with a stative base is given in (50). In such examples, the meaning of the construction marked with *pang-* *-keun* is both causative and benefactive; the bare stem *garing* refers to the state of being dry, while the affixed stem *pang-garing-keun* refers to the act of causing something to become dry on behalf of a beneficiary.

(50) Sundanese, Substitutive benefactive with stative base

- a. *Baju geus garing.*
 clothing already dry
 ‘The clothes are already dry.’ (BC)
- b. *Baju Asep di-pang-garing-keun dina mesin ku Ani.*
 clothing A. PV-CAUS.BEN.APPL-dry-CAUS.BEN.APPL LOC machine by A.
 ‘Asep’s clothes were dried (for him) in the dryer by Ani.’ (AC) (CT1-001)

Examples of benefactive construction with dynamic intransitive bases are given below. In such examples, the meaning of the construction marked with *pang-* *-keun* is again both causative and applicative. In example (51), the bare stem *kucur* in the BC means ‘to flow, as of liquids’, while the affixed stem *pang-(k)ucur-keun* in the AC means ‘to make (s.t.) flow for s.o.’. A similar example is shown in (52) with the root *jogéd* ‘to dance’.

(51) Sundanese, Benefactive construction with dynamic intransitive base

- a. *Cai hujan ng-(k)ucur=na ka solokan.*
 water rain AV-flow=3 to canal
 ‘Rain water flows into the drainage trench.’ (BC) (CT1-023)
- b. *Pang-(k)ucur-keun cai kana ember.*
 CAUS.BEN.APPL-flow-CAUS.BEN.APPL water to bucket
 ‘Please pour water into the bucket (for me).’ (AC) (CT1-024)

(52) Sundanese, Benefactive construction with dynamic intransitive base

- a. *Euis jogéd keur abi.*
 E. dance for 1SG
 ‘Euis danced for me.’ (BC)
- b. *Wayang abi di-pang-jogéd-keun ku Euis.*
 puppet 1SG PV-CAUS.BEN.APPL-flow-CAUS.BEN.APPL by E.
 ‘My shadow puppet doll was made to dance by Euis for me.’ (AC) (CT1-039)

An additional example of the benefactive construction with the transitive base *beuleum* ‘to roast, grill’ is given in (53). The AC shown in (53b) is ditransitive, compared to the BC in (53a), which is monotransitive.

- (53) Sundanese, Benefactive construction with transitive base
- a. *Bapa m-(b)euleum sampeu keur abi.*
 Father AV-roast cassava for 1SG.
 ‘Father roasted cassava for me.’ (BC)
- b. *Bapa m-(p)ang-meuleum-keun abi sampeu.*
 father AV-BEN.APPL-ROAST-BEN.APPL 1SG cassava.
 ‘Father roasted cassava for me.’ (AC) (CT1-002)

Example (54) below shows a benefactive construction with the transitive base *tinggali* ‘to see’. The base verb *ninggali* means ‘to see (AV)’ and selects a perceiver experiencer to map to A as in the BC in (54a). This verb is neutral for volitionality of A in AV.¹¹ The affixed verb, *pang-ninggali-keun*, which is shown in imperative form in the AC in (54b), means ‘watch (s.t.) for (s.o.)’, and A must be volitional, as reflected in the English translation using the active perception verb ‘watch’ rather than the experiential perception verb ‘see’. A similar example is given with the root *ambeu* ‘to smell, sniff’ in (55) below.

- (54) Sundanese, Substitutive benefactive with experiential transitive base
- a. *Déwi n-(t)inggali mobil Asep.*
 D. AV-see car A.
 ‘Dewi sees Asep’s car.’ (BC) (FM4-046)
- b. *Pang-ninggali-keun pasak-an abi.*
 BEN.APPL-see-BEN.APPL cook-NMLZ 1SG
 ‘Please watch my cooking (for me) (e.g. because I need to step away).’ (AC) (CT1-040)

- (55) Sundanese, Substitutive benefactive with experiential transitive base
- a. *Abi ng-ambeu kue haneut.*
 1SG AV-smell cookie warm
 ‘I smelled warm cookies.’ (BC) (CT1-039)
- b. *Pang-ambeu-an-keun minyak seungit éta.*
 BEN.APPL-smell-PLUR-BEN.APPL oil fragrant DIST.DEM
 ‘Please smell these perfumes (for me) (e.g. because I need to choose one).’ (AC) (CT1-015)

Taken together, the examples in this section show that the beneficiary-selecting constructions marked with *pang-* *-keun* are minimally monotransitive and always valency-increasing. With transitive base verbs, an AC marked with *pang-* *-keun* selects three semantic roles that may map to a core argument: an agent (i.e. a volitional participant that performs the action), a “primary affected target participant” role (e.g. a semantic patient, theme, or stimulus, among others), and a beneficiary. With intransitive base verbs, a *pang-* *-keun* marked AC is both causative

¹¹In PV, speakers may choose to use the voice prefix *ka-* to express non-volitionality, or the prefix *di-* to express volitional action. See also example (14) and the discussion of it in §2.2.2.

and applicative, and again selects three semantic roles that may map to core arguments: an instigating causer agent, a primary affected target participant, and a beneficiary argument. Of the three arguments selected in *pang-* *-keun* marked constructions, the beneficiary in particular shows variable mapping. If it is realized, the beneficiary can be expressed as an oblique PP in a monotransitive clause, the possessor of the P argument in a monotransitive clause, or the R argument of a ditransitive clause. The mapping of the affected target participant is determined in turn. In monotransitive clauses it is expressed as the P argument, while in ditransitive clauses, it is expressed as the T argument.

Even though the substitutive benefactive construction is extremely productive, not all verbal roots are frequently attested in usage with this construction. Due to the construction's substitutive benefactive meaning, in order for a clausal example to be felicitous, the action described by the base verb must be considered necessary or obligatory for some party to perform, and it must be considered possible and beneficial to relieve this party of that necessity by performing the action on this party's behalf. Thus the usage of this construction is influenced by cultural and contextual factors that condition an interlocutor's understanding of necessity and obligation, beneficiality, and performance in lieu of another party. For example, the primary speaker I worked with reported that the verb *pangdaharkeun* with the intended meaning 'to eat for (s.o.)' (cf. *dahar* 'to eat') would not generally be used because an agent's eating action would not relieve another party's need to eat. She also found it difficult to construct felicitous examples with the verb *pangliangkeun* 'to make a hole for (s.o.)' (cf. *ngaliang* 'to make a hole') as making a hole would not be considered the obligation of any person, and an animal which digs by instinct could not be said to benefit from or be relieved of the need to make a hole by some action of a human agent on its behalf.

2.8 Beneficiary-selecting constructions marked with *-keun*

Beneficiary-selecting constructions marked with *-keun*, as shown in (56) below, are reported by Hanafi (1997). However, in my data, such benefactives are rare; I found only a handful of these constructions during elicitation of meanings on the Leipzig valency questionnaire, in recorded natural texts, and in corpus resources.

(56) Sundanese, Beneficiary-selecting construction

a. *Udi m-(b)uka panto keur abi.*

U. AV-open door for 1SG
'Udi opens the door for me.' (BC)

b. *Udi m-(b)uka-keun abi panto.*

U. AV-open-BEN.APPL 1SG door.
'Udi opened the door for me.' (AC)

(CT1-039, based on Hanafi 1997: 23)

It is possible that the benefactive meaning associated with *-keun* is related to constructions in which marking with *-keun* indicates that the action is performed by a third party. Consider example (57) below, showing the verb *injeum* 'to borrow' in PV. In the BC in (57a), the theme (the item borrowed) is selected as P, and the agent which is selected as A is the borrower. In

the construction marked with *-keun* in (57b), the theme is still the P argument, but the agent is a third-party (neither owner nor borrower) who facilitates the borrower's (temporary) possession of the theme. I interpret this construction as a type of causative, that is, the third party causes or lets the item be borrowed by the borrower.

(57) Sundanese, Third-party causative construction

- a. *Buku téh bisa di-injeum ku manéh...*
 book PRT can PV-borrow by 2SG
 'The book can be borrowed by you... [but you must return it tomorrow].' (BC)
 (Hardjadibrata 1985: 142)
- b. *Sapédah Asep di-injeum-keun ku Ema ka Udi.*
 bicycle A. PV-borrow-CAUS.THM.APPL by E. to U.
 'Asep's bicycle was lent out by Mother to Udi (i.e. Mother let Udi borrow Asep's bicycle).' (AC) (CT1-003)

In other cases, transfer verbs marked with *-keun* show semantic effects from apparent lexicalization. Consider the examples below showing the verb *kirim* 'to send'. The example in (58a) shows that both the base form of the verb and the verb affixed with *-keun* can be used to mean 'send (s.t.)' without any detectable change in meaning. However in a sentence like (58b), where it is explicitly specified that a third-party service is used to perform the sending action, the suffixed form of the verb is preferred. It appears that the construction marked with *-keun* can indicate involvement of a third-party who carries out the action, but this has undergone semantic bleaching because it is not usually feasible to send an item to a recipient without the involvement of a third-party carrier.

(58) Sundanese, Semantic bleaching with *-keun*

- a. *Abi ng-(k)irim(-keun) buku ka adi.*
 1SG AV-send(-THM.APPL) book to younger.sibling
 'I sent the book to (my) younger sibling.' (No effect)
- b. *Abi ng-(k)irim-keun buku maké pos.*
 1SG AV-send-CAUS.THM.APPL book using post
 'I sent the book by mail (lit. had the book sent by the postal service).' (AC) (CT1-003)

2.9 Other ACs

In this section, I will discuss constructions in which the argument selected as P in a AM-marked clause is a semantic content, addressee (or other type of perceiver), stimulus, or performance.

With verbs of communication, especially verbs of speaking, suffixation with *-keun* often marks an content-selecting construction, while suffixation with *-an* often marks an addressee-selecting construction. In non-AM marked clauses with verbs of this type, the content of an utterance is expressed either as direct speech, in a clause headed by the complementizer *yén*, or

as the complement of the verb *ngeunaan* ‘concerning, with regards to’ cf. *keuna* ‘to hit, affect’. The addressee is typically expressed with a PP.

An example is given below in (59) with the root *béja* ‘to inform; news’. In (59b), we see that when the verb is marked with *-keun*, the content (or topic) of the communication event may be expressed as the P argument and realized as an unmarked NP in immediate postverbal position. In (59c), the verb is marked with *-an*, and the addressee may be expressed as the P argument and realized as an unmarked NP in immediate postverbal position.

(59) Sundanese, Content- and addressee-selecting constructions

- a. *Acéng teu bé-béja ka sa-saha ngeunaan k<al>bur=na b<ar>udak santri*
 A. NEG RDP-inform to RDP-who concerning PL-escape=DEF PL-child student
ti pasantrén.
 from religious.school
 ‘Acéng did not tell anyone about the students’ running away from the religious school.
 (BC) (Kurniawan 2013: 321)]
- b. *Udi nga-béja-keun pa-damel-an Asep ka Pak RT.*
 U. AV-inform-CONT.APPL NMLZ-work-NMLZ A. to mister neighborhood
 ‘Udi reported Asep’s job to the neighborhood officer (e.g. because it involves illegal activity).’ (AC) (CT1-044)
- c. *Pak RT nga-béja-an nu boga imah aya jelema di imah.*
 mister neighborhood AV-inform-LOC.APPL REL own house exist person in house
 ‘The neighborhood officer informed the homeowner that a person had been in the house.’ (AC) (CT1-044)

A second example is given in (60) below. The AC in (60b) shows that the semantic content may be expressed as P and realized as an unmarked NP when the verb is marked with *-keun*. The AC in (60c) shows that the addressee may be expressed as P and realized as an unmarked NP when the verb is marked with *-an*. There is also a semantic change in the meaning of the verb bearing *-an*; *carios-an* is used to mean ‘to scold, lecture’ while *carios* is more neutral and simply means ‘to talk’. The same meaning holds for *omong-an* ‘to scold, lecture’, which differs also in that *omong* ‘to talk’ represent a more casual speech register, while *carios* is more polite.

(60) Sundanese, Content- and addressee-selecting constructions

- a. *Abi ny-(c)arios ka mama, “Ma, abi hoyong miliarian damel.”*
 1SG AV-talk to mother, mom 1SG want AV.look.PLUR work
 ‘I said to my mother, “Ma, I want to look for work.”’ (BC) (FM4-027-B)
- b. *Hayang urang kempel ny-(c)arios-keun pa-damel-an.*
 let 1PL gather AV.talk-CONT.APPL NMLZ-work-NMLZ
 ‘Let’s meet up and talk about the job.’ (AC) (FM4-027-B)
- c. *Pa Guru keur ny-(c)arios-an Udi.*
 mister teacher PROG AV-talk-LOC.APPL
 The teacher is scolding/lecturing Udi. (CT1-040)

Other roots that are found with content- and addressee-selecting ACs include *carita* ‘to tell a story; story’, *harewos* ‘to whisper’, and *tanya* ‘to ask’.

With a number of verbs denoting performing arts, suffixation with *-keun* marks a performance-selecting construction. Example (61) below shows an AC with the verb *ng-ibing* ‘to dance’ c.f. *ibing* ‘a refined type of Sundanese dance’ (Hardjadibrata 2003: 327). The base verb is intransitive, as shown in (61a). A noun denoting the type of performance cannot be expressed as the P argument of the base verb, as shown in (61b). When the verb is marked with *-keun*, however, the type of performance can be expressed as the P argument and realized as an unmarked NP, as shown in (61a).

(61) Sundanese, Performance-selecting construction

- a. *Budak keur ng-ibing.*
child PROG AV-dance
‘The child is dancing’. (BC)
- b. **Budak keur ng-ibing dangdut.*
child PROG AV-dance k.o.dance
Intended: ‘The child danced *dangdut*.’ (Ungrammatical)
- c. *Lilis ng-ibing-keun jaipong Kawarang.*
L. AV-dance-THM.APPL traditional.dance Kawarang
‘Lilis danced *jaipong* in the style of Kawarang (place name)’ (AC) (CT1-031)

Example (62a) below shows an AC with the verb root *nyanyi* ‘to sing’. In the BC in (62a), we see that the NP expressing the performance may be realized as an unmarked NP in postverbal position when the verb is in AV. However, it is not grammatical for the performance to be selected as the P argument in PV, as shown in (62b). The performance is only licensed to map to P if the verb is suffixed with *-keun*, as shown in (62a). Base verbs like *nyanyi* ‘to sing’ may be considered semi-transitive because we do not see a full voice alternation in the mapping for two core arguments (A and P), as is otherwise observed for transitive verbs in Sundanese.

(62) Sundanese, Performance-selecting construction

- a. *Abi keur nyanyi lagu “Curug Cinulang.”*
1SG PROG sing song waterfall Cinulang
‘I am singing the song, “Waterfall of Cinulang”. (BC) (CT1-006)
- b. **Lagu di-nyanyi ku abi.*
song PV-sing by 1SG.
Intended: ‘The/a song was sung by me.’ (Ungrammatical) (CT1-003)
- c. *“Indonesia Raya” di-nyanyi-keun ku kabéh.*
Indonesia great PV-sing-THM.APPL by all
‘“Indonesia Raya” was sung by everyone.’ (AC) (CT1-003)

Communication events of the type described by content- and addressee-selecting ACs and performance events of the type described by performance-selecting constructions are both se-

mantically similar to other kinds of caused perception events.¹² Thus we can characterize a communicator as a type of agent, an addressee as a type of perceiver, and the content of an utterance as a type of perceived phenomenon, which I include under the semantic role label *stimulus*, following Van Valin (2005). In the same way, we might characterize a performer as a type of agent, and a performance as a type of perceived phenomenon, or stimulus.

Other types of verbs in Sundanese in fact likewise show constructions in which marking with *-keun* is associated with selection of a perceiver as a core argument and/or marking with *-an* is associated with selection of a stimulus as a core argument and the meaning ‘to cause to perceive’.

For instance, with a number of verbs describing expressions or gestures, the addition of the *-an* suffix is associated with the mapping of a core argument position to a semantic role that may be described as an (intended) perceiver, target, or goal.

In the BC, the verb *seuri* simply means ‘smile’. When the same verb is suffixed with *-an* in the AC in (63b), *seurian* means ‘to smile at (s.o.)’. It is difficult to precisely categorized the role of the participant that maps to P, i.e., the entity that is smiled at. However, it is possible to consider this entity an intended perceiver or a target in whose direction the gesture is perceivable.

(63) Sundanese, Perceiver-selecting construction

a. *Ujang seuri (ka abi).*

U. smile (to 1SG)

‘Ujang smiled (at me).’

b. *Ujang ny-(s)seuri-an abi.*

U. AV-smile-LOC.APPL 1SG

‘Ujang smiled at me.’

(CT1-002)

In (64), we see a similar construction in which *-an* is suffixed on the verb *kieup* ‘to blink’. When *kieup* is suffixed with *-an*, an intended perceiver is selected as the P argument. If winking is considered a type of communication (i.e., a signal), this participant can also be considered an addressee. Note that we also observe a semantic change in the suffixed verb; while *kieup* means ‘to blink’, *kieup-an* means ‘to wink at (s.o.)’.

(64) Sundanese, Perceiver-selecting construction

a. *Ujang ng-(k)ieup.*

U. AV-blink

‘Ujang blinked.’

b. *Ujang ng-(k)ieup-an Icih.*

U. AV-blink-LOC.APPL I.

‘Ujang winked at Icih.’

(CT1-007)

Verbs of perception in Sundanese typically select a perceiver to map to A and a stimulus to map to P, as shown in (65a) with the verb *tinggali* ‘to see’. When such verbs are marked

¹²For instance, in the Frame Semantics framework, communication events are considered to be a kind of caused perception event (see ICSI 2001).

with *-keun*, we see an interesting construction as shown in (65b). The meaning of the verb is now causative and an instigating causer is selected to map to A. However, perhaps contrary to expectation, the role that maps to A of the BC is not “demoted” to P. Instead, the semantic stimulus continues to be selected to map to P. The perceiver is expressed as an oblique argument marked with the preposition *ka* ‘to’. This type of construction is syntactically causative but also similar in meaning to stimulus-selecting ACs.

(65) Sundanese, Stimulus-selecting causative construction

- a. *Déwi n-(t)inggali mobil Asep.*
 D. AV-see car A.
 ‘Dewi sees Asep’s car.’ (BC)
- b. *Asep n-(t)inggali-keun mobil=na ka Dewi.*
 A. AV-see-CAUS.APPL to D.
 ‘Asep showed his car to Dewi.’ (FM4-046)

A similar construction is found with the stative root *tembong* ‘visible’, which refers to the quality of being perceivable by sight.

(66) Sundanese, Stimulus-selecting causative construction

- a. *Gunung tembong ti imah abi.*
 mountain visible from house 1SG
 ‘The mountain is visible from my house.’
- b. *Baju anyar di-tembong-keun ku Asep ka Icih.*
 clothing new PV-visible-CAUS by A. to I.
 ‘The new clothing was shown by Asep to Icih.’ (CT1-032)

When the same verb is suffixed with *-an* we observe a different type of causative construction, which is exemplified in (67).

(67) Sundanese, Perceiver-selecting causative construction

- Manuk n-(t)embong-an ka Udi.*
 bird AV-visible-CAUS to U.
 ‘The bird showed itself to Udi.’ (CT1-032)

Here, *nembongan* means ‘to show oneself (to s.o.)’. An agent is selected to map to A, as we expect in a causative construction, however, the same participant is understood to be stimulus, i.e. that which is perceived. A perceiver participant is selected to map to a verbal argument that is realized as an oblique PP. This PP is not an adjunct because its indefinite form can be inferred from the reduced clause. That is, from the reduced clause *Manuk nembongan*, we can infer that ‘The bird showed itself to someone.’

On the other hand, with many perception verbs that take a stimulus as a core argument without AM-making, we see no change in argument structure at all under suffixation with *-an*. With such roots, suffixation with *-an* may result in semantic changes to verbal meaning. In

example (68) below, when the suffix *-an* appears on the perception verb *téang* ‘to see (s.o.), look for (s.t.)’, the clause may take a durative or iterative reading, i.e. ‘looking for (s.t.) for some time, looking repeatedly for (s.t.)’.

(68) Sundanese, Pluractional stimulus-selecting construction

- a. *Abi keur n-(t)éang HP.*
 1SG PROG AV-look.for mobile.phone
 ‘I am looking for my mobile phone.’ (BC)
- b. *...lamun awak=na nu leumpang n-(t)eang-an ka-dahar-an.*
 while body=3 REL walk AV-look-PLUR NMLZ-eat-NMLZ
 ‘[the supernatural being was struck by someone] while its body was walking **looking around for food.**’ (AC) (CT1-019, based on MPI-SUN 18295)

With verbs of emotion, suffixation with *-an* or *-keun* may mark a stimulus-selecting construction. Examples are shown below in (69) and (70). Note that the verbs meaning ‘to cry’ and ‘to laugh’ also appear as reduplicated forms suggesting intensity or iterativity in these examples.

(69) Sundanese, Stimulus-selecting construction

- a. *Mariam ceurik lantaran indung=na maot.*
 M. cry because mother=3 die
 ‘Mariam cried because her mother died.’ (BC)
- b. *Mariam ny-(c)eung-ceurik-an indung=na.*
 M. AV-RDP-cry-STIM.APPL mother=3SG.POSS
 ‘Mariam cried intensely about her mother.’ (AC) (CT1-006, based on Hanafi 1997: 22)

(70) Sundanese, Stimulus-selecting construction

- a. *Abi keur seuri.*
 1SG PROG smile
 ‘I am smiling.’ (BC)
- b. *Abi keur seu-seuri-an.*
 1SG PROG RDP-smile-PLUR.
 ‘I am laughing.’
- c. *Ujang ny-(s)eung-seuri-keun Asep.*
 U. AV-RDP-laugh-STIM.APPL A.
 ‘Ujang laughed at Asep.’ (AC) (CT1-002)

With certain verbs denoting events of cognition, both *-keun* and *-an* are found to mark content-selecting constructions, as shown in (71) below. Both the construction marked with *-keun* in (71b) and the construction marked with *-an* in (71c) select the content role to map to P. The verb *pikiran* means ‘to reflect on (s.t.), to bear (s.t.) in mind’, while *pikirkeun* means ‘to think about (s.t.), think (s.t.) over’ (Hardjadibrata 2003: 626). These are quite close in meaning, but the former might be more durative.

(71) Sundanese, Content-selecting constructions

- a. *Abi keur m-(p)ikir.*
1SG PROG AV-think
'I am thinking.' (BC)
- b. *Abi keur m-(p)ikir-keun biaya sakola budak.*
1SG PROG AV-think-CONT.APPL fee school child
'I am thinking about the children's school fees.' (AC)
- c. *Abi keur m-(p)ikir-an hutang.*
1SG PROG AV-think-CONT.APPL debt
'I am reflecting on (my) debt.' (AC) (CT1-003)

In summary, in this section we have seen a variety of constructions marked with *-an* and *-keun* in which a peripheral semantic role is selected to map to a core argument. In one set of constructions marked with *-keun* the possible selected roles include the content of a communication event, the performance in a performance arts event, and the stimulus in caused perception events of various other types. In one set of constructions marked with *-an*, the possible selected roles include the addressee of a communication event, the target or intended perceiver of a gesture, and the perceiver in caused perception events of other types. These two sets of constructions both describe events in which one participant causes (or intends to cause) another participant to perceive some stimulus, though they differ in the suffixal form marked on the verb and mapping of participants observed. Aside from these sets, we also observe that both *-keun* and *-an* may be associated with the mapping of a core argument to the stimulus role of an experiential emotion event, and both are likewise associated with the mapping of a core argument position to the content role of a cognition event. These constructions describe the internal experiences of animate entities. The AM-marked constructions described in this section are almost always monotransitive, though some also exhibit an oblique PP that operates like an argument, which might be referred to as *extended monotransitive* following Dixon & Aikenvald's (2000) use of *extended intransitive* and *extended transitive*. For most verbal roots that occur with AM-marked constructions found in this section, the corresponding BCs are intransitive (or semi-transitive), and the addition the AM is valency-increasing. One exception is the set of transitive perception verbs that take a perceiver as A, and a stimulus as P in BCs. For such roots, the AM-marked clause generally is valency-preserving and retain mapping of the stimulus role to P, while also taking on a causative or pluractional meaning.

2.10 Other AM-marked constructions in Sundanese

2.10.1 Causative constructions

The suffixes *-an* and *-keun* both mark causative constructions in which an instigating agent is added to the argument structure and selected to map to A. Causative constructions with such AM-marking are quite productive.

Example (72) shows a construction marked with *-keun* on a stative base. The BC is intransitive, and the entity that exists in the state described by the root *pegat* 'snapped, broken off, cut off'

(or enters this state) is the sole core argument, S. In the AC, the entity now is expressed as the P argument, while an instigating causer agent maps to A. A similar example is given in (73) with the suffix *-an* and the root *kiruh* ‘turbid, muddy’.

(72) Sundanese, Causative construction with stative base

- a. *Tali pegat.*
 rope snap
 ‘The rope snapped. / The rope is broken.’
- b. *Tali seuseuh-eun ka-pegat-keun ku Bapa.*
 rope wash-NMLZ NVOL-snap-CAUS by father
 ‘The clothesline was caused to snap by Father on accident.’ (CT1-018)

(73) Sundanese, Causative construction with stative base

- a. *Cai éta kiruh.*
 water that turbid
 ‘The water is turbid.’
- b. *Kuring ng-(k)iruh-an cai éta.*
 1SG AV-turbid-CAUS water that
 ‘I made the water turbid.’ (Hanafi 1997: 18)

Example (74) show a construction marked with *-keun* on an intransitive root. The BC is intransitive, and the entity that engages in the activity described by the root *lumpat* ‘to run’ is sole core argument, S. In the AC, the entity that engages in the activity is expressed as the P argument, while an instigating causer agent maps to A. A similar example is given in (75) with the suffix *-an* and the root *dangdan* ‘to dress up, get ready’.

(74) Sundanese, Causative construction with dynamic intransitive base

- a. *Asep lumpat ka sakolah.*
 A. run to school
 ‘Asep ran to school.’ (FM4-032)
- b. *Budi nga-lumpat-keun motor.*
 B. AV-run-CAUS motorbike
 ‘Budi made the motorbike go fast (lit. run).’ (CT1-003)

(75) Sundanese, Causative construction with dynamic intransitive base

- a. *Enéng keur dangdan.*
 E. PROG dress
 ‘Eneng dresses up.’
- b. *Kuring nga-dangdan-an Enéng.*
 1SG AV-dress-CAUS E.
 ‘I made Eneng get dressed.’ (Hanafi 1997: 23)

With a number of roots, both *-an* and *-keun* mark causative constructions. An example is given in (76) below with the root *tiis* ‘cool’. Note that there are subtle semantic differences in the meanings of the affixed verbs; *tiis-keun* means ‘to let (s.t.) cool down’ while *tiis-an* means ‘to make (s.t.) cool’ and indicates more active intervention on the part of the agent.

(76) Sundanese, Causative construction with *tiis* ‘cool’

- a. *Ntéh geus tiis.*
tea already cool
‘The tea is already cool.’
- b. *Abi n-(t)iiis-keun ntéh.*
1SG AV-cool-CAUS tea
‘I let the tea cool down (e.g. by setting it aside before use).’
- c. *Icih n-(t)iiis-an cai maké és.*
I AV-cool-CAUS water using ice
‘Icih cooled the water using ice.’

(CT1-002)

2.10.2 Pluractional aspect

In addition to marking causative and ACs, the suffix *-an* often indicates pluractional aspect. Verbs marked with *-an* frequently describe iterative, durative, or habitual action, and events with distributed action and/or multiple P participants. The pluractional meaning of *-an* is highly productive, while the causative and applicative functions of *-an* are more restricted. Thus, with a large number of roots, *-an* freely occurs with no change in argument structure. Examples of this type are given below.¹³ As reflected in the free translations for these examples, in the clauses shown below without *-an*, the event is viewed as a singular whole. In the clauses marked with *-an*, the event may be indicated to be repeated, as in (77b), to involve many participants, as in (78b), or to be habitual, as in (79b).

(77) Sundanese, Pluractional aspect

- a. *Drona m-(p)énta bantu-an ti ba-batur-an=ana.*
D. AV-ask.for help-NMLZ from RDP-friend-NMLZ=3
‘Drona asked for help from his friends.’
- b. *Abi m-(p)énta-an bantos-an ti pamaréntah.*
1SG AV-ask.for-PLUR help-NMLZ from government
‘I asked for help from the government a number of times.’ (Pluractional)

(CT1-016)

¹³In example (77) the root *bantu* is the standard Sundanese word meaning ‘help’, and is neutral in speech register, while *bantos* is a non-standard variant used by some younger speakers as the equivalent of *bantu* in the *lemes* (‘refined’) register.

(78) Sundanese, Pluractional aspect

a. *Budak di-béré buku ku abi.*

child PV-give book by 1SG

‘The child was given a book by me.’

b. *Budak nu sunat-an di-béré-an co-coo-an.*

child REL circumcision-NMLZ PV-give-PLUR NMLZ-play-NMLZ

‘The child who was circumcised was given many toys.’ (Pluractional) (CT1-026)

(79) Sundanese, Pluractional aspect

a. *Abi m-(b)éré buku ka budak.*

1SG AV-give book to child

‘I gave the book to the child.’

b. *Pak Haji sok m-(b)éré-an duit ka Lilis.*

mister haji often AV-give-PLUR money to L.

‘Pak Haji often gives money to Lilis.’ (Pluractional) (CT1-026)

2.10.3 Marking with multiple AMs

In a benefactive AC marked with *pang-* *-keun* the verb stem may consist of a root plus *-an*. In such cases, the resulting clause selects a beneficiary as an argument and in addition has a pluractional, causative, and/ or locative applicative meaning. Marking with multiple AMs in this manner does not appear to be used frequently in natural discourse but this process is fairly productive as speakers will readily generate such forms over a good number of lexical bases when asked.

(80) Sundanese, Multiple AM-marking

Abi di-pang-neundeun-an-keun balanjaan di dapur ku Udi.

1SG PV-BEN.APPL-place-PLUR-BEN.APPL in kitchen by U.

‘I had many shopping purchases placed in the kitchen for me by Udi.’ (CT1-022)

(81) Sundanese, Multiple AM-marking

Udi m-pang-eusi-an-keun ember maké jambu.

U. AV-BEN.APPL-full-CAUS.LOC.APPL-BEN.APPL in kitchen by

‘Udi filled up the bucket with guavas (for me).’ (CT1-024)

(82) Sundanese, Multiple AM-marking

Tipi di-pang-maéh-an-keun ku Bapa.

television PV-BEN.APPL-dead-CAUS-BEN.APPL by father

‘The TV was turned off (for me) by Father.’ (CT1-003)

2.10.4 Comparatives and other degree constructions

With stative roots expressing gradable attributes, suffixation with *-an* may mark a comparative degree construction. As shown in the examples below, the root suffixed with *-an* is the predicate, and is followed by an unmarked NP expressing an entity. A phrase headed by *tibatan* (or its short form *batan*) ‘than’ expresses the standard to which the entity is compared.

(83) Sundanese, Comparative degree construction

- a. *Jangkung-an adi=na tibatan lanceuk=na.*
 tall-COMP younger.sibling=3 than older.sibling=3
 ‘The younger sibling is taller than his/her older sibling.’
- b. *Baguer-an Pak RT tibatan palura.*
 kind-COMP mister neighborhood than village.head
 ‘The neighborhood officer is kinder than the village head.’
- c. *Bodas-an baju ieu tibatan éta.*
 white-COMP clothing this than that
 ‘This clothing is more white than that (clothing).’ (CT1-044)

This use of *-an* shows semantic similarities similar to causative constructions where *-an* is suffixed on stative roots, e.g. the use of *niisan* ‘to cool (s.t.)’ in example (76c) above. Furthermore, there is a related intransitive construction in which *-an* appears on stative roots. This is shown in the following examples from Kurniawan (2013: 31).

(84) Sundanese, Comparative degree achievement construction

- a. *Orok téh nga-lintuh-an.*
 baby PRT AV-fat-COMP
 ‘The baby gets fatter.’
- b. *Manéhna rada nga-jangkung-an.*
 3SG rather AV-tall-COMP
 ‘She gets a bit taller.’ (Kurniawan 2013: 31)

However, note that not all stative roots in Sundanese require affixation with *-an* to be used with inchoative meanings. For instance, in example (72) above, the stative root *pegat* operates as a verbal predicate and can have a stative meaning e.g. ‘The rope is **broken**,’ or an inchoative meaning, e.g. ‘The rope **snapped**.’ On top of this, the sentences in (84) have meanings that indicate a greater degree of an attribute, rather than simply the beginning of a state. For this reason, I consider clauses like those in (84) to constitute a comparative degree achievement construction (see Hohaus & Bochnak 2020 on degree achievement verbs as part of a typology of degree constructions) and I have used the gloss comp for ‘comparative’. So while Kurniawan glosses *-an* in clauses like (84) as ‘inchoative’, I use ‘comparative’.

2.11 Summary and discussion

In this chapter, I have described and exemplified the applicative system of Sundanese, using the definition of ACs outlined in the introductory chapter. These include: (i) ACs marked with *-an* in which the peripheral role selected is a location, goal, or path, (ii) ACs marked with *-keun* in which the peripheral role selected is a theme or instrument, and (iii) ACs marked with *pang-* *-keun* (or rarely, *-keun*) in which the peripheral role selected is a beneficiary. Some ACs marked with *-an* or *-keun* in which the selected role is an addressee, the content of a communication or cognition event, a stimulus, or a type of performance were also presented. Aside from this, I have also described a number of other clausal constructions that are not applicatives, but are marked with the same three morphemes. In Sundanese, causative constructions in which an instigating causer is introduced as an A argument may be marked with any of the three AMs, while constructions with pluractional aspectual meanings and comparative degree constructions are marked with *-an*. In Sundanese, the AMs *-an* and *-keun* are also used with the category-changing function, that is, the formation of verbal stems from non-verbal roots, and some examples of these were also presented for *-an*.

Throughout the chapter, I have endeavored to show a number of advantages to adopting a constructional approach in analyzing the applicative system of Sundanese. For emphasis, some of these are repeated here with, using *-an*-marked constructions for illustration.

One advantage of adopting a constructional approach lies in the fact that an AC may be defined on the basis of its observed form and meaning alone. This allows ACs that are similar in meaningful ways—that is, having shared formal properties, and/or shared semantic properties—to be classified alike, regardless of whether a corresponding BC exists, or what structure such a BC takes. To illustrate, goal-selecting applicatives marked with *-an*, are listed for four possible bases in Table 2.3, along with representations of the argument structure in the form of a *frame* for the AC and BC (if one exists). Full clausal examples for these were given above in §2.4.2, §2.4.4, and §2.6.

Table 2.3: Sundanese goal-selecting applicative constructions compared

Base	AC Frame	BC Frame	Notes
<i>teundeun</i> 'to place'	[GOAL] <i>di</i> -[X]- <i>an</i> [THM] <i>ku</i> [AGT] 'place at (s.w.) (many items)'	[THM] <i>di</i> -[X] <i>ka</i> [GOAL] <i>ku</i> [AGT] 'place (s.t.) to s.w.'	[+pluractional]
<i>alung</i> 'to throw'	[GOAL] <i>di</i> -[X]- <i>an</i> [THM] <i>ku</i> [AGT] 'throw at (s.t.) (s.t.)'	[THM] <i>di</i> -[X] <i>ka</i> [GOAL] <i>ku</i> [AGT] 'throw (s.t.) to s.w.'	*BC uncommon
<i>anclom</i> 'touch water'	[GOAL] <i>di</i> -[X]- <i>an</i> [THM] <i>ku</i> [AGT] 'put into (some liquid) (s.t.)'	[THM] [X] 'touch water'	[+causative]
<i>uyah</i> 'salt (n.)'	[GOAL] <i>di</i> -[X]- <i>an</i> <i>ku</i> [AGT] 'put salt on (s.t.)'	(none)	[+cat.-changing]

We can identify the four ACs shown in the table as examples of one generalized construction in which different lexical bases may fill the verbal slot. The consistent elements of form include the appearance of the suffix *-an* on the verbal predicate, an argument structure with (maximally) three arguments, and in this particular set of examples, the occurrence of the PV prefix *di-* on the verb. A consistent meaning which may be generalized across bases also is observed. This

includes the mapping of a goal role to the preverbal P argument, and a clausal meaning that may be expressed as “an agent causes motion of a theme to a goal”, or in short “caused motion”.

If we were to define an AC in terms of its correspondence to—or derivation from—a BC, the examples with *uyuh-an* ‘to put salt on (s.t.)’ and *alung-an* ‘to throw at (s.t.) (s.t.)’ would be problematic, as the former cannot be used as a verbal predicate without *-an*, and the latter is not used as a verbal predicate without *-an* or *-keun* by some present-day speakers.

Furthermore, we observe in these four examples that the non-applicative meanings or functions of the AM *-an* are not distinct from its applicative functions. Thus, while *-an* sometimes indicates pluractional aspect without selection of a peripheral role as a core argument, as in examples with *béré-an* ‘to give many items’ or ‘to give habitually’ in §2.10.2 above, in the case of *teundeun-an* ‘to place at (s.w.) (many items)’ the meaning is pluractional and at the same time the argument structure is consistent with that of an applicative. Likewise, *-an* sometimes forms causative constructions in which an instigating causer is introduced as A, while P is a patientive causee as in examples with *dangdan-an* ‘to make (s.o.) get dressed’ in §2.10.1 above. Still, in the case of *anclom-an* ‘to touch to (some liquid) (s.t.)’, the function of *-an* and the argument structure observed with this verb are consistent both with that of a causative construction, and with that of an applicative construction. In Sundanese, we do not find strictly causative nor strictly applicative verbal constructions with the base *anclom*.

As stated in the introductory chapter, one important goal of this study is to describe the function and usage of AMs in the languages of West Nusantara and to understand them on their own terms. Therefore, I find it important to acknowledge, and at times center, the polyfunctional nature of morphemes like Sundanese *-an*, *-keun*, and *pang-keun*. This cannot be done effectively if non-applicative functions (e.g. aspectual meanings, the formation of causative constructions, category-changing functions) are treated as entirely separate and unrelated to applicative constructions with the same morphological marking. To capture meaningful similarities in the syntactic and semantic properties observed across various types of clauses in which these affixes appear on the verb, these various functions should be considered together, especially since the convergence of such functions on one form of morphological marking (i.e., an AM) is pervasive in the languages of this region (see §4.6.4 for more details).

To this end, the adoption of a constructional approach allows all these types of clauses to be meaningfully classified and related together through a constructional network. For Sundanese we might for example define a larger construction type of [Goal-selecting Applicative], of which there are various subtypes, including [Caused Motion Goal-selecting Applicative] and [Self-Motion Goal-selecting Applicative], each attracting bases according to the compatibility of their lexical semantics. Some clausal examples of the caused motion subtype may also be said to belong to another [Morphological Causative] construction type, e.g. *anclom-an*, ‘to put into (some liquid) (s.t.)’. Some examples of either the caused motion subtype or the self-motion subtype may belong to another [Pluractional Aspect] construction type, e.g. *teundeun-an* ‘to place to (s.w.) (many things)’ and *anjang-an* ‘to visit (many places)’, respectively.

While I do not attempt to formally represent such a constructional network, the functional typology of applicative constructions and other AM-marked constructions that I have laid out in this chapter (and will further develop in Chapter 7) lends itself neatly to this type of categorization, while also allowing for inclusion of clausal constructions that would be considered problematic under other approaches to applicatives, due to the nature of their relationships with corresponding BCs (or lack thereof). In Sundanese, such problematic examples occur with non-

trivial frequency. Thus, having explored the Sundanese applicative system as a featured case study in this chapter and identified some advantages to the use of a constructional approach in its analysis, in the following chapter, I examine some of the same issues at a much broader level. In it, I consider at length other common approaches to applicatives in the literature and constructions which are problematic for these, as found in other language families and other languages of West Nusantara.

Chapter 3

Approaches to applicatives in the literature and this study

In this chapter, my goals are to situate the present study in the context of previous research and approaches to applicatives in the linguistic literature, and to explain the approach to applicatives that I have taken in this study. In §3.1, I discuss definitions of applicatives that have been used by linguists, especially in functional and typological linguistic research, exploring commonalities in these definitions, as well as differing understandings of applicatives which are found. In §3.2, I present a number of attested constructions that have proved problematic for certain commonly adopted approaches to applicatives, first looking broadly at examples from diverse languages families around the world, and then more narrowly at examples from previous research on languages of West Nusantara. In §3.3, I discuss how these types of problematic constructions might be better accounted for with adoption of a constructional approach to applicatives, and explain how I will do this in this study. In §3.4, I discuss one particular challenge in developing a typology of applicative constructions (ACs) in West Nusantara languages, namely, the Philippine-type voice system and voice alternations in Philippine-type languages that have proved difficult to classify in typological studies of grammatical voice and applicatives. In §3.5, I discuss the relationship between applicatives and serial verb constructions, outlining the criteria by which I distinguish the two. Serial verb constructions are relevant to a study of applicatives in West Nusantara languages (and Austronesian languages more broadly) because of their similar functional and formal properties. In §3.6, I conclude the chapter by demonstrating how ACs and other constructions marked with applicative morphemes (AMs) in West Nusantara languages may be situated in a broader typology of verbal constructions in the framework that I use in this study.

3.1 What are applicatives?

3.1.1 Origins of the term *applicative*

The term *applicative* comes from the Latin *verbos applicativos*, ‘applicative verbs,’ which was first used by Jesuit missionaries in descriptions of benefactive and malafactive constructions in Nahuatl and other Uto-Aztecan languages beginning in the 1500s (see del Rincón 1885 and Carochi 1981). In Bantu languages, similar verbal constructions have been described in grammars dating

to the 15th century, but only in the 1800s did the terms *applied form* or *applicative* come into use (see Pacchiarotti 2020). The traditions of research in these two language families have most influenced the study and definition of applicatives.

More recently, the term applicative has been applied to constructions found in many other languages and language families. Besides Uto-Aztecan and Bantu, Polinsky (2013), based on a survey of more than 180 languages, identifies Salishan, Mayan, and Austronesian as language families where applicatives are commonly found. Peterson (2007) identifies languages which have applicatives in a wide variety of genetic groupings and geographic areas, spanning the Americas, Africa, Australia, South and Southeast Asia, and Papua New Guinea—but only two languages of Eurasia, where he reports applicatives are rare. As Dixon (2012: 334–335) points out, what are now called applicatives have been described using a variety of other labels, and while the term applicative is in widespread use today, it is not used uniformly by all authors.

3.1.2 Commonalities in definitions of applicatives

Despite this lack of uniformity, examination of commonly-accepted definitions of applicatives in the literature reveals some major areas where authors are in agreement. In this section, I aim to describe these commonalities in terms which are broadly inclusive.

First, the term applicative describes a type of clause-level or predicate-level construction (i.e., an AC). It appears to be the mainstream view, particularly in functional and typological approaches, that ACs are marked by morphology or morphological processes on the verbal complex.¹ The morphemes that signal these constructions are also commonly called applicatives.

Second, in addition to morphological marking, ACs are characterized by the presence of a phrase that shows marked properties, which I will refer to here as the *applied argument* or *applied phrase*.² The AC shows unusual mapping of roles to argument structure, and the applied argument is often said to be “introduced,” “added,” or “allowed” in some fashion in this structure. Most authors agree that the applied argument in an AC must represent a different grammatical relation than that characteristic of A, where A is the most agent-like argument in a basic transitive clause. In many typologies of voice and valency, this is in explicit contrast to causative constructions, which are defined by the introduction of an A argument (or a similar grammatical category) (e.g. Haspelmath & Müller-Bardey 2004; Dixon & Aikhenvald 2000).

Third, definitions commonly point to differences in the argument structure of ACs as compared to corresponding verbal constructions that lack the same distinctive morphological marking, that is, base constructions (BCs).³ In particular, two observations have been made about

¹However, many authors working in generative grammar consider unmarked ditransitive alternations, such as English ‘dative shift’ to also constitute applicatives (e.g. Baker 1988a). In generative frameworks, applicatives are commonly understood as a syntactic element that introduces a non-core argument, and this element may be overt or non-overt (see McGinnis 2008). In the typological literature, a few authors do not distinguish between unmarked and formally marked verbal constructions with similar functions, but this is a minority view for applicatives. In practice, unmarked verbal constructions are not commonly identified as examples of applicatives, even by authors who would allow this in principle (e.g. Kulikov & Song 2010).

²Following Zuñiga & Creissels (2024), I often employ the term *applied phrase* because it allows for inclusion of clausal constituents expressing a peripheral semantic role, for which status as a clausal argument is unclear.

³Authors may describe such differences in reference to a “basic construction” (Polinsky 2013), “underlying clause” (Dixon 2012; Dixon & Aikhenvald 2000; Aikhenvald & Dixon 2011), “base predicate” (Zuñiga & Kittilä 2019), “base verb” (Bresnan & Moshi 1990), “original verb” (Payne 1997), and similar. A few make reference to non-applicative

ACs: (i) the applied argument is syntactically similar to core arguments in BCs, and (ii) the applied argument shows semantic roles similar to non-core arguments in BCs. Authors typically describe the syntactic similarities in terms of syntactic coding and grammatical relations, and the applied argument is most commonly equated or likened to P, the most patient-like argument in a transitive construction. Authors commonly describe possible roles for the applied argument using lists or by equating them to roles characteristic of obliques or adjuncts in BCs.

To illustrate, let us consider the definition used in Peterson (2007), the most comprehensive systematic cross-linguistic study of applicatives to date. Peterson characterizes applicatives by stating: “Applicative constructions are a means some languages have for structuring clauses which allow the coding of a thematically peripheral argument or adjunct as a core argument. Such constructions are signalled by overt morphology” (2007: 1). He goes on to define ACs as “sentential structures which involve a participant that normally would not be instantiated in a core object relation, but rather as an oblique of one or another sort, in a core (usually direct object) instantiation” and stipulates that applicatives must be “highly productive across a significant portion of the verbal lexicon (all verbs, all transitive verbs, etc.)” (2007: 39). Peterson’s definition of ACs adheres closely to the commonalities I have outlined above, with the exception that he explicitly addresses productivity, while other authors in general do not.

Table 3.1 below summarizes a number of definitions of applicatives which are widely cited and commonly-accepted in the linguistic literature on this topic. Columns in the table from left to right indicate: (i) source citations, (ii) whether the authors consider overt morphological marking to be characteristic (or essential) for applicatives, (iii) wording used to characterize the relationship of the applicative to a special argument in the clause, i.e. the applied argument, (iv) wording used to describe the syntactic properties of the applied argument, and (v) descriptions used to characterize the semantic properties of the applied argument. Note that some authors specify semantic properties of the applied argument directly, with lists or labels for groups of semantic roles. Others do so indirectly, by equating possible participant roles for the applied argument with participants encoded as non-core arguments in BCs, or by excluding participants encoded as core arguments in the same. Such indirect specifications of semantic properties are listed with a preceding [=] symbol in the rightmost column of Table 3.1.

Here, I observe that if we look at both the definitions represented in Table 3.1 and the data used to exemplify them in the same sources, there is widespread agreement that possible roles for the applied argument include beneficiary, recipient, goal, instrument, and location, and usually also comitative (also called associate) and purpose (also called motive). On the other hand, agent is almost never included among possible or prototypical roles for the applied argument, and patient is sometimes included but very uncommonly. A few definitions (notably Peterson 2007, Zúñiga & Kittilä 2019, and perhaps Comrie 1985a) appear to allow patient as a role for the applied argument, at least in principle, provided that a corresponding patient participant is normally realized as a non-core argument in a corresponding BC.

In summary, definitions of applicatives in the literature share important commonalities. On this basis, we can describe applicatives in terms of a consensus definition as follows:

constructions by describing how arguments are “otherwise” (Baker 1988b; Creissels 2014) or “normally” realized (Peterson 2007). Some claim or imply that the applicative clause or predicate has a formal derivational relationship to a basic clause or predicate. Here, I use ‘base construction’ as a neutral, descriptive term and make no such claim.

Table 3.1: Commonly-accepted definitions of applicatives

Source	Morph. Mark.	Rel. to Applied Arg.	Syntactic Properties	Semantic Properties
Baker (1988b, 1988a)	Not req'd	may “assign Case” to argument	“looks like” direct object	= “otherwise” oblique
Alsina & Mchombo (1990; 1993)	Yes	“introduces” argument	is an “internal argument”	= “expressed as obliques”, “optional”
Bresnan & Moshi (1990; 1993)	Yes	“introduces” argument	is an “object argument”	= “expressed as oblique, if at all”
Dixon & Aikhenvald (2000)	Yes, prototypically	“adds” argument, which is “taken into the core”	is “in O function”	= “peripheral argument/function”
Mithun (2001)	Yes	“adds” to core arguments	role of “object, absolutive, or grammatical patient”	lists usual roles
Peterson (2007)	Yes	“allows encoding of” argument	“in a core (usually direct object) instantiation”	= “normally” oblique, “peripheral”
Polinsky (2013)	Yes, customarily	“increases” no. of arguments	is an “object”	= different than “base object”
Haspelmath & Müller-Bardey (2004)	Yes	argument “is supplied”	“status of a direct object”	“oblique roles”
Payne (1997)	Yes	“brings onto center stage”	is a “direct object”	“peripheral roles”
Foley (2007)	Yes	“adds” argument	“[-oblique] [-A]”	= “erstwhile [+ oblique] NPs”
McGinnis (2008)	Not req'd	“adding” argument	is an “object”	lists roles
Kulikov & Song (2010)	Not req'd	“adding” argument	is a “direct object”	“non-core semantic relations”
Creissels (2010; 2014)	Yes	may “promote,” is “way to encode” argument	“identical or similar” to patient of transitive	= otherwise adjuncts, “cannot be core”
Zúñiga & Kittilä (2019)	Yes, prototypically	“allows” predicate to take argument	“corresponds to primary/direct object”	= different from “base predicate”

(85) Consensus definition of applicatives

The term applicative describes clausal constructions in which overt morphological marking on the verb coincides with differences in argument structure. In such clauses, an applied argument shows syntactic encoding and properties similar to core arguments in corresponding clauses that lack such morphological marking, especially P in a basic transitive construction. Despite this, possible semantic roles found for the applied argument are similar, not to roles found for core arguments, but those found for non-core arguments, such as obliques and adjuncts, in corresponding unmarked clauses. As such, an applied argument typically has the role of a non-agent and non-patient participant, such as a beneficiary, recipient, goal, or instrument, among others.

3.1.3 Differences in definitions of applicatives: What do applicatives do?

While it is clear that ACs differ in argument structure from patterns observed for corresponding BCs, we have already begun to see that authors vary in the way that they frame these patterns, and in the criteria used to identify the relevant differences. These in turn reflect key differences in authors' conceptual understanding of applicatives.

Peterson (1999; 2007) has addressed parameters of variation among applicatives at length, though he focuses on *constructions* themselves, rather than on *definitions*. He identifies five parameters describing this variation. Pacchiarotti (2020) builds on this work by identifying how various definitions of applicatives in the literature address Peterson's parameters. In this section, I will refer to Peterson's parameters and a number of related concepts which are discussed by both these authors. But to describe how definitions of applicatives vary from each other, I organize my presentation of this material around the question of function, and one particular sense of function, i.e. "What do applicatives do?" I identify five conceptual approaches to this question, which are not necessarily mutually-exclusive with one another.

3.1.3.1 Approach 1: Applicatives as argument-promoting

Some definitions frame applicatives as a device that functions to promote an argument from the periphery of a clause into the core. This approach is related to Peterson's (1999: 41–45) parameter, "Optionality/obligatoriness of the construction." In this view, an applicative is primarily defined by a relationship between two clausal constructions that shows particular correspondences. The main verb in both constructions represents the same stem, and is considered underived in the BC, and morphologically derived in the AC. Between the two clauses, the following criteria generally apply, (i) a participant is realized as a syntactically oblique argument in the BC and (ii) the same participant is realized as a core syntactic argument, usually P, in the AC. This relationship may be thought of as a transformation that applies to the "underlying" BC, producing the AC as its output. Schema or diagrams of clause structure are often used to represent this process. Definitions of applicatives that reflect this approach include Dixon (2012); Aikhenvald & Dixon (2011); and Foley (2007). In the literature on applicatives, the term 'optional' is used if speakers of a language have a choice between two constructions for expressing a particular participant role—as an oblique in a BC or as a core argument in an AC. In definitions which prioritize the argument-promoting function, only optional constructions are considered applicatives.

3.1.3.2 Approach 2: Applicatives as argument-introducing

In contrast to Approach 1, some definitions frame the function of applicatives as the introduction of an argument to the argument structure of a clause or predicate. This view is broader with respect to the parameter of “Optionality/obligatoriness”. Definitions that reflect this approach include Alsina & Mchombo (1990; 1993), Bresnan and Mochi (1990; 1993), and Kulikov & Song (2010). In the literature on applicatives, the term ‘obligatory’ means that the AC is the only means in a language to express a particular participant role in a monoclausal construction describing a verbal event. Authors who view applicatives as argument-introducing typically include both optional and obligatory ACs. But to distinguish applicatives from causatives, and sometimes also other verbal transitivizers, authors typically stipulate additional criteria related to syntactic properties or semantic roles of the introduced argument. Because of this, many definitions feature a combination of argument-introducing view with another approach, especially the syntactic licensing function (Approach 4) or the role-mapping function (Approach 5) as described below.

3.1.3.3 Approach 3: Applicatives as valency-increasing

Some definitions frame applicatives as a morphological process that primarily functions to increase the valency of a clause. This view is related to Peterson’s (1999: 46–55) parameter, “Treatment of base and applicative objects”. This approach is especially reflected in Polinsky (2013), which states, “In an applicative construction, the number of object arguments selected by the predicate is increased by one with respect to the basic construction”, and “[a]n applicative construction is a particular instance of a double object construction.” Like Approach 1, Approach 3 relies on comparison of a BC with an AC. Here the primary criterion is an increase of one in the value of syntactic valency over the BC. As seen in Polinsky (2013), there might be no explicit reference to participant role or correspondence between arguments in the two clauses. Applied strictly, definitions that prioritize an increase in valency are narrow with respect to “Treatment of base and applicative arguments”. Under such definitions, a core argument must be added in an AC, and none of the original core arguments of the BC should be suppressed or “demoted” to non-core status.

In practice, however, the criterion of increased valency is rarely applied strictly for all types of BCs. In some typologies, applicatives are clearly considered a type of valency-increasing construction but certain ACs showing no change in the value of syntactic valency over corresponding BCs are also allowed to be classified as applicatives. For example, Dixon & Aikhenvald (2000: 12–14) describe applicatives and causatives as “two valency increasing derivations.” They go on to describe two schema for applicatives split by transitivity of the BC. An AC derived from an intransitive BC always expresses an O argument⁴ where there was none before. In an AC derived from a transitive BC, however, “the argument which was in O function is moved out of the core into the periphery of the clause (and may be omitted)”. Under this type of definition, applicatives only necessarily increase syntactic valency with intransitive verbal bases (see also Haspelmath & Müller-Bardey 2004). Payne (1997) addresses observed differences in effect on syntactic valency explicitly. He writes, “In most cases, an applicative can be insightfully be described as a valence increasing operation.... For verbs that already have one direct object, the applicative

⁴The terms P argument and O argument are equivalent, with the former following Comrie (1989), and the latter following Dixon (1972).

either results in a three-argument (ditransitive) verb, or the ‘original’ direct object ceases to be expressed. In the latter case, the applicative cannot be considered a valence-increasing device, since the original and the resulting verb have the same number of arguments; rather the applicative simply ascribes a new, formerly, peripheral semantic role to the direct object” (186-187). In summary, while applicatives are framed as a valency-increasing device in a number of definitions, the valency-increasing function might warrant qualification. An increase in valency might be observed only with verbal bases of lower transitivity values (e.g. intransitive bases only, or intransitive and monotransitive bases only but not ditransitives). This in turn might be explained by limits on the maximum transitivity value for verbal constructions; it appears that in some languages verbal clauses are maximally monotransitive, but in others they are maximally ditransitive, or may not obey a strict maximum. These two factors—transitivity of the base and maximum clausal transitivity—are described in greater detail in Peterson’s (1999: 56–60) discussion of his parameter labelled “Transitivity restrictions”.

3.1.3.4 Approach 4: Applicatives as syntactic licensing

Some definitions frame applicatives as devices which license or allow an argument to appear in a particular syntactic realization, as reflected in syntactic coding, and/or access to syntactic operations or structural position. Like Approach 3, this view relates to the parameter, “Treatment of the base and applicative arguments”. However, definitions which prioritize syntactic licensing of one argument are in general broader, because they do not require that the participant expressed by P in the BC retain its original syntactic status when the same role is expressed in the AC. I refer to the constituent expressing this participant in the AC as the *companion phrase* because it is a counterpart of the applied phrase, and may not itself be encoded as a clausal argument (see McDonnell & Truong 2024).

There is some variation in what exactly applicatives are said to license in this type of approach. In generative grammar, applicatives are often said to allow a non-core argument to move in particular ways in syntactic structure. Coding and syntactic properties of the applied phrase and companion phrase are explained as the result of this movement (see Pylkkänen 2002; McGinnis 2008). Peterson (2007: 1) states that applicatives “allow the encoding of a thematically peripheral argument or adjunct as a core argument.” Creissels (2004: 6) writes, “In its canonical use, the applicative suffix licenses the presence of a direct object with a semantic role that the same verb devoid of the applicative suffix cannot assign to a direct object.” In a similar fashion, Zúñiga & Kittilä (2019: 53) characterize applicatives as operations that “allow” the predicate “to take a direct or primary object... bearing a semantic role different from the one the base predicate would normally take (if any).” As seen in these last three examples, views that prioritize the licensing function of applicatives commonly also make reference to the possible semantic or participant roles of the licensed argument. Thus it is common to combine this approach with the view of applicatives as role-mapping (Approach 5), which will be described next.

3.1.3.5 Approach 5: Applicatives as role-mapping

Sometimes the function of applicatives is framed as selecting or signalling the mapping of semantic roles to particular arguments in the clause. For example, Payne (1997: 186–187) introduces applicatives by stating, “Some languages have operations whereby a verb is marked for the se-

semantic role of a direct object.” Later, regarding non-valency-increasing ACs (as discussed above), he also writes, “the applicative simply ascribes a new, formerly, peripheral, semantic role to the direct object.” Also, a number of definitions discussed under Approach 4 above stipulate that the licensed argument in an AC bears a different semantic role than the argument that bears the same grammatical relation in the BC (e.g. Creissels 2010; Zúñiga & Kittilä 2019). Thus, a change in mapping of semantic role is implicit in these definitions.

Admittedly, selecting or signalling the mapping of semantic roles does not appear to be commonly identified as the single primary function of applicatives. Still, a description of possible semantic roles is quite commonly included in definitions that emphasize the argument-introducing (Approach 2) or syntactic licensing (Approach 4) functions. For example, Pacchiarotti & Zúñiga (2022: 1) hold that one of three “fundamental attributes” of applicatives is that “there are multiple, typically ‘peripheral’ semantic roles that can be mapped onto the applied phrase (e.g. Beneficiary, Instrument, Location, etc.)” Similar statements and lists of usual semantic roles for the “added” argument are found in Mithun (2001) and Kulikov & Song (2010). This use of semantic roles in definitions of applicatives is related to Peterson’s (1999: 36-40) parameter “Semantic Role of the Applied Argument.” In general, while there is consensus about which semantic roles are most commonly found for the applied argument in ACs, there is less agreement as to whether the mapping of particular semantic roles to the applied argument is definitional of applicatives. Observed effects on the mapping of semantic roles to syntactic arguments might instead be described as the consequence or result of another primary function of applicatives, e.g. promoting an argument in a base structure (Approach 1) or introducing a new argument to an existing structure (Approach 2).

3.2 Problematic constructions and their significance

Aside from conceptual issues, research in linguistic typology is also particularly concerned with the classification of language structures and describing and explaining structural diversity found across languages. In this section, I present a number of attested constructions in languages representing diverse language families that prove problematic for classification according to various definitions of applicatives, and similar constructions in West Nusantara languages.

3.2.1 Problematic constructions in previous research on other languages

Since there are many commonalities in definitions of applicatives, for a good number of constructions, there is some consensus. As a starting point, consider example (86) below from Hakha-Lai [cnh, haka1240], a Tibeto-Burman language of western Burma.

(86) Hakha Lai, commitative applicative

a. *kay-ma?*=*hee ka-law ?an-thlaw*
 1SG-PRON=COM 1SG.POSS-field 3PL.SUBJ-weed₁
 ‘They weeded my field together with me.’ (BC)

b. *ka-law ?an-la-thlo?*-*pii*
 1SG.POSS-field 3PL.SUBJ-1SG.OBJ-weed₂-APPL
 ‘They weeded my field (together) with me.’ (AC)

(Peterson 2007: 45)

Hakha Lai has fairly rigid Object-Verb word order. Evidence of grammatical relations in Hakha Lai includes word order, nominal case marking, and agreement marking for subject and object through affixation on the verb.⁵

The BC in (86a) is monotransitive; there are two core arguments. The A argument is the agent participant, ‘they’, which is evident from the third plural subject agreement prefix *?an-* on the verb. The P argument is the patient participant, which is realized as the full NP *ka-law* ‘my field’ in preverbal position in unmarked absolutive case. The verb also shows zero-marking for third person singular object agreement. The comitative participant, *kay-ma?=hee* ‘with me’ appears with the comitative case marker *hee* indicating that it is a clausal oblique.

The AC in (86b) is ditransitive. The verb is marked with *-pii*, the comitative applicative suffix and the clause now has three core arguments. The A argument, ‘they’, remains the same. The companion phrase *ka-law* ‘my field’ is again realized in preverbal position in absolutive case. This case marking and position indicates that ‘my field’ remains a core argument. The applied argument is the comitative participant ‘me’, which is shown to be an additional core argument by the first person singular object agreement prefix *la-* on the verb. Besides object agreement marking, Peterson (2007: 37) reports that comitative applied arguments show other syntactic properties characteristic of P (object) arguments in corresponding BCs; when an applied argument is expressed as a full NP, it also appears in preverbal position in absolutive case.

This example meets the consensus definition of applicatives outlined in §3.1.2. There is overt morphological marking on the verb that coincides with the expression of a peripheral semantic participant (in this case, a comitative) as an argument with the syntactic properties characteristic of P. In addition, it would qualify as an applicative under all of the five approaches described in 3.1.3: (i) it involves optional “promotion” of an oblique in the BC, (ii) it shows introduction of a new core argument, (iii) it shows an increase in syntactic valency as evident in number of core arguments, (iv) it shows licensing of certain syntactic encoding for the applied argument, and (v) it signals mapping of a semantic role to a syntactic argument. I will call this type of construction a ‘canonical applicative’, because there is broad consensus that it constitutes an AC.

In contrast, there is no consensus on whether a number of other constructions found in the languages of the world should be classified as applicatives. I summarize some such constructions below in Table 3.2. The first column in Table 3.2 gives labels for each type of construction. Most of these represent accepted terms in the literature on applicatives, but for the example from Abaza, I have coined a descriptive label, “indexing” applicative. The second column lists one language in which each type of construction is exemplified. These are largely drawn from Pacchiarotti (2020). The third column gives a short description of the construction, formulated in the terminology established thus far in this section.

In light of the preceding discussion, the constructions in Table 3.2 challenge one or another approach to the conceptual understanding of applicatives presented in §3.1.3, and in some cases, also the consensus definition of applicatives presented in §3.1.2:

- **Obligatory applicatives:** These constructions are indistinguishable from canonical applicatives except for the fact that there exists no possible BC from which a thematically peripheral oblique argument can be “promoted” (see Creissels 2004; Post & Modi 2022). These present complications for the view that applicatives are primarily argument-promoting

⁵Hakha Lai shows split alignment: NP arguments are marked ergative (A) or absolutive (S or P), while the agreement marking paradigm distinguishes subject (S or A) and object (P).

Table 3.2: Constructions which pose challenges to definitions of applicatives

Label	Example	Description
Obligatory applicative	Tswana [tsn, tswa1253]	A peripheral participant is expressed as a core argument in an AC. This participant represents a semantic role which cannot be expressed as an oblique in a BC.
Redirective applicative	Halkomelem [hur, halk1245]	A participant with a peripheral semantic role is expressed as P—a core argument—in the AC. The companion phrase, i.e. participant expressed as the P of the BC is now encoded as an oblique.
Asymmetrical applicative	Chichewa [nya, nyan1308]	The applied phrase in an AC exhibits some or all syntactic properties usually observed for P in a BC. The companion phrase shows core encoding in the AC but may either retain or lose some such syntactic properties
“Indexing” applicative	Abaza [abq, abaz1241]	A peripheral participant is indexed on the verb. There is no change to argument structure. The companion phrase retains the syntactic properties observed for P in a BC.
Registration applicative	Otomí [otl, tila1239]	A peripheral participant takes on a greater discourse salience in the AC, but is still encoded as an adjunct.

(Approach 1); they call into question what constitutes a BC and if correspondence with a BC is truly definitional for applicatives.

- **Redirective applicatives:** A peripheral participant is expressed as a core argument but this coincides with the “demotion” of another participant to oblique, resulting in no change in the value of syntactic valency (see Gerdtts 1988). These challenge the view that applicatives are primarily valency-increasing devices (Approach 3) and show that syntactic valency is not inseparable from other key attributes of canonical applicatives.
- **Asymmetrical object applicatives:** In languages like Chichewa, it not clear that all constructions formed with applicative morphology are truly ditransitive—representing an increase in syntactic valency. Instead, some such constructions, appear to remain monotransitive, with object properties simply “dispersed” across the companion phrase and the applied phrase, depending on the semantic role of the latter (see Peterson 1999: 56). This highlights that it is non-trivial to evaluate syntactic valency (as required in Approach 3) and syntactic status (as practiced under some definitions reflecting Approach 4) and it may be more appropriate to decompose these concepts. It also shows that the syntactic effect of an applicative can be attenuated by lexical or semantic factors.
- **“Indexing” applicatives:** In Abaza, the presence of certain verbal prefixes coincides with

pronominal cross-referencing of a peripheral participant on the verb. The argument expressing this participant no longer receives prepositional marking, however, it does not acquire access to syntactic operations consistent with core argument status—in fact O’Herin (2001) argues that there is no change in argument structure at all compared to the BC. These constructions highlight that syntactic coding may not be a reliable indicator of core argument status, which challenges some definitions that use syntactic coding as criteria for applicatives (a subset of Approach 4). Additionally “indexing” constructions in Abaza do not fit neatly under the consensus definition of applicatives, as the peripheral participant, or applied phrase, is not an absolutive argument (consistent with P) nor an ergative argument. It is perhaps not a syntactic core argument at all.

- **Registration applicatives:** In these constructions, morphological marking on the verb coincides with greater discourse salience of a peripheral participant, but the argument expressing this participant remains a syntactic adjunct (Hernández-Green 2016). Like “indexing” constructions, these do not fit under the consensus definition of applicatives. They show some similarity to certain characterizations of applicatives discussed under Approach 5 (role-marking). But instead of signalling or marking an unusually prominent syntactic mapping for the peripheral participant (i.e., mapping to a core argument), they indicate an unusually prominent pragmatic status (i.e., topicality or salience). These constructions raise the question of whether discourse function must be considered in an adequate conceptual understanding of applicatives. In fact, while definitions of applicatives do not commonly make reference to discourse function, a number of authors recognize that discourse considerations are an important factor in the use and function of applicatives (Dixon & Aikhenvald 2000; Peterson 2007; Foley 2007).

In summary, there exist a number of constructions in the languages of the world which are quite similar to canonical applicatives in many ways, but do not qualify for inclusion under one or another definition of applicatives. We might call these ‘non-canonical’ applicatives. Examining the reasons that these fall outside of particular definitions reveals complications for developing reliable and explanatory criteria for applicatives. For example, these data call into question the necessity of a strict correspondence relationship between a BC and an AC. Also, in the absence of such correspondence relationships, the range of possible semantic roles for the applied phrase in an AC must be otherwise explained and distinguished. Furthermore, these data show that indicators of grammatical relations observed in BCs do not always pattern neatly in ACs. Thus, a useful definition of applicatives should address the fact that indicators of grammatical relations, e.g. agreement marking, case marking, access to syntactic operations, might vary by construction or even be split across arguments in various manners. Finally, these data point to some other ways in which the consensus definition of applicatives may be inadequate, as it does not account for factors outside of semantic role and syntactic realization of the applied argument which appear to influence observed properties of ACs, such as lexical semantics, language-specific restrictions on clausal transitivity, and discourse considerations. In light of this, we see that there is a need for a more nuanced understanding of applicatives—one that accounts for the diverse properties and variation observed for canonical applicatives and that explains the relationship of these canonical applicatives to non-canonical applicatives. Coming to such an understanding in turn rests on the development of an adequately descriptive, nuanced, and representative typology of canonical and

non-canonical ACs in the world's languages. In the next section, I will argue that the languages of West Nusantara are particularly deserving of such study and that they represent an important test case for this due to the broad range of ACs they exhibit, showing diverse structural and functional properties.

3.2.2 Polyfunctionality and problematic constructions in languages of West Nusantara

Based on a review of previous research on applicatives in West Nusantara languages (see §1.6), it is clear that ACs are found in a good number of languages of this region. In such studies, the examples of ACs most often reported and discussed would largely fall within the common definition of applicatives outlined above; these may be considered canonical applicatives. One example of a canonical AC is given below in (87) from *Tukang Besi*, a language of Southeastern Sulawesi.

(87) *Tukang Besi*, Beneficiary-selecting applicative

- a. *No-wila i daoa kene no-waliako.*
 3.RLS-go OBL market and 3.RLS-return
 'She went to the market and came back.' (BC) (Donohue 1999: 431)
- b. *No-wila-ako te ina-no i daoa.*
 3.RLS-go-APPL CORE mother-3.GEN OBL market
 'She went to the market for her mother.' (AC) (Donohue 1999: 232)

Donohue (1999) reports that the applied argument in such ACs, that is, the beneficiary, shows coding and syntactic properties largely equivalent to that of P in a corresponding BC in terms of word order, agreement, case marking, control properties, and access to relativization and passivization. This type of applicative would also be considered 'optional', as it is possible to express a beneficiary as a non-core, dative marked phrase in a BC in *Tukang Besi* (see Donohue 1999: 333-335, and discussion in §3.5 below).

In Standard Indonesian, beneficiary-/recipient-selecting constructions marked with the applicative suffix *-kan* also appear to show canonical properties, as in the example given below from Chung (1976). In such ACs, the beneficiary applied argument shows coding and syntactic properties like that of P in a corresponding BC. It is encoded as an unmarked NP, and can be the privileged syntactic argument in a PV construction marked with *di-*.

(88) Standard Indonesian, Beneficiary-selecting applicative

- a. *saya mem-bawa surat itu kepada Ali.*
 1SG AV-bring letter DIST to A.
 'I brought the letter to Ali.' (BC)
- b. *saya mem-bawa-kan Ali surat itu.*
 1SG AV-bring-BEN.APPL A. letter DIST
 'I brought Ali the letter.' (AC) (Chung 1976: 41)

It is no wonder then, that tidy examples like these from the better-known cases (e.g. Indonesian, *Tukang Besi*) are commonly used to exemplify Austronesian languages in cross-linguistic studies of applicatives (alongside examples from Austronesian languages of the Oceanic subgroup such as Fijian, and sometimes East Nusantara languages closely related to Oceanic, such as *Taba*). Often missing from discussions of applicatives, however, are examples showing the breadth of variation of non-canonical ACs in Austronesian languages and the diversity of functions associated with AMs in such constructions. Preliminary investigation suggests that a large number of non-canonical ACs are attested in languages of West Nusantara. A sample of these are listed in Table 3.3 below. Most of these are described and exemplified in Truong & McDonnell (2022), on the basis of original fieldwork, investigation of corpora, and existing grammatical descriptions (see also van den Berg & Busenitz 2012; Donohue 1999).

Table 3.3: Constructions in West Nusantara languages that pose challenges to definitions of applicatives

Example	Description
Balantak reason applicative	The complement of the verb in an AC marked with the AM <i>-kon</i> may be an entire clause expressing a reason (or motive). There is, of course, no alternative monoclausal BC.
“Lexically obligatory” applicative (e.g. Sundanese <i>alung</i> ‘throw’)	Certain verbal stems cannot function as a predicate in base form, only when affixed with an AM. In ACs, P expresses a peripheral participant. This appears to be specific to certain lexical stems, and has been found in Sundanese, Indonesian, and Sasak, among others.
Sundanese locative applicative	A location or goal participant is expressed as P in the AC. The participant expressed as the P of the BC (theme participant) is encoded with prepositional (oblique) marking when it appears as the companion phrase in the AC.
Tukang Besi purpose applicative	When the applied phrase expresses a purpose (motive), the companion phrase in the clause lacks access to most syntactic operations usually available to objects.
Tukang Besi double applicative	If two applicative suffixes are added the verb, the first applied object may bear core case marking, but has no syntactic object properties.
Pragmatic use of applicative <i>-ka</i> in S. Barisan Malay	For some verbal stems, there is no change in argument structure between the BC and the AC. However, the P argument in the AC tends to show certain pragmatic properties, e.g. greater individuation or affectedness.
Indonesian benefactive applicative	A beneficiary participant is semantically entailed by the AM-marked verb, but may either be realized as a P argument or with prepositional (oblique) marking.

Besides the non-canonical ACs given here, Truong & McDonnell (2022) also identifies several additional types of constructional meanings often marked with AMs in languages of West

Nusantara. These include formation of morphological causative constructions, formation of comparative degree constructions, and semantic effects, such as indication of pluractional aspect and greater intensity. All of these except the intensive function were exemplified in Sundanese in Chapter 2. While these functions or meanings are observed with great frequency on AM-marked clauses in the languages of West Nusantara, and are foundational to the role that AMs play in these verbal systems, they are not always viewed as particularly relevant to the study of applicatives in these languages.

Returning to the constructions listed in Table 3.3 above, these non-canonical ACs also generally show verbal affixation with the same AMs that mark canonical ACs like (87b) and (88b) above. This notwithstanding, these problematic constructions show a different argument structure or relationship with a BC than seen in with the canonical ACs and do not fit neatly into one or another conceptual approach to applicatives. As such, there are many parallels between these constructions and the problematic constructions presented in Table 3.2 earlier:

- Like obligatory applicatives in Tswana, reason applicatives in Balantak (a language of eastern Sulawesi) and “lexically obligatory” applicatives found in a number of West Nusantara languages pose a challenge to the use of BCs in definitions of applicatives and the concept of “promotion” .
- Like asymmetrical applicatives in Chichewa, *Tukang Besi* purpose applicatives appear to allow object properties to be dispersed over two arguments.
- Like “indexing” applicatives in Abaza, *Tukang Besi* double applicatives show that indicators of grammatical relations observed in BCs may not neatly coincide as expected in ACs.
- Like “indexing” applicatives in Abaza and registration applicatives in Halkomelem, certain *-ka* marked constructions in South Barisan Malay (a language of southern Sumatra) and *-kan* marked benefactive constructions in Indonesian show no apparent change in argument structure. Instead AM-marking appears to signal a special pragmatic or semantic status for a selected participant.

In summary, we see that both canonical and non-canonical ACs are represented in West Nusantara languages. These show diversity in key properties—both across and within individual languages—similar to that which is observed across languages of the world representing many different genetic families. Thus, study of West Nusantara applicatives—and crucially, inclusion of non-canonical ACs in this—is fertile ground for exploring a number of unresolved questions about applicatives as a cross-linguistic phenomenon. Some of these are listed below:

1. How can the function(s) of applicative morphemes be adequately and insightfully characterized, especially in light of observed variance in associated structures and observed polyfunctionality?
2. How may syntactic and semantic properties be distributed across arguments in ACs? To what extent do these diverge from expected patterns?
3. How are properties of applicatives influenced by lexical and semantic factors, and usage? To what extent does this occur?

These issues will be explored through investigation of West Nusantara ACs in this study. In light of the preceding discussion, I have adopted a constructional approach to applicatives for this study, and in the next section, I show why this is well-suited to addressing the issues identified here.

3.3 Constructional approaches as a lens for West Nusantara applicative systems

As outlined in Chapter 1, an important goal of this study is to broaden available description and cross-linguistic comparison of applicatives in languages of West Nusantara. By doing so, I seek to understand the applicative systems of West Nusantara in a typological context, but also on their own terms, in the context of the diachronic and synchronic systems of language in which they developed and are used (see Haspelmath 2010 on the difference between cross-linguistic comparative concepts and language-specific descriptive categories). Given the diversity of non-canonical ACs and other AM-marked constructions found in the languages of West Nusantara, it is my view that the consensus definition of applicatives presented in §3.1.2 and the previous approaches to applicatives discussed in §3.1.3 above do not serve the goals of the study well.

In this study, I have proposed a streamlined definition of applicatives as in (1), repeated again below. This definition clearly draws in large part on the consensus definition of applicatives given in (85) above in §3.1.2. However, in response to the issues identified above, to augment this definition, I have adopted a constructional approach in explaining what it means for an applicative to be a clausal construction. Specifically, in this approach, an AC is understood as a conventionalized pairing of fixed form (i.e. AM-marking, clausal argument structure) and consistent meaning (i.e. mapping a peripheral role to a core argument position, clausal meaning) (see Perek 2015; Goldberg 2006). The form of a construction is assumed to be only partially specified or fixed. Thus for an AC, the association between form and meaning described above holds across variable lexical items that fill in slots for non-specified content in examples, especially the verbal bases that bear AM-marking (see §1.4.1).

- (1) An applicative is a clausal construction in which overt morphological marking on the verbal complex coincides with the selection of a non-agent, non-patient semantic role to map to a core argument in the clause.

Using this approach allows me to delineate the object of study in Part II and Part III of this dissertation without imposing distinctions that are not well-supported in the grammatical systems of West Nusantara languages. These include: the distinction between AM-marked constructions which show an increase in syntactic valency and those which do not, the distinction between AM-marked constructions in which the applied phrase is encoded as a core argument and those in which it is not, and the distinction between ACs that select a peripheral participant and other types of constructional meanings marked with the same AMs in these languages (e.g. causative constructions, pluractional aspect, intensive semantics). To illustrate, I return to different types of Indonesian *-kan* marked benefactive constructions mentioned in the previous section, which are exemplified in (89).

(89) Standard Indonesian, Beneficiary-selecting applicative

- a. *Saya mem-(p)anggang-kan Eric roti.*
1SG AV-bake-APPL E. bread
'I baked Eric bread.' (AC)
- b. *Saya mem-(p)anggang-kan roti untuk Eric.*
1SG AV-bake-APPL bread for E.
'I baked bread for Eric.'
- c. *Saya mem-(p)anggang-kan roti.*
1SG AV-bake-APPL bread
'I baked bread for someone.'

(Cole & Son 2004: 341–342)

In an approach centered around ACs as conventionalized pairings of form and meaning, any-time two constructions show shared elements of form or shared elements of meaning, there is a clear way to identify and classify the relationship between the two, even if they differ in other observed properties. Across examples (89a)–(89c), a shared element of form is the marking of the verb with *-kan*, and a shared element of meaning is the accruing of benefit to some beneficiary. Differences across these are also observed. There is an alternation in the realization of the beneficiary participant, which is encoded as an unmarked NP immediately following the verb in (89a), is encoded as a prepositional phrase in (89b), and is unrealized in (89c). The relationships between these three AM-marked clauses leads to questions about how to explain this observed alternation: How did *-kan* come to be used in these three different structures? Is the coding of the beneficiary influenced by other properties of the *-kan* marked clause (e.g. length of applied phrase, (non)realization of the companion phrase, voice, mood, etc.)? Across alternating constructions, are there observed differences in the distribution of base lexemes, e.g. are some bases more likely to be used with non-realized participants? These questions are framed differently than questions might arise when focusing on the relationships between BCs and ACs.

While other approaches to applicatives may yield meaningful typological classification, here I want to demonstrate how adopting a constructional approach has three main advantages. First, viewing ACs in this way leads to broad identification of the types of clauses that are relevant in understanding West Nusantara applicatives, with the result that diverse types of AM-marked constructions may then be described and considered for inclusion in a typology of applicatives. The constructional alternations seen with Indonesian *-kan* marked benefactives above are just one small example of the kind of diversity I seek to capture. Second, viewing ACs in this way aids in developing a functional typology in which AM-marked constructions may be classified with emphasis on distinctions that are meaningful in the internal linguistic systems of West Nusantara languages. Such distinctions might include differences observed in properties of ACs in A-oriented vs. P-oriented transitive clauses, or the distribution observed for types of ACs across available forms of AM-marking, and how this patterns across West Nusantara languages. Third, this type of constructional approach emphasizes the fact that speakers' linguistic knowledge includes knowledge about specific items and knowledge of patterns of varying levels of generalizability across items or exemplars as observed in usage (see e.g. Goldberg 2006: 45–62). Thus, viewing ACs in this way aligns with the observation that AM-marked constructions in these languages are not equally compatible with all kinds of bases, and in frequent usage may be observed with a large number of bases, or only a small number of bases, or at the extreme end of lexicaliza-

tion, perhaps with only one lexical item. In the Sundanese applicative system, we see this with ACs that select a content, addressee, stimulus, or performance applied phrase (see §??), with each of these constructions being found with just a limited number of semantically similar bases, even in a corpus of a not insignificant size, like the MPI Sundanese corpus of 30,000 sentences.

For these reasons, the constructional understanding of applicatives that I have outlined here and in §1.4.1 is the underpinning for this study, and informs the types of data that have been compiled, and the categories used in presentation of summary results and descriptive details in Part II and Part III of this study. In the remaining sections of this chapter, I explore two specific challenges that remain in relating applicatives in West Nusantara languages to their larger context, and show how I will address these in this study.

3.4 On the problem of Philippine-type voice in a study of applicatives

The languages of West Nusantara—alongside Philippine and Formosan languages—are a subset of western Austronesian languages. Because of this, any discussion of verbal constructions in these languages requires frequent and substantive reference to the distinctive grammatical systems known as *symmetrical voice*. These systems have also been referred to as western Austronesian voice, ‘focus’, or ‘trigger’, and show important differences with well-known voice systems in the world’s languages, such as active-passive systems (see Chen & McDonnell 2019; Riesberg 2014b; Himmelmann 2002). To be clear, not all languages of West Nusantara show symmetrical voice systems. Even so, because a symmetrical voice system was present in Proto Malayo-Polynesian and Proto Austronesian before that, systems of voice and diathesis in these languages are either symmetrical in their synchronic form, or “asymmetrical” but nonetheless known to be derived from an earlier symmetrical system.

In western Austronesian symmetrical voice systems, each clause has one argument that is syntactically privileged, by virtue of access to syntactic operations (e.g. relativization, word order, quantifier float). I will call this argument the *pivot* (Chen 2017). Affixes on the verb signal the selection of the semantic role (or possible roles) that maps to the pivot. In these systems, there is a paradigm of (symmetrical) voice alternations, each distinguished by its morphological marking, and the set of possible semantic roles that map to the pivot. Crucially, voice alternations in a symmetrical voice system include multiple transitive constructions, which generally take one pivot argument, and one or more *non-pivot* core arguments. An example with the verb ‘buy’ is given below from Balinese, which has a two-way symmetrical voice system.

(90) Balinese, Voice alternations

a. *Buku beli tiang di toko ento.*

book pv.buy 1SG at shop DIST

‘I bought **the book** in that shop.’ (PV, with patient pivot)

b. *Tiang m-(b)eli buku di toko ento.*

1SG AV-buy book at shop DIST

‘I bought the book in that shop.’ (AV, with agent pivot)

(Artawa 1998: 48)

In (90a) the verb takes no overt morphological marking, and the semantic patient, i.e., the thing bought, is the pivot argument (bolded). This represents the patient voice or P-Voice (PV) construction. In (90b), the verb is marked with the prefix *N-* (underlined), and the semantic agent, i.e., the buyer, is the pivot argument (bolded). This represents the actor voice or A-Voice construction (AV). In Balinese, the P argument in PV and the A argument in AV have unique access to syntactic operations. Only the pivot can be the head of a relative clause, and in non-finite subordinate clauses, only the referent of the pivot is controlled by an argument of the matrix clause, among other properties (see Arka 2003; Riesberg 2014b). The pivot appears in preverbal position in unmarked word order, with some variation possible due to pragmatic factors.

Besides the pivot argument, each clause in (90) has one additional unmarked NP, which is the non-pivot core argument. In the PV clause in (90a), the non-pivot core argument is the semantic agent, *tiang* ‘me’. In the AV clause in (90a), the non-pivot core argument is the semantic patient, *buku* ‘the book’. This argument has strict word order in Balinese, and must immediately follow the verb (Wechsler & Arka 1998: 404–405).

Syntactically transitive clauses in Balinese must have a value for voice; each clause is either in AV or PV, but never both. Thus, AV and PV are contrastive values in the voice system of Balinese.

Balinese also shows applicative morphology in the form of suffixes *-in* and *-ang*. Examples of *-in* in a locative-selecting AC are shown in (91).

(91) Balinese, Locative-selecting applicative + Voice alternations

- a. **Toko ento beli-in** *tiang buku.*
shop DIST PV.buy-LOC.APPL 1SG book
‘I bought the book in that shop.’ (AC in PV, with location pivot)
- b. **Tiang m-(b)eli-in** *toko ento buku.*
1SG AV-buy-LOC.APPL shop DIST book
‘I bought the book in that shop.’ (AC in AV, with agent pivot) (Artawa 1998: 55)

In corresponding BCs, as in (90), a semantic location is realized as an oblique PP, i.e., *di toko ento* ‘in that shop’. In locative-selecting ACs marked with *-in*, as in (91), the semantic location is selected as a core argument, and encoded as an unmarked NP, i.e. *toko ento* ‘that shop’. On this basis, this construction meets the definition of an applicative as given in (1). Furthermore, just as in non-applicative transitive clauses, transitive clauses marked with *-in* in Balinese show an alternation between AV and PV. In PV applicative clauses, as in (91a), the location applied phrase shows encoding and syntactic properties consistent with status as the pivot argument. In AV applicative clauses, as in (91b), the location applied phrase shows encoding and syntactic properties consistent with status as a non-pivot core argument. The pivot in AV applicative clauses is the semantic actor, just as in base clauses.

In the same manner, in beneficiary-selecting ACs marked with *-ang*, the semantic beneficiary is selected as a core argument. These constructions again show an alternation between AV and PV, which determines whether the beneficiary is realized as the pivot argument, or a non-pivot core argument. This is shown in (92).

(92) Balinese, Beneficiary-selecting applicative + Voice alternations

- a. *Cai beli-ang tiang buku-ne ene.*
 2SG PV.buy-BEN.APPL 1SG book-DEF PROX
 ‘I bought you this book.’ (AC in PV, with beneficiary pivot)
- b. *Tiang m-(b)eli-ang cai buku-ne ene.*
 1SG AV-buy-BEN.APPL 2SG book-DEF [PROX]
 ‘I bought you this book.’ (AC in AV, with actor pivot) (Artawa 1998: 55)

In examples (90)–(92), we consistently see that AV and PV contrast with one another, and they substantiate a basic voice category in Balinese for which every transitive clause must have a specified value. On the other hand, ACs exist outside of this paradigm. A verb marked with *-in* or *-ang* is neutral for the basic voice category in Balinese, which must be further specified, with AV and PV as possible values. For the purpose of simplicity, I have not included examples of passive constructions in Balinese; passive constructions contrast with AV and PV but co-occur with ACs marked with *-in* and *-ang*. I also note that in some varieties of Balinese, *-in* and *-ang* can co-occur with each other on the same verb (see Clynes 1995b). Thus we can identify two categories of constructions in Balinese that specify the mapping of semantic roles to grammatical relations, as presented in Table 3.4. One category may be identified as “basic voices” because one value for it is obligatorily specified in a clause, and a second category is identified as “applicatives”, as they involve selection of a peripheral role as a core argument.

Table 3.4: Overview of Balinese voice system

	Basic voices	Applicatives
Possible values	AV, PV, passive	beneficiary-, locative-selecting
Pivot-selection	determines mapping	neutral
Core argument selection	neutral	determines mapping
Valency	valency neutral	increases valency
Clausal requirements	one value must be specified	requires additional specification
Meets def. of applicative in (1)	No	Yes

Table 3.5: Partial paradigm for Balinese verbs

	Base	Appl 1	Appl 2
PV	beli	beli-in	beli-ang
AV	m-(b)eli	m-(b)eli-in	m-(b)eli-ang
passive 1	ka-beli	ka-beli-in	ka-beli-ang
passive 2	beli-a	beli-in-a	beli-ang-a

While the voice system of Balinese supports a fairly straightforward division into two categories of verbal constructions—basic voice constructions and ACs—in many western Austronesian languages, we see a different type of organization. The Philippine-type symmetrical voice

languages raise complications in a study of applicatives, because the categories for basic voices and applicatives are not distinct.

An example of a Philippine-type voice system is given in (93) from Kimaragang, a language of the Sabahan subgroup spoken near the northeast tip of Borneo.

(93) Kimaragang, Voice alternations

- a. M-(p)anga-lapak **okuh** do niyuw.
 AV-TR-split 1SG.NOM GEN coconut
 ‘I will split a coconut/some coconuts.’ (AV)
- b. Lapak-on kuh **it** niyuw.
 split-PV 1SG.GEN NOM coconut
 ‘I will split **the coconut(s)**.’ (PV)
- c. Lapak-an kuh do niyuw **it** wogok.
 split-DV 1SG.GEN GEN coconut NOM pig
 ‘I will split some coconuts **for the pigs** (to eat).’ (DV)
- d. Nokurohtu n-i-lapak nuh do niyuw **inoh** dangol kuh?
 why PST-IV-split 2SG.GEN GEN coconut MED.NOM bush.knife 1SG.GEN
 ‘Why did you use **my bush knife** to split coconuts?’ (IV) (Kroeger 2005: 405)

Each voice construction in (93a–d), is distinguished by affixal marking on the verb (underlined) and the semantic role of the pivot argument (bolded). The pivot argument, if expressed, always appears in the case glossed as ‘nominative.’ In (93a), the verb is marked with the prefix *m-* and the semantic agent is the pivot; this represents the actor voice or A-Voice (AV) construction. In (93b), the verb is marked with the suffix *-on* and the semantic patient is the pivot; this represents the patient voice or P-Voice (PV) construction. In (93c), the verb is marked with the suffix *-an* and the semantic beneficiary is the pivot; this represents the so-called dative voice (DV) construction, and is also used when the pivot is a goal, recipient, or a certain type of distributed location, among other roles. In (93d), the verb is marked with the prefix *i-*, and the semantic instrument is the pivot; this represents the instrumental voice (IV) construction.

(94) Kimaragang, LV construction

- Siongoh** l-in-apak-on dit tayar nuh?
 where <PST>split-LV GEN 2SG.GEN
 ‘**Where** did your tire burst?’ (Kroeger 1988: 233)

A fifth voice alternation, locative voice (LV), is illustrated in (94). In this construction a semantic location is the pivot (bolded), and the verb is marked with *-on* (underlined). The affixal marking for LV is homophonous to that of PV in non-past tense, but is distinct in past tense, where PV is zero-marked, meaning that it is distinguished by lack of other voice morphology. Past tense is marked with the infix *-in-*. Hence, the verb *lapak* ‘split’ has the past tense form *l-in-apak* in PV but *l-in-apak-on* in LV.

In Kimaragang, these five voice categories operate as a single paradigm. All main clauses and most dependent clauses require the specification of one and only one basic voice category,

i.e., AV, PV, DV, IV, or LV. Each clause accordingly has one pivot argument, the semantic role of which is indicated by the morphological marking on the verb. However, some of the basic voice categories involve selection of a non-agent, non-patient peripheral role as a core argument—the pivot being a core argument; thus DV, IV, and LV meet the definition of applicatives in (1), but AV and PV do not.

Table 3.6: Overview of Kimaragang voice system

	Basic voices	
Possible values	AV, PV, DV, IV, LV	
Pivot-selection	determines mapping	
Clausal requirements	one value must be specified	
	Subset I	Subset II
Values	AV, PV	DV, IV, LV
Core argument selection	neutral	determines selection
Valency	valency neutral	increases valency
Meets def. of applicative in (1)	No	Yes

Properties of the voice constructions illustrated in the preceding Kimaragang examples are summarized in Table 3.6. In the context of the Kimaragang verbal system, there is one category of basic voice constructions with five contrastive values (see Foley 2008). This category is of major importance in the grammatical system of Kimaragang. Case marking, word order patterns, and syntactic operations all show sensitivity to the selection of the pivot argument. However, when a cross-linguistic definition of applicatives is applied in Kimaragang, for these five voice constructions we might draw a distinction between non-applicative constructions (AV and PV) in which a peripheral semantic role is not selected to map to a core argument, and ACs (DV, IV, and LV) in which a peripheral semantic role is selected to map to a core argument. This distinction is of interest in a cross-linguistic study of applicatives like this one, and is certainly relevant when comparing the Balinese grammatical system to that of Kimaragang (see Davies 2005 on similar comparisons for Madurese) but it is not of particular consequence in the internal organization of grammatical system of Kimaragang.

Thus, on the one hand, viewed within the internal grammatical systems of individual western Austronesian languages, symmetrical voice alternations and pivot-neutral applicatives belong to two independent dimensions of clausal organization (see also Himmelmann & Riesberg 2013, on criteria to distinguish symmetrical voice and applicatives). Yet, on the other hand, constructions like Kimaragang DV, IV, and LV show strong similarities to applicatives in other languages. Following Haspelmath (2010) then, all five of the Kimaragang alternations presented above (AV, PV, DV, IV, and LV) may be considered to belong to a single language-specific *descriptive category* of symmetrical voice, while only three of those (DV, IV, and LV) belong to the *comparative category* of applicative.

Nonetheless, while Kimaragang DV, IV, and LV constructions may be classified as ACs, they show a property that is cross-linguistically unusual. The applied phrase (peripheral role selected to map to a core argument) in Kimaragang DV, IV, or LV clauses, is always realized as the most syntactically privileged clausal argument, the pivot; in fact, the relevant peripheral roles generally

cannot be expressed as a non-pivot core argument. I will call constructions like Kimaragang DV, IV, and LV *pivot-selecting* applicatives. The fact that these constructions target the syntactically privileged pivot relation is a defining characteristic.

As discussed above, an important criterion in most definitions of applicatives is that the applied phrase has the coding and/or syntactic properties characteristic of P in transitive clauses. Whether this represents a syntactically privileged position in clausal argument structure is incidental, rather than defining. If a BC in a given language shows a subject relation and an object relation, the applied phrase in a corresponding AC is not expected to have subject properties. That is, the applied phrase is not expected to show evidence of unique syntactic privilege as a grammatical subject might, but it may show object properties. If a BC in a given language shows an absolutive and an ergative relation, the applied phrase in a corresponding AC may show properties of the absolutive. That is, the applied phrase may show evidence of unique syntactic privilege to the extent that absolutive P arguments generally do so in the language. In neither case do we have evidence that the syntactically privileged relation is specifically targeted by the applicative. We simply see that the applied phrase in an AC may take properties characteristic of P in basic transitive constructions, whether these constitute evidence of syntactic privilege or not. Such applicatives are neutral with respect to syntactic privilege for the applied phrase.

Constructions like the Balinese ACs exemplified above, I will call *pivot-neutral* applicatives. These fit into the more general cross-linguistic pattern. In PV, the applied phrase shows coding and behavioral properties characteristic of the pivot relation, as would P in a corresponding BC in PV. In AV, the applied phrase shows coding and behavioral properties characteristic of the non-pivot core argument relation, as would P in a corresponding BC in AV. These applicatives are *pivot-neutral* because whether the applied phrase is syntactically privileged or not is incidental, rather than defining.

Because they expressly target the grammatical relation that is uniquely syntactically privileged, the pivot-selecting ACs found in Philippine-type languages like Kimaragang are not well-accounted for in typological work on applicatives to date.

Peterson's (2007) typological survey of applicatives does not include Philippine-type languages, neither among the fifty languages of his sample with applicatives nor among the fifty without applicatives, and as far as I can tell, shows no examples of ACs that expressly target the privileged syntactic relation. Peterson does not discuss the pivot-selecting constructions of Philippine languages, except to speculate how morphological marking for LV and IV in Proto Austronesian may have been grammaticalized from earlier prepositions or auxiliaries, as attested for applicatives in other families.⁶

In the WALS feature on applicatives (Polinsky 2013), two Philippine-type languages are included in the map, and symbolized there as having applicatives. These are Tagalog, a language of the Philippines, and Paiwan, a Formosan language. However, there is no discussion in the text explaining how Philippine-type peripheral nonactor voices fit into the characterization of ap-

⁶Peterson (2007: 165–169) proposes that the Proto Austronesian voice system developed from applicative alternations in relative clauses that were extended to main clauses. Peterson's account is framed as an alternative to the proposed origin of Proto Austronesian nonactor voices as reanalyzed nominalization structures (Starosta, Pawley & Reid 1982), but it does not seriously interact with the historical evidence for the nature of the Proto Austronesian grammatical system or controversies surrounding this. See Chen (2017) for an deep discussion of evidence that Philippine-type voice morphology was present in both main clauses and relative clauses among the higher order groupings of Proto Austronesian as reflected in present-day Formosan languages.

plicatives as a “type of double object construction,” and a construction in which “the number of object arguments selected by the predicate is increased by one with respect to the basic construction.” These constructions expressly result in mapping of the applied phrase to the pivot relation, and it is not at all clear that the pivot relation is similar to an object relation, or that the term “object” is relevant in Philippine-type languages. Nonetheless, the inclusion of two Philippine-type languages is a small step in the right direction, as it allows them to be compared with the other 164 languages in the sample (64 with applicatives) for a great number of typological features.

Finally, Zúñiga & Kittilä (2019: 122–129) do discuss Philippine-type voice alternations, but do not place them firmly under the category of applicatives. Instead, they view LV, IV, and CV alternations as a special type of “subjective undergoer nucleative” or “Philippine undergoer nucleative.” They write that the effect of these constructions “is equivalent to applicativization (which introduces the new argument to the syntactic core) plus passivization (which promotes the nonagent to subject)” (125). Their view is not too far from my own; we agree that that Philippine-type peripheral nonactor voice constructions are similar to applicatives, and also have unique properties not usually found in applicatives in other language families, with respect to targeting of a syntactically privileged relation. I go one step further than they do, in explicitly calling these constructions applicatives, under the subcategory “pivot-selecting applicative.”⁷

This decision to include Philippine-type peripheral nonactor voices as applicatives in this study, and to distinguish them by the use of the term *pivot-selecting* as opposed to *pivot-neutral*, is desirable for multiple reasons. First, this approach clarifies the relationship of Philippine-type peripheral nonactor voices to cross-linguistic definitions of applicatives. There is a clear basis by which these constructions are included under the label applicative, especially if applicatives are defined in reference to core argument status for the applied phrase, rather than an “object” relation, which is less universally applicable. Second, this approach acknowledges that, as compared to ACs in many other language families and regions of the world, the pivot-selecting applicatives in western Austronesian languages have distinctive characteristics that do not necessarily have direct analogs in asymmetrical voice systems. These constructions can be classified as applicatives without reference to active-passive voice systems or ergative systems, and there is also no need to equate symmetrical voice systems with such. Third, even though I have exemplified here two voice systems of West Nusantara, which have *either* pivot-selecting ACs *or* pivot-neutral applicatives, there exists a whole spectrum of logical possibilities for configurations of symmetrical voice categories and applicative categories, and this terminology facilitates the description and comparison of this diversity as attested in West Nusantara and beyond. Finally, LV, CV, and IV affixes are in many cases cognate (or apparently cognate) with pivot-neutral applicative suffixes, and this approach allows for better recognition of potential diachronic relationships between pivot-selecting and pivot-neutral ACs accordingly.

⁷I also consider it important to distinguish the nonactor voice alternations in symmetrical voice languages from passive constructions. Unlike passives, in many symmetrical voice languages it can be demonstrated that the nonactor voice constructions are syntactically transitive rather than syntactically intransitive, and they are distinguished from intransitive voice categories in the same languages, including true passives, by morphological marking and syntactic properties of arguments.

3.5 On the relationship between serial verb constructions and applicatives

Another challenge to articulating the place of West Nusantara applicatives in typological perspective lies in the relationship between applicatives and serial verb constructions. In a serial verb construction, a sequence of verbs are members of a single clause and “act together as a single predicate” without being compounded or marked for coordination, subordination, or other types of syntactic dependency (Aikhenvald 2006: 1). The meaning of a serial verb construction expresses a complex event, with verbs in the sequence “expressing various facets” of this event (Payne 1997: 307).

In (1), I define an applicative as a clausal construction in which morphological marking on the verbal complex coincides with the selection of a peripheral role to map to a core argument. At face value, the use of the term “morphologically marked” in this definition excludes serial verb constructions, which are a type of analytic construction, rather than a morphologically marked construction. That being said, some serial verb constructions clearly express meanings similar to those of applicatives. Like applicatives, they may be used to express benefactive relationships, locative relationships, and instrumental relationships with respect to the event expressed by a clause. In addition, there are some shared formal characteristics. In both serial verb constructions and ACs, a peripheral semantic role may be realized as a core argument of a verb rather than an oblique or adjunct, though the fact that a serial verb construction contains more than a single verb suggests that the syntactic status of the phrase expressing the peripheral participant in the two types of constructions is not equivalent.

Complicating the relationship between ACs and serial verb constructions, in some cases, the distinction between morphological marking and analytic or syntactic means for installing an argument is not clear.

Tukang Besi, for example, as mentioned above has an applicative suffix *-ako*. Like other suffixes, *-ako* belongs to the same phonological word as the verb root, as reflected in word level stress patterns with primary stress on the penult (see example in Donohue 1994: 41). An example of *-ako* used as an applicative suffix is given in (95).

(95) Tukang Besi, Beneficiary-selecting applicative

- a. *No-ala te kau.*
3.RLS-fetch CORE wood
'She fetched the wood.' (BC)
- b. *No-ala-ako-'e na ina-su.*
3.RLS-fetch-APPL-3.OBJ NOM mother-1SG.POSS
'She fetched (it) as a favour for my mother.' (AC) (Donohue 1999: 231–232)

The suffix *-ako* is probably diachronically related to the independent morpheme *ako* ‘do for’, which, according to Donohue (1999: 333) shows ‘preposition-like behavior’ and is ‘best described as an atypical verb.’ In some uses, *ako* shows properties common to verbs, as it must take a subject indexing prefix when used as a main predicate as shown in (96). Also, in relative clauses in Tukang Besi, affixal marking is used to indicate the grammatical relation of the head noun to the dependent clause verb, and *ako* takes such affixal marking, just like other verbs.

(96) Tukang Besi, Use of *ako* as a main predicate

Mbea-do 'u-ako-naku wa?
not-yet 2SG.RLS-do.for-1SG.DAT.OBJ PRT
'Haven't you done it for me yet?'

(Donohue 1999: 333)

In other clauses with benefactive meanings, we see *ako* behave more like a light verb or preposition, as shown in (97). Here, *ako* is not part of the main verb stem. Unlike the example of suffixal *-ako* in (95b), here *ako* is non-adjacent to the main verb root *sai* 'make', and is positioned after the third person object suffix *-e* rather than before it. Unlike a main verb, however, in (97) *ako* is not fully inflected, as it bears no subject-indexing prefix, while *sai* bears the second person subject prefix *'u-*. This usage can be considered a serial verb construction with a benefactive meaning, or alternately be analyzed as simply as a BC showing prepositional marking of the beneficiary.

(97) Tukang Besi, Use of *ako* in a serial verb construction

Mbea-do 'u-sai-'e ako-naku wa?
not-yet 2SG.RLS-make-3.OBJ BEN-1SG.DAT.OBJ PRT
'Haven't you made it for me yet?'

(Donohue 1999: 334)

Thus, we can identify three constructional uses of a morpheme with the shape *ako* in Tukang Besi, each with similar meanings. We might also surmise that suffixal *-ako* has been grammaticalized from the independent morpheme *ako*. But we cannot necessarily determine whether *ako* in (97) behaves as a serialized verb or preposition.

Therefore, while I maintain the criterion "morphologically marked on the verbal complex" for applicatives, I acknowledge that other types of clausal constructions without morphological marking, especially serial verb constructions, may be closely related to applicatives and play a role in the diachronic development of applicatives. Not only can AMs develop from independent morphemes, including verbs or prepositions, at other times, morphological marking for applicatives can be lost while the alternation in argument structure and meaning remain active in the language (see discussion of unmarked applicative analogs in Mualang in §5.8.5, which in the present-day language resemble English 'dative shift').

When viewed in broader typological perspective, certain serial verb constructions are thus quite closely related to applicatives; they share aspects of meaning (semantic relationships expressed) and may share some aspects of form (mapping of peripheral role to a clausal argument position), while differing in other aspects of form (type of predicative marking). Of course, not all serial verb constructions are related to applicatives in this way. In addition to expressing relationships commonly found for peripheral participants, e.g. beneficiary, instrument, or goal, serial verb constructions may indicate a sequence of events, manner of action, a wide range of postural, spatial and directional relationships, and various aspectual categories (see Lovstrand 2021). Serial verb constructions also exhibit a wide range of structural characteristics which are outside the scope of the discussion here, and the inventory of lexical verbs used in serial constructions often far exceeds the typical range for the inventory of AMs in a given language. Furthermore as seen in Tukang Besi, compared to applicatives, serial verb constructions may be less highly grammaticalized. Within the Austronesian family, serial verb constructions are more prevalent in East Nusantara languages, especially Oceanic languages of Melanesia (see Crowley 2002), as

compared to West Nusantara languages. So, while I will discuss serial verb constructions and similar analytic clausal constructions at times, this will be mostly restricted to situations in which they are relevant to the diachronic development of morphologically marked ACs, in keeping with the scope of this study.

3.6 West Nusantara applicatives in typological perspective

In light of the discussion in the two previous sections of this chapter, as a summary of sorts, in this section, I show how West Nusantara ACs may be related to one another and other types of verbal constructions in broad typological context. I do this using the framework of the understanding of a construction as an association between a fixed form and a conventionalized meaning that I have adopted. Some dimensions of form and meaning distinguishing different types of clausal constructions with similar meanings to applicatives are given in Table 3.7.

Table 3.7: Typological dimensions relating applicatives to similar clausal constructions

Form: Marking of predicate	Form: Mapping of peripheral role	Meaning: Semantic relationship
morphological marking	privileged syn. arg.	benefactive
analytic marking	core argument	locative
no formal marking	oblique	instrumental
	adjunct	comitative
	possessor phrase	stimulus
		...

For applicatives, morphological marking on the verb or verbal complex is a component of the form of an AC. This morphological marking (e.g. affixation) can be seen as a subtype of a larger category, which may be called ‘marking of the predicate.’ Compared to ACs, serial verb constructions, may be said to make use of analytic marking of the predicate, because verbs used in such constructions generally operate as independent words and in some cases other elements may intervene, or typically intervene, between serialized verbs. Another possibility is simply lack of formal marking on the predicate. This is shown in languages that allow unmarked alternations in clausal argument structure like that observed in ‘dative shift’ in English or the aforementioned Mualang examples.

Another component of form of an applicative is the mapping of peripheral semantic roles to a core argument position in argument structure. Applicatives were originally conceptualized because such mapping was considered unusual and observed to covary with the marking of the predicate with applicative morphology. This was first noticed by linguists in clausal constructions in Uto-Aztecan and Bantu languages, and then in many other language families. In Philippine-type languages, a peripheral semantic role always maps to the privileged syntactic argument, i.e. the pivot, in certain symmetrical voice constructions, e.g. LV, CV, IV, which I have called pivot-selecting applicatives. This is not really unusual when viewed within the grammatical systems of these languages, because all basic transitive voice alternations in the languages expressly target the pivot relation. Nonetheless, in a broader typology, it is reasonable to classify these type of constructions as applicatives. There are also some resemblances between pivot-selecting

applicatives and other applicatives in which peripheral participants are incidentally mapped to a privileged syntactic argument, e.g. the absolutive relation in ACs in certain Mayan languages. Pivot-neutral applicatives, on the other hand, show mapping of the peripheral role to a core argument. These show some resemblances to applicatives found in both ergative and accusative languages, depending on voice selection. Mapping of the peripheral role to an oblique argument or adjunct phrase is found in some types of non-canonical ACs mentioned in §3.2. Turning to serial verb constructions, peripheral roles in these constructions may be realized as an object of a verb, which is a kind of core argument, though this might require further comment, as more than one verb is present in the clause. Beneficiary participants, expressed as the possessor of another core argument, as in certain Sundanese benefactive constructions, might also be included as another possible mapping for the peripheral participant.

Finally, all of these clausal constructions, that is ACs—whether pivot-selecting, pivot-neutral, more canonical, or less canonical—together with serial verb constructions and unmarked dative alternations, have in common that they show special treatment of one or more peripheral semantic roles, irrespective of the specific syntactic realization of this participant. Languages differ as to which of these roles are targeted in ACs, though there is general agreement on which roles are most canonical. However, serial verb constructions as mentioned above, show more diversity in the semantic relationships expressed as part of the constructional meaning, e.g. temporal relationships, logical relationships, and more diverse and finely differentiated participant roles.

Again, the framework that I have articulated here is not certainly not the only useful way to view applicatives and the relationships of ACs to other types of clausal constructions that express similar meanings. My purpose here is to demonstrate the ways that such clausal constructions might be related to applicatives, if constructions are understood as conventionalized pairings of fixed form and consistent meaning. Furthermore, as outlined here, this framework is particularly suited for my purposes in studying the applicatives of West Nusantara languages in typological perspective. Viewing ACs in this manner not only allows for inclusion of diverse types of AM-marked clauses in the object of study, it also helps to situate them in the context of the wider Austronesian family and in the context of other functionally similar alternative verbal constructions. These in turn leads us to look for diachronic processes by which applicatives may arise from, or be replaced with, such functionally similar constructions in a given language.

In Part II of this study, I conduct a large-scale typological survey of the languages of West Nusantara based on this understanding of applicatives, and I examine the patterns of distribution observed and the implications that these patterns have for understanding the nature and development of applicatives in this region.

Part II: Typological Survey

Chapter 4

Introduction to the typological survey of applicative systems in West Nusantara

As discussed in Chapter 3, there are a number of significant gaps in the existing literature on ACs in languages of West Nusantara. Relatively few languages of the focus area are represented in typological studies of applicatives to date, and for these, there is reason to believe that variation in the syntactic and semantic properties of constructions marked with AMs is not adequately represented. This chapter and the two following describe a typological survey of the Austronesian languages of West Nusantara that is undertaken to address these gaps. In the remainder of this chapter, I describe the goals of the survey and aspects of the methodology used, including sampling and design of questionnaires. Two chapters on results of the survey follow, with some additional information about methodology for particular statistical tests integrated into relevant sections. In Chapter 5, I present results showing the distribution of applicatives in languages of West Nusantara, and the general and structural properties of languages with and without applicatives. In Chapter 6, I present results showing the distribution and properties of applicative systems of languages of West Nusantara, including the number and form of applicative morphemes (AMs), possible semantic roles of the applied phrase, and selected syntactic and semantic properties of ACs).

4.1 Goals of the survey

My goals for the typological survey are threefold. First, the most basic goal is to describe the distribution of languages with and without applicatives and key properties of the forms and meanings of ACs in languages which exhibit them. These findings are then used to inform the typology of ACs that I develop in Chapter 7. Second, I seek to determine whether there is a typological profile for languages with applicatives in West Nusantara, and if so, whether this profile corresponds to tendencies and structural correlations identified in previous studies (e.g. Peterson 2007; Polinsky 2013). Third, I aim to investigate patterns in the basic properties of form and meaning for applicatives constructions in the languages of West Nusantara, especially implicational relationships that explain how semantic and syntactic properties of ACs are distributed across West Nusantara.

Throughout the chapters discussing the typological survey, I am also concerned with the historical development of applicative systems in these languages. Thus, in discussion of findings related to the survey's three goals, I identify implications for our understanding of this development. Here it is important to note that the languages of West Nusantara show two types of ACs as defined in this study. First, in some languages of West Nusantara we observe Philippine-type peripheral nonactor voice alternations—*pivot-selecting applicatives* as defined in §3.4 above. These constructions are inherited from Proto Malayo-Polynesian (PMP) and reconstructable to Proto Austronesian (PAN). Second, in many languages we observe applicatives that operate independently of the symmetrical voice paradigm—*pivot-neutral applicatives*, as defined in §3.4. These constructions are not found in PMP. Because I am particularly interested with ACs that developed within West Nusantara, I take care to distinguish between the two types in discussing survey results, and at many points I focus on the latter.

4.2 Selection of languages

In this section, I describe the design of the sample used in the typological survey, and the selection process for languages to be included. A final sample of 85 total languages was used. An overview map in Figure 4.1 shows languages of the sample (exclusive of Merina Malagasy and Suriname Javanese) by the primary location in which they are spoken.¹

As described in §1.5, the scope of the study includes the Austronesian languages of West Nusantara and languages most-closely related to them. Thus, the languages included belong to genetic affiliations within Malayo-Polynesian that are identified as indigenous to the West Nusantara geographic area, even if the present-day communities that speak them are situated outside of this region, as is the case for the Malagasy languages, Suriname and New Caledonia Javanese, and Chamic and Malay languages spoken in parts of Southeast Asia north of the Malayan peninsula.

One inherent difficulty in using genetic affiliation in the study is the fact that there are outstanding questions about the higher-level ordering of the Malayo-Polynesian languages of West Nusantara. Indeed, the higher-level ordering of all the languages belonging to Malayo-Polynesian outside of its Central-Eastern branch—or perhaps even all those outside of Eastern Malayo-Polynesian—is still unclear (see Donohue & Grimes 2008; Smith 2017), and these probably rightfully belong to numerous primary branches. Therefore, for the typological survey, I relied on genetic groupings at a lower level under Malayo-Polynesian that are well-established. This means that for each grouping, it can be said that most scholars generally agree that it is a valid subgrouping and, with minor exception, the individual languages and/or primary branches classified within it are clear. In order to maximize inclusion of the diversity of the languages of West Nusantara and their ACs, I erred on the side of over-differentiation in selecting genetic groupings.

The sample of languages was taken from the total group of 321 languages with an assigned ISO 639-3 code that belong to the following genetic affiliations, according to the Ethnologue (Eberhard, Simons & Fennig 2021): Northwest Sumatra-Barrier Islands (12), Enggano (1), Tomini-Tolitoli (10), Kaili Pamona (16); Saluan-Banggai (6); Bungku-Tolaki (15); Muna-Buton (12), Wotu-Wolio, (5), South Sulawesi (30), Chamic (12), Malayic (42), Bali-Sasak-Sumbawan (3), Madurese

¹Throughout this study, maps showing language data include geospatial data from Glottolog 4.7 (Hammarström et al. 2022), which is used under the CC BY 4.0 license (<https://creativecommons.org/licenses/by/4.0/>).

(2), Sundanese (2), Javanese (5), Rejang (1), Lampungic (3), Nasal (1), Greater Barito (35), Land-Dayak (15), Melanu-Kajang (12), North Sarawak (41), Sabahan (35), and Rejang-Sajau (5). For each affiliation, primary branches for internal subgrouping and geographic regions in which the present-day languages are spoken were recorded. Also recorded were languages for which genetic affiliation is disputed, based on comparison of genealogical classification in the *Ethnologue* and *Glottolog* 4.7 (Hammarström et al. 2022).

An initial screening was conducted to evaluate available data for the languages based on the OLAC languages resources catalog² and the *Glottolog* 4.7 references section (Hammarström et al. 2022). From the initial list, 137 of the 321 total languages were removed because there were no resources relevant to the morphology or syntax of the language. For the remaining 184 languages, descriptive and pedagogical materials were reviewed, and a confidence rating was assigned for the availability and reliability of information to be collected in the survey across all available source material. The rating was based on type and length of resource, publication status, publisher, category of subject matter assigned in the source catalog, and cursory review of the material whenever possible. Of the 184 languages for which resources were reviewed, 76 received a high confidence rating for source material, and all of these were advanced to the data compilation stage. However, four of the 76 languages were subsequently removed during data compilation; one because of incomplete information (Malayic Dayak [xdy, mala1480]), two because they are too similar to very closely related varieties that have separate ISO 639-3 codes (Tukang Besi South [bhq, tuka1249], Standard Malay [zsm, stan1306]), and one because of discrepancies between the code listed and variety described (Bara Malagasy, [bhr, bara1369]).³ An additional 58 languages were given a medium confidence ranking for source material. From these, 13 were added to the sample in order to provide representation of an as yet unrepresented primary branch of a genetic grouping, geographic region for a genetic grouping, or set of languages whose inclusion in a genetic grouping is disputed. Thus, 85 total languages are included in the final sample.

Due to incomplete available information, the following were not represented in the sample: Mentawai [mwv, ment1249] and Simeulue [smr, sime1241], isolates within Northwest-Sumatra Barrier Islands; Limolang [ley, lemo1243], an isolate within South Sulawesi; Keninjal [knl, keni1248], an isolate within Malayic; Sanggau [scg, sang1339], an isolate within Land Dayak; Bintulu [bny, bint1246], an isolate within North Sarawak; New Caledonian Javanese [jas, newc1244], a Javanese language spoken in the Pacific region; Kajang (7 languages), a primary-branch of Melanau-Kajang; and Rejang-Sajau (5 languages), a genetic grouping listed by Eberhard, Simons & Fennig (2021). To compensate for this, relevant qualitative and descriptive data for a good number of languages with partially available information has been included in the discussion of survey results presented in Chapters 5 and 6.

The languages of the final sample are listed in Appendix A along with bibliographic references used to compile the survey data. They are also presented in the map in Figure 4.1. In the following sections, I describe the design of questionnaires used to compile the survey data.

²<http://www.language-archives.org/>

³Due to coding discrepancies, data compiled from the description of Bazaar Malay as spoken in Singapore by Aye (2005) was reassigned from Sabah Malay [msi, saba1263], under which this resource is listed in *Glottolog* 4.7, to Malay [zlm, mala1479], which is a label used for Colloquial or local Malay varieties. Like Sabah Malay, Singapore Bazaar Malay is used as a lingua franca in inter-ethnic communication and its characteristics are highly influenced by the conditions of contact under which it originated.

Figure 4.1: Languages included in the sample for the typological survey



Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

4.3 General properties of languages considered

For each language in the sample, two general properties of languages were recorded: geographic location and genetic grouping. These patterns are included in Questionnaire A (Appendix B.1).

4.3.1 Geographic location

Geographic location was coded in the survey according to major island grouping within West Nusantara and general region outside of West Nusantara. Coding values used for this pattern include six major island groupings: the Barrier Islands, Sumatra, Borneo, Java & Madura, Sulawesi, and the Lesser Sundas. Also used were five additional regions: the Philippines, East Africa, the Americas, the Pacific region, mainland Southeast Asia (i.e., the region of continental Asia directly north of Peninsular Malaysia). Languages were coded according to the primary present-day location in which they are spoken. Two languages were classified as ‘Other’: Standard Indonesian and Singapore Bazaar Malay.

4.3.2 Genetic grouping

Languages were coded by genetic groups identified above in §4.2, and values were assigned according to assigned classification in Eberhard, Simons & Fennig (2021). This source tends to favor long-standing and conservative classifications in the linguistic literature (and is updated less frequently compared to the Glottolog catalog). The following genetic groupings were represented in the sample as coding values: Northwest Sumatra-Barrier Islands, Enggano, Tomini-Tolitoli, Kaili Pamona, Saluan-Banggai, Bungku-Tolaki, Muna-Buton, Wotu-Wolio, South Sulawesi, Chamic, Malayic, Bali-Sasak-Sumbawan, Madurese, Sundanese, Javanese, Rejang, Lampungic, Nasal, Greater Barito, Land-Dayak, Melanau-Kajang, North Sarawak, and Sabahan.

4.4 Structural features of language considered

For each language in the sample, structural properties of the language were considered and recorded as part of Questionnaire A. These are broken into six categories: word order, morphosyntactic alignment, system of voice or diathesis, morphological case marking, order of Noun-Possessor, and morphological causative constructions. Each category is described in turn below including the motivation for inclusion and coding used. Note that some categories are represented by more than one item on the questionnaire, and thus made be coded for multiple patterns. These patterns were included in Questionnaire A (Appendix B.1)

4.4.1 Word order

Word order refers to the preferred order of the verb relative to other constituents in a main clause, particularly those expressing the core argument relations, S, A and P. Siewierska (1999) finds a generalized tendency for languages with applicatives to show verb-object word order.⁴

⁴Peterson (2007: 212-214) does not find a general correlation between word order and the presence of applicatives, however, the coding used did not distinguish for the relative position of P in the clause, only whether the verb is

This parameter is relevant to West Nusantara languages, as common applicative suffixal forms separate from voice marking are suggested to be derived from captured prepositions (see Pawley & Reid 1979, Adelaar 1992). If this is the case we might expect that the verb will precede the P argument and that the typical position of the A argument is elsewhere, not immediately following the verb.

In this survey, data on transitive word order was collected in order to determine if a correlation between V-P word order and the presence of applicatives holds for languages of West Nusantara. This was recorded as Pattern 1: *In basic transitive clauses (A-oriented) is the predominant word order Verb-P?* Coding values used for this pattern are listed below.

- **Y:** source indicates predominant word order is Verb-P, NP expressing A generally does not intervene
- **N:** source indicates predominant word order other than Verb-P
- **free:** there is no predominant word order

For this and all other patterns represented in the survey, if it was not possible to determine the appropriate value from the source material, no value was entered and the pattern was reported to be undetermined in the accompanying notes field. When this occurs, it is noted in the corresponding results section of Chapter 6 or 7.

4.4.2 Morphological and syntactic alignment

Alignment here refers the morphosyntactic properties of constituents that express the core argument relations S, A, and P, and ways in which these properties pattern. A major typological distinction is drawn between systems where the primary morphosyntactic indicators of grammatical relations pattern together for S and A, i.e., accusative alignment, from systems where these indicators pattern together for S and P, i.e., ergative alignment. Peterson (2007: 214–219) found a correlation between the presence of applicatives and non-accusative alignment and some evidence of correlation between the presence of applicatives and ergative alignment in particular. He suggests that this can be explained by the fact that the absolutive grammatical function (representing S and P) typically has relatively high topicality, which is compatible with the function of some types of applicatives. He notes however, that these findings are complicated by areal patterns, particularly the pervasive nature of ergative alignment in Australian languages, and the absence of it in Africa.

Because many languages of West Nusantara exhibit mixed patterns of alignment, in this survey morphological alignment and syntactic alignment were coded separately. Each is described in turn below.

Morphological alignment refers to patterning of the morphological marking of nominal constituents expressing the core argument relations S, A, and P. Evidence of morphological alignment may include the use of case markers, articles, or other grammatical particles, as well as paradigmatic alternation in the forms of nouns and pronominal indexes when these co-vary with the core argument relations.

initial, medial, or final.

Data on morphological alignment was recorded as Pattern 2: *With respect to morphological marking of core arguments in basic intransitive and transitive clauses, how do S, A, and P pattern?* Data for this pattern was recorded on the basis of basic, declarative clauses. If marking of core relations differed by subclass of nominal argument or indexing (e.g. by animacy, common vs proper noun, pronominal vs. full NP), the class that showed the most distinctions was used. Coding values used are given below. Of these, the value, “mixed-NP.IV.A”, was added because a large number of languages in the sample showed such mixed morphological alignment with distinct marking of the non-pivot A compared to other core arguments (see discussion of Pendau pronominal sets in §1.4.2).

- **accusative:** S and A are marked alike, to the exclusion of P
- **ergative:** S and P are marked alike, to the exclusion of A
- **core-oblique:** S, A, and P are marked alike (and distinct from non-core arguments)
- **split-S:** S is marked like A when agentive and P when patientive (see Mithun & Chafe 1999: 578)⁵
- **pivot-non-pivot:** S shares distinct marking with A in A-oriented transitive constructions, and P in P-oriented transitive constructions
- **mixed-NP.IV.A:** A of P-oriented transitive constructions shows special marking, otherwise S, A, and P are marked alike.
- **mixed (other):** marking of S, A, and P otherwise varies according to construction

Syntactic alignment refers to patterning of syntactic properties that co-vary with the core argument relations S, A, and P. Evidence of syntactic alignment includes word order and access to syntactic operations like relativization, control, and raising (see Keenan 1976 on ‘behavioral’ evidence, Arka 2003 for examples of a number of these applied in analysis of Balinese).

Data on syntactic alignment was recorded as Pattern 3: *With respect to syntactic properties of core arguments in basic intransitive and transitive clauses, what is the most prevalent patterning of S, A, and P?* I use the words “most prevalent” here to acknowledge that, in a given language, different syntactic operations may point to different patterns of alignment, and syntactic alignment patterns also may vary according to construction. My goal here is to identify distinctions between syntactic properties of S, A, and P that are highly pervasive in a given language and its structures. Coding values used for this pattern are listed below.

- **accusative:** S and A share special syntactic properties to the exclusion of P
- **ergative:** S and P share special syntactic properties to the exclusion of A
- **core-oblique:** S, A, and P alike share special syntactic properties distinct from those of non-core arguments
- **split-S:** special syntactic processes generally apply to equally to A and S when agentive, and equally to P and S when patientive

⁵The term “semantically aligned” may be a more appropriate in describing languages in which some S arguments are encoded like A and others like P (see Holton 2010). The incidence of such systems of alignment was uncommon in the sample, see §5.2.2.

- **pivot-non-pivot:** S shares special syntactic properties with A in A-oriented transitive constructions, and P in P-oriented transitive constructions
- **mixed:** no one primary patterning can be said to hold for syntactic properties of S, A, and P across basic intransitive and transitive clauses

4.4.3 Symmetrical voice and diathesis

Following Zúñiga & Kittilä (2019), I will use diathesis to refer to particular mappings of semantic roles to grammatical relations, and grammatical voice to refer to types of diatheses that are coded and thus formally marked on predicates. Voice and diathesis are of particular interest in developing a typology of applicatives in the focus area because these languages are a subset of the western Austronesian languages, and as such, many of them show distinctive voice systems that exhibit alternations between multiple basic transitive constructions. In these systems, which have been called western Austronesian voice, symmetrical voice, focus, or trigger, one argument in every clause is syntactically privileged, and each basic transitive construction is distinguished by morphological marking that indicates the semantic role of this syntactically privileged argument in the clause (see Chen & McDonnell 2019; Riesberg 2014b).

Voice and applicatives are connected in western Austronesian languages in a number of important ways. Note that, in the literature on western Austronesian languages, a major typological division has been drawn between Philippine-type languages and Indonesian-type languages (Wolff 1996; Arka 2003; Himmelmann 2002, 2005). The Philippine-type languages, on the one hand, show alternations between one actor voice (AV) and two or more nonactor voices, e.g. patient voice (PV), benefactive voice (BV), locative voice (LV), instrument voice (IV), and/or circumstantial voice (CV). The Indonesian-type languages, on the other hand, show alternation between just two voices, one actor voice (AV) and one nonactor voice (PV). Philippine-type languages are relevant to applicatives because many of the nonactor voices are considered (pivot-selecting) ACs under the definition used in this study (see §3.4). Indonesian-type languages are relevant because the presence of (pivot-neutral) ACs that operate independently of symmetrical voice morphology has been identified as characteristic of this type in the literature (Himmelmann 2005). However, there may also be reasons to consider the Indonesian-type problematic and in general, too narrow to be appropriately descriptive of non-Philippine-type languages of West Nusantara (see §5.10 below).

In order to clarify the typological relationship between voice and the presence of applicatives in the languages of West Nusantara, data on grammatical voice and diathesis were recorded as Pattern 4: *What is the pattern of voice and/or diathetical alternations found in basic transitive clauses in the language?* Coding values for this pattern were defined according to the number of basic transitive constructions observed in the language, and are as follows.

- **Philippine-type:** Three or more basic transitive constructions are observed, distinguished by mapping of semantic role to grammatical relation. These include at least two distinct nonactor oriented constructions.
- **marginal-Philippine-type:** Three or more transitive constructions are observed, distinguished by mapping of semantic role to grammatical relation. However, one or more of these is non-productive or substantially restricted in distribution.

- **two-way-symmetrical:** Two basic transitive constructions are observed, one A-oriented, and one P-oriented.
- **marginal-two-way-symmetrical:** Two transitive constructions are observed, one A-oriented, and one P-oriented. However, one of these is non-productive or substantially restricted in distribution.
- **asymmetrical:** No alternations in basic transitive constructions are observed which are distinguished by mapping of semantic role to grammatical relation.

4.4.4 Morphological case marking

In previous research Polinsky (2013) observes that “applicatives are commonly found in those languages that have little or no case-marking of noun phrases in a clause” and Peterson (2007: 222–223) finds a tendency for languages with applicatives to lack case markers. Meanwhile for western Austronesian languages the use of preposed case marking particles is associated with Philippine-type languages (Himmelmann 2005), but it is not clear whether loss of morphological case-marking is related to the emergence of pivot-neutral applicatives.

To determine whether a relationship between lack of case-marking and the presence of applicatives holds in the focus area, data were recorded as Pattern 5: *Is morphological case-marking used in the language?* Coding values for this pattern are listed below.

- **Y:** Morphological elements (or morphological processes) that attach to nouns or NP constituents are generally used to indicate grammatical or spatial relations.
- **Limited:** Case-marking morphological elements are found to attach only to a small subclass of nominals, or case is marked morphologically by means of sets of pronominal elements, but is not marked on nouns generally.
- **N:** No morphological elements (or processes) are found that attach to nouns or NP constituents and indicate grammatical or spatial relations.

4.4.5 Order of Noun-Possessor

The linear order of possessed noun and possessor is one indicator of the extent of head or dependent marking in a language (see Nichols 1992). Peterson (2007) finds that languages with applicatives show a tendency to be “extremely dependent marking.” A preposed possessor construction in which the possessor precedes the possessed noun (with or without a linker element) has also been put forward as a defining characteristic of a typological category for western Austronesian languages that contrasts with the category for symmetrical voice languages (see Himmelmann 2005: 112–114, 163–165). However, note that most preposed possessor languages of this sort are found in East Nusantara and outside of the focus area to the east.

To investigate a possible relationship between order of noun and possessor and the presence of applicatives in the sample, data was compiled as Pattern 6: *What is the order of possessed noun and possessor in the language?* Coding values are given below.

- **N-Poss:** The possessed noun generally precedes the possessor.
- **Poss-N:** The possessor generally precedes the possessed noun.
- **mixed:** Both orders are observed.

4.4.6 Morphological causative constructions

The relationship between applicatives and other valency-modulating constructions has been of interest in the literature. Peterson (2007) found a tendency for languages with ACs to also show morphological causative constructions. Furthermore, AMs in the languages of West Nusantara also very commonly serve as markers of morphological causative constructions (see Truong & McDonnell 2022) and various explanations for this have been explored in previous literature (see e.g. Shibatani & Pardeshi 2002).

Here, causative refers to a clausal construction in which an A argument is introduced with the semantic role of causer, i.e., the participant that instigates the event (Dixon & Aikhenvald 2000). Morphological causative means that morphological marking on the verb or verbal complex is associated with such argument structure.

In addition to AMs that function as morphological causative markers, many languages of West Nusantara show a reflex of PAn and PMP *pa- ‘causative of dynamic verbs’ (R. Blust 2003), which is a well-established reconstructed form. Of potential interest is distributional evidence that may shed light on the relative recency of the causative functions of AMs (e.g. compared to applicative functions) and the internal history of causative constructions in these languages, including possible functional replacement of causative *pa- with AMs.

To investigate the relationship between morphological causatives and applicatives in the focus area, data was compiled in Patterns 7a, 7b, and 7c.

Pattern 7a: *Does the language have a morphological causative construction?*

- Y: The language has a morphological causative construction.
- N: The language does not have a morphological causative construction.

Pattern 7b: *Which of the following are sources of morphological causative markers in the language?*

- **appl**: Causative morphology shares the same form with an AM in the language .
- ***pa-**: Causative morphology in the language apparently derives from PMP *pa- ‘causative’.
- **other**: Causative morphology in the language apparently derives from some other source.
- **NA**: The pattern is not applicable because the language has no morphological causative construction.

Pattern 7c: *What is the productivity of morphological causatives derived from *pa-?*

- **high**: productive with a large number of lexical roots across multiple syntactic categories (transitive verbs, intransitive dynamic verbs, stative verbs, nouns, etc.)
- **medium**: productive with a large number of lexical roots that primarily belong to one syntactic category
- **low**: only attested with a limited number of lexical roots.
- **NA**: There is no morphological causative derived from *pa-.

4.5 Properties of applicative systems considered

For each language in the sample, I determined whether applicatives are present or absent based on available descriptions. If applicatives are present, information was collected about general properties of the applicative system (included in Questionnaire A, see Appendix B.1), and features of individual ACs (included in Questionnaire B, see Appendix B.2). The rationale and system of coding used for each pattern related to general properties of applicatives are described in this section, followed by properties of AMs in §4.6.

4.5.1 Presence of applicatives

The presence of ACs in a language was determined according to the definition given in (1) above, which is repeated here for convenience.

- (1) An applicative is a clausal construction in which overt morphological marking on the verbal complex coincides with the selection of a non-agent, non-patient semantic role to map to a core argument in the clause.

This feature was recorded in the survey as Pattern 8a: *Does the language have applicative constructions?* Coding values used are listed below. Note that if it was not possible to determine the presence of applicatives in a language from review of the source material during the initial screening, the language was not considered eligible for inclusion in the language sample (see §4.2).

- Y: ACs as defined in (1) are identifiable in the language based on the source material.
- N: ACs as defined in (1) are not identifiable in the language based on the source material.

4.5.2 Co-occurrence with other constructions

As mentioned above, certain constructions considered applicatives in this study represent alternations integral to the voice paradigms of Philippine-type languages (i.e., pivot-selecting applicatives), while other constructions considered applicatives are not part of the symmetrical voice paradigm (i.e., pivot-neutral applicatives). The former type generally is mutually exclusive with all other basic transitive voice constructions, e.g. AV, PV, and other peripheral nonactor voices, while the latter type is not.

To aid in development of a typology of possible interactions between applicatives and voice/diathesis in the focus area, data was collected in Pattern 8b: *Does applicative marking in the language freely co-occur with the basic transitive constructions of the language and the passive (if applicable)?* Coding values are given below.

- Y: At least one of the AMs in the language co-occurs with the major diathetical alternations of the language, including all the basic transitive constructions and any passive construction.

- **N-Partial:** None of the AMs in the language co-occurs with all the basic transitive constructions of the language and any passive constructions. However, at least one AM co-occurs with more than one such construction, under certain conditions.
- **N:** The AMs in the language do not co-occur with other basic transitive constructions of the language or the passive construction.
- **no-appl:** There are no applicatives, so this pattern is not applicable.

4.5.3 Indexing of the applied phrase

Of particular interest in the literature has been properties characteristic of P (often called object properties) and whether they are exhibited by the applied phrase in an AC (see Baker 1988b; Bresnan & Moshi 1990, 1993; Alsina & Mchombo 1993; Peterson 2007: 51–60).

In languages which show agreement marking or argument indexing on the verb for P, the presence or absence of such marking for the applied phrase may be one criterion or parameter used to distinguish types of ACs (see Peterson 2007: 51–56; Pacchiarotti 2020). Here, I will primarily use the term (person or argument) *indexing*, meaning the presence of bound forms on the verb or verbal complex that index arguments of the verb and express person and/or number features in the manner used by Haspelmath (2013). This terminology is more relevant for the languages of West Nusantara, as person indexing in these languages is usually non-obligatory, while agreement is defined by many authors as an obligatory marking. In addition, a small number of languages in the focus area make use of person indexing with optional conominals (i.e., cross-referencing), while the majority only show person indexing in the absence of a conominal.⁶

Data on patterns of argument indexing and applicatives were collected as Pattern 9: *In applicative constructions, does the applied phrase show person-indexing on the verb in the manner generally characteristic of P in monotransitive clauses?* Here, applied phrase refers to the peripheral semantic role selected as a core argument in an applicative construction. Coding values are given below.

- **Y:** The applied phrase generally shows person indexing on the verb in the same manner as other P arguments in the language.
- **N:** The applied phrase generally does not show person indexing on the verb, even though other P arguments in the language do show such indexing.
- **mixed:** Under certain conditions but not others, the applied phrase shows indexing of person and number features on the verb in the same manner as other P arguments in the language are indexed.
- **no-index:** The pattern is not applicable because the language does not generally make use of person indexing for P arguments.
- **no-appl:** The pattern is not applicable because the language does not have applicatives.

⁶In West Nusantara languages, indexing on the verb for number features apart from person features is found in some languages in certain constructions, e.g. Sundanese *-ar-* construction marking plural actor (Kurniawan 2013: 23–24). Such constructions are used fairly infrequently, and indexing that expresses number only is never a general feature of verbal clauses.

4.5.4 Syntactic properties of the applied phrase

Beside argument indexing, access of the applied phrase to syntactic operations that are otherwise allowed for P has been used to distinguish types of ACs (Peterson 2007).

Since it is not particularly common for syntactic diagnostics to be reported for ACs in descriptive grammars, syntactic properties of the applied phrase are addressed in this survey using criteria that are possible to evaluate based on surface forms and grammatical categories that are more likely to be overtly discussed in available sources. Data on syntactic properties of the applied phrase were compiled in Patterns 10 and 11, which are presented in turn below.

Pattern 10: *Does the applied phrase show evidence of status as a syntactically privileged argument across basic clause types which co-occur with applicative constructions?* Evidence here refers to morphological marking, word order, and/or access to syntactic operations as discussed in the source material or possible to evaluate from data presented in the sources. Coding values are given below.

- **Y-obl:** Yes, there is evidence that the applied phrase always holds a syntactically privileged relation across basic clause types with which the AC co-occurs.
- **Y-opt:** There is evidence that the applied phrase generally holds a syntactically privileged relation in P-oriented basic clause types.
- **mixed:** There is evidence that the applied phrase may hold a syntactically privileged relation in basic clause types; this co-varies according to the form of the AM and/or semantic role of the applied phrase.
- **N:** The applied phrase generally does not appear to hold a syntactically privileged relation in the clause, or does so only in marked, infrequent constructions.
- **no-appl:** The pattern is not applicable because the language does not have applicatives.

Pattern 11: *Does the applied phrase have access to relativization?* Coding values are given below.

- **Y:** Yes, the applied phrase may generally be the head noun of a relative clause, subject to the same conditions as other P arguments.
- **N:** No, the applied phrase generally may not be the head noun of a relative clause.
- **mixed:** The applied phrase sometimes may be the head noun of a relative clause, but this is subject to additional conditions not observed for other P arguments.
- **no-appl** This pattern is not applicable because the language does not have applicatives.

4.6 Properties of applicative morphemes considered

Features in this section were recorded by language and AM. The features included here were compiled in large part on the basis of prior studies on much smaller samples languages of West Nusantara (Truong & McDonnell 2022, McDonnell & Truong 2024). In some languages, the form of an AM was reported to undergo morphological alternation, e.g. co-variance of the form with certain TAM categories in the morphological paradigm for verbs (see 5.2.3 for examples of some

such allomorphy). In such cases, the allomorphs were considered to represent one AM for the purpose of this section of the survey.

4.6.1 Applicative morphemes and source morphology

As mentioned above, some applicative suffixal forms common in West Nusantara are thought to be historically derived from captured prepositions (see Pawley & Reid 1979, Adelaar 1992). However, voice markers are another possible source of such morphology in the focus area (Adelaar 2011; Truong & McDonnell 2022), and causative markers may also play a role (McDonnell & Truong 2024). Cross-linguistically, verbs are also identified as a common source (Peterson 2007: 124), however, in West Nusantara, to my knowledge these are primarily seen in analytic constructions, see discussion of analytic benefactive constructions in §5.7). To investigate patterns of development of applicative morphology in the focus area, information identifying source morphology and its function was collected as follows below. Coding values have been taken from commonly identified source morphology for valency modifying morphology in PMP (see Sirk 1996 for discussion of various proposals for source morphology and R. Blust 2003: 472–474 for a list of reconstructed affixes and clitics for PMP).⁷

Pattern A: *What is the apparent source morphology from which the applicative morpheme is derived?*

- ***akən**: The AM in question appears to derive from an earlier preposition *akən.⁸
- ***i**: The AM in question appears to derive from PMP *i.⁹
- ***-an/*-ən**: The AM in question appears to derive from PMP voice morphology *-an ‘locative voice’ or *-ən ‘patient voice’.
- ***pa-**: The AM in questions appears to derive from PMP *pa- ‘causative’ or a combination of *pa- and another prefix, e.g. *paka-, popa-, etc.
- **other**: The AM in question appears to derive from another known historical source.
- **undetermined**: The historical source of the applicative marked in question is not clear.

Pattern B: *What was the category of the historical source morphology from which the applicative morpheme is derived?*

- **adpos**: The source morphology was an adposition.
- **verb**: The source morphology was a verb.
- **caus**: The source morphology was a causative marker.
- **case-mrkr**: The source morphology was a case marker.

⁷The phoneme representing schwa in PAN and PMP is traditionally written as *e, though some authors also use *ə. For clarity, I use *ə in PAN and PMP reconstructions.

⁸Very few sources gloss the proposed reconstructed form *akən. Pawley & Reid (1979: 14) treat *akən as a preposition marking ‘accessory case’. Malay *akan*, which means ‘concerning, regarding, about’, is another point of reference as it is commonly discussed as a possible cognate form of *akən (see Sirk 1996; Adelaar 2011).

⁹Such forms may be attributed either to the locative preposition *i or the locative voice imperative/negative suffix *-i in PMP (see further discussion in §6.2.2).

- **voi-mrkr:** The source morphology was a voice marker.
- **other:** The source morphology is known to have a different category or function than that listed above.
- **undetermined:** The category or function of the source morphology is not known.

4.6.2 Semantic role of the applied phrase

For each AM, information was collected about the possible semantic role of the applied phrase in constructions marked with the morpheme. Each semantic role attested in the source material, including textual examples and descriptive accounts, was coded accordingly.

Pattern C: *What are the semantic roles of the applied phrase in constructions bearing the applicative morpheme?*

- **BEN:** Beneficiary, i.e., a participant who accrues a benefit through an event or state of affairs.
- **REC:** Recipient, i.e., an entity that receives possession (physical or otherwise) of an entity
- **LOC:** Static locative, i.e. the static or generalized location of a state or event
- **GOAL:** Goal, i.e., the end point of an entity that changes location in a motion event.
- **THM:** Theme, i.e., an entity undergoes a change of location or is located in space
- **INST:** Instrument, i.e., an inanimate entity manipulated to some effect in a event
- **STIM:** Stimulus, i.e., the object of an act of perception, stimulus of an emotional response
- **CONT:** Content, i.e., the content of an act of speaking or cognition.
- **CIRC:** Circumstantial, i.e., the reason or purpose for a state or event.
- **COM:** Comitative, i.e., a participant that accompanies an actor or mover.
- **ADDR:** Addressee, i.e., a participant that is the intended receiver of some communication

4.6.3 Co-occurrent marking with other morphology

Under certain conditions, it has been noted that the AM obligatorily combines with another morphological marking on the verb, besides that normally found on basic transitive constructions. This typically marks only a subset of ACs, i.e. with certain modes, certain semantic roles for the applied phrase, or both (for an example, see the discussion of Pendau applicatives constructions marked with both applicative suffixes and ‘stem-forming’ morphemes in §5.2.3). This information was recorded in the survey as follows.

Pattern D: *Must the applicative morpheme co-occur together with another morpheme on the verbal complex in certain applicative constructions?*

- **Y-SF:** Yes, the AM appears together with a “stem-former” in certain ACs. The stem-former does not have a semantic meaning of its own, but it may have a syntactic or morphological function.
- **Y-CAUS:** Yes, the AM appears together with a causative morpheme in certain ACs. The causative morpheme has an independent causative function in other constructions.
- **Y-other:** Yes, the AM appears together with some other morphological marking on the verbal complex that is not generally found in basic transitive constructions.
- **N:** No, the AM generally does not co-occur with other morphological marking on the verb necessary to form the ACs.

4.6.4 Other functions and constructional meanings

It has been widely noted that AMs in West Nusantara languages are polyfunctional, and may also mark aspectual and causative meanings, among others. To track patterns of distribution of such functions, the following information was collected for each AM in a given language.

Pattern E: *Which of the following non-applicative functions are also marked by the applicative morpheme?*

- **CAUS:** The AM also forms causative constructions, in which an instigating causer participant is introduced and selected to map to A.
- **CAUS-combo:** The AM co-occurs with another morphological causative on the verbal complex in certain (non-applicative) causative constructions.
- **PLUR:** The AM indicates pluractional aspect, e.g. durative, habitual, iterative, multiple actors or undergoers.
- **INTENS:** The AM indicates greater intensity, e.g. greater volitionality, greater application of force, etc.
- **COMP:** The AM indicates comparative degree of a gradable quality.
- **sem-change:** The AM is associated with some other semantic change in the meaning of a verb, e.g. ‘discard’ cf. ‘throw’, ‘elope’ cf. ‘run’.
- **none:** No non-applicative functions are indicated from the source material.

4.7 Summary

In this chapter I have provided an overview of methodology used in sampling languages for the typological survey, and design and implementation of the questionnaires used for collection of data. An example of the coding scheme for properties of language and properties of applicative systems is given below in (98) for Sundanese. The coding scheme for properties of AMs is illustrated below in (99) for the Sundanese applicative suffix *-an. In the following two chapters, results of the survey are presented, along with additional notes on methodology used in analysis of data, as relevant.

(98) Sundanese

Genetic group: Sundanese

Location: Java

Word order: Y (AVP)

Morphological alignment: mixed-NPIV.A

Syntactic alignment: pivot-non-pivot

Symmetrical voice: two-way symmetrical

Morphological case-marking: N

Order of noun and possessor: N-Poss

Presence of morphological causative: Y

Source of morphological causative: appl

Productivity of reflex of *pa-: NA

Presence of applicatives: Y

Co-occurrence with other constructions: Y

Indexing of the applied phrase: Y

Syntactic privilege and the applied phrase: Y-opt

Access to relativization: Y

(99) Sundanese applicative morpheme *-an*

Source morphology: *-an/*-ən

Source morphology type: voi-mrkr

Semantic roles of the applied phrase: REC, LOC, GOA, STIM, CONT

Co-occurrent marking with other morphology: N

Other functions: CAUS, PLUR, COMP, sem-change

Chapter 5

Results: Distribution and properties of languages of West Nusantara with and without applicatives

In this chapter, I present results of the survey showing the distribution and properties of languages of West Nusantara with and without applicatives by type. In §5.1, I present an overview of results. In §5.1.1, results of statistical tests of non-random association between individual features of the languages and the presence or absence of applicatives are summarized. Then in §5.1.2, I present results of a multivariable modeling analysis using the random forest classification algorithm to show the relative importance of each feature in predicting whether a language has applicatives or not. In §5.2, I show detailed results for individual structural properties of languages. These are generally of less importance for predicting the presence or absence of applicatives in West Nusantara languages, though morphological alignment and type of voice system do show statistically significant non-random association with occurrence and type of applicatives. Following this, I consider the features of greatest importance in predicting the presence or absence of applicatives in West Nusantara: location and genetic affiliation. A summary of the distribution of applicatives by location and genetic affiliation is given in §5.3, followed by detailed presentations of results for languages of Sumatra and the Barrier Islands (§5.4), Java and Madura (§5.5), the Lesser Sundas (§5.6), mainland Southeast Asia and Peninsular Malaysia (§5.7), Borneo and the Southern Philippines (§5.8) and Sulawesi (§5.9). The chapter concludes with a summary of major findings in §5.10.

On the whole, these results show significant areas of difference with some previous cross-linguistic studies of applicatives, and suggest that pivot-neutral applicatives are an areal feature of West Nusantara associated with the breakdown of the inherited four-way Philippine-type voice system. Nonetheless, the development of pivot-neutral applicatives may be blocked or eroded by special factors, including phonological and morphological changes affecting word structure, and associated shifts towards analytic syntactic structures. Such changes leading to a complete lack of applicatives may be triggered or spread by language contact, and are found primarily in two sets of languages in survey, one centered on mainland SE Asia, and a second centered on Borneo south of Sabah.

5.1 Overview of results and multivariable analysis

In order to determine whether there is a typological profile for languages of West Nusantara with applicatives, and whether this profile is consistent with tendencies and correlations with properties of language found in previous studies, I considered nine features. Two of these, geographic location, and genetic affiliation, are general properties of language (see §4.3 above). The other seven features represent structural properties of language: verb-P word order, pattern of morphological alignment, pattern of syntactic alignment, type of voice system, presence of case-marking, order of noun and possessor, and presence of a morphological causative construction (see §4.4 above). In this section, I give an overview of results for two types of analysis done to determine the relationship of these nine features vis-a-vis the distribution of applicatives in languages of West Nusantara: (i) tests of nonrandom association for individual features (§5.1.1), and (ii) multivariable analysis using the random forest algorithm (Breiman 2001) for classification problems (§5.1.2). All statistical analyses described in this section were performed using R Statistical Software (v4.2.3, R Core Team 2023).

5.1.1 Testing of individual features for association with applicatives

For individual features, Fisher's exact test was used to determine if there is a nonrandom association between the feature and (i) the presence of applicatives of any type in the language, and (ii) the presence of pivot-neutral applicatives in the language. Fisher's exact tests were used rather than chi-squared tests because of low expected values ($n < 5$) for at least one categorical variable, i.e., coding value, for each feature. In such cases, the approximation of frequencies used in chi-square tests may be inadequate. Except for the voice system feature, which does not have an independent relationship with the general presence of applicatives, two Fisher's exact tests were performed for each of the nine features of language listed above.¹ An alpha level of .05 was used for all statistical tests. Results of Fisher's exact tests are reported in Table 5.1 as p -values (two-sided).

The results of Fisher's exact tests conducted indicate that there is a statistically significant association between the presence of applicatives of any type and location ($p < .001$), genetic group ($p = .002$), verb-P word order ($p = .014$), and morphological alignment ($p = .036$). Likewise, the results indicate a statistically significant association between the presence of pivot-neutral applicatives and location ($p < .001$), genetic group ($p < .001$), and morphological alignment ($p < .001$). However, no statistically significant association was found between the presence of pivot-neutral applicatives, and verb-P word order ($p = .753$). In addition to these features, a statistically significant association is also indicated between the presence of pivot-neutral applicatives and type of voice system ($p = .009$). Detailed results for individual features representing structural properties of language are given in §5.2 below.

¹Note that no test was performed for the relationship between the voice system feature and the general presence of applicatives. This is because the voice system feature is coded according to the number of transitive voice constructions distinguished by mapping of semantic role in the language, and all languages coded as 'Philippine-type' for this feature have more than two distinct undergoer voices and thus always have pivot-selecting applicatives as defined in this study.

Table 5.1: Results of Fisher’s exact tests by feature

Feature	Presence of applicatives (any type) <i>p</i> -value (2-sided)	Presence of pivot-neutral applicatives <i>p</i> -value (2-sided)
Location	< .001	< .001
Genetic group	.002	< .001
Verb-P word order	.014	.753
Morph. alignment	.036	< .001
Syn. alignment	.873	.806
Voice system	—	.009
Case marking	.158	.202
Order of N + Poss	.091	.167
Morph. causative	.081	.155

5.1.2 Multivariable modeling of features: Random forest analysis

The statistical tests described above were conducted to determine whether a nonrandom association exists between each individual feature and the presence of applicatives in the survey data. Results from such testing are useful for comparison with previous studies identifying tendencies or patterns of correlation for the presence of applicatives based on analysis of similar individual features. However, to develop an adequate explanation of the distribution of applicatives in languages of West Nusantara, we also want to know the relative importance of features surveyed, and the extent to which they are collectively useful for classifying languages into relevant types or groups (i.e., languages with and without applicatives generally, or languages with and without pivot-neutral applicatives specifically). For this purpose, a random forest analysis was performed to model the distribution of languages with and without applicatives in the survey using classification trees based on subsets of the features of language listed above. Random forest analysis uses a machine-learning algorithm to yield estimates of overall and relative importance across many variables in classification tasks, as well as error estimates for accuracy of models using these variables in correctly classifying observations into categories (Breiman 2001). This analysis was performed using the `randomForest` R package (Liaw & Wiener 2002).

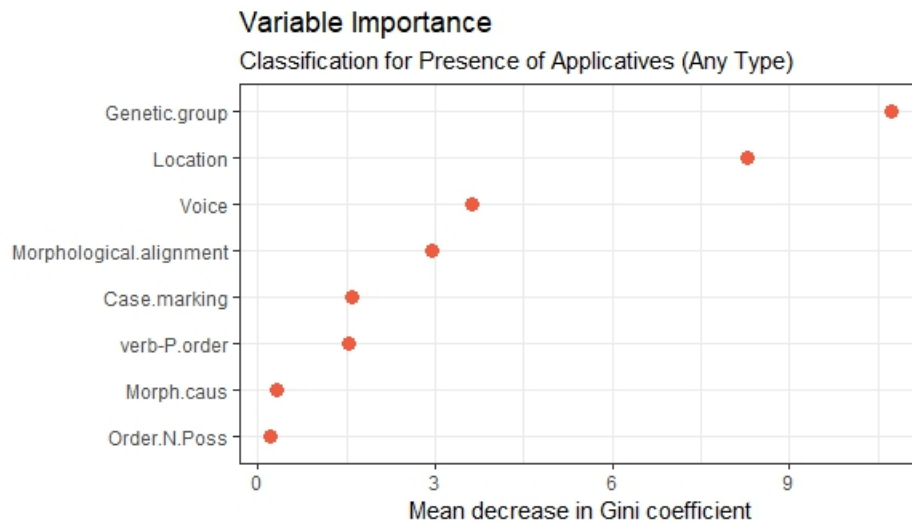
The feature for syntactic alignment was excluded from the random forest analysis because of a large number of missing values (undetermined for 13 of 85 languages) and large *p*-values on the Fisher’s exact tests indicating low probability of nonrandom association with the presence of applicatives of any type ($p = .873$) and with the presence of pivot-neutral applicatives specifically ($p = .806$). Accordingly, eight features of language were included in the random forest analysis: location, genetic grouping, verb-P word order, morphological alignment, voice system, case marking, order of Noun + Possessor, and presence of a morphological causative. As with the tests of individual features described above, two target variables were tested: (i) the presence of applicatives of any type in the language, and (ii) the presence of pivot-neutral applicatives with marking separate from symmetrical voice morphology.

Parameters for the random forest analysis were set based on typical default values identified in

the literature on statistical methods and machine learning (see Probst, Wright & Boulesteix 2019). Because the target variable, i.e., presence of applicatives generally, or presence of pivot-neutral applicatives, is categorical, a classification type tree, rather than regression type tree, was selected for use. For the parameter *mtry*, the number of variables to randomly sample as candidates at each split in the tree, the value 3 was used, based on the default value for classification (square root of the total number of variables, in this case $\sqrt{8} = 2.82\dots$). The standard node size for classification, 1, was also used. The number of trees to generate was set to 2000 as a generally high value as 500 or 1000 is normally the default value.

For the presence of applicatives of any type in a language, the following results were found. The out-of-box estimate of error rate was 16.46%, meaning that the trees generated by the random forest algorithm on average misclassify 16.46% of the languages, and accurately classify the other 83.54%. Variable importance is reported in Figure 5.1.2 as mean decrease in Gini coefficient, which is a measure of how each variable contributes to homogeneity in nodes below when chosen for a split in the decision tree. Variables with larger Gini coefficients produce splits in the tree that result in more accurate classification of observations into groups by the target variable; in this study, such variables thus have higher explanatory power for the presence or absence of applicatives. The two most important variables by far are genetic group and location, with mean decreases in Gini coefficient of 10.723 and 8.274 respectively. Of somewhat less importance are voice and morphological alignment, with mean decreases in Gini coefficient of 3.613 and 2.929 respectively. The four other variables are of relatively low importance, with mean decrease in Gini coefficient of less than 2.

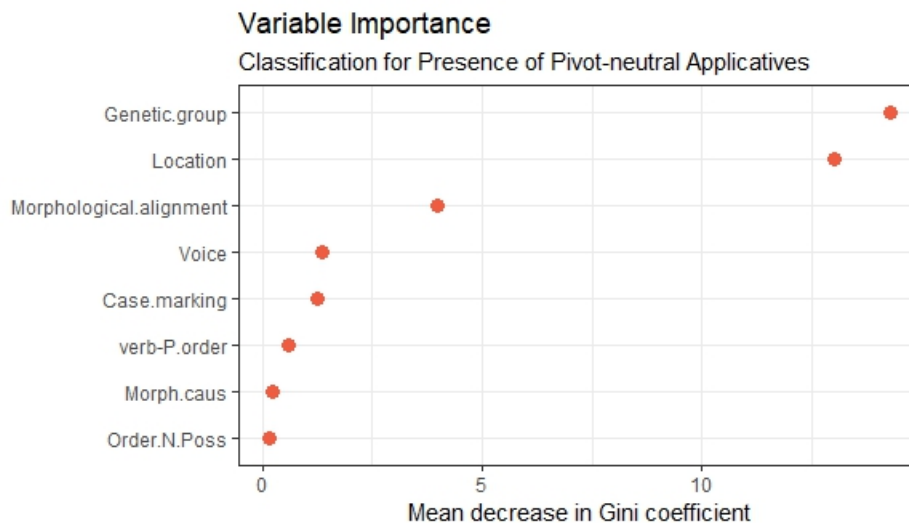
Figure 5.1: Importance of features for classification by presence of applicatives



For the presence of pivot-neutral applicatives, the following results were found. The out-of-box estimate of error rate was 18.99%, meaning that the trees generated by the random forest algorithm on average misclassify 18.99% of the languages, and accurately classify the other 81.01%. Variable importance is reported in Figure 5.1.2. Again, the two most important variables are genetic group and location, though this time genetic group is the higher of the two. The mean decreases in Gini coefficient for location is 13.023, while for genetic group it is 14.280. The feature

for morphological alignment shows a mean decrease in Gini coefficient of 3.991. The other five variables are of relatively low importance, with mean decrease in Gini coefficient of less than 2.

Figure 5.2: Importance of features for classification by presence of pivot-neutral applicatives



Overall, the random forest analysis shows that both the presence of applicatives generally and the presence of pivot-neutral applicatives specifically can be predicted in large part on the basis of the location in which a language is presently spoken, i.e., major island group within West Nusantara or general region outside of West Nusantara, and genetic affiliation within Malayo-Polynesian. While testing of individual features showed that morphological alignment and voice have a statistically significant non-random relationship with the presence of applicatives in the sample, these features are less useful for predicting if a given language of West Nusantara will show applicatives of any type or lack them, and only morphological alignment is moderately useful for predicting the presence of pivot-neutral applicatives. Of even less utility are the features for case marking, verb-P word order, presence of morphological causatives, and order of Noun + Possessor. In the following sections, the patterning of distribution of languages with and without applicatives is broken down in greater detail by sets of individual features. Structural properties of language are considered in §5.2, while location and genetic affiliation are considered at length in §5.3 through §5.9

5.2 Detailed results for structural properties of language

In this section, detailed discussion of the results is presented for individual features of languages, especially those that show a statistically significant nonrandom association with the presence of applicatives (all types) or the presence of pivot-neutral applicatives.

5.2.1 Word order

In contrast to previous studies, the results of the survey do not support a tendency for languages with applicatives to show verb-P word order. In fact, as shown in Table 5.2, languages of the

sample with predominant word order other than verb-P are more likely to lack applicatives of any type, and this association is statistically significant ($p = .014$). For four languages of the sample, predominant word order was undetermined. Of the remaining 81 languages for which this pattern was recorded, 36% of languages with Verb-P word order lack applicatives altogether (25 of 69), while no languages with other word orders lack applicatives (0 of 12). On the other hand, there is no statistically significant association between verb-P word order and the presence or absence of pivot-neutral applicatives specifically ($p = .753$).

Table 5.2: Verb-P word order and presence of applicatives

Word order (A-oriented)	Presence of applicatives (any type)		Presence of pivot- neutral applicatives		TOTAL
	Yes	No	Yes	No	
Verb-P	44	25	39	30	69
Other than verb-P	12	0	8	4	12
TOTAL	56	25	47	34	81
%	69	31	58	42	

Further review of the source material reveals that languages that lack applicatives are far more likely to have predominant AVP order in A-oriented transitive constructions. In fact, as shown in Table 5.3, AVP is the predominant word order in A-oriented transitive constructions for 23 of 25 languages lacking applicatives (92%), compared to 28 of 56 languages with applicatives of any type (50%). There are 19 languages of the sample showing predominant verb-initial word order, and these overwhelmingly have applicatives (95%, 18 of 19). No languages of the sample show predominant verb-final word order, while for 11 languages, the position of the verb can be either initial or medial in normal word order for A-oriented constructions.

Since we know that Philippine-type voice alternations—and thus pivot-selecting applicatives—were present in PAn and PMP, these results cannot indicate that languages with AVP word order are simply less likely to develop applicatives of any type. Instead, I interpret this pattern to show that in languages of West Nusantara that underwent certain types of restructuring, a shift from the verb-initial word order prevalent in PAn to AVP word order coincided with loss of Philippine-type voice, as well as other factors that mitigated against the development or maintenance of pivot-neutral applicatives. That is, while the presence of applicatives is the norm for most languages of West Nusantara, AVP word order in the A-oriented construction may form part of one or more typological profiles for languages that lack applicatives in the sample. For further discussion of word order and lack of applicatives, see §5.7 on mainland Southeast Asia and Peninsular Malaysia and §5.8 on Borneo.

5.2.2 Morphological and syntactic alignment

As reported in §5.1.1, the results of Fisher’s exact tests show a significant nonrandom association between the feature for morphological alignment and the presence of applicatives of any type ($p = .036$), as well as the presence of pivot-neutral applicatives ($p = < .001$). In contrast, no significant relationship was indicated between the feature for syntactic alignment and the presence of

Table 5.3: Clausal word order and presence of applicatives

Word order (A-oriented)	Presence of applicatives (any type)		Presence of pivot- neutral applicatives		TOTAL
	Yes	No	Yes	No	
Verb-medial (AVP)	28	23	27	24	51
Verb-initial (VAP and/or VPA)	18	1	11	8	19
Position of verb varies	10	1	9	2	11
TOTAL	56	25	47	34	81
%	69	31	58	42	

applicatives, either generally ($p = .873$), or for pivot-neutral applicatives specifically ($p = .806$). Table 5.4 below shows languages of the sample by pattern of morphological alignment.

Table 5.4: Morphological alignment and presence of applicatives

Morphological alignment	Presence of applicatives (any type)		Presence of pivot- neutral applicatives		TOTAL
	Yes	No	Yes	No	
Accusative	4	1	4	1	5
Ergative	9	0	9	0	9
Neutral	5	7	4	8	12
Pivot-Nonpivot	13	5	6	12	18
Mixed-NPIV.A	19	12	18	13	31
Mixed (other)	9	1	9	1	10
TOTAL	59	26	50	35	85
%	69	31	59	41	

Previously, Peterson (2007: 214–219) found a correlation between the presence of applicatives and non-accusative alignment, especially ergative alignment. The survey results differ quite a bit from those findings. The presence of applicatives in languages of the sample is overall fairly high across types of morphological alignment. An especially strong tendency to have pivot-neutral applicatives is found both in languages with ergative morphological alignment (100%, 9 of 9) and those with accusative morphological alignment (80%, 4 of 5). Languages with neutral morphological alignment show a greater proportion of languages lacking applicatives of any type in the sample; only 41% of these languages (5 of 12) have applicatives.

5.2.2.1 Ergative alignment

As in previous research, in the sample, we observe a strong tendency for languages with ergative morphological alignment to have applicatives, and specifically pivot-neutral applicatives. However, the nine languages of the sample that show ergative morphological alignment and pivot-neutral applicatives—Uma, Mori Bawah, and all seven South Sulawesi languages of the sample—show little evidence of ergative syntactic alignment. For example, in Uma, absolutive

enclitics index S and P in transitive clauses (both in AV and PV), while ergative proclitics index A (in PV only, A is not indexed in AV) (Martens 1988b: 172–175). Nevertheless, when it is the pivot, A (of AV) shows access to relativization, just as other pivot arguments do (i.e. P of PV, S of intransitives, S of non-volitional passives), as shown in examples in Martens (1988b). As a result, I categorize Uma as a language with pivot-non-pivot syntactic alignment rather than ergative syntactic alignment. (For discussion of alignment in South Sulawesi languages, see §5.9.1).

Thus, as a group, the languages of the sample with ergative morphological alignment are not necessarily comparable to the languages coded for ergative alignment in Peterson’s (2007) study. For this reason, and because languages with ergative morphological alignment in the sample are concentrated in Sulawesi geographically and in the South Sulawesi genetic group in particular, the observed relationship between ergative alignment and the presence of applicatives in the focus area must be taken with a grain of salt.

5.2.2.2 Accusative alignment

Turning to languages in the sample with accusative morphological alignment, we do not observe evidence that these languages are less likely to have applicatives, as might be predicted based on results of previous studies. Languages with accusative morphological alignment in the sample include Muna, Busoa, Laiyolo, Wolio, and Sumbawa. For this relatively small group of languages, 80% (4 of 5) show applicatives, compared to 69% (59 of 85) of languages in the total sample that show applicatives of any type. However, again it should be stated these languages are not necessarily comparable to the group of accusative languages surveyed in Peterson (2007); at least one of these languages—Laiyolo—shows non-accusative syntactic alignment (see Belding, Laidig & Maingak 2001). Areal patterns are also at play here. Sumbawa, which does not have applicatives, is spoken the transition area between West and East Nusantara languages (see §5.6). The other four languages, which have pivot-neutral applicatives, are all spoken in Sulawesi, where the presence of such constructions is by far the norm (see §5.9).

5.2.2.3 Neutral alignment

There are also languages of West Nusantara which morphologically distinguish between core and oblique arguments, but otherwise mark S, A, and P in the same manner. For most such languages in the sample, there is generally no distinction of separate sets of pronouns that co-vary with grammatical relations, and NPs are simply bare when they encode core arguments and marked with prepositions when oblique.² In the sample, such languages are more likely to show no applicatives of any type; 58% (7 of 12) do not have applicatives, compared to 31% of the larger sample (26 of 85). Of these, five languages (Bih, Eastern Cham, Tsat, Urak Lawoi’, Singapore Bazaar Malay) show extensive contact with non-Austronesian languages (see §5.7 on languages of mainland Southeast Asia and Peninsular Malaysia) and the remaining two languages, Matéq (Land Dayak) and Mualang (Malay) are both spoken in Borneo south of Sabah, where lack of applicatives is the norm (see §5.8).

²Enggano, also classified as a language with “core/oblique” morphological alignment, differs from this characterization, as it makes use of noun-marking prefixes, see §5.4.

5.2.2.4 Mixed alignment

Finally, I will note that languages coded as “mixed (other)” for morphological alignment do not form one coherent category. Some of these languages, like Moronene and Sasak, show complex and variable use of pronominal clitics for person indexing (see Andersen & Anderson 2005; Khairunnisa 2022). Others, like Nias and Kerinci, make use of morphophonological alternations in the form of nominal arguments which shows different patterns of co-variance with grammatical relations across different types of constructions (see Brown 2001; Ernanda 2017). Still others, like Brunei Malay, show different patterns of person indexing with free pronouns, as well as pronominal clitics, across A-oriented and P-oriented constructions (see Clynes 2001). Note that during analysis, Acehnese, the only language coded for split intransitive morphological alignment, was collapsed into the “mixed (other)” category. Languages coded as “mixed (other)” show a strong tendency in the data to have pivot-neutral applicatives (90% or 9 of 10 languages). However, because these languages are so different from one another, not much can be taken from this fact, except to say that for these languages, as in the sample as a whole, a great diversity of patterns of alignment are observed across which applicatives may be found.

5.2.3 Symmetrical voice and diathesis

The results of Fisher’s exact tests show a significant nonrandom association between type of voice (or diathesis) system and the presence of pivot-neutral applicatives ($p = .009$). Table 5.5 below shows languages of the sample by type of voice system.

Table 5.5: Voice system and presence of pivot-neutral applicatives

Voice system	Presence of pivot-neutral applicatives		TOTAL
	Yes	No	
asymmetrical	7	6	13
marginal two-way	8	0	8
two-way symmetrical	30	20	50
marginal Philippine-type	3	1	4
Philippine-type	2	8	10
TOTAL	50	35	85
%	59	41	

Unsurprisingly, languages that show Philippine-type voice systems are far less likely than other languages of West Nusantara to show pivot-neutral applicatives. Of the languages in the sample coded as Philippine-type for this feature, 80% (8 of 10) lack pivot-neutral applicatives. This tendency is likely due to functional overlap. Languages that retain the four-way voice distinctions found in PMP (or expand on it), already have a way to select peripheral semantic roles like location, beneficiary, instrument, and theme as a core argument in the form of location voice (LV) and circumstantial voice (CV) constructions. Such languages therefore usually do not develop pivot-neutral applicatives. Specifically, the eight languages of the sample that retain productive Philippine-type voice and lack pivot-neutral applicatives all (8 of 8) show at least one voice

category consistent with the functions of CV in PAN and PMP—that is, selecting a beneficiary, instrument or theme as the pivot—and most (7 of 8) also show LV.

More unexpected then, is the fact that a number of Philippine-type languages do have a system of pivot-neutral applicatives. In the sample, two languages of Sulawesi coded as Philippine-type voice systems also show pivot-neutral applicatives: Balantak (Saluan-Banggai) and Totoli (Tomini-Tolitoli). Notably, these two languages retain an LV voice category, but have entirely lost CV. In addition, three languages coded as marginal-Philippine-type (3 of 4) also show pivot-neutral applicatives: the Sama-Bajaw languages Central Sama and Yakan, and the Saluan-Banggai language Bobongko. These languages show a restricted distribution for both LV and IV, indicating that Philippine-type voice is being lost while pivot-neutral applicatives have developed. All together, these five languages of the sample thus show a clear transition between Philippine-type voice and pivot-neutral applicatives. Totoli, Balantak, and Bobongko are discussed in greater detail in §5.9.4 on transitional languages of Sulawesi, while Central Sama and Yakan are discussed in §5.8.4.2 on Sama-Bajaw languages.

Turning to two-way symmetrical systems, in the data we observe that 60% of such languages in the sample (30 of 50) have pivot-neutral applicatives. Given that 59% of languages in the entire sample show pivot-neutral applicatives, languages with two-way symmetrical voice systems cannot be said to show a special affinity for pivot-neutral applicatives in the West Nusantara context. Areal patterns are also at play. Of the two-way symmetrical voice languages with no applicatives of any type in the sample, 18 are spoken in Borneo (see §5.8), and only two—Rejang, and Kerinci—are spoken in Sumatra (see §5.4.1 on outliers in Sumatra).

As another point of comparison, for languages of the sample with asymmetrical voice systems (those with no systematic contrast between A-oriented and P-oriented transitive constructions), about 54% (7 of 13 languages), have pivot-neutral applicatives. For languages with asymmetrical voice systems, the lack of applicatives follows genetic and geographic patterns. The six languages with asymmetrical voice systems and no applicatives of any type (Tsat, Bih, Eastern Cham, Urak Lawoi', Singapore Bazaar Malay, Sumbawa) are primarily found in mainland Southeast Asia (4 of 6), and within the Chamic genetic group (3 of 6).

Pivot-neutral applicatives are more likely to be found in marginal two-way systems than any other coding category for voice. All eight of the languages with marginal two-way voice systems in the sample show such applicatives. However, six of the languages in this category are from the South Sulawesi genetic grouping (see §5.9.1). Because of this large degree of overlap with one genetic grouping, the strong relationship between marginal two-way symmetrical voice systems and pivot-neutral applicatives may be primarily due to common inheritance.

Given these observations, a number of conclusions can be drawn. First, pivot-neutral applicatives are much less likely to appear in Philippine-type languages where LV and especially CV are found as basic transitive voice constructions. Second, two-way symmetrical voice systems in West Nusantara do not show any special attraction to pivot-neutral applicatives, compared to other coding categories for type of voice system. The survey results undercut the idea that the two-way symmetrical voice system and pivot-neutral applicatives are two characteristic features of an Indonesian-type category of western Austronesian languages in opposition to the Philippine-type category (see also McDonnell & Chen 2022; Kaufman 2009; Ross 2002 for reasons that 'Indonesian-type' is not a coherent typological category for western Austronesian languages cf. Himmelmann 2002). Pivot-neutral applicatives in fact cut across West Nusantara languages of all types of symmetrical and asymmetrical voice systems. Third, the patterns that we observe

for pivot-neutral applicatives in languages that show signs of impending loss or diminished distribution for CV, or both LV and CV, indicates that the development of pivot-neutral applicatives is related to the breakdown of Philippine-type voice. Finally, while pivot-neutral applicatives are found across languages with all types of voice systems, there are nonetheless clear pockets of languages that did not develop pivot-neutral applicatives—or did but lost them—following geographic and genetic patterns. These are primarily located in mainland Southeast Asia and Borneo. Accordingly, the lack of applicatives in these geographic areas is treated as a distributional fact requiring an explanation in the remainder of this chapter.

5.2.4 Other structural features

For remaining structural features—case marking, order of Noun + Possessor, and presence of morphological causatives, the results of statistical testing did not show any significant non-random relationship with the presence of applicatives generally, or the presence of pivot-neutral applicatives specifically. In the random forest multivariable analysis, these features were also of relatively low importance. Here I will note that for order of Noun + Possessor, and presence of morphological causatives, there was little diversity in the sample. In almost all languages of the sample (82 of 85), the noun precedes the possessor, while in just three—Kimaragang, Tsat and Singapore Bazaar Malay—sources report multiple possessor constructions representing both the order Noun + Possessor and the order Possessor + Noun. Similarly, almost all of the languages of the sample (81 of 85) have a morphological causative construction. For two of the 85 languages there was no available information to determine this, and for another two—Tsat and Singapore Bazaar Malay—it appears that only analytic causative constructions are used and there are no productive morphological causative constructions. These results indicate that case marking, order of Noun + Possessor (as a proxy for head-initial vs. head-final structure) and the presence of morphological causatives are not important features that help to distinguish between languages of West Nusantara that have applicatives from those that lack them, and this holds for both applicatives in general and pivot-neutral applicatives specifically.

5.3 Location and genetic affiliation: Summary of results

As described in §5.1.2, based on the results of multivariable modeling through random forest classification analysis, location and genetic affiliation are by far the two most important features in the survey that may be used to classify languages by the presence of applicatives generally or the presence of pivot-neutral applicatives specifically. Because location and genetic affiliation show a large degree of overlap for many of the genetic groupings used in the survey, they are treated together in this section, which presents a summary of the distribution of applicative by these two features, as well as in following sections that give detailed results by major island grouping (§5.4–5.9).

Table 5.6 shows languages of the sample by geographic location and presence of applicatives. The languages of the sample spoken in mainland Southeast Asia (north of Peninsular Malaysia) are characterized by a lack of applicatives; all five show no applicatives of any type. The languages of the sample spoken in Borneo also show a strong tendency to lack applicatives, especially pivot-neutral applicatives (23 of 27 lack such constructions). In contrast, languages of the

sample spoken in Sulawesi show a strong tendency to have applicatives, especially pivot-neutral applicatives (25 of 26 languages).³

The geographic distribution of languages with and without applicatives in the sample is represented visually in the map in Figure 5.4. Not shown in the map but included in the sample are Merina Malagasy, spoken in central Madagascar in east Africa, which has pivot-selecting applicatives only, and Suriname Javanese, spoken in Suriname in northern South America, which has pivot-neutral applicatives only. As seen in the map, languages with pivot-selecting applicatives that represent Philippine-type voice alternations are concentrated in northern Borneo (Sabah and adjacent parts of northeast Sarawak) and nearby areas of the southern Philippines. A handful of languages with Philippine-type voice alternations are also found in the northern third or so of Sulawesi. Languages with pivot-neutral applicatives are spread throughout West Nusantara, and are quite prevalent in Sulawesi, Java, and Sumatra.

Table 5.6: Location and presence of applicatives

Location	Presence of applicatives (any type)		Presence of pivot- neutral applicatives		TOTAL
	Yes	No	Yes	No	
Africa	1	0	0	1	1
Barrier Islands	2	0	2	0	2
Sumatra	9	2	9	2	11
MSEA	0	5	0	5	5
Java	4	0	4	0	4
Borneo	11	16	4	23	27
Philippines	3	0	2	1	3
Lesser Sundas	2	1	2	1	3
Sulawesi	25	1	25	1	26
Americas	1	0	1	0	1
Other	1	1	1	1	2
TOTAL	59	26	50	35	85
%	69	31	59	41	

Table 5.7 summarizes languages of the sample by genetic group and presence of applicatives. Many genetic affiliations considered in the survey are homogeneous with respect to the presence of applicatives generally and the presence of pivot-neutral applicatives specifically. For example, Land Dayak languages consistently lack all applicatives, while the Northwest Sumatra-Barrier Islands and South Sulawesi groups consistently show pivot-neutral applicatives only. A smaller number of genetic groups represent varied patterns for the presence of applicatives and/or type of applicatives attested, especially Malayic, Greater Barito, and to a lesser extent, Sabahan, and North Sarawak. Description of these patterns and possible explanations for them are taken up in greater detail in following sections of this chapter, organized by major island grouping.

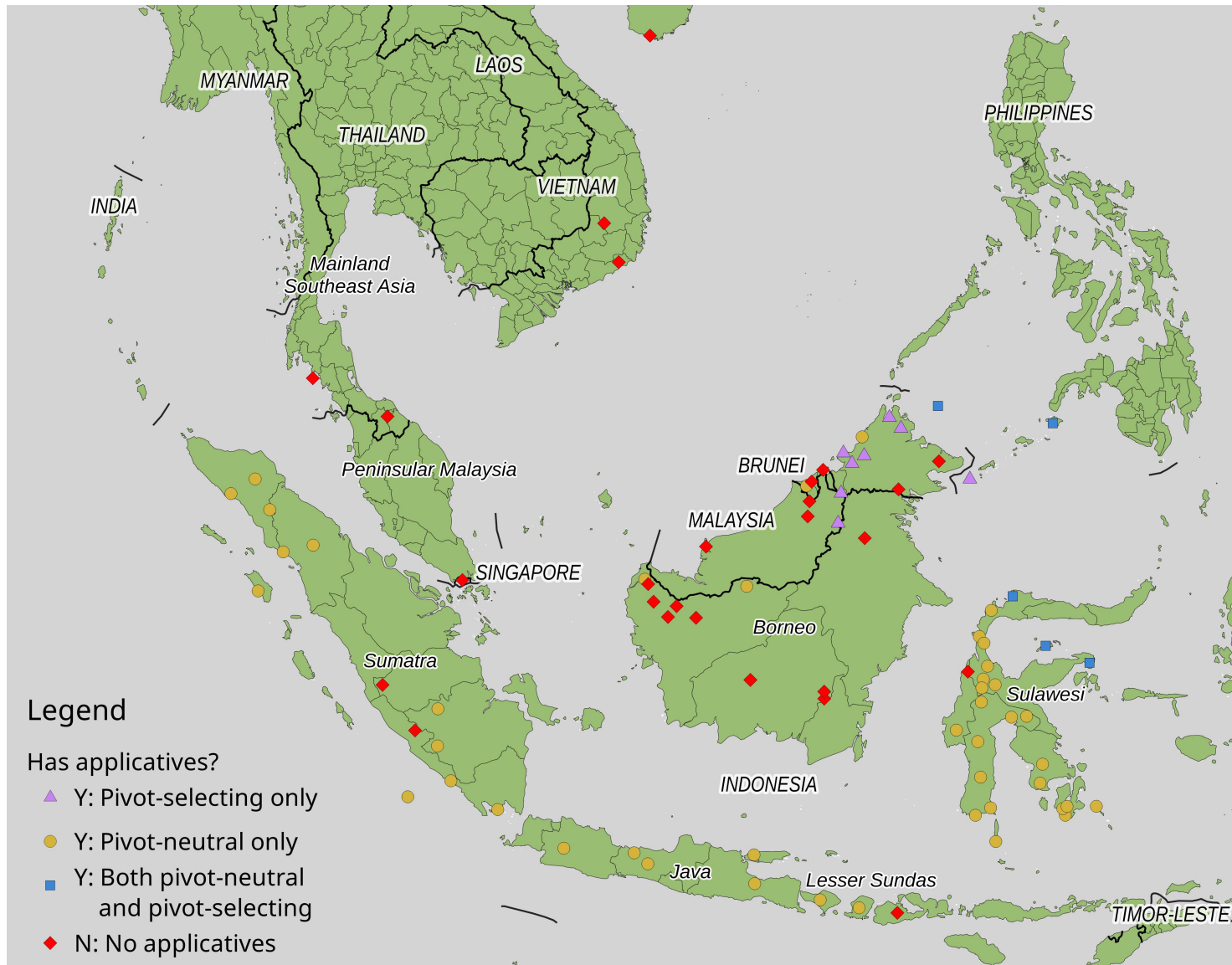
³Note that Gorontalo-Mongodow, Sangiric, and Minahasan languages, which are spoken in the northernmost parts of Sulawesi and smaller islands off its northern coast, were not included in the study because they are more closely related to languages of the Philippines than those of West Nusantara proper. These languages generally show Philippine-type voice systems.

Table 5.7: Genetic affiliation and presence of applicatives

Genetic grouping	Presence of applicatives (any type)		Presence of pivot- neutral applicatives		TOTAL
	Yes	No	Yes	No	
<i>Indigenous to Sumatra & Barrier Isl.</i>					
NW Sumatra-Barrier Islands	5	0	5	0	5
Enggano	1	0	1	0	1
Lampungic	1	0	1	0	1
Rejang	0	1	0	1	1
Nasal	1	0	1	0	1
<i>Indigenous to Java & Madura</i>					
Javanese*	3	0	3	0	3
Madurese	1	0	1	0	1
Sundanese	1	0	1	0	1
<i>Indigenous to Lesser Sundas</i>					
Bali-Sasak-Sumbawa	2	1	2	1	3
<i>Indigenous to Borneo</i>					
Greater Barito*	5	3	3	5	8
Sabahan	5	2	0	7	7
North Sarawak	2	5	0	7	7
Land Dayak	0	4	0	4	4
Melanau-Kajang	0	1	0	1	1
<i>Indigenous to Sulawesi</i>					
South Sulawesi*	7	0	7	0	7
Kaili-Pamona	4	1	4	1	5
Tomini-Tolitoli	4	0	4	0	4
Bungku-Tolaki	3	0	3	0	3
Muna-Buton	3	0	3	0	3
Wotu-Wolio	3	0	3	0	3
Saluan-Banggai	2	0	2	0	2
<i>Other genetic groupings</i>					
Malayic*	5	5	5	5	10
Chamic*	1	3	1	3	4
TOTAL	59	26	50	35	85
%	69	31	59	41	100

* These genetic groupings show a large degree of geographic dispersion from their original homeland for one or more members of the group.

Figure 5.3: Map of languages of the sample by presence of applicatives and type



Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

5.4 Sumatra and the Barrier Islands

In this section, I present and discuss results of the typological survey for the languages of Sumatra and the Barrier Islands. Sumatra represents the westernmost of the major islands in West Nusantara, and the Barrier Island chain runs from north to south off its west coast. The languages of Sumatra and the Barrier Islands are diverse with respect to genetic affiliation within Malayo-Polynesian. They are comprised of (i) Malay languages (i.e. the Malay subbranch of Malayic), (ii) the Northwest Sumatra-Barrier Islands group, (iii) Acehnese, which belongs to the Chamic group, and (iv) a number of small outlier groupings and isolates including Enggano, Nasal, Rejang, and Lampungic.

As shown in Table 5.8, 13 languages of Sumatra and the Barrier Islands are included in the sample. Due to lack of adequate descriptive material, languages of the Barrier Islands are not well represented in the sample. To address this, in the remainder of this section, when partial information is available I will make mention of two additional languages spoken in the Barrier Islands: Mentawai and Sigulai (Sikule) (Adriani 1928; Morris 1900; Kähler 1955; see also McDonnell & Truong forthcoming). Both belong to the Northwest Sumatra-Barrier Islands group. An overview map of the languages of the sample in this geographic area and other languages discussed in this section is presented in Figure 5.4.

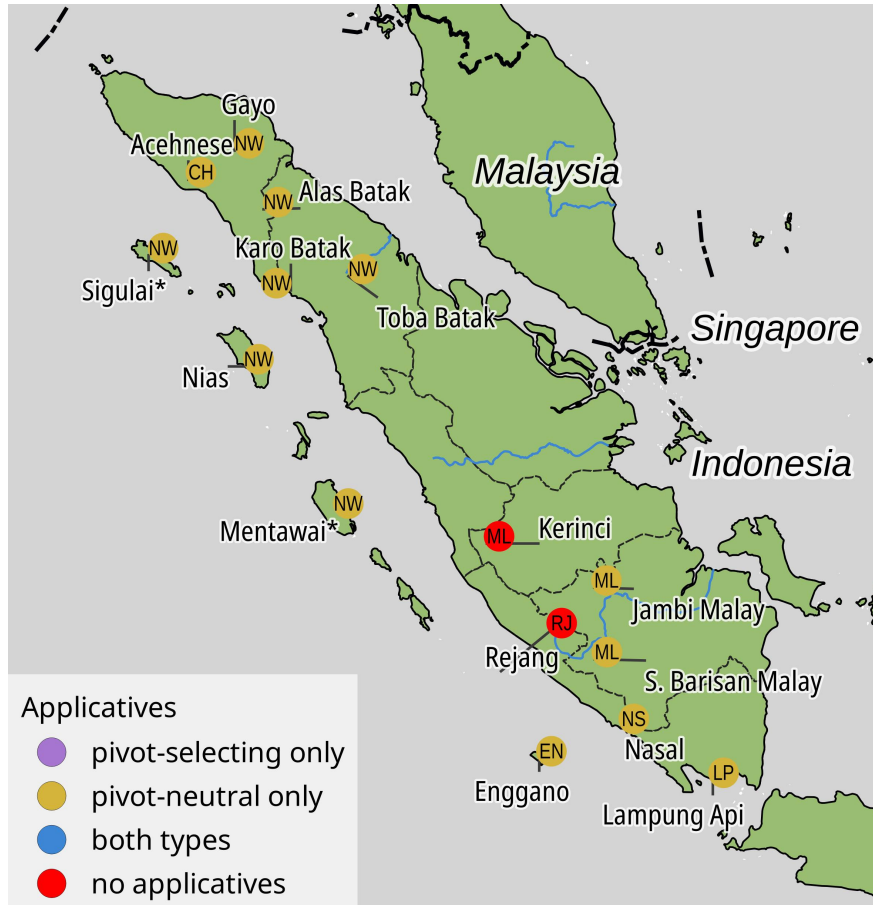
Table 5.8: Typological survey results for Sumatra and the Barrier Islands

Gen. Grp.	Name	Applicatives?		Voice	Morph. align.	Order
		pivot-selecting	pivot-neutral			
Chamic	Acehnese	N	Y	asymmetrical	split-S	free
NWS-BI	Alas Batak	N	Y	two-way	mixed-NPIV.A	AVP
NWS-BI	Karo Batak	N	Y	two-way	pivot-nonpivot	AVP
NWS-BI	Toba Batak	N	Y	two-way	mixed-NPIV.A	VPA
NWS-BI	Gayo	N	Y	two-way	mixed-NPIV.A	AVP
NWS-BI	Nias	N	Y	marg. two-way	mixed:other	VPA
Enggano	Enggano	N	Y	asymmetrical	neutral	AVP
Nasal	Nasal	Y	Y	two-way	pivot-nonpivot	AVP
Malayic	Kerinci Malay	N	Y	two-way	mixed:other	AVP
Malayic	Jambi Malay	N	N	two-way	mixed:other	AVP
Malayic	S. Barisan Malay	N	Y	two-way	mixed-NPIV.A	AVP/VPA
Rejang	Rejang	N	N	two-way	mixed-NPIV.A	AVP
Lampungic	Lampung Api	N	Y	two-way	mixed-NPIV.A	AVP

Across all genetic groupings in Sumatra, the presence of pivot-neutral applicatives is the norm (11 of 13 languages of the sample),⁴ and pivot-neutral applicatives are also found in Mentawai and Sigulai. Pivot-selecting applicatives are not found in Sumatra and the Barrier Islands (0 of 13), and there is no clear evidence that pivot-selecting applicatives can be reconstructed for any of the

⁴Acehnese is the only Chamic language spoken primarily outside of mainland Southeast. Unlike other Chamic languages (see §5.7), Acehnese does have one construction that may be considered an applicative: when the verb is marked with prefix *peu-*, the clause may mean ‘administer ROOT onto/at an undergoer’ (Durie 1985: 79). This construction meets the definition of applicative used in this study because it selects a goal as a core argument. However, most other meanings associated with the *peu-* verbal prefix are causative and not applicative, and the applicative usage is of limited productivity.

Figure 5.4: Overview map of languages of Sumatra and the Barrier Islands



* = Not included in language sample but discussed in §5.4. Abbreviations: [EN] Enggano, [LP] Lampungic, [ML] Malayic, [NS] Nasal, [NW] Northwest Sumatra-Barrier Islands. Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

genetic groupings represented. Sporadic cases of lack of applicatives are found in two languages of Sumatra (2 of 13); Rejang and Kerinci do not have applicatives of any type (see §5.4.1 below for greater detail).

Pivot-neutral applicatives are found across languages with diverse types of voice systems in Sumatra and the Barrier Islands. They are found in languages with asymmetrical voice systems (2 of 2; Acehnese, Enggano) and one marginal two-way voice system (1 of 1; Nias),⁵ as well as the large majority of two-way symmetrical voice systems (8 of 10), which represent the most common type of voice system in languages spoken on the island of Sumatra.⁶

Pivot-neutral applicatives are found across languages with diverse systems of morphological alignment as well. Quite common are systems of morphological alignment in which non-pivot A arguments have distinctive forms compared to all other core arguments, especially for languages spoken on the island of Sumatra. In the sample, almost all such languages (5 of 6) show pivot-neutral applicatives. The remaining seven languages in the sample are split across four other coding categories for morphological alignment: split intransitive (1), pivot-non-pivot (2), neutral (1), and mixed (other) (3). Nonetheless almost all of these languages (6 of 7) also have pivot-neutral applicatives.

The majority of languages of Sumatra and the Barrier Islands have AVP unmarked word order in A-oriented transitive constructions, with most of these (7 of 9) also showing pivot-neutral applicatives. Verb-initial word order is preferred in Nias and Toba Batak, which both have pivot-neutral applicatives (2 of 2), and this also appears to hold for Sigulai (see McDonnell & Truong forthcoming: 422). In Acehnese, word order is relatively free, with the majority of clauses being verb-initial (Durie 1985: 191). In South Barisan Malay, the non-pivot immediately follows the verb, while the pivot has freer word order (AVP/VPA in AV). Both Acehnese and South Barisan Malay have pivot-neutral applicatives.

On the whole, it is clear that the presence of pivot-neutral applicatives is the norm in Sumatra and the Barrier Islands, and these constructions show a broad distribution, being found across all five genetic groupings, all four types of voice systems, and all five coding categories for morphological alignment represented, as well as both languages with verb-initial unmarked word order and those with verb-medial word order. Survey results for Sumatra and the Barrier Islands thus support the finding that the presence of pivot-neutral applicatives is not associated with a particular typological profile in West Nusantara.

In particular, the presence of pivot-neutral applicatives in Mentawai and Enggano, is telling. Mentawai has two pivot-neutral applicative suffixes, *-ake*, which forms instrument- and theme-selecting constructions, and *-i*, which forms locative- and goal-selecting constructions (Adriani 1928: 69–70, 81–83). Mentawai speakers, as inhabitants of Siberut, Sipura, and the Pagai islands west of Sumatra, for many centuries had minimal contact with outsiders prior to first visits by British and Dutch colonists in the late 1700s. There are little signs of prior Indic or Islamic cultural influence in the Mentawai islands (Loeb 1928), including that of the powerful Malay-speaking kingdoms established in western Indonesia and Malaysia from the 7th century until the 16th century. Therefore, it is highly unlikely that applicatives were introduced to Mentawai via

⁵In main clauses and most dependent clauses, Nias does not show an alternation between A-oriented and P-oriented transitive constructions. A relic of symmetrical voice is found only in certain P-oriented relative clauses marked with *ni-*, in which the actor is frequently overtly expressed (see Brown 2001: 421).

⁶Sigulai probably also represents an asymmetrical voice system, while Mentawai appears to be a two-way symmetrical system (see McDonnell & Truong forthcoming: 417–420).

language contact.

Enggano, spoken on the southernmost of the Barrier Islands, is considered to be one of the most aberrant Austronesian languages due to its unusual lexicon, and has even been argued to be non-Austronesian though it is now fairly well-established that it belongs to Malayo-Polynesian (see Edwards 2012). Enggano also shows unusual characteristics for West Nusantara in its grammatical system, including its system of case and alignment. Enggano makes use of nominal prefixes to distinguish core arguments from obliques (neutral alignment) and the forms of these prefixes co-vary with grammatical number, human vs. nonhuman referent, and status as a proper vs. common noun (Crowley n.d.). Enggano also has two pivot-neutral applicative suffixes, which are illustrated in examples (100) and (101).

(100) Enggano, Locative-selecting applicative

- a. *kia ki-hēkū i-kuə eʔana.*
 3SG VERB-sit LOC.CN-tree MED
 ‘He is sitting in that tree.’ (BC)
- b. *kia ki-hēkū-hūĩ e-kuə eʔana.*
 3SG VERB-sit-LOC.APPL CORE.CN-tree MED
 ‘He is sitting in that tree.’ (AC)

(Kähler 1940: 195, English translation & glosses added)

(101) Enggano, Instrument-selecting applicative

- a. *kia ki-pudu e-kəyɔ iʔiɔ u-bəhɛ*
 3SG VERB-kill CORE.CN-spear LOC.PREP OBL.CN-pig
 ‘He killed the wild pig with a spear.’ (BC)
- b. *kia ki-pudu-ʔa e-bəhɛ iʔiɔ u-kəyɔ*
 3SG VERB-kill-INST.APPL CORE.CN-spear LOC.PREP OBL.CN-pig
 ‘He killed the wild pig with a spear.’ (AC)

(Kähler 1940: 196, English translation & glosses added)

Edwards (2012) offers two possibilities for the aberrant nature of Enggano: “(i) that Enggano suffered contact effects with a non-Austronesian language present in the region prior to the appearance of MP [Malayo-Polynesian] languages, and/or (ii) that Enggano Island was relatively isolated from the rest of the Austronesian world.” If the latter case is true, the presence of pivot-neutral applicatives in Enggano further bolsters the argument that the distribution of pivot-neutral applicatives in West Nusantara is not primarily explained by language contact with culturally prominent languages. But even if it is not, Enggano’s unusual case markers and system of alignment still underscore the other major finding of this section, that the presence of pivot-neutral applicatives is not associated with a particular typological profile in West Nusantara.

5.4.1 Outliers in Sumatra

In Sumatra and the Barrier Islands, only Kerinci (Malayic) and Rejang (isolate) lack applicatives completely. The 11 other languages of the sample spoken in this geographic area—plus Mentawai

and Sigulai—have pivot-neutral applicatives and these are almost exclusively marked with verbal suffixes. Only Acehnese has an AC marked with a prefix, i.e. *peu-*, and this is of limited productivity.

Unlike other languages of Sumatra and the Barrier Islands, Rejang lacks suffixes of any type and makes use of only prefixes (McGinn 1982). This is likely the reason for lack of applicatives in Rejang. However, we do not have evidence by which to determine whether (i) Rejang never had applicative suffixes—neither pivot-selecting nor pivot-neutral, or (ii) that some such suffixes were previously present but subsequently lost (cf. the case of Mualang, where clausal alternations persist after the loss of applicative suffixes, see §5.8.5). Rejang is currently classified as an isolate within Malayo-Polynesian, and not much is known about the pre-history of the language, including whether it arrived in Sumatra after a period of development in another part of West Nusantara, such as Borneo (see McGinn 2009). Therefore, I consider this an open question, awaiting more historical evidence linking Rejang genetically or geographically with other languages of West Nusantara.

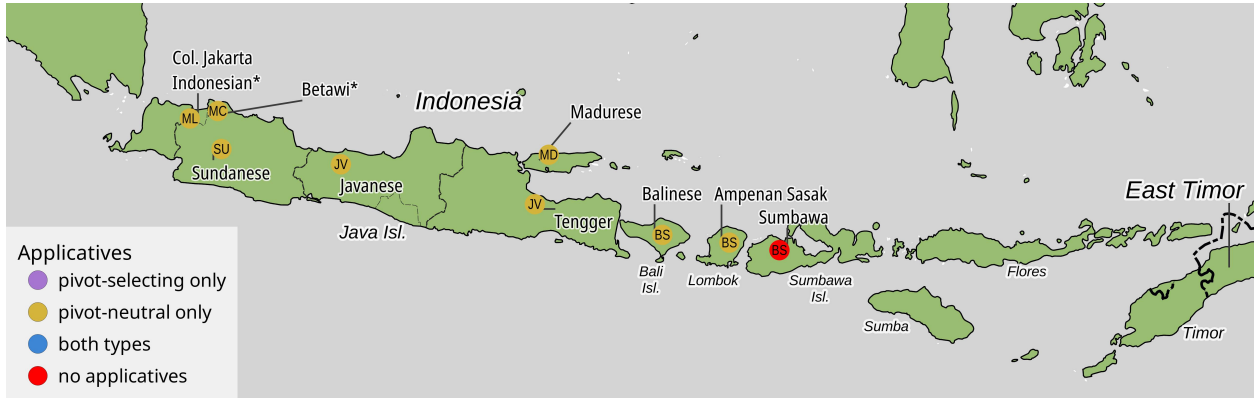
In Kerinci, it was not loss of suffixation per se, but phonological changes involving fusion of certain suffixes (and pronominal enclitics) to stems, that have resulted in a lack of applicatives (Yanti et al. 2018). For Kerinci, it is clear that pivot-neutral applicative suffixes were among those lost through fusion, as one such suffix, identified as cognate with Standard Malay/Indonesian *-kan* is found in remnant forms in the Tanjung Pauh variety of Kerinci, e.g. *kato-ka* ‘say (s.t.)’ (Yanti et al. 2018: 469). This process of fusion has resulted in radical changes in the grammar of the language that have been of some interest in the literature including the genesis of “absolute” and “oblique” forms of stems that color its unusual system of agreement and mixed system of morphological alignment for core grammatical relations (McKinnon, Cole & Hermon 2011; Yanti 2010).

5.5 Java and Madura

In this section, I present and discuss survey results for the languages of Java, the most populous of the major islands in West Nusantara, and Madura, a smaller island off the northeast coast of Java. These islands are home to the Sundic languages, Javanese languages, and Madurese languages. An overview map of Java, Madura, and the Lesser Sundas is presented in Figure 5.5. Javanese languages are also found outside of West Nusantara in New Caledonia (Pacific region), and Suriname (a Caribbean nation of South America, see Figure 5.6), where they are spoken by diaspora communities of ethnically Javanese people.

The survey sample includes one Sundic language, three Javanese languages, and one Madurese language, as shown in Table 5.9, which includes Standard Indonesian as well for reference. These languages share a similar typological profile. They have two-way symmetrical voice systems with pivot-neutral applicatives that are marked with verbal suffixes. In unmarked word order, the non-pivot argument immediately follows the verb, and the pivot argument is preverbal. Thus most have predominant AVP word order in AV, with VPA also possible. These languages have one distinctive morphological marking for the non-pivot A argument and one for all other core grammatical relations. The type of marking may be dependent on the person category of the non-pivot A argument. In Sundanese and Madurese, the verb in a PV construction is marked with a voice prefix, and the non-pivot A argument is marked with a preposition meaning ‘with, by’ (i.e.

Figure 5.5: Overview map of languages of Java and the Lesser Sundas



* = Not included in language sample but discussed in §5.5. Abbreviations: [BS] Bali-Sasak-Sumbawa, [JV] Javanese, [MC] Malay-based Creole, [MD] Madurese, [ML] Malayic, [SU] Sundic. Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

Figure 5.6: Map showing location for Suriname Javanese in South America



Abbreviations: [JV] Javanese. Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

ku in Sundanese and *bi'* in Madurese). The preposition is optional if A immediately follows the verb as in unmarked word order, but obligatory in other word orders. In Javanese languages, third-person non-pivot A arguments follow the pattern for optional prepositional marking found in Sundanese and Madurese, while first- and second-person non-pivot A arguments are expressed with special proclitic pronominal forms that replace the PV voice prefix.⁷

Table 5.9: Typological survey results for Java and Madura

Gen. Grp.	Name	Applicatives?		Voice	Morph. align.	Word Order
		pivot-selecting	pivot-neutral			
Sundic	Sundanese	N	Y	two-way	mixed-NPIV.A	AVP
Javanese	Javanese	N	Y	two-way	mixed-NPIV.A	AVP
Javanese	Tengger	N	Y	two-way	mixed-NPIV.A	AVP
Javanese	Sur. Javanese	N	Y	two-way	mixed-NPIV.A	AVP
Madurese	Madurese	N	Y	two-way	mixed-NPIV.A	AVP/VPA
Malayic	Std. Indonesian	N	Y	two-way	mixed-NPIV.A	AVP

Also spoken on Java but not included in the survey sample are Betawi Malay and Colloquial Jakarta Indonesian, both spoken in and around the special administrative region of Jakarta, the capital of Indonesia. Betawi is the language spoken by the original inhabitants of the Jakarta area. It is not included in the survey due to its classification as a Malay-based creole by Eberhard, Simons & Fennig (2021). Colloquial Jakarta Indonesian is a variety of standard Indonesian influenced by Betawi. It is not included in the survey because it does not have a separate ISO-639-3 code, but shares one with Standard Indonesian, which is included in the sample. Like other languages of Java, both Betawi and Colloquial Jakarta Indonesian have two-way symmetrical voice systems and pivot-neutral applicatives marked with verbal suffixes (Ikranagara 1975, Sneddon 2006: 30–34).

Overall, the languages of Java and Madura show a fairly coherent typological profile, one that is also associated with Standard Indonesian and that is consistent with the defining features of the proposed Indonesian-type of western Austronesian languages. However, when viewed in light of the full results of the typological survey, this profile is too limited and does not explain the full distribution of pivot-neutral applicatives in West Nusantara.

5.6 The Lesser Sundas

The Lesser Sundas are a chain of islands stretching eastward from the east coast of Java (see map in Figure 5.5). In the Lesser Sundas, only the westernmost islands of Bali and Lombok are considered to constitute part of West Nusantara, with Sumbawa, Sumba, Flores, and Timor generally being defined as part of East Nusantara (see Klamer & Ewing 2010). But because Balinese, spoken on Bali, Sasak, spoken on Lombok, and Sumbawa, spoken on Sumbawa Island, together form a

⁷The proclitic construction is reported to be rare with second-person non-pivot actors and relatively uncommon with first-person actors in Tengger, probably due to avoidance of these forms for politeness (Connors 2008: 146–147).

well-defined lower-level subgroup of Malayo-Polynesian, all three are eligible for inclusion in the survey. Of the three languages, Sasak and Sumbawa form one primary branch, while Balinese represents another. As shown in Table 5.10, all three languages are included in the sample. In the case of Sasak, survey data was compiled based on the Ampenan variety (Khairunnisa 2022), which shows some differences with other Sasak varieties described by Austin (2001).

Though they represent a small, closely-related genetic grouping, Balinese, Sasak, and Sumbawa show important differences in their typological profiles. These languages show a cline of features, with Balinese being typologically similar to the languages of Java and Madura, Sumbawa being similar to other East Nusantara languages, and Sasak occupying an intermediate position.

Table 5.10: Typological survey results for the Lesser Sundas

Gen. Grp.	Name	Applicatives?		Voice	Morph. align.	Word Order
		pivot-selecting	pivot-neutral			
Bali-Sasak-Sumbawa	Balinese	N	Y	two-way	mixed-NPIV.A	AVP
Bali-Sasak-Sumbawa	Ampenan Sasak	N	Y	two-way	mixed:other	AVP
Bali-Sasak-Sumbawa	Sumbawa	N	N	asymmetrical	accusative	undet.

The voice and applicative systems of Balinese were discussed in §3.4 above. The language shows a two-way symmetrical voice system, with AV marked with the prefix *N-*, as well as pivot-neutral applicatives marked with two suffixal forms. In Balinese, third-person non-pivot A arguments have a special morphological form that distinguishes them from other core arguments; these argument may be expressed with the enclitic *=na*. In all other cases, core argument are realized as bare NPs or full pronouns, when expressed. Normally, the pivot precedes the verb followed by the non-pivot argument; accordingly, the unmarked word order in AV is categorized as AVP. Balinese shows both prefixes and suffixes, with “a relatively small number of total affixes” (Shiohara & Arka forthcoming).

Like Balinese, Ampenan Sasak contrasts A-oriented transitive constructions with P-oriented transitive constructions, however the distinction between them is usually not morphological marked on the verb (Khairunnisa 2022). Consequently, Ampenan Sasak may be considered a two-way symmetrical diathesis system. Like Balinese, Ampenan Sasak also shows pivot-neutral applicatives, but only one suffixal form is used to mark these. Ampenan Sasak shows a complex, mixed system of morphological alignment. In P-Diathesis (PD), a non-pivot A argument expressed as an NP has special marking with *siq*. However, the use of clitic pronouns, which are extremely common, show a different pattern of alignment: P arguments expressed as clitics must be enclitics regardless of diathetical construction, while A arguments expressed as clitics may be either proclitics or enclitics. In canonical word order, the pivot appears in preverbal position; accordingly, the unmarked word order in A-oriented constructions is AVP. While Ampenan has a number of prefixes, its sole suffix is *-an*, which is highly polyfunctional, and is used to mark the verb in ACs (Khairunnisa & McDonnell in prep).

Finally, Sumbawa shows even greater differences with Balinese while sharing some characteristics with Ampenan Sasak. Sumbawa shows an asymmetrical diathesis system, and has “only one type of transitive construction with invariably bare verbs” (Shiohara & Arka forthcoming). It does not have applicatives of any type. It shows accusative morphological alignment, with

pronominal proclitics expressing the S and A core relations, but not P.⁸ In these respects, Sumbawa resembles other Austronesian languages of East Nusantara, which do not show morphological marking on verbs for voice, and make use of proclitics or prefixes indexing the clausal subject or agent (Klamer & Ewing 2010: 10). Many East Nusantara languages also do not have applicatives (though a few do, e.g. Taba, Bowden 2001). There is no identified predominant word order for transitive clauses in Sumbawa, however verb-initial clauses are most common in narrative texts (Shiohara & Arka forthcoming). Sumbawa shows only prefixes and no suffixes.

Bali-Sasak-Sumbawa represents a transition zone between the languages of Java and Madura to the west and East Nusantara languages to the east (Klamer & Ewing 2010). Balinese has the most verbal morphology including AV prefixal marking and two applicatives suffixes. Ampenan Sasak shows reduced morphological marking for voice and one applicative suffix. Sumbawa has the least verbal morphology and no applicatives. Because pivot-neutral applicatives are found in both primary branches of the subgroup, they can be reconstructed to Proto Bali-Sasak-Sumbawa. Pivot-neutral applicatives are unlikely to have been borrowed wholesale from Balinese into Sasak, as the applicative suffix **-an*, reflected in Balinese and Ngenó-Ngené Sasak as *-ang* (Austin 2001) and Ampenan Sasak as *-an*, is quite old and also found with applicative functions in Proto-Malayic, which has been proposed to subgroup with Chamic and then Bali-Sasak-Sumbawa at some higher level in PMP (Adelaar 2005a). Most likely then, Sumbawa represents a case of loss of pivot-neutral applicatives, apparently due to loss of suffixation and under the influence of neighboring East Nusantara languages.

5.7 Mainland Southeast Asia and Peninsular Malaysia

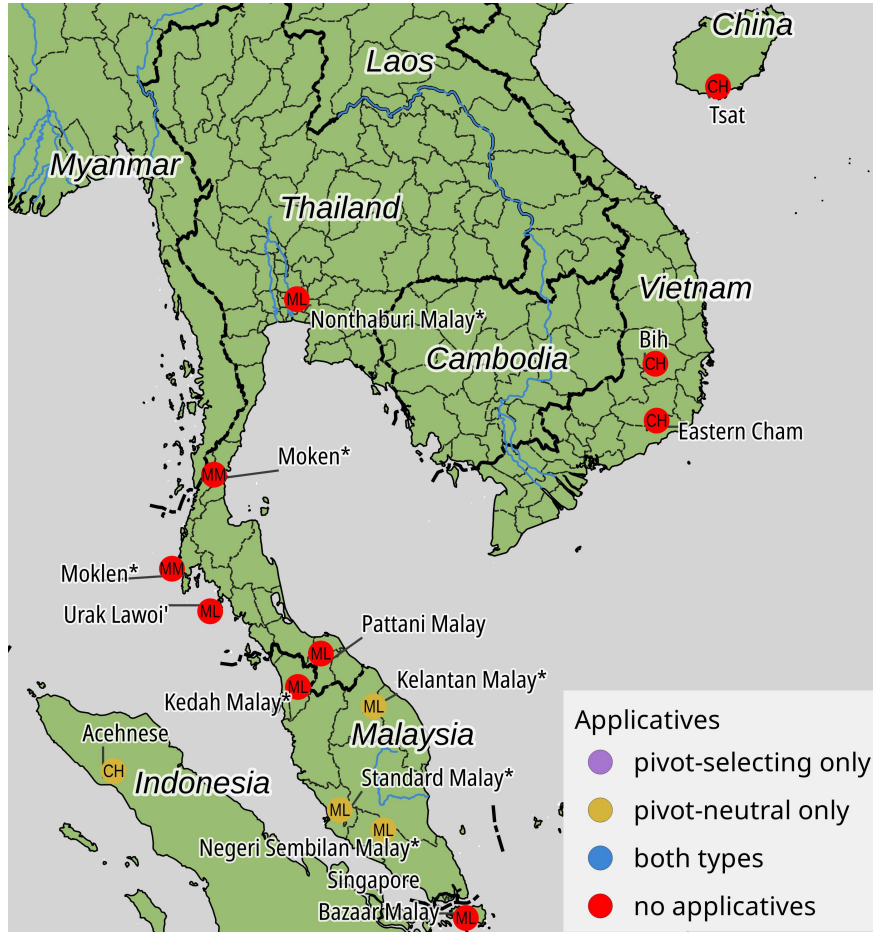
In this section, I present and discuss results of the typological survey for mainland Southeast Asia and Peninsular Malaysia. This geographic area includes (i) the Malay peninsula, which is considered part of West Nusantara and politically belongs to Malaysia, and (ii) other parts of Southeast Asia to the north of the Malay peninsula, which are not part of West Nusantara proper, and are governed by Myanmar, Thailand, Laos, Cambodia, Vietnam, and China (Hainan Island). Austronesian languages spoken in mainland Southeast Asia and Peninsular Malaysia include Chamic languages, Malay languages, and the Moken-Moklen languages, which are not included in the sample, but are discussed as a point of comparison in this section. An overview map of languages of the sample in this geographic area and other languages and varieties discussed in this section is shown in Figure 5.7. A summary of results for languages of the sample is given in Table 5.11.

Chamic is made up of 12 total languages, with four included in the sample. One of these, Acehnese, is spoken in Aceh Province in northern Sumatra, and is discussed in §5.4 above, while the other three, Eastern Cham, Bih, and Tsat, are spoken in mainland Southeast Asia, and are discussed in this section. Within Malayo-Polynesian, Malayic languages are the most closely related to Chamic and are considered “next-of-kin” (Thurgood 1999).

Malay is a primary branch of the larger Malayic genetic group. Of the 42 total Malayic languages, 33 are classified as Malay. Two Malay languages are spoken in mainland Southeast Asia and both are included in the sample: Pattani Malay and Urak Lawoi’. In addition, about nine

⁸The third-person proclitic patterns slightly differently, as it is used only in transitive clauses, and thus only for A (Shiohara & Arka forthcoming).

Figure 5.7: Overview map of languages of mainland Southeast Asia and northern Peninsular Malaysia



* = Not included in language sample but discussed in §5.7. Abbreviations: [CH] Chamic, [ML] Malayic, [MM] Moken-Moklen. Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

Table 5.11: Typological survey results for mainland SE Asia and Peninsular Malaysia

Gen. Grp.	Name	Applicatives?		Voice	Morph. align.	Word Order
		pivot-selecting	pivot-neutral			
Chamic	Eastern Cham	N	N	asymmetrical	neutral	AVP
Chamic	Bih	N	N	asymmetrical	neutral	AVP
Chamic	Tsat	N	N	asymmetrical	neutral	AVP
Malayic	Urak Lawoi'	N	N	asymmetrical	neutral	AVP
Malayic	Pattani Malay	N	N	two-way [†]	mixed-NPIV.A	AVP
Malayic	Singapore Bazaar Malay	N	N	asymmetrical	neutral	AVP

[†] The Nonthaburi variety of Pattani Malay shows an asymmetrical voice system.

Malay languages are spoken on the Malay peninsula, however, for many of these, there is little available grammatical description, aside from studies of phonological differences with Standard Malay. Four additional Malay languages and varieties not included in the sample will be discussed in this section on the basis of available information: Kedah Malay, Negeri Sembilan Malay, and the Nonthaburi and Kelantan varieties (which share one ISO-639-3 code with Pattani Malay). Standard Malay, the national language of Malaysia, is not included in the sample because of its similarities to Standard Indonesian, but will also be discussed here. Singapore Bazaar Malay, a Malay language used in interethnic communication in Singapore, just to the south of the Malay peninsula, is included in the sample (but was coded as “other” for location), and will be discussed in this section as a point of comparison.

The Austronesian languages of mainland Southeast Asia and the northern part of Peninsular Malaysia show a strong tendency to lack applicatives altogether. Chamic languages spoken in mainland Southeast Asia all lack applicatives (3 of 3), as do Moken and Moklen (Larish 2005). Malayic languages spoken in mainland Southeast Asia (2 of 2; Pattani, Urak Lawoi’) again lack applicatives, and this holds also for the Nonthaburi variety spoken in Central Thailand. In northern Peninsular Malaysia, Kedah Malay lacks applicatives (Omar 1981: 7), while Kelantan Malay appears to retain some pivot-neutral ACs formed with the suffix *-i*, but with very limited productivity (Mahmood 1994: 56–58).

Moving further southward, Negeri Sembilan Malay, spoken in southern Peninsular Malaysia, has productive pivot-neutral ACs formed with both *-i* and *-kan* (Hendon 1966: 61–68), just like Standard Malay and Indonesian. On the other hand, Singapore Bazaar Malay lacks applicatives altogether.

Austronesian languages of mainland Southeast Asia and Peninsular Malaysia that lack applicatives tend to show typological features different from most other languages of West Nusantara. Chamic languages have asymmetrical voice systems (3 of 3), do not show morphological marking that distinguishes core arguments by grammatical relation (neutral alignment, 3 of 3), and show AVP word order in transitive constructions (3 of 3), with word order being important for interpreting grammatical relations. Urak Lawoi’ and Singapore Bazaar Malay share these characteristics, as do Nonthaburi Malay (Tadmor 1995), Moken and Moklen (Larish 2005).

Pattani Malay as spoken in southern Thailand and the Kelantan dialect both appear to retain an alternation between an A-oriented transitive construction and a P-oriented transitive construction (two-way symmetrical voice), with special marking of the A argument in the P-oriented construction (Tadmor 1995: 249–250; Mahmood 1994: 201–207). The P-oriented construction in these varieties differs somewhat from the PV construction in Standard Malay/Indonesian. The PV prefix *di-* is not used to mark the verb; instead the morpheme *(a)no* precedes the verb, while another particle, *di/d* or *kə/kə/k* precedes the A argument.⁹ This construction shows some formal similarities to analytic PV constructions found in some languages of Borneo (e.g. Land Dayak languages, see §5.8.2, and North Sarawak languages, see §5.8.3).

Negeri Sembilan Malay, on the other hand, shows morphological marking for PV with *di-* and optional prepositional marking for the non-pivot A argument (Hendon 1966: 68–69), and thus is similar to Standard Malay/Indonesian and the languages of Java and Madura described above.

⁹The particle marking the A argument may also be realized as gemination on the initial consonant of the NP expressing A.

5.7.1 Language contact, restructuring, and the loss of applicatives

We know that restructuring of Chamic languages happened in mainland Southeast Asia due to contact with Austroasiatic languages. This restructuring is marked by phonological changes, i.e., reduction of multisyllabic words and development of final word stress, morphological changes, i.e., reduction of affixation and shift toward analytic or isolating word structure, and syntactic changes, i.e., development of more fixed word order, loss of case distinctions, use of periphrastic constructions to express meanings similar to morphologically marked causatives and ACs in other West Nusantara languages (Thurgood 1999, 2010; Brunelle 2020).

An example of a periphrastic construction with a benefactive meaning is given in (102). Here, the particle *brei*, which elsewhere may function as a full lexical verb meaning ‘give’, precedes the beneficiary, *kơ ñu* ‘for him’, which is always expressed as prepositional phrase in this type of construction. This construction is not an applicative as defined in this study because *brei* is an independent morpheme and its distribution is not closely tied to the verb or verbal complex; here the noun expressing the P argument, *brăm* ‘arrow’, and its modifiers intervene between the closest verb and *brei*. Instead, this type of clause may be classified as a serial verb construction (see §3.5).

(102) Bih, Periphrastic benefactive construction with ‘give’

Thô gơ magĩr ngă ana năn

T. 3 PFX.try make crossbow DIST

ngă leh ngă ana rĩ brăm,
make PFV make crossbow whittle arrow

dua tlâu urăt, brei kơ ñu.

two three CL BEN DAT 3

‘Thô tried to make a crossbow and some arrows for him.’

(Nguyen 2013: 90)

The pattern of restructuring observed for Chamic is also borne out for Moklen and Moken. Both lack applicatives, lack productive affixation, and show word final stress (Larish 2005). Moklen is also reported to make “heavy use of serial verb constructions” (Larish 2005: 527).

Malay varieties of Thailand and northern Peninsular Malaysia show similar changes to those found in Chamic at lower time depths. There is no productive affixation at all in Nonthaburi Malay. Pattani Malay retains a single affix, nominalizing */-ε/*, and all other inherited affixes are found only in loanwords or remnant forms (including causative *per-*, which is rare and has been reduced to initial gemination) (Tadmor 1995: 228–232). Kelantan Malay is similar to Pattani, except it retains limited use of *-i* as a causative or applicative suffix (Mahmood 1994: 56–60). These varieties also form benefactive constructions by use of the verb meaning ‘give’, as shown in example (103) below. Urak Lawoi’, a Malay language spoken on islands in Southern Thailand, has no applicatives and shows several productive prefixes but no suffixes. Hogan (1988) writes that, “when [Urak Lawoi] is compared with Bahasa Malaysia [Standard Malay], it is evident that it has a much smaller inventory of affixes, and that these are used much less frequently than the corresponding affixes in that language.” While Urak Lawoi’ has not developed word final stress, it shows signs of phonological reduction in non-final syllables, which include stressless “pre-syllables” and “minor syllables” that may bear stress but show a reduced number of phonemic

contrasts in coda position compared to the final syllable (Hogan 1988: 11–12). A pattern of reduced affixation may also be found in Kedah Malay. Omar (1981: 7) writes, “in the Kedah dialect of North Malaya there is hardly any verbal suffix, while the nominal suffix *-an* has a very low frequency. The causative function of *-kan* is, in this dialect, fulfilled by the prefix *pər-* or a verbal phrase consisting of two verbs, while its benefactive function fulfilled by a prepositional phrase.”

(103) Nonthaburi Malay, Periphrastic benefactive construction with ‘give’

məʔ bli tpoŋ byi anəʔ makin
 mother buy snack give child eat

‘The mother bought snacks for her children.’

(Tadmor 1995: 261)

Finally, Singapore Bazaar Malay, used as a language of interethnic communication in Singapore, has also been shaped by language contact with non-Austronesian languages, especially Hokkien (Sinitic). It employs only compounding and reduplication in word formation, and shows no use of affixation (Aye 2005: 62). Serial verb constructions with the verb *kasi* ‘give’ are used to express recipient and beneficiary participants (Aye 2005: 290–292).

In summary, the languages of mainland Southeast Asia and northern Peninsular Malaysia show a strong tendency to lack applicatives, and a distinct typological profile compared to most other West Nusantara languages. These changes were caused by language contact with non-Austronesian languages, which resulted in changes in word structure, a shift away from morphological processes of affixation, and a shift towards analytic structures, including the use of serial verb constructions to express some meanings elsewhere associated with applicatives.

5.8 Borneo and the Southern Philippines

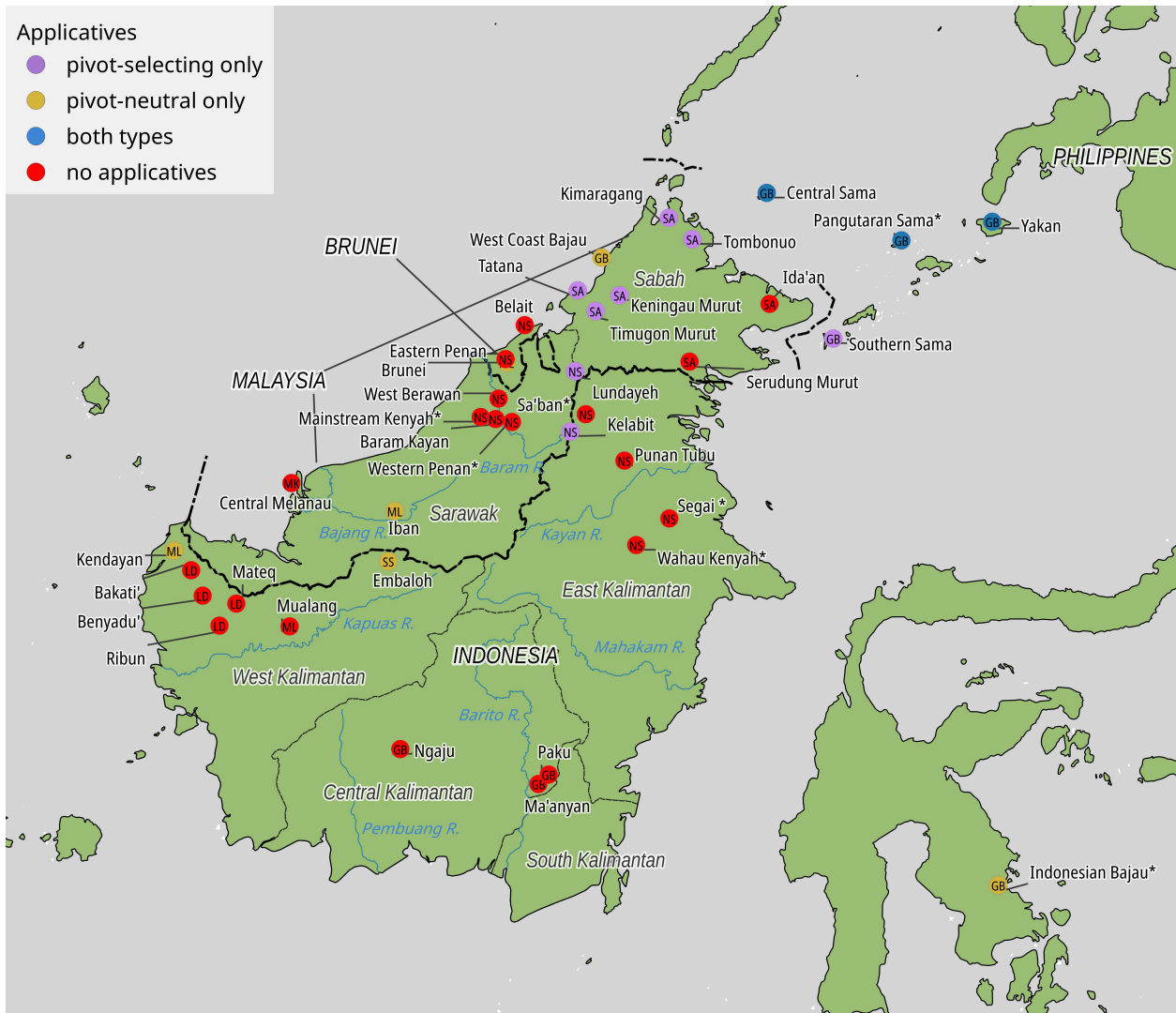
In this section, I present and discuss results of the typological survey for languages of Borneo and nearby areas of the Southern Philippines. Borneo is the largest major island in West Nusantara and the most linguistically diverse, with about 200 total languages spoken on the island. This section is organized by genetic grouping, and covers Sabahan languages (§5.8.1), Land Dayak languages (§5.8.2), North Sarawak and Melanau-Kajang (§5.8.3), Greater Barito languages, including Malagasy and Sama-Bajaw languages which are geographically dispersed outside of Borneo (§5.8.4), and other languages spoken in Borneo, but belonging to genetic groupings primarily located in other parts of West Nusantara (§5.8.5). The map in Figure 5.8 gives an overview of the distribution of applicatives in the languages of Borneo and the Southern Philippines.

5.8.1 Sabahan Languages

As reflected in their name, Sabahan languages are spoken across the Malaysian state of Sabah in northeastern Borneo, though some members of the group are also found in Brunei and parts of North Kalimantan Province adjacent to Sabah. As shown in Table 5.12, seven Sabahan languages are included in the sample.

Sabahan languages for the most part show productive Philippine-type voice systems (5 of 7 languages in the sample), and lack pivot-neutral applicatives (7 of 7). Sabahan languages are predominantly verb-initial (6 of 7) and typically exhibit systems of morphological alignment that

Figure 5.8: Overview map of languages of Borneo and the Southern Philippines



* = Not included in language sample but discussed in §5.8. Abbreviations: [GB] Greater Barito, [ML] Malayic, [MK] Melanau-Kajang, [NS] North Sarawak, [SA] Sabahan, [SS] South Sulawesi. Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

mark the pivot grammatical relation distinctly from non-pivot core arguments (6 of 7 languages). They also tend to show complex systems of verbal morphology, as morphological marking for several TAM categories appear on the verb (e.g. tense, volitionality, and non-indicative moods). In these respects Sabahan languages are more similar to Philippine and Formosan languages than other languages of Borneo. The common features of Sabahan languages noted here are broadly distributed, and are found across primary branches of Sabahan (i.e. Greater Dusunic [Kimaragang], Murutic [Timugon and Keningau Murut], and Paitanic [Tombonuo]).

Table 5.12: Typological survey results for Sabahan languages

Gen. Grp.	Name	Applicatives?		Voice	Morph. align.	Word Order
		pivot-selecting	pivot-neutral			
Sabahan	Kimaragang	Y	N	Philippine-type	pivot-nonpivot	VAP/VPA
Sabahan	Keningau Murut	Y	N	Philippine-type	pivot-nonpivot	VAP
Sabahan	Timugon Murut	Y	N	Philippine-type	pivot-nonpivot	VAP/VPA
Sabahan	Tombonuo	Y	N	Philippine-type	pivot-nonpivot	VAP
Sabahan	Tatana	Y	N	Philippine-type	pivot-nonpivot	VPA
Sabahan	Ida'an	N	N	two-way	pivot-nonpivot	AVP
Sabahan	Serundung Murut	N	Y	two-way	mixed-NPIV.A	VPA

5.8.1.1 Pivot-selecting applicatives in Sabahan

Most Sabahan languages show between four and six contrastive voice categories, of which two or three are pivot-selecting applicatives (see discussion of voice and applicatives in Kimaragang in §3.4). Tatana shows six contrastive voices, presented in example (104).¹⁰ Of these six constructions, four can be considered pivot-selecting applicatives, namely, theme voice (TV) in which the pivot is a theme or object undergoing transfer, instrument voice (IV), “referent voice” (RV), in which the pivot is a semantic beneficiary, recipient, or content item, and finally, locative voice (LV), in which the pivot expresses the time or location of the clausal event.¹¹ Morphological marking for pivot-selecting applicatives in Sabahan is closely integrated into the complex systems of verbal morphology in these languages. Table 5.13 presents a partial paradigm for verbal morphology in Tatana, which distinguishes volitive and non-volitive moods, and indicative and imperative moods, though not all TAM categories may be marked in all of the six voices.¹²

(104) Tatana, Philippine-type voice alternations

- a. *Mom-(p)opot aku do daging karabau.*
 AV.TR.NPST-chop 1SG.NOM DAT meat buffalo
 ‘I am chopping buffalo meat.’ (AV) (Dillon 1994: 22)

¹⁰Tatana is considered an unclassified Sabahan language by Eberhard, Simons & Fennig (2021), but alternately, may subgroup with Murutic (Hammarström et al. 2022).

¹¹The construction that I call locative voice here is called “setting voice” by Dillon (1994). It should not be confused with the separate locative adversative construction marked with *-on* in Tatana. The locative adversative is not a basic voice category in Tatana (see Dillon 1994: 61–64).

¹²Intransitive AV shows distinct marking from transitive AV in Tatana. For the sake of simplicity, the former is not included in Table 5.13.

- b. *Boli-in ku dudungu' diti*
 buy-PV 1SG.GEN banana this
 'I am buying **these bananas**.' (PV) (Dillon 1994: 44)
- c. *I-taak ku sita' ku dokou*
 CV-give 1SG.GEN shirt 1SG.GEN 2SG.DAT
 'I'll give **my shirt** to you.' (TV) (Dillon 1994: 47)
- d. *Duit pam-(b)ali ku do kana'*
 money IV-buy 1SG.GEN DAT fish
 'I buy fish **with money**.' (IV) (Dillon 1994: 49)
- e. *Bali-an ku okou do dudungu.*
 buy-RV 1SG.GEN 2SG.NOM DAT banana
 'I am buying bananas **for you**.' (RV) (Dillon 1994: 52)
- f. *Kadai diti andang-andang pam-(b)ali-an ku*
 shop this RDP-usual LV-buy-LV 1SG.GEN
 '**This shop** is where I usually buy things.' (LV) (Dillon 1994: 60)

Table 5.13: Partial paradigm for Tatana verbs

Voice category	Volitive		Non-volitive		Imperative
	Non-past	Past	Non-past	Past	
Actor voice (AV TR)	moN-	noN-	moko-	noko-	poN-
Patient voice (PV)	-on	-in- -∅	mo- -∅	no- -∅	-o'
Theme voice (TV) [†]	i-	—	moko-	noko-	—
Instrument voice (IV) [†]	poN-	pinoN-	—	—	—
Referent voice (RV) [†]	-an	-in- -an	mo- -an	no- -an	-i'
Locative voice (LV) [†]	poN- -an	pinoN- -an	—	—	—

[†] These voice categories constitute pivot-selecting applicative constructions.

5.8.1.2 Outliers in Sabahan

Two Sabahan languages in the sample, Ida'an and Serudung Murut, are outliers in the group. Both have lost Philippine-type voice, and thus the pivot-selecting applicatives. Neither shows evidence of development of pivot-neutral applicatives.

The voice system for Ida'an has been reduced to AV and PV for basic transitive constructions (Goudswaard 2005). The language shows other morphological and phonological characteristics unusual for Sabahan languages and more like that of the North Sarawak languages to the south, i.e., final stress, ablaut, and loss of all suffixation (Goudswaard 2005: 67–69). Ida'an also shows a shift away from verb-initial word order towards verb-medial word order (AVP in AV). Word order is an important indicator of grammatical relations, as only the pivot argument can precede the verbal predicate. Together, these are clues that Ida'an may have undergone a path of grammatical restructuring similar to that many Borneo languages south of Sabah that completely

lack applicatives, such as Land Dayak (§5.8.2) and North Sarawak languages (§5.8.3). It has also been suggested that Ida'an and Bonggi—both Northeast Sabahan languages—should be classified separately from the other (Southwest) Sabahan languages (Lobel 2013).

Serudung Murut is also analyzed as a two-way symmetrical voice system, with one A-oriented transitive construction (A-Voice) and one contrasting undergoer-oriented transitive construction (P-Voice) (Townsend 2017). Unlike other Sabahan languages, in Serudung Murut pivot arguments and non-pivot P arguments are marked alike morphologically (use of bare NP or nominative free pronoun), while non-pivot A arguments have distinctive marking (use of genitive linker *nu* or genitive enclitic pronoun), see example (105) below.

It is not entirely clear why Philippine-type voice was lost in Serudung Murut. However, it appears that extreme lexicalization of verbal suffixes has occurred, obscuring the distinction between patient voice and some former circumstantial voice (CV) category, that is, a voice alternation in which an instrument-, theme- or beneficiary- is the pivot (see Wolff 1973; Ross 2002, 2009; Chen 2017 on CV as a reconstructed voice category in PAn that encompasses these functions). Verb forms in undergoer-oriented transitive clauses in Serudung Murut may take either (i) suffixal marking *-on/-oʔ* or (ii) suffixal marking *-in/-iʔ* in imperfective aspect and *-an* in perfective aspect (Townsend 2017: 27–29). The former is noted to occur with verbs that take patient-like P arguments, and the latter with verbs that take theme-like P arguments. Elsewhere in Sabahan, *-on* is the normal marker of patient voice, while *-an/-in/-iʔ* mark a type of CV in various TAM categories (as with Tatahan beneficiary-selecting RV discussed above). Occurrences of *-an/-in/-iʔ* as verbal suffixes in Serudung Murut might thus be considered a type of degraded or non-productive pivot-selecting CV construction. However, the alternation between the two suffixal sets in the present-day language is completely lexically-conditioned; no single verb may take both sets, as shown in (105) for the verb *taak* 'give'. The two suffixal sets therefore are not contrastive, and neither would mark a fully productive voice category on its own. For this reason, I consider Serudung Murut to represent a case of attrition of applicatives that can be attributed—at least in part—to lexicalization. Remnants of a former CV construction remain only as an apparent irregularity in the paradigm for morphological marking of P-Voice.

(105) Serudung Murut, Voice alternations with 'give'

a. *aku an-(t)aaak lamun sokou*
 1SG.NOM AV-give rice 2SG.OBL
 'I give rice to you.' (AV)

b. *lamun taak-in=ku sokou*
 rice give-PV=1SG.GEN 2SG.OBL
 'Rice is given to you by me.' (PV)

c. **taak-on*
 *ungrammatical, not a possible word

(Townsend 2017: 29–30)

5.8.2 Land Dayak languages

On the other end of Borneo from Sabah, we find the Land Dayak languages, a group of about 15 languages spoken in southwestern Borneo, including the far western extent of Sarawak state

and bordering parts of West Kalimantan Province, stretching south into inland West Kalimantan. There is little linguistic description available for most Land Dayak languages. Four languages are included in the survey: Benyadu', Bakati', Matéq and Ribun. Two additional languages for which short sketches are available—Bidayuh Serian and Biatah Bidayuh—will also be referenced (Omar 1983).

Table 5.14: Typological survey results for Land Dayak

Gen. Grp.	Name	Applicatives?		Voice	Morph. align.	Word Order
		pivot-selecting	pivot-neutral			
Land Dayak	Benyadu'	N	N	two-way	pivot-nonpivot	AVP
Land Dayak	Bakati	N	N	two-way	mixed-NPIV.A	AVP
Land Dayak	Matéq	N	N	two-way	neutral	AVP
Land Dayak	Ribun	N	N	two-way	mixed-NPIV.A	AVP

All Land Dayak languages show no applicatives of any type (4 of 4 languages of the sample, plus Bidayuh languages). There is also no evidence that pivot-selecting applicatives can be reconstructed for Land Dayak, and no evidence that these languages ever had pivot-neutral applicatives.

Land Dayak languages generally have two-way symmetrical voice systems (4 of 4 languages of the sample). AV constructions have unmarked AVP word order in AV constructions (4 of 4). In all four languages of the sample, word order is an important indicator not just of grammatical relations, but also the voice category of a transitive clause, as morphological marking for voice on the verb no longer clearly distinguishes AV and PV, with the standard word order for PV being PAV. Morphological marking for arguments distinguishing core grammatical relations also shows signs of reduction. Benyadu regularly makes use of pronominal sets distinguishing non-pivot A arguments from other core arguments. But in other Land Dayak languages this distinction is lost—as in Matéq, which has neutral alignment with all core argument marked the same—or almost entirely lost, being found only for third plural pronominal actors in Bakati, and optionally with second-person singular pronominal actors in Ribun. In addition, Land Dayak languages show complete loss of suffixation.¹³ Here, as in Sumatra and mainland Southeast Asia, loss of suffixation appears to be one major factor associated with lack of applicatives.

Related to reduction of verbal morphology, the four Land Dayak languages of the sample show a shift towards analytic means of marking the distinction between AV and PV. The verb in AV clauses is marked with the prefix *N-* in Land Dayak, as is common in West Nusantara. Unusual however, is the marking of PV clauses, in which the verb also may bear *N-*. This is shown in (106) below from Matéq. Here the morpheme marking PV is not a verbal affix, but a particle *ni* that precedes both the A argument (underlined), if expressed, and the verb bearing *N-*. In Matéq both analytic PV and morphologically marked PV are available, but the latter only with certain verbs. Similar analytic PV constructions are found in Ribun, Benyadu, and Bakati' (Sommerlot 2020), as

¹³Suffixation is not found in the four Land Dayak languages of the sample (Mateq shows a discourse marker =*éh*, which behaves as a clitic). There is also no suffixation in Bidayuh Serian [sdo] (Omar 1983: 447–451), Biatah Bidayuh [bth] (Omar 1983: 446–469), nor in Sungkukng and other Bidayuhic languages in the Bengkayang, Landak and Sanggau regencies of West Kalimantan, Indonesia (Adelaar 2006: 81 citing personal fieldnotes). Rensch et al. (2012) find only enclitics, prefixes, and a few infixes—some fossilized—in their study of over 25 Bidayuh varieties.

well as Kendayan (Salako), a Malayic language spoken nearby in West Kalimantan and western Sarawak (Adelaar 2002). The particle marking PV in Ribun and Salako can be omitted if clausal word order is PAV, showing that word order is a primary indicator of PV in these languages.

(106) Matéq, Analytic PV

- a. *pingàt aiq yoh ni koq moruh*
 plate that PRT PV 1SG AV.smash
 ‘I smashed the plate’
- b. *ni ular aiq degeq nyora ruba turuaq=ng*
 PV snake that constantly AV.attack hole dibbling.stick=3
 ‘the snake kept on attacking their dibbling holes’ (Connell 2013: 113)

In summary, Land Dayak languages show a shift away from morphologically marking for voice and grammatical relations and towards analytic means of indicating the same, especially in PV constructions. This trend and the loss of suffixation probably account for the loss of pivot-selecting applicatives and lack of pivot-neutral applicatives in Land Dayak.

5.8.3 North Sarawak and Melanau-Kajang languages

This section covers the North Sarawak group and the Melanau-Kajang group of languages, which are spoken across north central Borneo to the south of Sabah. Melanau-Kajang languages are concentrated in central Sarawak state, while North Sarawak languages are spoken across a broad area covering in northern and central Sarawak state, North Kalimantan and East Kalimantan. North Sarawak as defined in the study includes 41 total languages, with eight selected for the sample. Melanau-Kajang includes 12 total languages, of which just one (Central Melanau, Mukah dialect) was able to be included in the sample.

Many languages of these two groups lack adequate linguistic description and even basic documentation. Making matters more complicated, a number of subbranches of North Sarawak as defined in the sample have a disputed classification, e.g. Kayanic languages, Kenyah languages, and the Punan-Mueller Schwaner languages (Eberhard, Simons & Fennig 2021, c.f. Hammarström et al. 2022). On top of this, even languages within the same subbranch of North Sarawak are known to show important differences in the verbal system (see for example B. Clayre 2005 on voice in the Dayic subbranch of North Sarawak). For this reason, in this section I will make mention of a number languages not included in the survey sample for which some sketch material and limited description of verbal constructions are available. These are (i) for Kayanic languages, Segai (Soriente 2013), (ii) for Kenyah languages, the Lepo Ké dialect of Mainstream Kenyah, the Lebu’ Kulit dialect of Wahau Kenyah, and the Penan Benalui dialect of Western Penan (B. Clayre 1996; Omar 1983; Soriente 2013); and (iii) for Dayic languages, Sa’ban (B. Clayre 2005). Unfortunately, there are no grammatical descriptions for Punan-Mueller-Schwanner languages nor for Melanau-Kajang languages beside Central Melanau; these remain unrepresented, or nearly so, in the sample.

The large majority of languages of North Sarawak (6 of 8 languages in the sample) lack applicatives of any type, as does Central Melanau (1 of 1).¹⁴ Other Kenyah languages and the Dayic

¹⁴For Punan Tubuh, Soriente 2013: 184 makes brief mention of a possible benefactive function for the verbal

Table 5.15: Typological survey results for North Sarawak and Melanau-Kajang

Gen. Grp.	Name	Applicatives?		Voice	Morph. align.	Word Order
		pivot-selecting	pivot-neutral			
N. Sarawak	Kelabit	Y	N	Philippine-type	neutral	VPA
N. Sarawak	Lun Bawang	Y	N	Philippine-type	pivot-nonpivot	AVP/VPA
N. Sarawak	West Berawan	N	N	two-way	mixed-NPIV.A	AVP
N. Sarawak	Belait	N	N	two-way	mixed-NPIV.A	AVP/VPA
N. Sarawak	Baram Kayan	N	N	two-way	pivot-nonpivot	AVP
N. Sarawak	Eastern Penan	N	N	two-way	mixed-NPIV.A	AVP
N. Sarawak	Punan Tubuh	N	N	two-way	mixed-NPIV.A	AVP
Melanau-Kajang	Central Melanau	N	N	two-way	mixed-NPIV.A	AVP

language Sa'ban also appear to lack applicatives (B. Clayre 2005; Soriente 2013). Within North Sarawak, just two languages (2 of 8) show applicatives; both Kelabit and Lun Bawang (Lundayeh) exhibit a pivot-selecting instrumental applicative (IV).

Of the North Sarawak languages lacking applicatives in the sample, all six have two-way symmetrical voice systems, as does Central Melanau. This appears to hold for Sa'ban (Dayic), Segai (Kayanic), and Western Penan (Kenyah). However, B. Clayre (1996: 74) reports that the Lepo Ké dialect of Mainstream Kenyah does not have an undergoer-oriented construction, which indicates that its voice system is asymmetrical. On the other hand, Kelabit (Philippine-type), shows AV, PV, and IV voice categories (Hemmings 2016: 201–204), and Lun Bawang (marginal Philippine-type), shows AV, PV, and an archaic IV category that is rare in current usage (Mortensen 2021: 115); for more detail see §5.8.3.1 below.

Most languages of North Sarawak show a system of morphological alignment in which the non-pivot A argument has distinctive marking compared to other core arguments, which are marked alike (5 of 8). This is also found in Central Melanau. The distinctive form for non-pivot A is generally limited to pronominal arguments, and often found only in certain number and person categories, as in Punan Tubu, where there are distinct pronominal forms for first- and second-person singular non-pivot A arguments, and third singular non-pivot A can be expressed either with a distinct form or with the free pronoun form used for all other core arguments. In Sa'ban, distinctive forms of the pronouns for non-pivot A have been lost (neutral alignment), these being found only in fossilized remnants (B. Clayre 1996: 57). Two languages of North Sarawak (2 of 8; Lun Bawang, Baram Kayan) show one distinct marking for pivot arguments, and other types of marking for non-pivot core arguments (pivot-non-pivot). It looks like Kelabit formerly had such a system, but it has been collapsed such that core arguments have distinct marking distinguishing them from obliques, but are now not otherwise distinguished for grammatical relation (neutral alignment, Hemmings 2016: 331–332).

In terms of word order, most languages of North Sarawak show a preference for AVP unmarked word order in A-oriented constructions (verb-medial; 5 of 8), which also holds for Central Melanau. A number of these languages also allow VPA order in AV, but it is not always reported whether or not this is possible. In Belait and Lundayeh, word order is reported to be split between

prefix *pe-*. However, this is not illustrated with lexical or clausal examples by which it can be verified that clauses marked with this morpheme meet the definition of an AC used in this study. I have classified Punan Tubah as lacking applicatives for this reason.

AVP and VPA in AV, with no indication of preference. In Kelabit, VPA is the most natural word order for AV (verb-initial), but VAP and AVP are also possible. Thus, it appears that most North Sarawak languages show a shift away from preference for verb-initial word order as is typical in Philippine-type languages and most common in Sabahan languages. Strict ordering of verb + non-pivot core argument is common here, with this ordering being important for signalling grammatical relations, especially in AV. For example, in West Berawan, word order is more strict in AV than in PV; in AV only AVP is possible, while in PV, both PVA and VAP are found (B. Clayre 1997: 234).

Regarding verbal morphology and general morphological complexity, Lun Bawang and Kelabit are the most conservative. Only these two languages retain productive use of verbal suffixes, including PV *-en* and imperative suffixes, though the latter are reportedly rarely found in Lun Bawang (Hemmings 2016; Mortensen 2021). For all other languages of North Sarawak there is no evidence of productive suffixation, verbal or otherwise. In Central Melanau, there is one reported suffix, *-ai*, but it is found only on a very small number of directional adverbs (I. F. C. S. Clayre 1972: 210–211). Thus, as in languages of mainland Southeast Asia and Land Dayak languages, loss of suffixation is associated with complete lack of applicatives in North Sarawak and Melanau-Kajang.

A number of North Sarawak languages are reported to make use of an analytic or periphrastic PV construction, most commonly marked with the verb ‘do/make’. This is reported in Lun Bawang, Sa’ban, and Dayic varieties of East Kalimantan, as well, as Eastern Penan (Kenyah), Central Berawan and some Kayan varieties (see Mortensen 2021; B. Clayre 1996: 75–81, 2005).¹⁵ An example of this is given in (107) from Sa’ban, where this construction is replacing morphologically marked PV (B. Clayre 2005: 30). This construction can be considered a type of serial verb construction (see §3.5), but not an applicative, as it does not select a peripheral semantic role as a core argument. It appears that no construction of this type is found in Central Melanau.

(107) Sa’ban, Periphrastic PV with ‘make’

Ayeu noknai an ieh m-paeng.
 tree this make 3SG AV-cut.down
 ‘He will cut down this tree.’

(B. Clayre 1996: 78)

In summary, Central Melanau and most languages of the North Sarawak group show loss of Philippine-type voice (pivot-selecting applicatives) and lack of pivot-neutral applicatives. Languages in this area that lack applicatives altogether also lack suffixation, and tend to show reduced verbal morphology. In some of these languages, we also see a shift towards analytic structures with word order becoming important for signalling grammatical relations and increasing use of periphrastic PV constructions, a type of serial verb construction.

5.8.3.1 Outliers in North Sarawak

The only North Sarawak languages that retain Philippine-type voice alternations are Lun Bawang and Kelabit. Geographically, the areas where Lun Bawang and Kelabit are traditionally spoken

¹⁵This construction may not necessarily be an innovation, but its frequent use in place of morphologically marked PV, especially in indicative clauses, does appear to mark a diachronic shift. The form of the verb meaning ‘do/make’ in North Sarawak languages is often *an*, or similar, which is possibly a reflex of a preverb **an*, held to be used in CV in non-indicative moods in the PAn voice system by Ross (2009). For more, see the discussion in §6.1.6 below.

stand in the northeast corner of Sarawak state, with Lun Bawang to the north of Kelabit.

Both Kelabit and Lun Bawang show an instrument-selecting voice construction (IV) that contrasts with AV and PV (Hemmings 2016; Mortensen 2021), as in example (108). In this construction, shown in (108c), the instrument is always the clausal pivot (bolded); thus, this construction constitutes a pivot-selecting applicative. The IV construction in Lun Bawang is reportedly rarely used in spontaneous, spoken language, but it is still found in some written documents and “many speakers are aware of it and can produce it and judge its usage upon request” (Mortensen 2021: 115).

(108) Kelabit, Transitive voice alternations

- a. *La'ih sineh ne-nekul nubaq nedih ngen seduk.*
man DIST PFV-AV.spoon.up rice 3SG.POSS with spoon
 ‘That man spooned up his rice with a spoon.’ (AV)
- b. *Sikul lai'h sineh nubaq nedih ngen seduk.*
 PV.PFV.spoon.up man DIST **rice** 3SG.POSS with spoon
 ‘That man ate his rice with a spoon.’ (PV)
- c. *Seduk pe-nekul la'ih sineh nubaq nedih.*
spoon IV-spoon.up man DIST rice 3SG.POSS
 ‘That man used a **spoon** to spoon up his rice.’ (IV) (Hemmings 2016: 303)

Neither Lun Bawang nor Kelabit show a productive LV voice construction. In Kelabit, however, the suffix *-an* is retained in fossilized locative nominalizations, and a remnant of LV is found with a single verb, *tu'an* ‘to do/put, LV’ (Hemmings 2016: 145–146). In this construction, shown in (109), the semantic location is the clausal pivot (bolded), as is indicated by its access to relativization in (109b).

(109) Kelabit, Remnant locative voice construction

- a. *Lidung tu'an neh babeh nedih.*
corner do/put.LV 3SG.CORE.II bag 3SG.POSS
 ‘He put his bag in **the corner**.’ (LV) (Hemmings 2016: 145)
- b. *Seni'er kuh lidung [suk tu'an neh babeh nedih].*
 PV.PFV.see 1SG.CORE.II **corner** [REL put.LV 3SG.CORE.II bag 3SG.POSS]
 ‘I saw the **corner** [where he put his bag].’ (LV, relative clause)
 (Hemmings 2016: 218)

Kelabit and Lun Bawang belong to the Dayic primary branch of North Sarawak languages. Other Dayic languages, including Sa'ban, Lengilu', and the Kerayan dialects, whether spoken to the south of Kelabit in Sarawak or to the east in North Kalimantan, have lost IV and now show a two-way voice system (B. Clayre 2005), which is also the norm for the rest of the North Sarawak group. Thus, it appears that North Sarawak languages show decay of the Philippine-type voice system, but this process is progressing slower in Kelabit and Lun Bawang than the rest of the group, as Kelabit and Lun Bawang have lost LV, while retaining IV (at least to some extent). This retention of IV should be viewed in light of a general pattern of linguistic conservatism for

Lun Bawang and Kelabit, which is likely due to the two languages’ geographic position in the mountainous highlands, affording a degree of separation from other North Sarawak languages, including other Dayic varieties as well as Kenyah and Kayan languages present in the region (see Mortensen 2021: 232–256 on contact and dialect-mixing leading to spread of innovative features in Dayic outside of Lun Bawang and Kelabit).

5.8.4 Greater Barito languages

Greater Barito is a linkage representing 35 languages that originated in the Barito River basin encompassing much of southwestern Borneo (Smith 2018). About 14 of these languages are spoken today near their homeland, in the Indonesian provinces of Central Kalimantan and East Kalimantan, while the others show an extremely broad geographic dispersion. Eight Greater Barito languages are included in the sample, as shown in Table 5.16. In the survey results, Greater Barito languages fall into three types. These are (i) the Malagasy languages of Madagascar in eastern Africa, which show Philippine-type voice and thus pivot-selecting applicatives, (ii) the Sama-Bajaw languages found in the Philippines and across West Nusantara, which show a combination of pivot-selecting and pivot-neutral applicatives, and (iii) languages spoken in the Barito River basin, which show no applicatives of any type.

Table 5.16: Typological survey results for Greater Barito languages

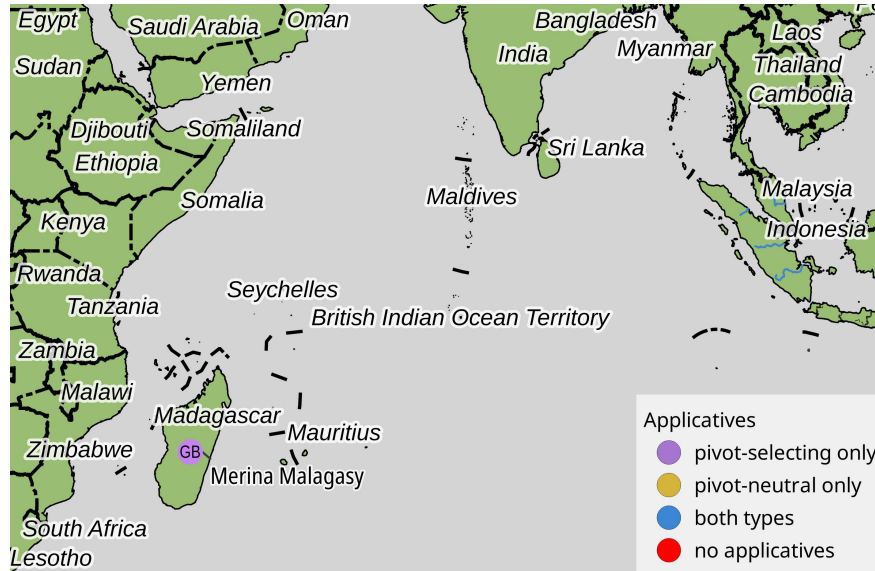
Gen. Grp.	Name	Applicatives?		Voice	Morph. align.	Word Order
		pivot-selecting	pivot-neutral			
Gr. Barito	Merina Malagasy	Y	N	Philippine-type	pivot-nonpivot	VPA
Gr. Barito	Southern Sama	Y	N	Philippine-type	mixed-NP.V.A	VPA
Gr. Barito	Central Sama	Y	Y	marg. Philippine-type	mixed:other	VPA
Gr. Barito	Yakan	Y	Y	marg. Philippine-type	pivot-nonpivot A	VAP/VPA
Gr. Barito	West Coast Bajau	N	Y	two-way	pivot-nonpivot A	AVP
Gr. Barito	Ngaju	N	N	two-way	pivot-nonpivot	AVP
Gr. Barito	Ma’anyan	N	N	two-way	mixed-NP.V.A	AVP
Gr. Barito	Paku	N	N	two-way	mixed-NP.V.A	AVP

5.8.4.1 Malagasy

The Malagasy languages represent a group of closely-related languages spoken on the island of Madagascar, some 3,800 miles from West Nusantara, off the eastern coast of Africa. All belong to Malagasic, a lower-level subgroup of the Greater Barito linkage, which falls under the same branch as Paku and Ma’anyan, which are spoken in the Barito River basin. Only one Malagasy language, Merina Malagasy, is represented in the sample, as it has the most available resources due to its status as the most prestigious variety. While the Malagasy languages are listed under 12 separate ISO-639-3 codes, they are by all accounts very similar to one another, such that they are commonly referred to as “dialects,” and Rasoloson & Rubino (2005) state that they are “often so closely related to one another that a clear group classification is uncertain” (456). In some cases, the identification of the variety represented in available source material is not straightforward.

For the purposes of this survey, most Malagasy languages would be quite similar in profile to Merina. A map showing the location for Merina Malagasy is presented in Figure 5.9.

Figure 5.9: Map showing location for Merina Malagasy



Abbreviations: [GB] Greater Barito. Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

Compared with other West Nusantara languages, Merina, like other Malagasy languages, is relatively syntactically conservative, while showing phonological and lexical innovations due to contact with Bantu languages in Coastal East Africa (Adelaar 2012). Merina retains a Philippine-type voice system with five contrastive voice categories (Pearson 2001). Three of these represent pivot-selecting applicatives: theme (or conveyance) voice (TV), dative voice (DV) and circumstantial voice (CV). The language is verb-initial, with the pivot in final position in unmarked word order. In Merina, only pronominal forms are distinguished morphologically for core grammatical relations. The pivot relation is marked with the nominative set of pronouns, while other non-pivot core arguments are marked with dative or genitive pronouns (pivot-non-pivot morphological alignment). With respect to these typological characteristics, Malagasy is quite similar to most Sabahan languages and quite different from other Greater Barito languages.

5.8.4.2 Sama-Bajaw languages

Sama-Bajaw languages also show geographic dispersal outside of Borneo, though most are still spoken inside West Nusantara and nearby in the southern Philippines. In total there are nine Sama-Bajaw languages, which together constitute one exclusive subgroup. Four of these are represented in the sample, but in order to capture patterns in the diverse typological features of this group, I will also refer to three others in the discussion here: Pangutaran Sama, Sama Balangingi, and Indonesian Bajau (Walton 1986; Gault 1999; Donohue 1996).

Sama-Bajaw languages represent a cline of typological characteristics showing gradual decline of Philippine-type voice (and thus pivot-selecting applicatives) and development of pivot-neutral applicatives. The voice and applicative systems of these languages fall into three types,

as represented in Table 5.17. One set of languages (Southern Sama, Sama Balangingi) has productive Philippine-type systems with five voice categories (AV, PV, IV, BV, LV). A second set (Central Sama, Yakan, Pangutaran Sama) shows marginal Philippine-type systems with reduced productivity or restricted distribution of LV and IV, and no use or only minimal use of BV. This set also has productive pivot-neutral applicatives. The third set (West Coast Bajau, Indonesian Bajau) shows two-way symmetrical systems and productive pivot-neutral applicatives only.

Table 5.17: Applicatives in Sama-Bajaw languages

Pivot-selecting only	Both types*	Pivot-neutral only
Southern Sama Sama Balangingi'	Central Sama Yakan Pangutaran Sama	West Coast Bajau Indonesian Bajau

* These languages have marginal Philippine-type pivot-selecting constructions and productive pivot-neutral applicative constructions.

Central Sama will be used to illustrate Sama-Bajaw languages with marginal Philippine-type alternations and pivot-neutral applicatives. The voice systems of Yakan and Pangutaran Sama, are for the most part, very similar to Central Sama.

Central Sama has four transitive voice alternations, which I will call AV, PV, IV, and LV. However, stems marked for the latter two voices—the pivot-selecting applicatives—have a different distribution that those marked for the former two. IV and LV stems appear mostly in subordinate clauses, often appear in coordination with nouns and as the object of a preposition, and may be lexicalized as locative or instrumental nominalizations. At the same time, stems marked with the IV prefix *paN-* and LV suffix *-an₁* show evidence of status as verbs, including co-occurrence with aspectual, modal, and imperative morphology (James 2017: 59–67). For these reasons, James considers such constructions to be voice alternations but calls them ‘minor voices’ in contrast to the much more common and broadly distributed AV and PV constructions, as well as a separate passive construction in Central Sama.

Central Sama also has an applicative suffix *-an₂*, which marks the verb when a beneficiary, recipient, goal, stimulus, or addressee is selected as core argument. In example sentences, we see that this suffix co-occurs with AV, marked with *aN-* on the verb as in (110b), PV, which has no overt voice morphology and shows obligatory actor indexing on the verb, as in (110c), and the passive construction, marked with *-in-* without actor indexing, as in (110d).

(110) Central Sama, Applicative *-an*

- a. *Am-(b)uwan akū badju' ma abagay=ku.*
 AV-give 1SG.I shirt to friend=1SG.II
 ‘I will give a shirt to my friend.’ (AV) (James 2017: 48)
- b. *Am-(b)uwan-an akū bagay=ku badju'*
 AV-give-APPL 1SG.I friend=1SG.II shirt
 ‘I will give a shirt to my friend.’ (AV + APPL) (James 2017: 49)

- c. *B'lla-han=ta ka buwas itu.*
 PV.COOK-APPL=1SG.II 2SG.I rice this
 'I will cook you this rice.' (PV + APPL) (James 2017: 47)
- d. *B<in>uwan-an kamī kaldero.*
 <PASS>give-APPL 1PL.I pot
 'We were given a cooking pot.' (Passive + APPL) (James 2017: 49)

On the other hand, this applicative *-an*₂, which selects one of the aforementioned semantic roles as a core argument, never co-occurs with the 'minor voices', IV or LV. Conversely, it is clear that IV and LV cannot co-occur with AV, PV, or passive constructions. Instrument- or theme-selecting *paN-* and location-selecting *-an*₁ never co-occur with AV *aN-* or passive *-in-*. Also, while PV is not marked with distinct verbal morphology (i.e., it is zero-marked), it has obligatory actor-indexing on the verb, while IV and LV do not, as shown in (111) and (112) below. Thus, an analysis of *paN-* and *-an*₁ as applicatives that combine with (only) PV is not supported. Central Sama therefore represents a language in transition between a robust Philippine-type voice system to a two-way symmetrical voice system plus pivot-neutral applicatives.

(111) Central Sama, IV construction

Sīn, limangibu bay pam-(b)uwan aku.
 money 5,000 PST IV-give 1SG.III
 'Money, five thousand (pesos) was given me.' (IV) (James 2017: 63)

(112) Central Sama, LV construction

Wai al'ssu' kaldero,
 PFV cracked pot,
mbal na ta-pam-(b)'lla-han daing.
 NEG now ABIL-PUNCT-COOK-LV fish
 'The pot has cracked, it can no longer be used for boiling fish' (lit. place for boiling fish).
 (LV) (James 2017: 61)

If we compare languages like Central Sama with the more productive Philippine-type systems in the group, i.e. Sama Balangingi and Southern Sama, it becomes clear that pivot-neutral applicatives marked with *-an* in Sama-Bajaw are in complementary distribution with the voice category that selects a beneficiary pivot (BV) in Sama-Bajaw. This voice construction is also marked with *-an* in indicative mood. Thus, the development of pivot-neutral applicatives in Sama-Bajaw is directly related to the breakdown of the Philippine-type voice system. This is discussed in further detail in Chapter 6, where I argue, based on the distributions of forms of applicative markers and their functions, that many pivot-neutral applicatives of West Nusantara showing selection of beneficiary-, theme-, and/or instrument applied phrases are derived from former Philippine-type circumstantial voice (CV) alternations.

5.8.4.3 Greater Barito languages of the Barito River basin

Besides Malagasy, and the Sama-Bajaw languages, the remaining Greater Barito languages are still spoken near the Barito River basin in southeastern Borneo. In total, there are 14 such languages, but only 3 are represented in the sample, due to the general paucity of descriptive materials for Bornean languages, which is especially sparse for inland parts of the Malaysian state of Sarawak, and the Indonesian provinces of West, Central, and East Kalimantan.

In the sample, the Greater Barito languages of the Barito River area show similar typological characteristics to other Borneo languages spoken south of Sabah. All have no applicatives of any type (3 of 3). They show two-way symmetrical voice systems (3 of 3), systems of morphological alignment with distinctive marking for the non-pivot A argument (3 of 3), and predominant AVP word order in AV (3 of 3), with increasing reliance on word order for indicating grammatical relations (see Diedrich 2018: 182; Hardeland 1858: 161; Gudai 1985: 147–148). Greater Barito languages of southeastern Borneo also show reduced suffixation, though not as completely as Land Dayak and most North Sarawak languages. Ma’anyan shows a single suffix (i.e. nominalizing *-an*), and two productive circumfixes (i.e. adversative *kV-* *-an* and nominalizing *pVN-* *-an*), but 11 productive prefixes (Gudai 1985). Paku also shows a single suffix, again nominalizing *-an*, but 16 total prefixes (Diedrich 2018). In Ngaju the suffix *-an* is reported to be rare, found only on a small number of nominalizations, and a few causative verbs also bearing the prefix *ma-* (Hardeland 1858: 64–66). Including *ma-*, Ngaju has 16 prefixes.

Looking at the broader picture, the presence of Philippine-type voice in Malagasy and some Sama-Bajaw languages shows that Philippine-type voice was likely present in early stages for Greater Barito languages. In all likelihood, Philippine-type voice was only lost in Greater Barito languages of the Barito River basin after ancestral Malagasy speakers left Borneo and started their long and storied migration to Madagascar, sometime after AD 400 and most likely in the 7th century AD (Adelaar 2009). This loss probably also post-dates the departure of ancestral Sama-Bajaw speakers from the Barito River basin, sometime between AD 670–800 (R. A. Blust 2005). So while Greater Barito languages of the Barito River basin area show some similarities with Land Dayak and North Sarawak languages, and may have undertaken a similar path of grammatical restructuring resulting in the complete loss of applicatives, it is likely that these changes happened much later for Greater Barito languages like Paku, Ma’anyan, and Ngaju than for Land Dayak, and probably most North Sarawak languages as well.

5.8.5 Other languages spoken in Borneo

In addition to the indigenous languages of Borneo discussed already in this section, we also find languages spoken in Borneo which genetically belong to the Malayic and South Sulawesi genetic groups. Some of these kept what appear to be inherited systems of pivot-neutral applicatives after arrival at their present-day location in Borneo, including Brunei Malay, Salako, and Embaloh (see also discussion of Malayic languages of Sumatra in §5.4 and discussion of South Sulawesi languages in §5.9.1).

On the other hand, some Malayic languages spoken in Borneo appear to have lost inherited pivot-neutral applicatives and to have taken on some resemblance to other Borneo languages spoken south of Sabah in use of analytic constructions. These include Mualang, which is included in the sample, and Iban and the Belangin variety of Kendayan, which are not included in the sample,

but are discussed here on the basis of available descriptive material (Omar 1981; Adelaar 2006).

Mualang and Belangin, are both spoken in West Kalimantan province, in the area that is also home to the Land Dayak languages. Both show no applicatives of any type, and no suffixation, showing loss of the applicative suffixes *-i? and *-an in Proto-Malayic (Tjia 2007; Adelaar 2006).

While Mualang has lost applicative marking, it still shows alternations in clausal structure that I will refer to as “unmarked applicative analog” constructions. Examples are given in (113) and (114). For comparison, Indonesian ACs with suffixal AM marking are given in (115) and (116).

(113) Mualang, Unmarked benefactive construction

- a. *Inay N-beli kayin baju ka ia.*
mother AV-buy clothes shirt to 3SG
'Mother bought **clothes** for her.' (Prototypical AV)
- b. *Inay N-beli ia ka kain baju.*
3PL AV-buy 3SG to clothes shirt
'Mother bought **her** clothes.' (Beneficiary-selecting construction) (Tjia 2007: 174)

(114) Mualang, Unmarked locative construction

- a. *Ia N-isi' ay ka kualu.*
3SG AV-content water to cooking.pan
'She is putting **water** into the pan.' (Prototypical AV)
- b. *Ia N-isi' kualu ka ay.*
3SG AV-content cooking.pan to water.
'She is filling **the pan** with water.' (Goal-selecting construction) (Tjia 2007: 175)

(115) Indonesian, Beneficiary-selecting applicative

- a. *John Dul mem-beli buku itu untuk Mary Yem.*
J. D. AV-buy book that for M. Y.
'John bought **that book** for Mary Yem.' (BC in AV)
- b. *John Dul mem-beli-kan Mary Yem buku itu.*
J. D. AV-buy-BEN.APPL M. Y. buku itu
'John Dul bought **Mary Yem** a book.' (AC in AV) (Kaswanti Purwo 1995: 79)

(116) Indonesian, Goal-selecting applicative

- a. *Mary Yem men-(t)uang(-kan) air ke ember.*
M. Y. AV-pour(-THM.APPL) water to bucket
'Mary Yem poured **water** into the bucket.' (BC in AV)
- b. *Mary Yem men-(t)uang-i ember dengan air.*
M. Y. AV-pour-LOC.APPL bucket with water
'Mary Yem filled the **bucket** with water.' (AC in AV) (Kaswanti Purwo 1995: 80))

While many languages allow for unmarked alternations of this kind with a limited number of semantically ditransitive verbs, in Mualang, we see a quite large range of possible lexical verbs and verbal meanings used with such constructions. These examples show alternation in the realization of arguments for the semantic patient/theme and beneficiary/locative role. The existence of these alternations suggest that the morphological marking for applicative alternations in Mualang has been lost, but they remain as clausal constructions with distinctive argument structure, and these are still associated with the former meanings. I do not know of any other languages of West Nusantara that systematically exhibit this type of unmarked alternation.

In Belangin, the functions of ACs have been taken over, at least in part, by serial verb constructions (Adelaar 2006: 78). This is shown in example (117), where Belangin uses *beri* ‘give’ + *minjam* ‘borrow AV’ to mean ‘lend to (s.o.)’. For this same meaning, the closely-related variety Salako, uses *nginyapm-i?* from the same root with the applicative suffix *-i?*.

- (117) Belangin, Serial verb construction with ‘give’
Ba beri aku m-(p)injam sa-bantar ba!
 HORT give 1SG AV-borrow one-moment HORT
 ‘Why don’t you lend it to me for a while!’ (about a flute) (Adelaar 2006: 78)

Iban, another Malayic language spoken in Central Sarawak, has also lost the original applicative affixes from Proto-Malayic. According to Adelaar (2006), Iban subsequently developed a new causative/benefactive applicative suffix *-ka* from the preposition *ka*, noting “that in older Iban sources *-ka* is still written as a separate word” (79). Other than *-ka*, Iban has no other suffixes but shows nine prefixes (Omar 1981).

While Adelaar (2006) connects the loss of suffixation in Iban and Belangin to an areal pattern also found in Land Dayak languages, he does not mention Mualang, nor consider the many North Sarawak languages which also lack all suffixation. As shown in the results presented in §5.8.2-5.8.4 here, the areal pattern for reduced suffixation extends to a greater geographic area than previously recognized, and this pattern is broadly associated with complete lack of applicatives in Borneo south of Sabah. Possible explanations for this pattern are discussed in the next section.

5.8.6 Possible explanations for the lack of applicatives in Borneo south of Sabah

Based on the results of the typological survey and other available descriptive material, among the languages of Borneo, loss of suffixation looks to be a major precipitating factor for the complete lack of applicatives, which is by far the predominant pattern in languages spoken south of Sabah. These languages mostly retain a two-way symmetrical voice alternation (certain Kenyah varieties reported to lack PV are exceptions). Reduction of verbal morphology, and a shift to reliance on word order for signalling grammatical relations, as well as a shift towards periphrastic or analytic signalling of PV are also observed, as is use of analytic structures with benefactive, locative, or causative meanings in some languages. This shift looks to be related to the loss of suffixation, including restructuring following the loss of morphological marking of PV with suffixal *-en*.

In mainland Southeast Asia and northern Peninsular Malaysia (as well as Singapore Bazaar Malay), loss of suffixation was triggered by language contact with non-Austronesian languages

(Tadmor 1995; Thurgood 1999; Aye 2005). On the other hand, it is unknown why loss of suffixation occurred in so many Bornean languages, cutting across several genetic affiliations. Little is known of possible language contact for speakers of Malayo-Polynesian languages with non-Austronesian languages upon arrival in Borneo. The present-day pattern for lack of suffixation might be observed however, if some such contact was centered on the southwest part of Borneo, where Land Dayak languages are presently spoken, and these changes only later spread to North Sarawak, Melanau-Kajang and, finally, Greater Barito languages.

Contact-induced spread leading to loss of suffixation may have been chained, and might have had multiple centers, with possible differences in the precise path of historical change. In the case of Sa'ban (Dayic, North Sarawak), for example, it is argued that contact with Modang (Kayanic) set off the phonological changes that resulted in its highly reduced system of affixation (R. A. Blust 1999), in contrast to the relatively rich systems of Kelabit and Lun Bawang (also Dayic). Some Malayic languages whose speakers settled in southwestern Borneo also seem to have undergone loss of suffixes known to be present in Proto-Malayic, e.g. Belangin, Mualang, and Iban, but others retain some inherited suffixes, e.g. Salako (Adelaar 2005b). It is unclear whether the nature or extent of the language contact can explain these differences, and this further complicates a proposed timeline for the spread of these changes. Nonetheless, it is clear that reduced suffixation is an areal pattern for Borneo south of Sabah, and that this pattern was influenced by complex histories of migration, contact, and genetic inheritance for Bornean languages and peoples. Still, the possibility that the decline of suffixation and subsequent changes in the verbal system were initially set off by contact with non-Austronesian speakers is an enticing prospect, albeit a speculative one at present.

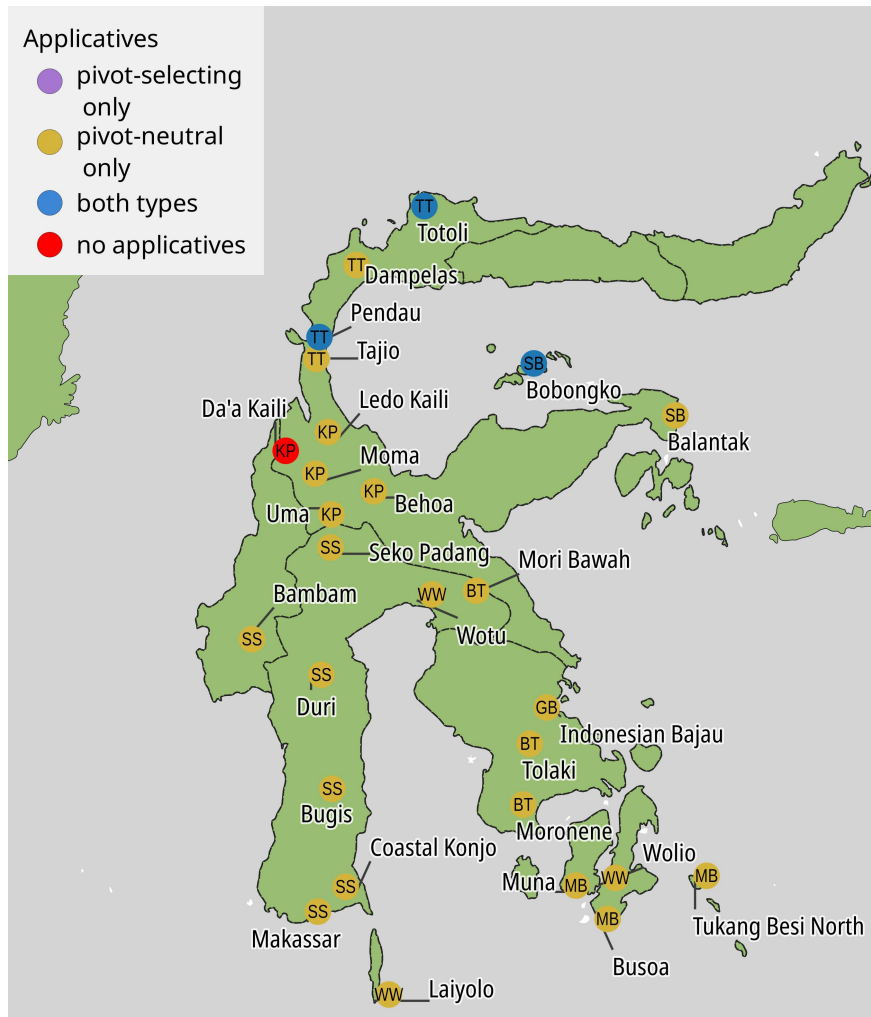
5.9 Sulawesi

This section presents and discusses results of the typological survey for languages of Sulawesi. Sulawesi is the second largest island in West Nusantara after Borneo, and also shows great linguistic diversity, with over 100 languages spoken on the island. Sulawesi is traditionally divided into microgroups, which I have generally adopted as genetic groupings for the purpose of the survey, though I have excluded three traditional microgroups that are thought to subgroup with Philippine languages (i.e. Gorontalo-Mongondow, Minahasan, and Sanggiric, see R. A. Blust 1991). In the remainder of this section, I will consider separately languages of the South Sulawesi microgroup in §5.9.1 and languages classified as Tomini-Tolitoli, Kaili-Pamona, Saluan-Banggai, Bungku-Tolaki, Muna-Buton, and Wotu-Wolio in §5.9.2. This second set of genetic groupings have more varied typological profiles and have been proposed to belong to a genetic supergroup, Celebic (Mead 2003), that excludes the South Sulawesi languages. An overview map showing languages spoken in Sulawesi in the sample is presented in Figure 5.10. In addition to the languages discussed in this section, Indonesian Bajau is spoken in Sulawesi (see §5.8.4.2).

5.9.1 South Sulawesi languages

The South Sulawesi genetic group includes 30 total languages, which are primarily spoken on the west and southwestern side of Sulawesi. Two other languages of this group are spoken in Borneo (i.e. Embaloh, Taman). In available documentation, these two appear to be quite similar

Figure 5.10: Overview map of languages of Sulawesi



Abbreviations: [BT] Bungku-Tolaki, [GB] Greater Barito, [KP] Kaili-Pamona, [MB] Muna-Buton, [SB] Saluan-Banggai, [SS] South Sulawesi, [TT] Tomini-Tolitoli, [WW] Wotu-Wolio. Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

to other South Sulawesi languages (Adelaar 1994). Seven South Sulawesi languages are included in the sample, as shown in Table 5.18. All of the South Sulawesi languages in the sample show pivot-neutral applicatives (7 of 7). They also share common typological features, especially in terms of voice system and morphological alignment.

Table 5.18: Typological survey results for South Sulawesi languages

Gen. Grp.	Name	Applicatives?		Voice	Morph. align.	Word Order
		pivot-selecting	pivot-neutral			
S. Sulawesi	Embaloh	N	Y	marg. two-way	ergative	VAP
S. Sulawesi	Bugis	N	Y	marg. two-way	ergative	AVP/VPA
S. Sulawesi	Coastal Konjo	N	Y	marg. two-way	ergative	VAP
S. Sulawesi	Makassar	N	Y	marg. two-way	ergative	VPA
S. Sulawesi	Bambam	N	Y	marg. two-way	ergative	AVP
S. Sulawesi	Seko Padang	N	Y	marg. two-way	ergative	undet.
S. Sulawesi	Duri	N	Y	two-way	ergative	AVP

Almost all South Sulawesi languages in the sample show what I have called a marginal two-way symmetrical voice systems (6 of 7). These languages show an alternation between A-oriented and P-oriented constructions, with the A-oriented construction showing reduced semantic transitivity (B. Friberg 1991: 105-112) or very limited distribution (Campbell 1989: 66). These A-oriented constructions are discussed in greater detail below. In the sample, only Duri is considered a two-way symmetrical voice language (1 of 7). Valkama (1993: 69, 79–89) shows that the A-oriented construction in Duri does not have lower transitivity than the P-oriented construction, and though it is less frequent in usage than the P-oriented construction, it does not appear to be rare.

All seven South Sulawesi languages of the sample are coded as showing ergative patterns of morphological alignment. This is because S and P in P-oriented transitive clauses pattern together in morphological encoding; both are typically indexed on the verb by means of a pronominal enclitic, while A in P-oriented clauses is typically indexed with a proclitic, as shown in (118a–b) from Makassar. In Seko Padang, two sets of proclitics are used instead of one enclitic set and one proclitic set (Payne & Laskowske 1997: 426–429). However, there are some complications for indexing of arguments in these languages that make them difficult to classify for type of morphological alignment (Pattern 2).¹⁶

(118) Makassar, Voice alternations

- a. *Tinroi i Ali*
tinro=i i Ali
 sleep=3 PN A.
 ‘Ali is sleeping.’ (Intransitive)

(Jukes 2020: 246)

¹⁶In most South Sulawesi languages, A of A-oriented constructions is indexed as an enclitic form on the verb and P is not indexed. If A-oriented clause are syntactically intransitive, this is consistent with ergative alignment patterns, though see below for some reasons that the transitivity value of these clauses is not straightforward. Furthermore, in some languages, the core argument of an intransitive verb (S) may show different patterns of indexing in certain irrealis or non-indicative clauses, complicating patterns of morphological alignment for core arguments (see D. Laskowske 2016: 43–47; K. Laskowske 1994, B. Friberg 1991: 120–121).

- b. *Kukanrei untia*
ku=kanre=i unti=a
 1=eat=3 banana=def
 ‘I eat the bananas.’ (Transitive, P-oriented) (Jukes 2020: 135)
- c. *Angnganrea’ unti*
aN(N)-kanre=a’ unti
 BIV-eat=1 banana
 ‘I eat bananas.’ (Semi-transitive, A-oriented) (Jukes 2020: 135)

South Sulawesi languages show some diversity of word order patterns. In three languages of the sample, verb-initial order is predominant, while in two languages (Duri and Bambam), pivot-initial order is predominant (AVP in A-oriented constructions). In Bugis, word order is described as split between AVP and VPA in the A-oriented construction. Word order is not discussed in available source material on Seko Padang.

As mentioned above, most South Sulawesi languages show an A-oriented construction with reduced semantic transitivity. An example is given in (118c) above from Makassar. Determining the syntactic transitivity of these constructions is not straightforward. While the P argument in such constructions typically is strictly indefinite, often it is overtly expressed, and it may even be obligatory to mention P overtly, as in Makassar and Bugis (Jukes 2020: 250–253; D. Laskowske 2016: 26). Furthermore, the P argument in A-oriented constructions may sometimes occur with modifiers not typically observed for incorporated nouns.¹⁷ These constructions have been labelled in various ways by authors, but it is not clear whether they should be considered intransitive (e.g. “antipassive” as used by B. Friberg 1991), transitive (e.g. “actor focus” as used by Campbell 1989) or somewhere in the middle (e.g. “semi-transitive” as used by Jukes 2020).

It is interesting, then, that in languages of this type, pivot-neutral applicative suffixes typically may co-occur with both the more prototypically transitive P-oriented construction and the A-oriented construction that shows lower semantic transitivity. Some examples of ACs in A-oriented constructions are given below from Coastal Konjo in (119) and Embaloh in (120).

(119) Coastal Konjo, Locative applicative in A-oriented clause

- a. *Ammalu’a (tappere).*
ang-halu’-a tappere.
 AV-roll.up-1SG.ABS mat
 ‘I roll up a mat.’ (BC)
- b. *Ammaliikia palungang.*
ang-halu’-i-a palungang
 AV-roll.up-LOC.APPL-1SG.ABS pillow
 ‘I roll up a pillow (in something).’ (AC) (B. Friberg 1991: 115)

¹⁷For example, the following A-oriented sentence in Bugis contains a P argument that is modified by a relative clause: *M-elli=ka’ [waju ia i-balu’-e’ ku Takkalala]*. ‘I bought [a shirt that was sold in Takkalala]’ (D. Laskowske 2016: 7). The NP meaning ‘a shirt that was sold in Takkalala’ does not appear to be consistent with an incorporated noun.

(120) Embaloh, Locative applicative in A-oriented clause

Urtka?-ak *i-asan* *marampas tahu?* *tau,* *ingka,*
 possibly-1SG.ABS ERG-name AV.snatch fiancé person said

i-asan *de?i* *man-jolo-i* *laki* *tau.*
 ERG-name aforementioned maN-take.away.place.next.to-LOC.APPL husband person
 ‘They might think that I’m after other people’s fiancés, she said, they might think that I
 tried to take away other women’s husbands’. (Adelaar 1995: 391)

In the Embaloh example above, the locative applicative suffix *-i* occurs on the verb *man-jolo-i* ‘to take away the place next to’ (cf. *jolo* ‘preceding, ahead of’). This verb also bears the *maN-* prefix associated with low referentiality of P or oblique-marking of the semantic undergoer.

Lastly, I will note that though these languages may show ergative morphological alignment, they are not necessarily syntactically ergative. While the ergative pattern for morphological alignment stems from the distribution of clitic pronominal forms, these forms do not co-vary neatly with syntactic behavior. For Makasar, Jukes (2020) writes: “the fact that in the majority of cases =ABS corresponds to S and P, while ERG= corresponds to A, suggests that clitics cross-reference core arguments, and that they do this according to an ergative-absolutive pattern... However, when examining the behaviour of certain Makasar clauses it becomes clear (a) that it is not always easy to correlate =ABS or ERG= clitics with particular grammatical relations, and (b) that some arguments which are arguably core are not always cross-referenced” (331). Other authors make similar observations for other South Sulawesi languages. Campbell (1989) writes: “Since, however, the ergative characteristics are limited to the pronominal system, PUS [Bambam] is morphologically ergative in a very limited sense. Like many languages which have ergative morphology, PUS does not have ergative syntax” (56). Of Coastal Konjo, Barbara Friberg (1991) writes “Only here [in clitic marking] is there an ergative system functioning; syntactically Konjo functions as an accusative system” (106–107). Thus, South Sulawesi languages may be morphologically ergative, but they do not necessarily provide evidence to support a correlation between non-accusative alignment and the presence of applicatives as found elsewhere (Peterson 2007).

5.9.2 Celebic languages

The remaining languages of Sulawesi in the sample belong to six microgroups, all treated as genetic groupings for the purpose of the study: Tomini-Tolitoli, Kaili-Pamona, Saluan-Banggai, Bungku-Tolaki, Muna-Buton, and Wotu-Wolio. These microgroups have been proposed to belong to a supergroup, called Celebic, which includes 64 total languages, spoken across central, eastern, and southeastern Sulawesi. Of these, 20 languages are included in the sample as shown in Table 5.19. Note that the genetic classification used is disputed for Totoli (representing the Tolitoli subgroup that may not be non-Celebic, rather than subgroup with Tomini) and Behoa (representing the Bada-Behoa-Napa subgroup, which may belong with the Seko languages of South Sulawesi rather than Kaili-Pamona).

Almost all of the languages of the sample in these microgroups show pivot-neutral applicatives (29 of 30). The only exception to this is Da’a Kaili, which is discussed in greater detail in §5.9.3 below. Geographically, these languages are broadly distributed in Sulawesi. The presence

Table 5.19: Typological survey results for other Sulawesi languages

Gen. Grp.	Name	Applicatives?		Voice	Morph. align.	Word Order
		pivot-selecting	pivot-neutral			
Tomini-Tolitoli	Totoli*	Y	Y	Philippine-type	mixed-NPIV.A	AVP/VPA
Tomini-Tolitoli	Dampelas	N	Y	two-way	mixed-NPIV.A	AVP
Tomini-Tolitoli	Pendau	N	Y	two-way	mixed-NPIV.A	AVP
Tomini-Tolitoli	Tajio	N	Y	two-way	mixed-NPIV.A	AVP/VPA
Kaili-Pamona	Behoa*	N	Y	two-way	pivot-nonpivot	AVP
Kaili-Pamona	Ledo-Kaili	N	Y	two-way	mixed-NPIV.A	AVP/VPA
Kaili-Pamona	Moma	N	Y	two-way	mixed-NPIV.A	AVP
Kaili-Pamona	Uma	N	Y	two-way	ergative	AVP
Kaili-Pamona	Da'a Kaili	N	N	two-way	mixed-NPIV.A	AVP
Saluan-Banggai	Balantak	Y	Y	Philippine-type	pivot-nonpivot	AVP
Saluan-Banggai	Bobongko	Y	Y	marg. Philippine-type	mixed-NPIV.A	undet.
Bungku-Tolaki	Mori Bawah	N	Y	asymmetrical	ergative	VAP/VPA
Bungku-Tolaki	Moronene	N	Y	asymmetrical	mixed:other	VAP/VPA
Bungku-Tolaki	Tolaki	N	Y	asymmetrical	split-S	undet.
Muna-Buton	Busoa	N	Y	asymmetrical	accusative	VAP/VPA
Muna-Buton	Muna	N	Y	asymmetrical	accusative	VAP/VPA
Muna-Buton	Tukang-Besi (N.)	N	Y	two-way	mixed:other	VPA
Wotu-Wolio	Laiyolo	N	Y	two-way	accusative	AVP/VPA
Wotu-Wolio	Wolio	N	Y	marg. two-way	accusative	VAP
Wotu-Wolio	Wotu	N	Y	two-way	mixed:other	AVP

* For these languages, the genetic classification listed is disputed.

of pivot-neutral applicatives is also observed to cut across other typological features of language in these microgroups, especially voice, morphological alignment and word order.

The languages of the sample in these affiliations are diverse in voice system. Most are two-way symmetrical systems (11 of 20), but asymmetrical voice systems (5 of 20)¹⁸ and Philippine-type systems (3 of 20) are also found. There is also one marginal two-way system (Wolio, which retains an alternation between A-oriented and P-oriented clauses only in certain verbal “participial” forms, see Anceaux 1952) and one marginal Philippine-type language (Bobongko, which appears to show limited usage of LV and IV constructions, see discussion in §5.9.3 below).

Languages of these six genetic groups have diverse systems of morphological alignment, with all seven coding categories for this feature represented, and pivot-neutral applicatives attested across all seven.

These languages also show diversity in word order patterns, with some languages preferring verb-initial orders (6 of 20), some favoring verb-medial orders (AVP in A-oriented constructions,

¹⁸Languages coded as asymmetrical for voice in Bungku-Tolaki and Muna-Buton, do generally show more than one (apparently) transitive construction, e.g., “ae-class marking” vs. “a-class marking” in Muna (van den Berg 1995), and zero-marked verbs vs. *moN*-marked verbs in Moronene (Andersen & Anderson 2005). These languages are not coded as two-way symmetrical voice systems because there is little evidence that A in one such construction is syntactically privileged, while P is syntactically privileged in the other. To the contrary, these languages generally show syntactic behavior indicating that S and A constitute a privileged syntactic relation to the exclusion of P (accusative alignment). Alternations between transitive constructions in these languages are often distinguished by definiteness constraints on P and semantic transitivity, in which sense they are similar to clausal alternations described above for South Sulawesi languages (§5.9.1).

8 of 20), and some reported to have split word order with flexibility for the pivot (AVP/VPA word order in AV, 4 of 20). For Bobongko and Tolaki, word order patterns could not be determined on the basis of available source material.

An unexpected finding in these data, is the fact that a few Philippine-type languages of Sulawesi also have a system of pivot-neutral applicatives which co-occur across different voice constructions, though not necessarily all such categories. This occurs in Balantak and Bobongko, two Saluan-Banggai languages of East Sulawesi province, and Totoli, a Tolitoli language of Central Sulawesi province. All three of these languages are located roughly in the northern third or so of Sulawesi, and are discussed in greater detail in §5.9.4 below, along with Pendau, which shows some similarities to these three its applicative system.

5.9.3 Outliers in Sulawesi

The only language of Sulawesi included in the sample that does not show applicatives is Da'a, a Kaili-Pamona language of Central Sulawesi. Da'a lacks the *-i* suffix (Martens 1988b: 193), which marks locative ACs in other Kaili-Pamona languages (e.g., Moma, Kaili-Ledo, and Uma). Additionally, while Barr (1988a) does report the use of the suffix *-(C)aka* as a causative marker with a limited number of stems, he does not mention or include data showing use of *-(C)aka* as an applicative marker.¹⁹ Martens (1990: 193) indicates that Proto Pamona-Kaili (PPK) had a suffix **-aka*, “a transitivizer or causativizer or intensivizer”, with Da'a *nan-taji-aka* ‘to throw away’ given as one reflex of the suffix. Martens (1990: 190) also tentatively reconstructs PPK **-uli'-ka* for ‘to tell, to inform’, which appears to be an addressee-selecting applicative verb marked with *-ka*, and he includes Da'a *nang-uli-ka* with the same meaning as a form supporting this. Thus it is possible that the applicative use of *-(C)aka* and/or *-ka* was lost or diminished in productivity in Da'a, but with only limited descriptive and textual material available it is hard to state definitively that this is the case.

5.9.4 The transition from Philippine-type voice to pivot-neutral applicatives in Sulawesi languages

Three languages of Sulawesi in the sample show Philippine-type voice alternations in addition to pivot-neutral applicatives: Balantak, Totoli, and Bobongko. Like the Sama-Bajaw languages discussed in §5.8.4.2 above, these languages generally show reduction in the four-way voice system, and can be considered to show a transition between Philippine-type voice (pivot-selecting applicatives) and pivot-neutral applicatives. In addition, Pendau, a Tomini language of Central Sulawesi, is coded as a two-way symmetrical voice system with pivot-neutral applicatives in the survey. However, it shows irregularities in marking of voice and applicatives that suggest it has taken a path of development similar to Balantak, Totoli, and Bobongko. For this reason, it is also discussed in this section. Together, these four cases indicate that the transition between Philippine-type voice and pivot-neutral applicatives takes place via reduced productivity of the

¹⁹For comparison, the suffixes *-ka* and *-(C)aka* are found in both Moma and Kaili-Ledo with applicative functions, though *-(C)aka* is noted to be rare in Moma (Adriani & Esser 1939: 31), and to occur with some but not all verbs in Kaili-Ledo (D. Evans 2003: 502). In Uma, *-(C)aka* is only found with one verb, *sapu* ‘to deny (s.t.)’ cf. *sapuaka* ‘to deny’ and a handful of fossilized verb stems (Martens 1988b: 235).

CV and/or LV voice categories, reinterpretation of certain affixal marking for voice (e.g. *pog-/paN-*, *-an*, and *-i*), and functional replacement of CV and/or LV with pivot-neutral ACs.

5.9.4.1 Balantak: Three-way voice system with pivot-neutral applicatives

Balantak is the most straightforward of these cases. Balantak has three basic transitive constructions, AV, PV, and LV, but no circumstantial voice (CV) construction.²⁰ It also has three applicative suffixes, which each may co-occur together with any of the three basic voices. The general applicative suffix *-kon* is shown in example (121) with the verb *bisara* ‘speak’. In clauses marked with *-kon*, one of a wide range of peripheral semantic roles is selected as a core argument. The applied phrase may be a semantic recipient, beneficiary, instrument, purpose, comitative, or content, among others (van den Berg & Busenitz 2012: 97–99). As shown in (121a), without applicative marking, *bisara* is intransitive. With applicative *-kon*, the verb selects a content applied phrase, and this may co-occur with all three transitive voice categories as shown in (121b–121d). Thus, Balantak has a productive Philippine-type system with a three-way distinction for voice, and shows pivot-neutral applicatives that are outside of the symmetrical voice paradigm. In Balantak, a number of functions of CV in other Philippine-type languages (selection of a beneficiary or instrument as a core argument) are filled instead by pivot-neutral applicatives.

(121) Balantak, Applicative *-kon*

- a. *Kai ba-bisara sang-ilio.*
1PL.EXCL INTR-speak one-day
‘We talked for a whole day.’ (Intransitive)
 - b. *Kai nim-bisara-kon parakala i-ya’a*
1PL.EXCL AV.RLS-speak-APPL issue DEIC-DEM
‘We talked about the issue.’ (AV + APPL)
 - c. *Parakala men bisara-kon-on-ta...*
issue REL speak-APPL-PV.IRR-1PL
‘The issue that we will talk about...’ (PV + APPL)
 - d. *Na laigan-mo ka-ni’i a bo pim-bisara-kon-an-ta parakala.*
LOC house-PFV DEIC-DEM ART for GER-speak-APPL-LV-1PL issue
‘Let this be the house where we will talk about the issue.’ (LV + APPL)
- (van den Berg & Busenitz 2012: 31)

5.9.4.2 Totoli: Three-way voice system with pivot-neutral applicatives

The voice and applicative systems of Totoli show some similarities with that of Balantak. Himmelmann & Riesberg (2013) analyze Totoli as a Philippine-type language with three basic transi-

²⁰In addition to a transitive PV construction in which A is overtly expressed, Balantak also has an agentless P-oriented construction, in which the actor argument is suppressed. The agentless construction is only found in realis mode and is more akin to the English passive. The agentless construction and “agented” PV are distinguished by verbal morphology and the fact that in the latter, a pronominal form indexing P precedes the realis form of the verb (see van den Berg & Busenitz 2012: 29).

tive constructions, AV, PV, and LV and two additional pivot-neutral applicatives. Like Balantak, Totoli does not show CV as a basic voice category.

Applicative 1 in Totoli selects a beneficiary, instrument, or theme as a core argument and Applicative 2 selects a goal or recipient as a core argument (while LV selects a static location role as the syntactic pivot).²¹ In Himmelmann and Riesberg’s analysis, Applicative 1 and 2 may only occur with AV and PV, never with LV. Example (122) shows the Applicative 1 construction with the verb *taip* ‘to peel’ and a beneficiary applied phrase.

(122) Totoli, Beneficiary-selecting construction

- a. *I Rinto man-(t)aip taipang.*
 HON R. AV-peel mango
 ‘Rinto peels mangoes.’ (AV nonrealis, BC)
- b. *I Rinto man-(t)aip-an aku taipang.*
 HON R. AV-peel-APPL1 1SG mango
 ‘Rinto peels mangoes for me.’ (AV nonrealis + APPL1)
- c. *Aku ko-doong pan-(t)aip-an Rinto taipang.*
 1SG POT-want SF-peel-APPL1 R. mango
 ‘For me, Rinto will peel a mango.’ (PV nonrealis + APPL1)
- d. *Aku ni-pan-(t)aip(-an) Rinto taipang.*
 1SG RLS-SF-peel(-APPL1) R. mango
 ‘For me, Rinto peeled a mango.’ (PV realis + APPL1)

(Himmelmann & Riesberg 2013: 400–401, 405)

Notably, there are some irregularities in the paradigm for the morphological marking of voice and applicatives in Totoli. The suffixal form *-an*, which is found with Applicative 1 constructions in AV and PV nonrealis mode, is missing in PV realis mode in varieties of the language spoken in Totoli City, but included in the Northern Totoli dialect. This is represented with parentheses around the suffix which I have added in example (122d). Furthermore, the analysis of voice and applicatives in Totoli relies on a number of nuanced interpretations of the functions of the affixal forms *po(g)/-poN-* (glossed as a stem-former above), *-an* (marker of Applicative 1, and Applicative 2 in PV realis), and *-i* (marker of Applicative 2 in nonrealis) (see Himmelmann & Riesberg 2013: 411).

While the detailed arguments for and against various parts of Himmelmann and Riesberg’s analysis are beyond the scope of the discussion here, it appears that Totoli has both a Philippine-type voice system with a productive LV alternation and pivot-neutral ACs that co-occur only with AV and PV. Irregularities in the paradigm of morphological marking for voice and applicatives in Totoli suggest that reinterpretation of *po(g)/-poN-*, *-an*, and *-i* have played a role in the development of the pivot-neutral applicatives. In many other Malayo-Polynesian languages with Philippine-type voice, these affixes are used to mark voice categories, including CV categories.

²¹Applicative 1 in Totoli is marked by *-an* except in PV realis, where the suffix is absent (see Himmelmann & Riesberg 2013: 410). This most likely reflects PMP morphology for CV, which is marked by **-an* in imperative/negative mood, and **Si-/Sa-* in indicative mood. Applicative 2 is marked by *-i* except in PV realis, where it is marked by *-an*. This most likely reflects PMP morphology for LV, which is marked by **-i* in imperative/negative mood and **-an* in indicative mood. See §6.2.2 and §6.2.3 for more on development of pivot-neutral AMs from LV and CV morphology.

Compare, for example, the morphological marking for the ‘referent voice’ (RV) construction in Tatana in which an instrument, beneficiary, goal or recipient is selected as the pivot (see Table 5.13 in §5.8.1 above): *poN-* marks IV, *-an* marks RV in indicative forms, and *-i* marks RV in imperative forms (Dillon 1994). Thus, it appears that changes in the interpretation of some voice affixes in Totoli has given rise to the development of pivot-neutral applicatives, and this is responsible for irregularities in the paradigm for voice and applicatives, at least in part. In Totoli, some such irregularities are now undergoing regularization in the northern dialect area, where the language is used more frequently and by a larger share of the local population, as compared to Totoli City.

5.9.4.3 Bobongko: Marginal four-way voice system with pivot-neutral applicatives

In Bobongko, we see a different type of transitional system. The Bobongko voice system includes two major transitive voice alternations (Mead 2001). Verbs in AV select the A argument as syntactically privileged, and are normally marked by *moN-* in realis mode, *noN-* in irrealis mode, and *poN-* in imperatives.²² Verbs in PV select the P argument as syntactically privileged, and are marked by *o-* in irrealis mode, *-in-* in realis mode, and no overt morphology in imperatives.

Bobongko also makes use of two suffixes that co-occur with AV and PV morphology, which I will analyze as applicative morphemes (AMs): *-i* (and its variant *-an*), and *-akon*.²³ The suffix *-i/-an* marks the verb when a clausal argument is a semantic goal or location; in such constructions, the applied phrase appears to have the syntactic properties and coding typical of P in AV and PV base constructions. In AV, it normally appears as an unmarked NP immediately following the verb, as in (123a) below. In PV, the applied phrase is the apparent pivot or syntactically privileged argument as in (123b) below. Note that when the goal- or location-selecting construction appears in PV and realis aspect, the suffix *-an* always appears on the verb instead of *-i*.

(123) Bobongko, Locative-selecting applicative

- a. *Jadi no-sangalu ka’a, no-tugal lampi’*
 so RLS-be.friends this, RLS.AV-plant **banana**.
 ‘So they being friends, they planted **bananas**.’ (BC) (Mead 2001: 87)
- b. *Ka’a-taa’ inaut anu t<in>ugal-an-ku bele-nu binte’*
this-that garden REL <RLS.PV>plant-LOC.APPL-1SG.GEN with-LNK corn
 ‘This is **the field** that I planted with corn.’ (AC) (Mead 2001: 78)

The suffix *-akon* marks the verb when a clausal argument is a semantic beneficiary or recipient. However, unlike constructions marked with *-i/-an*, in constructions marked with *-akon*, the

²²What Mead (2001) describes as an aspectual distinction between realized and unrealized aspect in Bobongko appears to correspond to realis and irrealis mode (or nonrealis mode) in Balantak, Totoli, Pendau, and other Philippine-type languages. In all four languages discussed in this section, this distinction encompasses differences in both tense/aspect and evidentiality/factuality. In Bobongko examples, the realis or realized aspect forms are used to describe past completed actions and states and the irrealis or unrealized aspect forms are used to describe ongoing and future actions and states, and in certain requestive and prohibitive constructions. For the sake of consistency, I have used realis [RLS] and irrealis [IRR] to distinguish these forms in discussion and glossed examples for Bobongko.

²³Mead 2001: 77 does not refer to *-i* or its variant *-an* as AMs, primarily because its compatibility with verbal roots is highly lexicalized. Nonetheless, the *-i/-an* marked construction in Bobongko meets the definition of applicative used in this study.

beneficiary applied phrase is not realized with the syntactic properties and coding observed for P in AV and PV base constructions. The beneficiary is instead marked with *bele* ‘with, for’, and does not appear to be a core argument. An example is shown in (124) below.

(124) Bobongko, Beneficiary-selecting applicative

- a. *Gunsing b<in>oa.*
 key <RLS.PV>carry
 ‘(Someone) has taken the key.’ (BC) (Mead 2001: 76)
- b. *B<in>a-kon-nyo bele-nu sangalu-nyo, anu to-pomangan.*
 <RLS.PV>bring-BEN.APPL-3S.GEN FOR.LNK companion-3SG.GEN REL person-chew.betel
 ‘She brought (the lime) for his companions, who were betel chewers.’ (AC, oblique beneficiary) (Mead 2001: 89)

In addition to AV and PV, Bobongko makes use of constructions that Mead calls “special inverse” constructions. These appear to be of low textual frequency, though available data is limited. In one such construction, the instrument is syntactically privileged, and the verb is marked with the prefix *poN-* in irrealis mode and *pinoN-* in realis mode. This is shown in (125). In another, the location of the event is syntactically privileged, and the verb is marked with the circumfix *poN-* *-an* in irrealis mode, and *pinoN-* *-an* in realis mode.²⁴ This is shown in (126).

(125) Bobongko, Instrumental Voice

- a. *Sapi’ taio’ kana’ o-kolot*
 cow that must IRR.PV-slaughter
 ‘The cow must be slaughtered.’ (PV) (Mead 2001: 76)
- b. *Ka’a-mo kapara anu ku-pong-kolot sapi’ ka’a.*
 this-PFV machete REL 1SG.NPIV.IRR-IRR.IV-slaughter cow this
 ‘Here is the machete with which I will slaughter the cow.’ (IV) (Mead 2001: 81)

(126) Bobongko, Locative Voice

- a. *N-una’-ku-mo kacamata.*
 RLS.PV-stow-1SG.GEN-PFV eyeglasses.
 ‘I’ve put the eyeglasses away.’ (PV)
- b. *Lamari p<in>ong-una’-an-ku*
 cabinet <RLS.PV>LV-stow-LV.RLS-1SG.GEN
 ‘The cabinet is where I put them. (LV) (Mead 2001: 82)

As described thus far, it appears that Bobongko has a marginal Philippine-type voice system, with two major transitive voice alternations (AV and PV), and two minor transitive voice alternations (IV and LV), in addition to pivot-neutral ACs.

²⁴There is no available data on imperatives in these two constructions in Bobongko.

On top of this, we also observe what appears to be co-occurrence of instrumental-selecting (IV) *poN-* with the locative-selecting applicative suffix *-i/-an*, shown in (127) below.

(127) Bobongko, Instrumental Voice + Locative-selecting applicative

- a. *Oko ku-bobal-i!*
 2SG 1S.NPIV.IRR-PV.hit-LOC.APPL.IRR
 ‘I will hit you!’ (PV + Locative-selecting AC) (Mead 2001: 78)
- b. *Ka’a kau’ anu pom-bobal-i-nyo ara*
 This wood REL IV-hit-LOC.APPL.IRR-3SG.GEN 3PL
 ‘This is the wood he will hit them with.’ (IV + Locative-selecting AC)
 (Mead 2001: 81)

In the BC in (127a), the verb is marked with *ku-*, the non-pivot first-person actor prefix in irrealis mode. This prefix replaces the prefix *o-* that is seen with other person categories for PV in irrealis mode. The verb also bears the locative-selecting applicative suffix *-i*. The clause is monotransitive, and the goal, *oko* ‘you’, is realized as the pivot core argument (P). In the AC in (127b), the verb is marked with *poN-* and *-i*. The clause is ditransitive. It takes two nonactor core arguments, an instrument and a goal, which are both realized as unmarked NPs. The instrument, *ka’a kau* ‘this wood’, is the pivot (R) and may be relativized, while the goal *ara* ‘them’, is a non-pivot core argument (T). It appears immediately following the verb, like P of AV, another non-pivot nonactor core argument, normally does. The semantic and syntactic properties of this type of clause are consistent with co-occurrence of IV and the locative-selecting pivot-neutral applicative. Like the IV base construction, the instrument is always the pivot in this type of clause. Like the locative-selecting applicative in AV and PV, the location or goal is a core argument, but not necessarily the pivot, as the semantic role selected as pivot depends on the voice alternation selected.

In contrast, I do not analyze the construction marked with *poN- -an/pinoN- -an*, shown in (126) above, as co-occurrence of instrumental voice (IV) and the locative-selecting applicative marked with *-i/-an*. This construction does not share semantic or syntactic properties that are characteristic of the IV base construction; an instrument is not seen to be part of the semantic meaning of the clause and is not the role selected as pivot. Furthermore, the locative-selecting applicative is marked with *-an* only in realis mode for PV, and *-i* everywhere else. But the construction in (126) always shows a suffixal *-an* component, both in realis and irrealis mode, according to Mead (2001: 83).

Finally, in a single example, we observe a construction marked with *pinoN- -akon-an* in realis mode, shown in (128).

(128) Bobongko, Beneficiary as pivot

- Aliali-um anu p<in>om-be’-akon-an-ku kapara-um*
 younger.sibling-2SG.GEN REL <RLS>LV-give-BEN.APPL-LV-1SG.GEN machete-2SG.GEN
 ‘It was your younger sibling to whom I gave your machete.’ Mead 2001: 83

This construction is also ditransitive, taking two nonactor core arguments coded as unmarked NPs, a recipient, and a theme. The beneficiary/recipient, *aliali-um* ‘your younger sibling’, is the

pivot (R), while the theme, *kapara-um* ‘your machete’, is a non-pivot core argument (T) appearing immediately following the verb. This type of clause could either be analyzed as a minor nonactor voice, that is, a BV construction, marked with *pinoN-* *-akon*. Alternately, it could be analyzed as co-occurrence of LV, (marked with *pinoN-* *-an*, and the beneficiary-selecting applicative, marked with *-akon*.

Evidence for analysis as BV includes the fact that a location is not the pivot in this clause, and therefore this construction does not share a key property of LV base constructions. Evidence for analysis as LV plus pivot-neutral applicative *-akon* includes the fact that this clause is ditransitive, while other instances of *-akon* in co-occurrence with AV and PV are not. In this second possible analysis, we might consider the function of *-akon* as adding a semantic beneficiary or recipient to the constructional meaning, and the function of *pinoN-* *-an* to be selecting the most location- or goal-like argument as the pivot—that being the semantic recipient rather than the semantic theme. Given the limited data, it is not possible to make a definitive assessment, though I lean towards analyzing this construction as LV plus a co-occurring pivot-neutral beneficiary-selecting applicative. Interestingly, the *-akon* suffix appears closer to the verb stem than does the suffixal portion of the circumfix *pinoN-* *-an*. This is consistent with analysis in which a beneficiary or recipient is added first, perhaps at a semantic level of verbal structure, and is subsequently selected to be realized as the syntactic pivot.

The paradigm of morphological marking for symmetrical voice alternations and pivot-neutral applicatives in Bobongko, as analyzed here, is shown in Table 5.20. This system has four Philippine-type voice alternations and two pivot-neutral ACs, plus an aspectual distinction marked with *-in-* (or its variant *n-*) for realis mode. The two pivot-neutral ACs co-occur with both major voice categories (AV and PV), but each is found with only one of the two minor voice alternations, that is, either IV or LV, but not both.

Table 5.20: Partial paradigm for voice and pivot-neutral applicatives in Bobongko

Voice	Mode	Base	Loc. Appl.	Ben. Appl.
AV	irrealis	<i>moN-</i>	<i>moN- + -i</i>	<i>moN- + -akon</i>
AV	realis	<i>noN-</i>	<i>noN- + -i</i>	(no data)
PV	irrealis	<i>o-</i>	<i>o- + -i</i>	<i>o- + -akon</i>
PV	realis	<i>-in-</i>	<i>-in- + -an</i>	<i>-in- + -akon</i>
IV	irrealis	<i>poN-</i>	<i>poN- + -i</i>	—
IV	realis	<i>pinoN-</i>	<i>pinoN- + -i</i>	—
LV	irrealis	<i>poN- -an</i>	—	(no data)
LV	realis	<i>pinoN- -an</i>	—	<i>pinoN- -an + -akon</i>

Note: The indication ‘no data’ is used when the form is predicted to be possible but is not found in the given mode in available data. The indication ‘—’ is used when there is no evidence that the form is possible in either mode.

Bobongko thus represents one type of transitional system in which Philippine-type voice and pivot-neutral applicatives are both in use, though IV and LV appear at much lower textual frequency and appear to be in the process of being lost. In a pure Philippine-type voice system, each peripheral nonactor semantic role, e.g. beneficiary, instrument, location, etc., can be selected only as the pivot by the use of voice morphology on the verb, that is, by means of what I have called pivot-selecting ACs. In Bobongko, we see the emergence of pivot-neutral ACs in

which such peripheral roles may alternately be selected to map to the pivot, a non-pivot core argument, or an oblique, depending on the combination of co-occurring constructions. There are some irregularities and gaps in the Bobongko paradigm; such as the marking of PV with *o-* only in irrealis mode and not realis mode. There are also differences between the locative-selecting applicative marked with *-i/-an* vs. the beneficiary-selecting applicative marked with *-akon*; each co-occurs with a different set of nonactor voices, and shows different patterns in the syntactic properties of the applied phrase across these voices. Perhaps such irregularities might be expected when a language exhibits a transitional state, as I am arguing that we see here, with the loss of Philippine-type voice and the development of pivot-neutral applicatives both in progress. It is possible that, over time, in Bobongko we might see regularization of the morphological marking for the locative-selecting applicative bearing *-i/-an*, as is reported for Totoli Applicative 1 as discussed above. Also possible in the future would be a regularization of the use of *o-* PV.IRR, perhaps by loss of the prefix in irrealis mode, resulting in no overt morphological marking on the verb for PV besides person-index affixes or clitics, as seen in Central Sama, and many languages of Borneo and Sulawesi.²⁵ Such a loss would make it more tenable for verbs in IV and LV constructions to be interpreted as PV forms to which co-occurring applicative affixes have been applied, e.g. *poN-*, *poN-* *-an*.

The exact type of system as found in Bobongko is rare or even unique in West Nusantara. However, there are a number of languages in the sample that show “special” ACs akin to certain Bobongko LV and IV forms, where an applicative suffix appears together with some other fossilized affixal marking on the verb under certain conditions. One example is Pendau, which is discussed in the following section.

5.9.4.4 Pendau: Pivot-neutral applicatives showing remnants of Philippine-type voice

Pendau is a Tomini language spoken in the northern half of Central Sulawesi province. Like Bobongko, Pendau shows two major voice alternations (AV, PV) and two applicative suffixes, *-a'* and *-i*. A partial paradigm for symmetrical voice alternations and ACs in Pendau is given in Table 5.21.

In Pendau, a so-called ‘stem former’, *pV(C)-* appears on the verb in certain constructions. The exact form of the stem former is lexically determined by verb class; for example, most transitive verbs take *pong-*, while factive verbs (e.g. ‘make’, ‘build’) take *po-*, dynamic verbs (e.g. ‘search’) take *pe-*, and postural verbs (e.g. ‘sit’) take *popo-*. With certain classes of verbs, such as factives, the stem former is always obligatory in finite verb forms. With others, such as the primary transitive and dynamic classes, the stem former is generally absent in PV.

The appearance of the stem-forming prefix shows irregularities in certain ACs with primary transitive verbs, which is represented in the bolded cells in Table 5.21. In ACs marked with *-a'* that take beneficiary applied phrases, the stem-forming prefix may be absent in PV forms, as is the usual pattern for this verb class. But in ACs marked with *-a'* that take instrument applied

²⁵It is possible that Bobongko *o-* derives from one or more person-indexing prefixes akin to the first person non-pivot actor *ku-* mentioned above, which would account for the complementary distribution observed for *o-* and *ku-*. Besides *ku=* in Balantak (van den Berg & Busenitz 2012), I could find no other attested pronominal prefixes or proclitics in Saluan-Banggai. But for Bungku-Tolaki, which is thought to subgroup with Saluan-Banggai at a higher level (see Mead 2003), Mead (1999) reconstructs a set of preverbal actor pronominal forms that include first singular *aku, second singular *ko, and third singular *io, which appear to primarily be used in irrealis clauses in examples.

Table 5.21: Partial paradigm for voice and pivot-neutral applicatives in Pendau

Class	Voice	Mode	Base	Goal/Loc. Appl.	Ben./Inst. Appl.
Prim. Trans.	AV	irrealis	<i>M-</i> + <i>pong-</i>	<i>M-</i> + <i>pong-</i> + <i>-i</i>	<i>M-</i> + <i>pong-</i> + <i>-a'</i>
Factive	AV	irrealis	<i>M-</i> + <i>po-</i>	—	<i>M-</i> + <i>po-</i> + <i>-a'</i>
Prim. Trans.	AV	realis	<i>N-</i> + <i>pong-</i>	<i>M-</i> + <i>pong-</i> + <i>-i</i>	<i>M-</i> + <i>pong-</i> + <i>-a'</i>
Factive	AV	realis	<i>N-</i> + <i>po-</i>	—	<i>N-</i> + <i>po-</i> + <i>-a'</i>
Prim. Trans.	PV	irrealis	<i>ro-</i>	<i>ro-</i> + <i>-i</i>	<i>ro-</i> + <i>-a'</i>
				<i>ro-</i> + <i>pong-</i> + <i>-i</i>	<i>ro-</i> + <i>pong-</i> + <i>-a'</i>
Factive	PV	irrealis	<i>ro-</i> + <i>po-</i>	—	<i>ro-</i> + <i>po-</i> + <i>-a'</i>
Prim. Trans.	PV	realis	<i>ni-</i>	<i>ni-</i> + <i>-i</i>	<i>ni-</i> + <i>-a'</i>
				<i>ni-</i> + <i>pong-</i> + <i>-i</i>	<i>ni-</i> + <i>pong-</i> + <i>-a'</i>
Factive	PV	realis	<i>ni-</i> + <i>po-</i>	—	<i>ni-</i> + <i>po-</i> + <i>-a'</i>

phrases, the stem-forming prefix always appears. So, the PV verb form *ni-pong-gabu-a'* ‘cooked with’ takes an instrument applied phrase while *ni-gabu-a'* ‘cooked for’ takes a beneficiary applied phrase. Similarly, in ACs marked with *-i* that take goal applied phrases, the stem-former prefix may be absent in PV forms. But in ACs marked with *-i* that take static location applied phrases, the stem-forming prefix is obligatory.

Furthermore, for both primary transitive verbs and factive verbs, the instrument and static location applied phrases are always realized as the pivot, and only found when the verb is morphologically marked like other PV constructions, i.e. with *ro-* in irrealis mode and *ni-* in realis mode. Beneficiary and goal-selecting ACs, on the other hand, are compatible with both the PV and the AV constructions.

I argue that the use of *pV(C)-* with primary transitive verbs in instrument- and static location-selecting ACs in Pendau represents residual marking for IV and LV constructions.²⁶ In Bobongko, there are formal differences in morphological marking for base forms in LV, base forms in IV, and PV + applicative suffixes. In Pendau, such distinctions are minimal. Because the *pV(C)-* morpheme has no clear function in many verbal forms, and its presence in PV is only unusual for certain verb classes, its presence in the special applicatives looks like an irregular use of the stem former rather than part of distinct affixes marking separate IV and LV constructions. In most verb classes in Pendau, the form of the verb in pivot-neutral goal-selecting ACs in PV is identical to the verb in the pivot-selecting locative constructions, and the form of the verb in pivot-neutral benefactive ACs in PV is identical to the verb in the pivot-selecting instrumental ACs. Still, there is admittedly a degree of irregularity and ambiguity in the paradigms, that allow for different analyses of both Bobongko and Pendau. Further discussion of applicatives that sometimes appear with fossilized morphology is found in §6.3.

In summary, Pendau shows remnants of IV and LV constructions in irregular clausal properties and morphological marking for instrument-selecting and locative-selecting ACs, as compared to beneficiary-selecting and goal-selecting ACs. Like Balantak, Totoli, and Bobongko, here we see that reduction in Philippine-type voice categories coincides with development of pivot-neutral applicatives. Together, these four cases indicate that the transition between Philippine-type voice and pivot-neutral applicatives in West Nusantara may take place via (i) paradigmatic reduction of

²⁶Instrumental and location nominalizations in Pendau are also formed with the *pV(C)-* prefix, and this occurs across several verb classes, including a number for which ACs of this type are not attested (Quick 2007: 99–100).

CV and/or LV voice categories, (ii) reinterpretation of CV and/or LV affixal marking as applicative markers or stem formers without independent function that may appear on verbs in formerly contrastive voice categories (or functional replacement of such affixal marking with new AMs), and (iii) emergence of productive use of the new applicative markers in co-occurrence with PV, and crucially, AV transitive constructions, such that a selected peripheral role may map to the pivot argument, or to a non-pivot argument of the clause, depending on voice.

5.10 Summary of major findings

This chapter has explored the distribution of languages of West Nusantara with and without applicatives. Overall, geographic location and genetic affiliation are the most important factors shaping this distribution, though they do not fully account for the observed patterns.

Retention of the Philippine-type voice system, in which the peripheral nonactor voices (LV, CV) constitute pivot-selecting ACs, is found primarily in and adjacent to the Philippine archipelago—with Malagasy being an exception, and mainly in the Sabahan languages and the Sama-Bajaw languages (Greater Barito), with some occurrence in other languages of northern Sulawesi. Thus, languages of West Nusantara, like other western Austronesian languages, generally show decay in the Philippine-type voice system as one moves outside of Taiwan and the Philippines to the south (see McDonnell & Chen 2022).

Pivot-neutral applicatives that co-occur across basic voice alternations, on the other hand, are found to be a broadly distributed areal feature for West Nusantara associated with loss or reduction of the four-way Philippine-type voice system and not primarily spread through language contact. Such applicatives are found in languages (e.g. Mentawai, Enggano) which had little to no contact with Malay and other influential languages of the region. The presence of pivot-neutral applicatives furthermore cuts across languages of West Nusantara with a wide variety of structural features: symmetrical and asymmetrical voice systems, different types of morphological and syntactic alignment, different systems of case-marking and different preferred word order patterns. Given this, I conclude that there is no cohesive typological profile for languages of West Nusantara showing applicatives that co-occur across voice categories, by which they may be distinguished from languages of West Nusantara without such applicatives.

Considering these findings, the lack of applicatives, rather than the presence of them, is treated as a pattern that requires explanation in languages of West Nusantara. Based on the survey results and additional examination of descriptive resources for languages of West Nusantara outside of the sample, two profiles were developed to summarize key typological features of the languages of the sample without applicatives, divided into two sets. One of these is centered geographically on languages of mainland Southeast Asia and northern Peninsular Malaysia, and the other on Bornean languages south of Sabah. For both sets, loss of suffixation appears to be a major factor in the absence of applicatives. We also observe general reduction of morphological complexity, a greater reliance on word order to signal grammatical relations, and a shift towards analytic structures as alternatives for certain voice and applicative alternations observed elsewhere in West Nusantara. In the mainland SE Asia set, these changes are more complete across the grammars of individual languages, especially those with a greater time depth of language contact with non-Austronesian languages. In the Bornean set, these changes might affect different verbal constructions, e.g. AV and PV, to different extents and in different manners. Because

of this, the Bornean set is less cohesive compared to the mainland Southeast Asia set. Due to this fact, plus the existence of gaps in our understanding of the history of various Bornean language groups and lack of basic linguistic description for many others, the profile developed for the Bornean languages lacking applicatives is more preliminary in nature and less certain.

On the whole, the findings presented in this chapter cast further doubt on the appropriateness of an Indonesian-type subcategory of western Austronesian symmetrical voice languages, defined by two- or three-way symmetrical voice alternations, applicatives, and use of special clitic pronominal forms for indexing non-pivot actors (Himmelman 2005: 175). While many of the better known languages of the region meet these criteria, e.g. Standard Indonesian, Balinese, and Javanese, when we consider a larger, more representative sample of languages of West Nusantara, other conflicting combinations of features are quite broadly attested. Many Bornean language cross-cutting a diverse number of lower-level genetic groupings have no applicatives, but otherwise show features consistent with the definition of Indonesian-type languages. Conversely, outside of Borneo and mainland SE Asia, languages with voice and alignment systems with diverse characteristics are all found to have pivot-neutral applicatives in the great majority of cases.

In conclusion, pivot-neutral applicatives that co-occur across basic voice categories should be considered an areal feature of languages of West Nusantara, though such applicatives may be lost or blocked from developing due to special factors, such as phonological and/or morphological changes affecting word structure, perhaps triggered or spread by language contact. Furthermore, evidence from broad patterns of distribution, as well as nuanced interpretation of descriptive data for verbal constructions in individual languages, show that this type of applicative arises in the face of devolution of the Philippine-type voice system. In the next chapter, I will look closer at patterns of development for applicatives in light of the forms and functions of AMs and ACs in the focus area, and will argue that many AMs marking pivot-neutral applicatives are in fact derived from AMs that marked Philippine-type voice alternations (LV, CV) in PMP and PAN.

Chapter 6

Results: Properties of applicative constructions and their distribution in West Nusantara

In this chapter, I present results of the typological survey related to properties of applicatives systems, applicative morphemes (AMs) and applicative constructions (ACs). I find that the distribution of applicative functions across forms of AMs in West Nusantara is typologically unusual. A predominant pattern is observed whereby one form of AM-marking is used for locative- and goal-selecting ACs and another is used for beneficiary-, instrument- and theme-selecting ACs. Based on distributional evidence and TAM-conditioned relic alternations in certain languages, I argue that pivot-neutral ACs selecting locations and goal roles are derived from earlier LV constructions in Proto Malayo-Polynesian (PMP), while the pivot-neutral ACs selecting beneficiaries, instruments, and/or themes are derived from earlier CV constructions. Earlier LV morphology gives rise to pivot-neutral locative/goal AMs, while many benefactive/instrumental AMs are reflexes of the Proto Austronesian (PAn) CV imperative suffix *-an. However, this *-an has been replaced with newer suffixes like -kan and -akən in a number of subgroups. I also show that the distributional pattern for the causative function of AMs tend to be influenced by geographic location. Sulawesi languages in particular, show less prevalence and lower productivity of causative functions for AMs, than are observed in other parts of West Nusantara. Some languages of Sulawesi also show varying coding and behavior for the applied phrase, in patterns that are influenced by transitivity of the base and animacy of clausal arguments.

The organization of this chapter is as follows. Patterns in the distribution of applicative functions across forms of applicative morphology are discussed for AMs marking pivot-selecting ACs in §6.1 and for AMs marking pivot-neutral ACs in §6.2, with incidence of irregular marking for ACs presented in §6.3. Polyfunctionality of AMs is then discussed, and the distribution of causative functions, aspectual and intensive functions, and comparative functions is presented in §6.4. Following this, I turn to syntactic properties of ACs. In §6.5.1, results related to the co-occurrence of ACs across major voice constructions is presented. In §6.5.2, syntactic properties of the applied phrase in pivot-selecting ACs are discussed. In §6.5.3, syntactic properties of the applied phrase in pivot-neutral ACs are presented. Then in §6.5.4, alternations in the realization of the applied phrase in non-canonical ACs are presented. The chapter concludes with a summary

of major findings in §6.6.

6.1 Applicative morphemes marking pivot-selecting constructions

Of the 85 languages of the survey, a total of 59 show applicatives of any type. Nine languages show only pivot-selecting applicatives, 45 show only pivot-neutral applicatives, and five show both types. Each of these languages has between one and four total distinct forms of morphological marking for ACs.

In this section, I summarize and discuss survey results related to the forms and meanings of AMs marking pivot-selecting constructions in West Nusantara languages with Philippine-type voice systems. These include constructions in which a locative role is selected as the pivot (LV) and constructions in which an instrument, theme, or beneficiary is selected as pivot (CV). Here, I include both productive constructions of this type, and constructions with a restricted distribution or limited productivity, i.e. in languages that were coded as having marginal Philippine-type voice systems. There are nine languages of the sample that show pivot-selecting ACs only, and five that show both pivot-selecting and pivot-neutral ACs, for a total of 14 languages of the sample included in this section. To facilitate discussion of the diachronic sources of AMs found in the survey, in the next section, I give a brief overview of voice morphology in Proto Austronesian and Proto Malayo-Polynesian, from which all the Austronesian languages of West Nusantara are descended.

6.1.1 Voice morphology in Proto Austronesian and Proto Malayo-Polynesian

PAn and PMP both had a four-way Philippine-type voice system, with AV, PV, LV, and CV as contrastive voice categories. Voice morphology reconstructed for PAn is given in Table 6.1 from Chen (2017) (following Wolff 1973; and Ross 2009, 2012). As shown in the table, a three-way distinction in mood is also reconstructed. Additional aspectual or modal distinctions were also marked on indicative verbs in PAn but are not shown in the table. These include the (realis) perfective aspect marked with the infix *-in- and irrealis mode marked with the reduplicant *Ca- (Ross 2009). In addition to the infix *-um- shown in the table, in many present-day Philippine-type languages, a reflex of the stative prefix PAn *ma- can be analyzed as an AV marker with certain classes of verbs (see Chen 2017: 27).

Table 6.1: Proto Austronesian voice morphology (Chen 2017: 151)

Mood	Actor Voice	Patient Voice	Locative Voice	Circumstantial Voice
Indicative	*<um>	*-ən	*-an	*Si-/Sa-
Optative, hortative	*-a	*-aw	*-ay	*-anay
Imperative, negative	*-∅	*-u	*-i	*-an

In PMP we see a number of innovations in verbal morphology including the use of the prefixes *paN- and *paR- (*maN- and *maR- in AV) to form verbal stems (Ross 2002) to which voice and

TAM morphology were affixed. Ross (2002) suggests that the meaning of PMP *paN- is ‘distributive’ and that of PMP *paR- is ‘durative’ but notes that these semantic reconstructions are “very tentative” (49). The potentive or abilitative prefix *paka- (*maka- in AV) is another innovation in PMP. Otherwise, the four-way voice system, and mood and aspectual distinctions found in PAn are maintained in PMP.

6.1.2 Distribution and functions of LV and CV in the sample

There are 14 languages of the sample that show Philippine-type voice systems; contrastive voice distinctions in these languages are shown in Table 6.2. As shown in the table, all 14 of these retain AV and PV voice constructions. Additionally, in these languages we see between one and four contrastive pivot-selecting ACs. These may include up to one LV construction and up to three CV constructions, which are differentiated by distinct forms of morphological marking, and different sets of possible semantic roles for the applied phrase.

Table 6.2: Morphologically-marked voice distinctions in Philippine-type systems of the sample

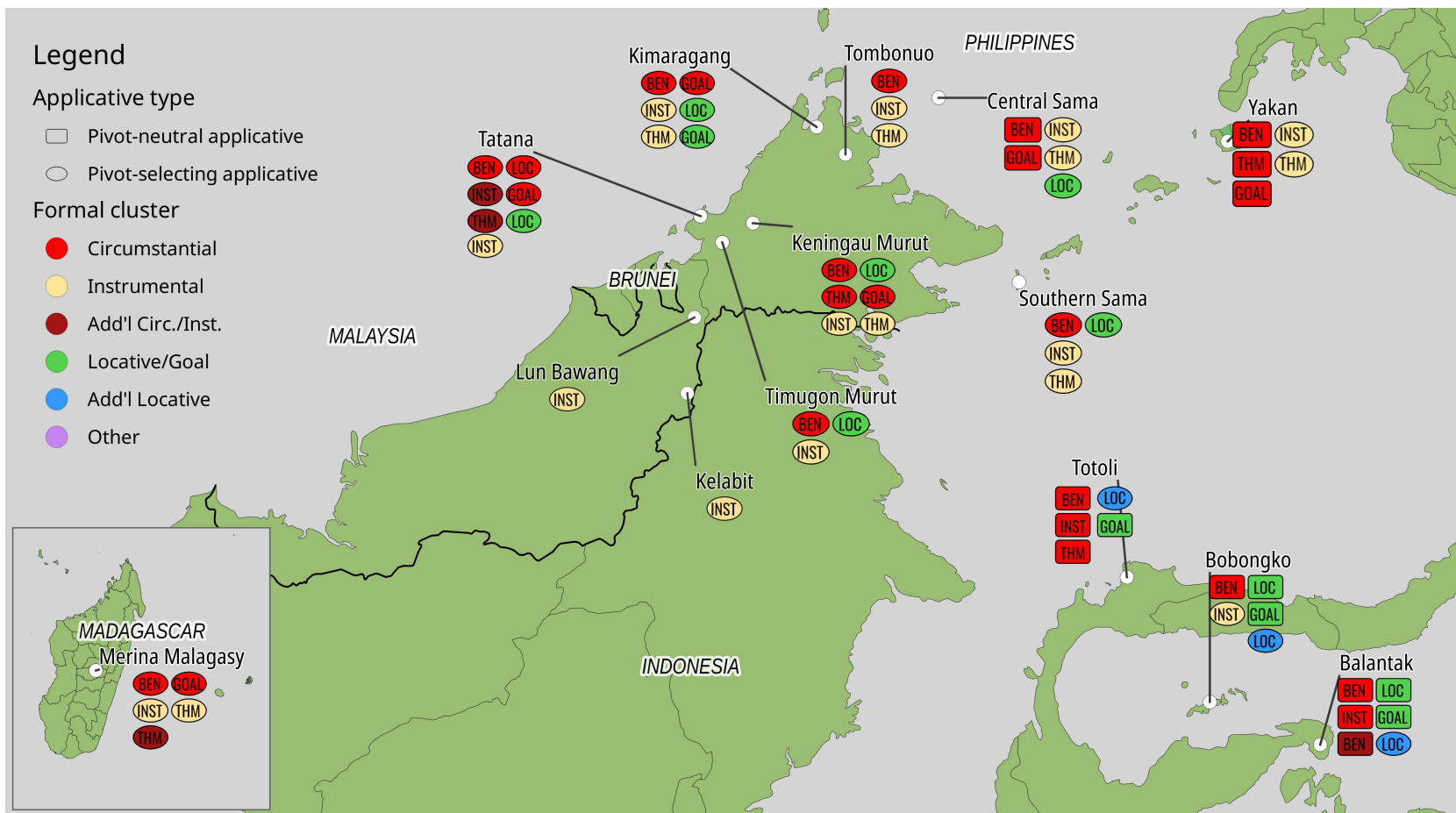
	AV	PV	LV	CV	No. Lgs.	Examples
3-way	✓	✓	✓	–	2	Totoli*, Balantak*
3-way	✓	✓	–	✓	3	Kelabit, Lun Bawang, Yakan*
4-way	✓	✓	✓	✓	2	Bobongko,* Central Sama*
4-way	✓	✓	–	✓✓	1	Tombonuo
5-way	✓	✓	✓	✓✓	4	S. Sama, Kimaragang
5-way	✓	✓	–	✓✓✓	1	Merina Malagasy
6-way	✓	✓	✓	✓✓✓	1	Tatana
TOTAL LGS.	14	14	9	12	14	

* These languages also show pivot-neutral applicative constructions.

Applicative systems of West Nusantara with pivot-selecting constructions are shown in the map in Figure 6.1. Here, each color of symbol on the map represents a distinct form of morphological marking for ACs, and each oval symbol represents a possible semantic role for the applied phrase in a pivot-selecting construction bearing such marking (while rectangular symbols pertain to pivot-neutral constructions in mixed applicative systems). For example, Central Sama has two AMs that marks pivot-selectings ACs: one marks an IV construction that takes an instrument or theme applied phrase as pivot (yellow ovals), and the other marks an LV construction that takes a location applied phrase as pivot (green oval). Central Sama also has an AM that marks pivot-neutral ACs; it marks ACs that select a beneficiary or goal applied phrase as a core argument (red rectangles).

LV constructions are represented in the map in Figure 6.1 as green or blue ovals. In an LV construction, the location of the clausal event is typically the pivot argument. The location may be the general location or setting in which the clausal event takes place. LV is also commonly used with a pivot expressing the location of a postural verb (e.g. sit, sleep). In a few languages, it is reported that the pivot in an LV construction may also express a goal role, i.e. Kimaragang, Central Sama. In addition to the location, in some languages, a phrase expressing the time that the clausal event takes place may be the pivot in an LV clause, with no difference in voice morphology

Figure 6.1: Pivot-selecting and mixed applicative systems of West Nusantara



Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

marked on the verb. In languages of the sample, the LV construction is retained in nine languages, but has been lost in five other languages that show at least one CV construction.

In CV constructions in PAn and PMP, the clausal pivot could be either (i) a beneficiary, (ii) an instrument, or (iii) a theme which undergoes a change of location, as well as (iv) a circumstantial role e.g. a phrase expressing the reason or purpose for an event. In the languages of the sample, we see that the single CV category of PAn and PMP may be narrowed or split into multiple contrastive voice distinctions based on the possible semantic roles of the pivot argument. In the map in Figure 6.1, CV constructions are represented by red, maroon, or yellow oval symbols, which each color representing one distinct form of morphological marking in a given language.

The most common CV voice category in the languages of the sample is one in which an instrument role—or either an instrument or a theme role—is selected as the pivot. As mentioned in §2.5.2, themes and instruments are inherently semantically related, as many instruments are directed into motion by an agent in order to achieve an effect. This category is commonly labeled instrumental voice (IV) or conveyance voice, and is found in all 12 languages of the sample with a CV construction. If a language of the sample has only one CV category, it is IV. In the sample, applicative systems with only one CV category are found in two North Sarawak languages (Kelabit and Lun Bawang), two Sama-Bajaw languages (Yakan and Central Sama), and Bobongko (Saluan-Banggai).

The second most common CV category is a construction that selects a beneficiary or recipient (sometimes also a goal) as the pivot. In many languages, a phrase expressing the circumstances under which an event occurs, such as a reason or purpose may also be selected as the pivot with no difference in voice morphology marked on the verb. This category is commonly labeled beneficiary voice (BV), and is also referred to as dative voice (DV), or referent voice (RV) by some authors. If a language of the sample has two CV categories, they are IV and BV. In the sample, applicative systems with two CV categories are found in Southern Sama, and four Sabahan languages (Tombonuo, Keningau Murut, Timugon Murut, and Kimaragang).

Finally, there are two languages of the sample in which three contrastive CV voice categories are found. In Tatana (Sabahan), the IV category is split into two constructions. In clauses marked with the verbal prefix *i-*, the pivot is generally a theme in a transfer event (e.g. the thing given, bought, or returned); while in clauses marked with *poN-* the pivot is may be an instrument or theme. In Merina Malagasy, there are three contrastive CV voice categories (Pearson 2001; Rasoloson & Rubino 2005). In clauses marked with the verbal prefix *a-*, the pivot is an instrument or a theme which is conveyed. In clauses marked with the verbal suffix *-ana* (underlying form */-an/*) alone, the pivot is a beneficiary, recipient, or goal. In clauses marked with both a prefixal form (e.g. *a-*, *an-* or *i-*) and suffixal *-ana*, the pivot can be any of a number of peripheral semantic roles, including an instrument, recipient, and in certain existential structures or nominalizations, the purpose, reason or circumstance of the event. This construction marked with prefix + *-ana* is not fully productive, and is primarily found in relative clauses. In both Tatana and Merina Malagasy, there is some functional overlap of the CV voice categories as constructions that take certain roles as pivot can be marked in more than one way.

On the basis of these data, an implication hierarchy may be formulated. In the languages of the sample showing pivot-selecting applicatives, if an applicative system has an AC that selects a beneficiary applied phrase, it also has an AC that selects an instrumental applied phrase. This pattern is cross-linguistically unusual, which will be discussed in §6.2.1.4 below following

presentation of data on distribution of functions across pivot-neutral AMs.

6.1.3 Forms of morphemes marking LV in the sample

The form of the morpheme marking LV in the languages of the sample is quite consistent. It generally has a suffixal component *-an* (*-on* in Kimaragang), that can be analyzed as a reflex of PAn LV indicative **-an*. In a few languages, we also see mood alternations reflecting use of earlier PAn LV imperative **-i*. In Totoli, LV is marked with *pog- -an/poN- -an* in realis forms and *pog- -i/poN- -i* in non-realis forms, with the latter apparently reflecting earlier LV imperative **-i*.¹ Thus, the suffixal element of AMs marking LV in these languages is clearly inherited from PAn.

In most languages of the sample, an additional prefixal component with the form *paN-* or *poN-* (and sometimes also *pag-* or *pog-*) commonly appears on the verb alongside the suffixal element. This component is not found in Kimaragang or Timugon Murut for LV, which do make use of similar prefixal forms in other voice constructions, but apparently not LV (see Prentice 1995: 385–389; Kroeger 2005: 419–422). Prefixal forms marking LV in these languages are not reconstructable as voice/applicative markers in PMP, but appear to be reflexes of PMP **paN-* ‘distributive’ or **paR-* ‘durative’. In the present-day languages, these prefixal forms have been alternately analyzed as stem formers, markers of semantic transitivity, aspectual markers, or simply part of a circumfix with suffixal *-an* indicating LV.

In summary, these observations are note for AMs marking LV construction in the sample: (i) inherited LV voice suffixes are observed with some relics of earlier mood distinctions, and (ii) prefixal elements are observed which are not derived from earlier voice markers, but verbal prefixes with possible aspectual meanings (e.g. PMP **paR-* ‘durative’, **paN-* ‘distributive’). Similar characteristics are also observed for AMs marking pivot-neutral locative/goal-selecting ACs in the survey, as discussed below.

6.1.4 Forms of morphemes marking CV in the sample

The forms of the morphemes marking CV categories in languages of the sample are more varied than those for LV. IV constructions, which are found in 12 languages of the sample, are most often marked with a verbal prefix. In just three languages of the sample—Kimaragang, Tombonuo, and Tatana—this form is *i-* reflecting the PMP CV indicative prefix **Si-*. In Merina Malagasy, the IV prefix is *a-*, as mentioned above, possibly reflecting the alternate form of the PMP CV indicative prefix, **Sa-*. In remaining cases, the prefix is not reconstructable to PMP as a CV marker. The most common form of the IV prefix in the survey data is *paN-* or a similar form (e.g. *peN-*, *piN-*, *poN-*), which is found in seven languages represented in the sample: Yakan, Central Sama, Southern Sama, Kelabit, Lun Bawang, Bobongko, and Tatana (which shows both *poN-* and *i-* in two contrastive categories). In Timugon Murut, the IV construction is marked with partial reduplication of the verb root plus a similar prefix, e.g. *pVN-* or *pVg-*, which is analyzed as a transitivity marker. Thus, Timugon has *pam-ba-bali* ‘buy with s.t.’ from root *bali* ‘buy’ (Prentice 1969: 15–17). Partial reduplication is seen elsewhere in West Nusantara as a TAM marker, which is the

¹In Southern Sama, LV is marked with *paN-* *-an* in indicative forms, and *pan-* *-in* in imperative forms, but the latter form is explained in Sama-Bajaw as a phonological reduction of *-an* + a single imperative suffix *-un*, held to mark imperatives in all voices except AV (see James 2017: 58).

probable origin for the Timugon form as well.² In Keningau Murut, the construction analyzed as IV is marked with *pa-* *-on* in imperfective forms, *pa-* \emptyset in perfective forms and *pa-* *-o'* in imperatives and other non-finite forms (Cohen 1999: 40–41). This construction is shown in (129) and its marking appears to derive from a combination of the causative prefix *pa-* and the PV voice marking *-on/- \emptyset /-o'* which notably also has a zero-marked perfective form in Keningau Murut. A similar construction is found in Tatana, where the causative construction in PV is an alternative means of expressing conveyance of a theme with the theme role as pivot, as shown in (130).

- (129) Keningau Murut, Instrument expressed as pivot with *pa-* *-on*
Pa-pidis-on ni Jon pais kuno du tataun.
 IV-cut-IV PN.NPIV J. knife 1S.POSS.DIST NM.NPIV wood
 ‘John is hacking wood with **that knife** of mine.’

Cohen 1999: 41

- (130) Tatana, Theme expressed as pivot

- a. *I-taak ku ani' dokou bua' diti.*
 TV-give 1SG.GEN only 2SG.DAT fruit PROX
 ‘I am giving **this fruit** to you.’ (IV)
- b. *Pa-taak-on ku ani' dokou bua' diti.*
 CAUS-give-PV 1SG.GEN only 2SG.DAT fruit PROX
 ‘I am giving **this fruit** to you.’ (Causative + PV)

Dillon 1994: 48

Thus, while a few languages of the sample mark IV with a prefix inherited from PMP indicative CV *Si-/*Sa-, most instead show reanalysis of aspectual markers—i.e. PMP *paN- ‘distributive’ and partial reduplication, which marked ‘imperfective’ in PMP (Ross 2002: 33—or use of causative prefixes, e.g. PMP *pa- ‘causative’.

BV constructions, which are found in seven languages of the sample, are generally marked with the suffix *-an* in indicative forms. Some alternations in the form of the suffixal component are observed based on TAM distinctions. Usually, perfective indicative verbs in BV are marked with a combination of *-in-* and *-an*, however Timugon Murut and Keningau Murut show *-in* for perfective indicative BV and *-an* for imperfective indicative BV. All seven languages have a different suffixal form marking BV in imperatives and other atemporal forms. This suffix is *-i* or *-i'* in most Sabahan languages, but *-ai* in Kimaragang. As with LV imperative forms, BV imperative verbs in Southern Sama show *-in*, explained as a contraction of BV *-an* + imperative *-un*. In Malagasy, the shape of the imperative suffix is *-y/-o*, depending on the phonological shape of the stem, for verb stems in both PV and BV. As mentioned above, in Merina Malagasy, verbs show both suffixal *-an* and a prefixal component *a-/an-/i-* when marking certain semantic roles as pivot, including reason, purpose, and sometimes instrument or recipient.

The marking for BV in the languages of the sample generally thus appears to show use reflexes of PAn CV imperative *-an in indicative clauses, with imperatives marked with the suffix *-i*. PAn indicative *Si-/*Sa- is also found in some of these languages, but as an IV marker. The

²For example, in Tatana, the verb in an IV construction is marked with the prefix *paN-* but often bears CV-reduplication in non-past tense, thus the root *patoi* ‘kill’ may appear in an IV construction in non-past tense as either *pamatoi* or *pamamatoi* Dillon 1994: 99–100.

use of a suffix derived from *-an in indicative clauses is also observed in many languages of West Nusantara with pivot-neutral applications, and I argue below that PAn CV imperative *-an is the source morphology for the earlier form of the pivot-neutral AMs marking ACs with beneficiary, instrument, or theme applied phrases in West Nusantara languages.

6.1.5 Implications for diachronic development

Two important observations can be made regarding the distribution of affixal forms marking pivot-selecting constructions in the languages of the sample. Firstly, we observe that aspectual or modal prefixal marking in earlier stages of the language family, such as PMP *paN- ‘distributive’, *paR- ‘durative’, are being reinterpreted as voice markers in languages of West Nusantara. While it has been noted elsewhere that this re-purposing of *paN- and *paR- occurs for AV verbal forms in Malayo-Polynesian languages (see Ross 2002), here, it is of particular interest that this is also found for LV constructions and IV constructions that select an instrument/theme pivot. Prefixal marking observed for certain pivot-selecting ACs may also reflect causative *pa-, as in Keningau IV *pa- -on/pa- -o/pa- -o*. This is also relevant to the development of pivot-neutral applicatives in West Nusantara, as we have already seen that certain ‘stem-former’ prefixes with shapes consistent with *paN-, *paR- and *pa- irregularly co-occur with pivot-neutral applicative suffixes in languages like Totoli and Pendau that are transitioning from Philippine-type systems to two-way symmetrical voice systems with pivot-neutral applicatives (see §5.9.4). Use of such aspectual and valency-modulating prefixal marking together with applicative morphology thus may represent one important way by which causative and aspectual meanings have come to be commonly associated with pivot-neutral AMs in West Nusantara languages. I do note, though, the aspectual prefixes are not found in AMs that mark BV categories, such AMs generally are suffixes with the shape *-an* in indicative clauses.

Secondly, in addition to the reinterpretation of other types of verbal prefixes as applicative markers, in the survey data we observe another innovation—or rather shift—in the marking of CV constructions. Only a few languages surveyed retain *i-* or *a-* to mark CV, reflecting the PMP prefixes *Si-/*Sa for CV indicative forms. However, we see considerable use of the verbal suffix *-an* to mark CV constructions in indicative forms, especially constructions that select beneficiary or recipient pivots. In contrast, in PAn, *-an was only found on non-indicative CV forms; CV was marked with *-an in imperative/negative mood, and marked with *anay (probably decomposable to *-an + *-ay (Ross 2009) in hortative/optative mood, while *Si-/*Sa- was used in indicative mood. The use of *-an to mark pivot-selecting applicatives where the applied phrase is a semantic theme, instrument, or beneficiary, especially in indicative mood, is again highly relevant to the development of pivot-neutral applicatives in West Nusantara languages. Later in this chapter, I argue that the pivot-selecting CV construction from PMP is the original source that explains the broadly distributed pattern of association of beneficiary-, instrument-, and theme-selecting applicatives with a single form (or proto-form) in languages of West Nusantara showing pivot-neutral applicatives. In addition, I argue that this single form was probably *-an in earlier stages, with a pattern of replacement that gives rise to the present-day distribution of forms found for these functions. Because a historical shift in the use suffixes reflecting *-an from the marker of CV in imperative mood in PAn to CV in indicative mood in certain languages descendent to PMP is an important foundation for this argument, in the next section, I present further evidence for this interpretation of the synchronic distribution of AMs and their functions presented above.

6.1.6 On the emergence of *-an as the marker of CV in indicative mood

While the marker for CV in imperative mood in PAn is reconstructed as *-an in many sources (and shown in Table 6.1 above accordingly), it is possible that this morpheme was more complex. Ross (2009) reconstructs *ani for CV in imperative mood on the basis of evidence from a broad set of Formosan languages in which morphemes with the form *ani* or similar mark CV in imperative and/or negative clauses (e.g. Saisiyat, Seediq Saaroa, Mayrinax Atayal). He holds that *ani originally occurred preverbally, though it takes the form of a verbal suffix in most present-day languages that show a reflex. He also holds that *ani is probably decomposable into *an + *-i, with *an representing a preverb used with CV in certain non-indicative forms, and *-i representing the imperative suffix in LV and CV. This fits into a reconstruction of the paradigm for voice and the non-indicative moods as shown in Table 6.3 based on Ross (2009: 306). Here, non-indicative mood as reflected in imperatives is zero-marked in AV, marked with *-u in PV, marked with *-i in LV, and marked with the preverb *an + *-i in CV. The optative/hortative mood is characterized by marking with suffixal *-a across all four voices, with *-aw being composed of *-a + *-u in PV, *ay being composed of *-a + *-i in LV, and *an-ay being composed of the preverb *an + *-a + *-i in CV.

Table 6.3: Proto Austronesian non-indicative verbal morphology revisited

	Actor voice	Patient Voice	Locative Voice	Circumstantial Voice
Optative, hortative	STEM-a	STEM-aw	STEM-ay	*an-ay + STEM
Imperative	STEM	STEM-u	STEM-i	*an-i + STEM

In such a paradigm, we see both the morpheme *an that is reflected as *-an* in CV indicative marking in many Philippine-type languages of West Nusantara of the sample, and the morpheme *-i, that is reflected as CV imperative marking in some of these languages. However, for my interpretation to be correct, *an must undergo a diachronic change from a voice marker in non-indicative mood in PAn to a voice marker in indicative mood in some West Nusantara languages. This precise shift occurs in a number of other Philippine-type languages outside of West Nusantara, and is found in both Formosan and Philippine languages, which will be taken up in turn below.

In Formosan languages, we see a number of instances where a reflex of *an (or *an-ay or *an-i) is used as a verbal suffix marking CV in indicative mood. This occurs in Puyuma, where the marker for CV in main clause indicative mood is *-anay*, while imperatives are marked with *-an* (Tsukida 2009). In Tsou, which does not show morphologically marked mood distinctions, the marker for CV is *=(n)eni*, which may attach to verb stems (Tsuchida 1976). In Saaroa, the marker of CV in both indicative main clauses and imperative main clauses is *-ani* (Zeitoun & Teng 2016). Chen (2017) interprets these patterns as a diachronic change whereby verbal morphology in non-indicative mood—and originally, dependent clauses, as these moods were marked with a preverb in PAn—comes to be used in indicative, main clauses. She cites a variety of evidence, including distribution of the relevant forms across Formosan subgroups, and the observed cross-linguistic tendency for main clause structures to be more innovative, while dependent clauses often retain older structures and remnant forms.³ Though I have chosen to describe the pattern

³This pattern of older voice constructions and voice morphology being primarily used or only retained in certain

we are seeing for marking of CV in West Nusantara in terms of this interpretation, for my purposes it does not matter whether *an (or *an-ay, or *an-i) begins in PAN as a marker of CV in non-indicative/dependent clauses and later replaces *Si-/*Sa in indicative/main clauses in certain daughter languages, or whether—as others have claimed—CV *an was used with all types of verbs in earlier stages of PAN and later was replaced by the circumstantial nominalizing prefix *Si-/*Sa in some clause types. For my purposes it only matters that it is plausible that *-an was used frequently to mark CV constructions in some relevant stage(s) below PMP, whereby this form—or a replacement for it—could be inherited in West Nusantara languages that lack pivot-selecting applicatives but show a pivot-neutral applicative marker that can select a beneficiary, instrument, or theme, among other semantic roles, as the applied phrase, e.g. Proto-Malayic applicative *-an (Adelaar 2006).

The use of *-an* as a marker for CV outside of the non-indicative moods is also noted to occur in Philippine languages. Based on a broad review of languages of the Philippines representing all accepted subgroups, Reid & Liao (2004: 460) observe that verbs that select an instrument/theme undergoer (irrespective of mood) are commonly marked with the prefix (?)*i-* derived from earlier *Si-. Verbs that select beneficiary undergoers in Philippine languages (irrespective of mood) are marked with either (i) only a prefix (?)*i-* (e.g. Ivatan), (ii) only a suffix *-an* (e.g. Maranao), (iii) a circumfix (?)*i-* *-an* (e.g. Balangaw), or (iv) a combination of more than one of these options, with the selection being lexically determined (e.g. Mamanwa, Tagalog, Ilokano). The first category here is of particular interest, and includes “Casiguran Dumagat, Bashiic languages (such as Ivatan, Yami, Itbayay, etc.), Kapampangan, and Sambalic languages, Bikol, etc.” (Liao 2004: 142). These languages are not particularly closely related, and include Baashiic (Batanic) languages, Central Luzon languages, Greater Central Philippine languages, and Northern Luzon languages, which are each classified as a separate primary branch of Malayo-Polynesian by Eberhard & Fennig (2023) and Hammarström et al. (2022).

I interpret this evidence to show that it is quite likely that there was an *-an marker for CV constructions that was inherited in languages descended from PMP, and this marker was associated with both beneficiary-selecting and instrument-/theme-selecting ACs. Furthermore, I hold that as the symmetrical voice and mood systems of these languages underwent diachronic change, for example in the collapsing of the three-way mood distinction between indicative/optative/imperative, and the loss or reinterpretation of certain verbal affixes indicating TAM distinctions, this marker *-an has frequently also been reinterpreted such that it is used normally in indicative main clauses. In both Philippine and Formosan languages, we see that reflexes of *-an are commonly used to mark CV in indicative mood. Because the Philippine languages, like the West Nusantara languages, are descendants of PMP, we know that the use of *-an to mark CV is not ruled out by some loss of a necessary ancestral construction or form prior to PMP. Second, the distribution of forms for the CV marker in Philippine languages, suggests that at some stages of development, in branches of PMP that retain a reflex of *an as a CV marker, both *Si- and *-an were in use as CV markers at the same time, and in some cases they were even used on the same stem at the same time, as in the use of *i-* *-an* in languages like Balangaw and possibly *a-* *-an* in Merina Malagasy. Thirdly, the use of only reflexes of *Si- to mark constructions with beneficiary applied

types of dependent clauses is found in a number of West Nusantara languages of the sample that I have classified as marginal Philippine-type voice systems (e.g. Central Sama) or marginal symmetrical voice systems (e.g. Nias), see §5.2.3, and is also related to observed differences in marking of clausal arguments in different clause types in a number of Sulawesi languages (e.g. Mori Bawah, Seko Padang) as mentioned in §5.2.2.

phrases alongside those with instrument applied phrases is found synchronically in Philippine languages of diverse subgroups, suggesting that the split in morphological marking for CV categories that select an instrument/theme applied phrase from CV categories that select a beneficiary applied phrase, which is observed both in Philippine languages and West Nusantara languages with Philippine-type voice systems, is unlikely to have occurred early on in the development of Malayo-Polynesian languages, as the daughter languages in which this split does not occur are fairly broadly distributed.

A full picture of the diachrony of morphemes marking pivot-selecting constructions in West Nusantara languages necessarily must take into account likely cognate forms that show different related functions, that is, those which do not presently function as pivot-selecting applicative markers. These related functions would include those that were present alongside voice marking for the relevant morphemes in earlier stages of the languages (e.g. nominalization, modal modulation, dependent clause marking, etc.), and related innovative functions. Neither of these have been systematically included in my discussion of morphological markers of LV and CV thus far, however, the latter is addressed in part in the following section, in which I turn to morphemes that mark pivot-neutral ACs in West Nusantara languages, and their functions. As I will show, many of these pivot-neutral applicative markers appear to derive from earlier pivot-selecting applicative markers, or to represent replacement forms for the same.

6.2 Applicative morphemes marking pivot-neutral constructions

In this section, I summarize and discuss survey results related to the forms and meanings of AMs marking pivot-neutral constructions in West Nusantara languages. There are 45 languages of the sample that show pivot-neutral applicatives constructions only, and five that show both pivot-selecting and pivot-neutral ACs, for a total of 50 languages of the sample that are included in results presented in this section. An overview of the applicative systems of these languages is shown in the maps in Figures 6.3 and 6.2.

These languages show between one and four applicatives with distinct morphological marking, which are represented in the maps as distinct colors for the symbols shown under the name of each language. Paradigmatic alternations in the form of the morphological marking of ACs, e.g. alternations in the forms of applicative affixes in clauses with different TAM categories, and lexically conditioned allomorphy are not counted as distinct here, as long as semantic characteristic of the marked construction are not otherwise observed to co-vary with such alternations.

6.2.1 Applicative systems by number and function of applicative morphemes

Table 6.4 below shows a summary distribution of the languages of the sample with pivot-neutral applicatives by the number of distinct forms of applicative marking observed for such constructions in each language. As shown in the table, applicative systems with two distinct forms of applicative marking are by far the most common in West Nusantara, with 36 languages of the sample representing this category (out of 50 total). Also observed but much less common are

systems with one distinct form of applicative marking (6 of 50), and systems with three distinct forms (7 of 50). Finally, only one language of the sample—Mori Bawah—shows four distinct forms of applicative marking (1 of 50).

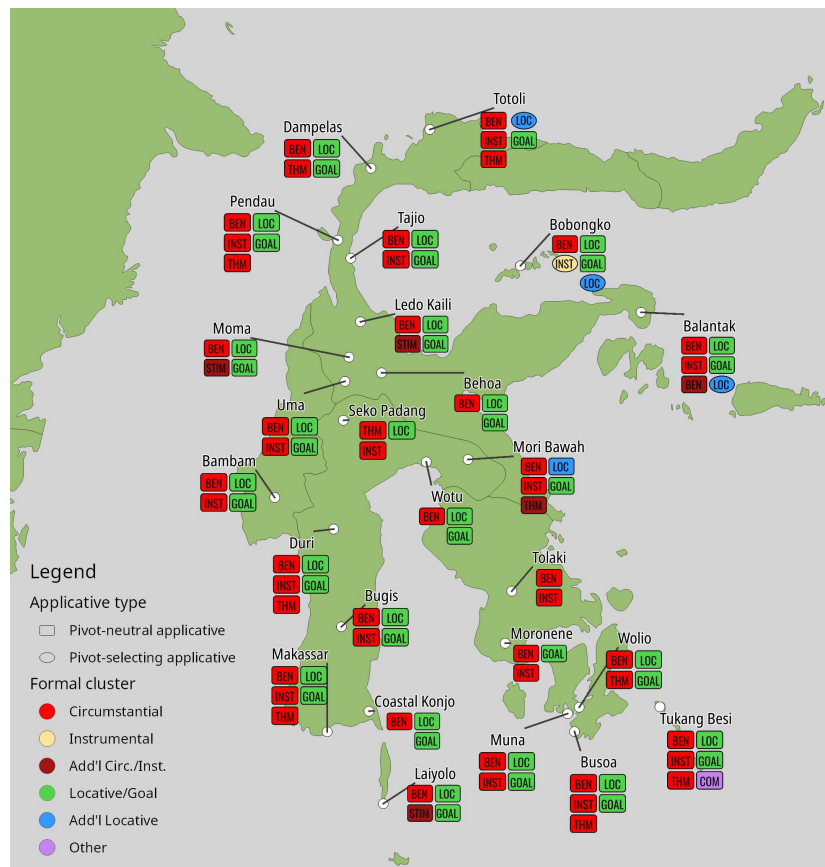
Table 6.4: Morphological marking for pivot-neutral applicatives

No. Forms	No. lgs.	Examples
1 form	6	Central Sama, Ampenan Sasak, Tolaki
2 forms	36	Balinese, Std. Indonesian, Pendau, Toba Batak, Bugis
3 forms	7	Sundanese, Balantak, Tukang Besi, Laiyolo, Moma
4 forms	1	Mori Bawah
TOTAL LGS.	50	

As mentioned above, five languages of the sample show both pivot-neutral and pivot-selecting ACs. With respect to pivot-selecting applicatives, these five languages either have just one locative-selecting (LV) construction (2 of 5; Balantak, Totoli) or one locative-selecting (LV) and one instrument-selecting (IV) construction (3 of 5; Bobongko, Central Sama, Yakan). Compared to languages of the sample with only pivot-selecting applicatives, this is on the lower end of the observed range for number of contrastive pivot-selecting constructions (see §6.1.2). Bobongko, Central Sama, and Yakan also show a restricted distribution and/or limited frequency for both LV and IV. In addition, none of the five languages with both types of applicatives show a pivot-selecting construction in which the beneficiary role is selected as pivot. With respect to pivot-neutral applicatives, these five languages either show one distinct form of morphological marking (2 of 5; Central Sama, Yakan), two distinct forms (2 of 5; Bobongko, Totoli), or three distinct forms (1 of 5, Balantak). This is generally comparable to the observed range for the number of distinct forms for morphological marking of applicatives for languages of the sample with only pivot-neutral applicatives.

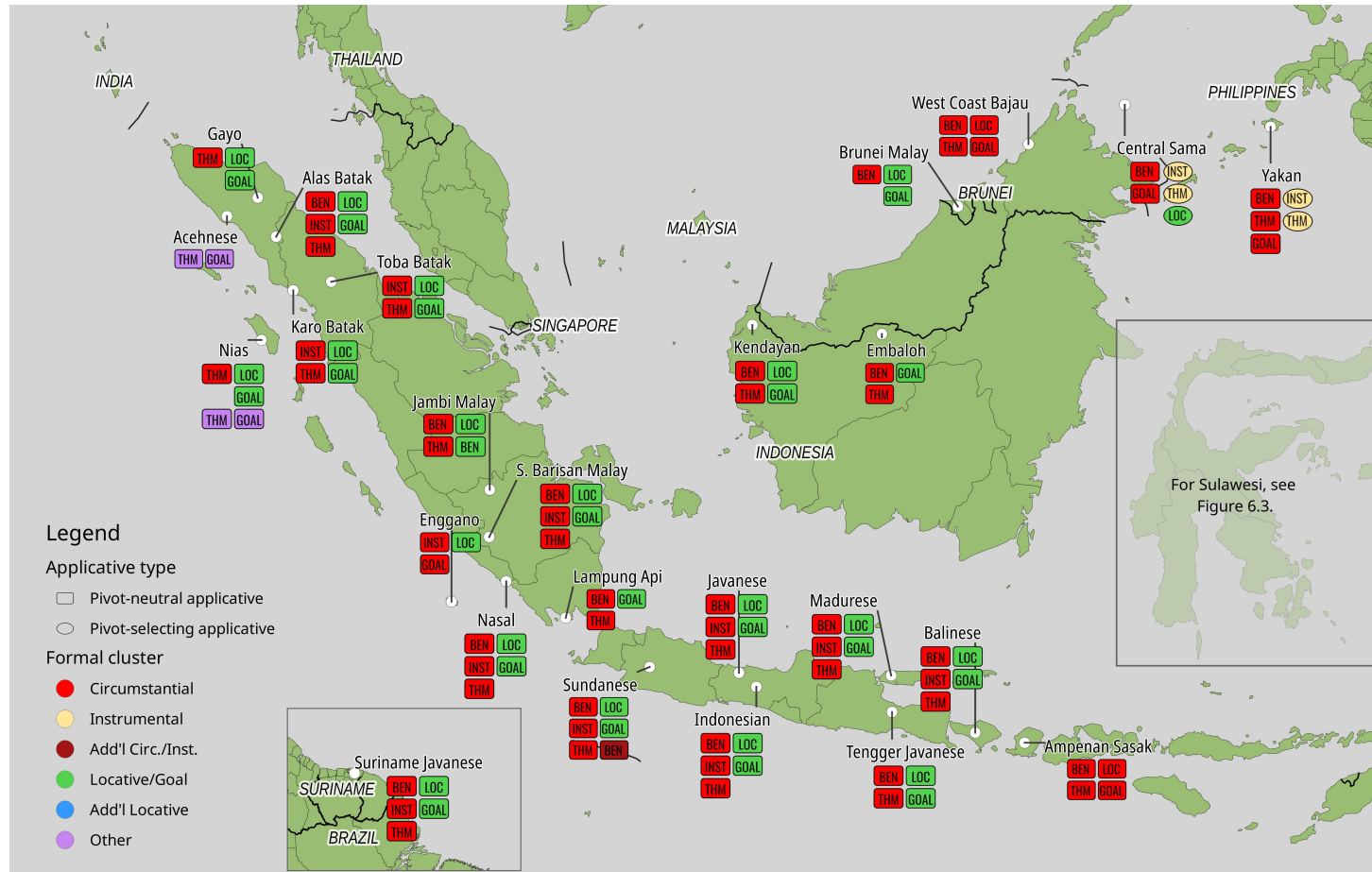
In summary, languages of West Nusantara in the sample show between one and four pivot-selecting constructions, and between one and four distinct forms of morphological marking for pivot-neutral constructions, but never more than four across the two types. Languages with both Philippine-type voice systems and pivot-neutral applicatives tend to have fewer contrastive pivot-selecting (or voice) constructions than observed in other Philippine-type languages and lack a construction in which the beneficiary role is selected as the pivot. These patterns suggest that pivot-neutral applicatives serve as functional replacements for lost pivot-selecting constructions, first for selection of the beneficiary role as the applied phrase, then for selection of instrument/theme roles as the applied phrase, and lastly or not at all for selection of locative roles (or perhaps just the static location role) as the applied phrase. Because there are relatively few languages with both types of applicatives in the sample, we must also look at other types of evidence to ascertain whether such a pattern of replacement is supported in the larger group of languages with only pivot-neutral applicatives. In the following sections, more detailed information on the distribution of forms of applicative markers and associated applicative functions are given for different types of systems observed, as categorized by number of distinct forms marking pivot-neutral applicatives in a given language.

Figure 6.2: Pivot-neutral and mixed applicative systems of Sulawesi



Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

Figure 6.3: Pivot-neutral and mixed applicative systems of West Nusantara



Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

6.2.1.1 Languages with two distinct forms marking pivot-neutral applicatives

The most common type of applicative system in West Nusantara shows two distinct forms of morphological marking for pivot-neutral applicatives; this is found in 36 languages of the sample (of 50 total languages with pivot-neutral applicatives). As a proportion of all languages in the sample with pivot-neutral applicatives, the languages with two distinct forms of marking for pivot-neutral applicatives include all Malayic languages (5 of 5), most non-Malayic languages of Sumatra and the Barrier Islands (7 of 8), most languages of Sulawesi (19 of 24), and all three Javanese languages (3 of 3), as well as Madurese, and Balinese. Languages with a two-way distinction in morphological marking of pivot-neutral applicatives are shown in Table 6.5. All of these languages (36 of 36) show a distribution of functions in which one form marks constructions that select a beneficiary, instrument, and/or theme role to map to the applied phrase, while the the second form marks constructions that select a locative and/or goal role to map to the applied phrase.

Table 6.5: Languages with two distinct forms marking pivot-neutral applicatives

Genetic Grp.	Language(s)	Form 1			Form 2	
		BEN	INST	THM	LOC	GOAL
NWS-BI	Karo Batak, Toba Batak, Alas, Gayo	(✓)	(✓)	✓	✓	✓
NAS	Nasal	✓	✓	✓		✓
ENG	Enggano		✓		✓	
LAM	Lampung Api	✓	✓		✓	✓
T-T	Dampelas, Pendau, Tajjo	✓	(✓)	(✓)	✓	✓
T-T	Totoli*	✓	✓	✓		✓
K-P	Uma	✓	✓		✓	✓
K-P	Behoa*	✓			✓	✓
B-T	Moronene	✓		✓		✓
M-B	Busoa, Muna	✓	✓		✓	✓
W-W	Wolio, Wotu	✓	(✓)	(✓)	✓	✓
SSUL	Bambam, Bugis, Duri, Embaloh, Co. Konjo, Makassar, Seko Padang	✓	(✓)	(✓)	✓	(✓)
MAL	Brunei, S. Barisan Malay, Jambi, Kendayan, Std. Indonesian	✓	✓	✓	✓	✓
JAV	Javanese, Sur. Jav., Tengger	✓	(✓)	✓	✓	✓
MAD	Madurese	✓	✓	✓	✓	✓
BBS	Balinese	✓	✓	✓	✓	✓

* An asterisk indicates a language for which inclusion in the listed genetic grouping is disputed or doubtful.

() Parentheses indicate that the function is found in this set of languages, but is not reported or marginal in at least one language represented.

6.2.1.2 Languages with one distinct form marking pivot-neutral applicatives

Systems that show one distinct form of morphological marking for pivot-neutral applicatives have a moderately low frequency of occurrence in West Nusantara. This type of system is found in seven languages of the sample (6 of 50 total languages with pivot-neutral applicatives). Table 6.6 below shows the distribution of common applicative functions across AMs in these languages.

This system is common among the Sama-Bajaw languages (Greater Barito linkage) and is otherwise found sporadically in West Nusantara, especially where one AM has been lost.

The system with one pivot-neutral AM is found in Sama-Bajaw languages that show loss of productivity for CV and/or LV constructions. These include West Coast Bajau, Central Sama, and Yakan in the sample (see also §5.8.4.2 on Indonesian Bajau and Pangutaran Sama). In these languages, the pivot-neutral AM *-an* is always observed to mark constructions where a beneficiary role is the applied phrase (and typically also goal-selecting constructions, though this may be much less productive). In Yakan and West Coast Bajau, theme-selecting constructions are also marked with *-an*, and in West Coast Bajau, locative-selecting constructions are as well. In conservative Sama-Bajaw languages *-an* marks CV constructions with a beneficiary as pivot (e.g. Southern Sama and Balangini’, see Akamine 2003; Gault 1999). Therefore, the pattern of meanings seen with *-an* as a single pivot-neutral AM appears to result from the extension of a benefactive voice marker to a generalized applicative marker.

A similar pattern is observed in Ampenan Sasak (Bali-Sasak-Sumbawa), which shows a single pivot-neutral AM that marks selection of beneficiaries, locations, themes, and goals, among other roles, as the applied phrase. By comparison with Balinese and some other Sasak varieties (e.g. Ngenó-Ngené) that have beneficiary-selecting *-ang* and locative-/goal-selecting *-in* (Arka 2003; Austin 2001), in the absence of *-in*, Ampenan Sasak *-an* functions as a more generalized applicative marker.

Loss of one AM without subsequent extension is also observed. In Tolaki (Bungku-Tolaki), the sole remaining AM *-Cako* marks the selection of a beneficiary or instrument role as the applied object. The applicative marker *-i* that marks locative-/goal-selecting ACs in other Bungku-Tolaki languages, e.g. Moronene and Mori Bawah—and widely elsewhere in West Nusantara—has been lost.

Finally, in Acehnese, we see limited use of a causative prefix as an AM. The prefix *peu-* (< PMP *pa- ‘causative marker’) can be used with a limited number of root with the meaning ‘administer ROOT onto/at an undergoer’ or ‘to place s.t. at ROOT’ (Durie 1985). The prefix *peu-* in Acehnese is highly productive, and the semantic meaning of roots bearing *peu-* is broad. In this light, the applicative functions of *peu-* may be considered of limited importance, though they illustrate a connection between causative and applicative meanings.

Table 6.6: Languages with one distinct form marking pivot-neutral applicatives

Genetic Grp.	Language(s)	Form 1				
		BEN	INST	THM	LOC	GOAL
GRB	West Coast Bajau, C. Sama, Yakan	✓	(✓)	(✓)	(✓)	✓
BSS	Ampenan Sasak	✓		✓	✓	✓
B-T	Tolaki	✓	✓			
CHAM	Acehnese			✓	✓	

() Parentheses indicate that the function is found in this set of languages, but is not reported or marginal in at least one language represented.

6.2.1.3 Languages with more than two forms marking pivot-neutral applicatives

Applicatives systems in West Nusantara that show three distinct forms of morphological marking for pivot-neutral applicatives have a moderately low frequency of occurrence; this type of system is found in seven languages of the sample (7 of 50 total languages with pivot-neutral applicatives). In a single language—Mori Bawah—four distinct forms of morphological marking for pivot-neutral applicatives are observed.

Table 6.7: Languages with more than two distinct forms marking pivot-neutral applicatives

Genetic Grp.	Language(s)	Form 1			Form 2		Form 3	Form 4
		BEN	INST	THM	LOC	GOAL		
S-B	Balantak	✓	✓	✓	✓	✓	BEN	
SUN	Sundanese	✓	✓	✓	✓	✓	BEN	
M-B	Tukang Besi	✓	✓	✓	✓	✓	COM	
K-P	Kaili Ledo, Moma	✓		(✓)	✓	✓	(THM), (STIM)	
W-W	Laiyolo	✓			✓	✓	STIM	
NWS-BI	Nias			✓	✓	✓	THM, GOAL	
B-T	Mori Bawah	✓	✓		✓	✓	THM	GOAL, STIM

() Parentheses indicate that the function is found in this set of languages, but is not reported or marginal in at least one language represented.

In Balantak, Sundanese, and Tukang Besi, systems with three distinct AMs marking pivot-neutral constructions originate from the addition of one unusual, innovative AM in addition to two AMs found more widely. These systems will be discussed in turn below.

In Balantak, the innovative AM is *-ii*, and it exclusively marks beneficiary-selecting constructions, while *-ako* marks the selection of a beneficiary, instrument, or theme, among other roles as the applied phrase. The suffix *-ii* is thought to derive from a combination of the locative-/goal-selecting AM *-i* and the personal article *i* that precedes personal names, titles, kinship terms, and pronouns (van den Berg & Busenitz 2012: 109). In this example, it appears that animacy effects—with beneficiaries and personal nouns both being animate and usually human—has driven the diachronic development of a new AM.

In Sundanese, the innovative beneficiary-selecting AM is the circumfix *pang- -keun*, which has a substitutive meaning ('to do s.t. in lieu of s.o. to their benefit') that distinguishes it from the use of *-keun*, which forms instrument-/theme-selecting ACs in addition to some beneficiary-selecting applicatives (see §2.8–2.7). While the origin of *pang- -keun* is not entirely clear, the shape of the prefixal component suggests some possibilities. On the one hand, it could be related to the PMP verbal prefix **paN-*, which, as mentioned above, plays in marking voice (especially IV) and TAM categories in other western Austronesian languages. On the other hand, a prefix **paŋ-* is also found as a marker of agent or instrument nominalizations in Philippine languages (e.g. Ilokano, Tagalog, Bikol) and West Nusantara languages (e.g. Standard Indonesian, Toba Batak, Balinese). In my estimation, the nominalization prefix is more likely to have given rise to the prefixal component of *pang- -keun*, because in agent nominalizations, it means 'one who performs an action', while Sundanese constructions marked with *pang- -keun* always mean 'to perform an action in lieu of some beneficiary'. The suffixal component is clearly cognate with *-keun*, the AM that marks selection of beneficiary-, instrument-, and theme- applied phrases. As

with Balantak *-ii*, here we see the innovation of a new AM that exclusively selects beneficiaries as the applied phrase.

In the case of *Tukang Besi*, the innovative AM is *-ngkene* which marks comitative-selecting constructions. This morpheme is related to the morpheme *kene*, which functions in *Tukang Besi* as an instrumental preposition meaning ‘with’, a conjunction meaning ‘and, with’, a verb meaning ‘to accompany’ with the form *kene* or *ngkene*, and the marker of an existential predicate with the form *ke(ne)*. As a verb *kene* is also found in serial verb constructions. Donohue (1999: 187-99) analyzes this as a case of grammaticalization as it appears that this morpheme is verbal root that has come to also behave as a preposition and/or bound AM. I also note that in *Tukang Besi* constructions marked with *-ngene*, the comitative applied phrase is “an equal and voluntary participant in the action indicated by the verb” (Donohue 1999: 228), and thus always animate.

In these three languages, the underlying pattern of one AM marking selection of beneficiary, instrument, and theme applied phrases and one AM marking selection of locative/goal applied phrases remains in place. However, in each case a new innovative AM has arisen and, though the mechanisms for this innovation are different in each case, all three of the new AMs is exclusively used with animate applied phrases.

In *Kaili-Pamona* and *Wotu-Wolio*, two Sulawesi microgroups that are probably most closely related to one another (Zobel 2020), applicative systems with three pivot-neutral AMs are also observed. In the sample, this is found in *Kaili Ledo* and *Moma* for *Kaili-Pamona*, and *Laiyolo* for *Wotu-Wolio*. In these systems, we see that the usual functions of the AM selecting beneficiary, instrument, and theme applied phrases are split across two morphemes, *-(C)aka* and *-ka*. The AM *-ka* consistently marks beneficiary-selecting constructions across the three languages, and was reported to rarely mark instrumental-selecting constructions in *Moma*, (Adriani & Esser 1939). The AM *-(C)aka* on the other hand, in these three languages marks certain theme-selecting constructions (especially with bodily function verbs, e.g. ‘to spit out s.t.’) and stimulus- or target-selecting constructions (e.g. ‘to laugh on account of s.t.’, ‘to be afraid of s.t.’). In languages that show both *-ka* and *-(C)aka*, *-ka* (used with beneficiary-selecting constructions) is more common, and *-(C)aka* may be rare in occurrence or of limited productivity in the set of verbal roots to which it attaches.

As a point of comparison, the functions of *-ka* and *-(C)aka* described above are both associated with a single AM in closely related languages in both microgroups. In *Wolio* (*Wotu-Wolio*) these are associated with *-aka* (Anceaux 1952) and in *Uma* (*Kaili-Pamona*), these are associated with *-ki*, from fusion of earlier *-ka* and the third singular person index *-i* (see Martens 1988b: 210–212). Based on these patterns, it is not tenable to reconstruct two separate AMs for these functions at the level of Proto *Kaili-Wolio*, and the observed split into *-ka* and *-(C)aka* with different applicative functions must be a later development.

The system that we observe in *Mori Bawah* (*Bungku-Tolaki*), which is analyzed as having four morphemes that mark pivot-neutral ACs, may result from not one but two splits in the form and meaning of earlier applicative markers. Instead of one AM *-ako* marking selection of beneficiary, instrument, and/or themeapplied phrases (as seen in other *Bungku-Tolaki* languages like *Moronene* and *Tolaki*), *Mori Bawah* has both *-ako* and *-Cako*, with only the latter showing an ex-crescent consonant. In *Mori Bawah* *-ako* marks beneficiary-, instrument-, and stimulus-selecting applicative constructions, while *-Cako* marks constructions in which the applied phrase is a theme that is carried, manipulated, or escorted (Mead 2005). These are almost certainly both derived from a single source morpheme, but in present-day *Mori Bawah*, both *-ako* and *-Cako* may co-

occur on the verb, as in *lua-rako-akune* ‘bring it outside for me’ (Mead 1998: 237), where *-akune* is the fused form of the AM *-ako* and a first person indexing suffix **(ko)na* (Mead 1998: 208, see also §6.5.3.2). This co-occurrence give credence to the interpretation that a split into two separate morphemes has occurred. Where most other West Nusantara languages have a single AM marking locative-/goal-selecting constructions, Mori Bawah again has two—the suffixes *-Ci* and *-Cari* (Mead 2005). The suffix *-Ci* marks goal-selecting applicatives with both nominal bases denoting inherent themes and verbal bases indicating self-motion, while also forming causative constructions and sometimes indicating intensive action or iterative aspect. The suffix *-Cari* marks the selection of a static location, source, or goal role as a core argument with a variety of verbal bases, e.g. ‘stay’, ‘eat’, ‘plant’, ‘place’. I am not aware of any other West Nusantara languages that show a split of this sort for the locative-/goal-selecting applicative; in the large majority of West Nusantara languages, all of the functions of Mori Bawah *-Ci* and *-Cari* are associated with a single AM *-i* or *-an*.

Finally, three affixes in Nias may be considered AMs. The suffix *-’ö* marks theme- and stimulus-selecting ACs with a small number of stems, including the verb meaning ‘to laugh’ and the bodily function verb meaning ‘to vomit’. This suffix is cognate with the AM elsewhere in Northwest-Sumatra Barrier Islands that selects a beneficiary, instrument or theme, among other roles as the applied phrase (Karo Batak *-ken*, Toba Batak *-hon* and Mentawai *-ake*). The suffix *-(C)i* in Nias marks locative-, goal- and stimulus-selecting constructions with a only few verbs as well. Lastly, Nias makes use of a verbal prefix *fa-* (< PMP **pa-* ‘causative’), that is used with verbs of throwing to indicate the selection of a additional verbal argument, either a theme or goal. Therefore, Nias looks to uphold the general pattern seen in languages with two distinct pivot-neutral applicative markers (albeit with very limited productivity), while also showing limited use of a causative prefix as an AM, as observed in Acehnese.

6.2.1.4 Cross-linguistic comparisons

In this section, for languages of West Nusantara with pivot-neutral ACs, I have described basic aspects of their applicative systems including the number of distinct forms of applicative marking and the distribution across these forms of common applicative functions (i.e. selection of various semantic roles as the applied phrase). On the whole, the predominant pattern observed is that of two distinct AMs marking pivot-neutral ACs; one for beneficiary-, theme-, and instrument-selecting constructions, and a second for locative- and goal-selecting constructions. In only Acehnese and the Sama-Bajaw languages is there no evidence that such a pattern previously existed. Aside from this, languages of the sample may vary from this pattern in a number of minor ways, including, (i) the loss of one AM form, as in Tolaki, (ii) the split of one AM into two derived forms, as with Mori Bawah *-Cako* and *ako*, and (iii) the innovation of unique third AM, as with Balantak, Sundanese, and Tukang Besi.

If we compare languages of the sample with pivot-neutral ACs to languages with applicatives in other languages families, this observed distribution of applicative functions across forms of AMs is unusual in a number of ways.

Cross-linguistically, Peterson (2007: 204) finds a notable relationship between instrumental applicatives, locative applicatives (any type), and circumstantial applicatives (e.g. selecting reason and purpose applied phrases). Having one of these constructions is strongly correlated with having the others, and this is “indicative of the trend for all of these constructions to be marked

by a single, generalized applicative marker” (204). In West Nusantara languages, while most languages with applicatives do have both an instrumental and a locative AC, these two functions are almost never marked with the same AM, neither in pivot-selecting nor pivot-neutral constructions.

Of the relationship between benefactive ACs and other types of applicatives, Peterson (2007) finds that “having a benefactive applicative construction correlates negatively with having an instrumental applicative construction” (204) and “if a language has a benefactive (or comitative) applicative construction, it is probably not especially likely that it will extend the marker to other applicative constructions types” (207). However, in West Nusantara languages with pivot-neutral ACs, there are 45 languages with benefactive ACs, and of these 39 mark theme-selecting and/or instrument-selecting constructions with the same AM.

Thus, the pattern whereby one AM form marks benefactive, instrumental, and theme-selecting ACs and another marks locative-selecting ACs is typologically unusual. This pattern however, is reconstructed for PAn and PMP in the form of the voice marking for CV and LV constructions. For these reasons, the most likely explanation for the observed distribution of applicative functions to AMs forms in West Nusantara languages with pivot-neutral ACs, is that this pattern is inherited.

In the following sections, I look more closely at the specific forms of pivot-neutral AMs observed in West Nusantara languages, the source morphology from which they derive, and the patterns of development that can be inferred from their observed distribution. I focus on AMs related to the two main patterns of functional distribution described in this section: morphemes that are associated with locative- and goal-selecting ACs (§6.2.2) and morphemes that are associated with beneficiary-, instrument-, and theme-selecting ACs (§6.2.3).

6.2.2 Forms and development of pivot-neutral locative/goal applicatives

As shown in §6.2.1, across West Nusantara languages with pivot-neutral ACs, we observe a stable, pervasive pattern of association in which one AM marks constructions in which a locative or goal role is selected as the applied phrase. In this section, I examine the forms of the morphemes that mark these construction and implications for our understanding of their historical development. I argue that this set of constructions derives from earlier Philippine-type LV constructions found in languages with pivot-selecting applicatives, and that the morphological marking for the present-day constructions show a pattern of shift and/or replacement of the original voice morphology for LV.

Table 6.8 below shows pivot-neutral AMs for which the primary applicative functions are to select locative and/or goal roles as the applied phrase. While the predominant pattern here is to have one AM that marks the selection of locative and goal roles (e.g. static location, source, path, endpoint of directed motion), a few other semantic roles are also observed as the applied phrase when the verb bears these affixes. The recipient role (e.g., give to s.o., send to s.o.) is attested with around one quarter of these morphemes in the survey (16 of 45), no doubt because of its semantic overlap with the goal role, as both express the endpoint of a theme. Similarly, about one quarter of locative-/goal-selecting AMs in the survey (18 of 45) are also attested in constructions where the applied phrase is an addressee in an act of communication (e.g. to speak to, to ask of). In addition, for more than half of locative/goal AMs in the survey (27 of 45), a stimulus role is also attested for the applied phrase (e.g. to be angry at, to laugh at, to cry for). In a relatively small number of

cases (11 of 45), a content role is attested for the applied phrase (e.g. to think about, to remember s.t.). For more than half of these morphemes (28 of 45), a causative function is also attested. For about one-third of these morphemes, a pluractional aspectual meaning (e.g. iterative, habitual) (17 of 45) is observed. Less commonly found is an intensive meaning (e.g. with great force, with greater thoroughness or completeness) (10 of 45) when the same affix appears on the verb.

In West Nusantara, the marker of locative-/goal-selecting pivot-neutral applicatives is most commonly a suffix with the form *-i*. This form occurs with an extremely high frequency in the survey data. In the sample a total of 44 languages distinguish locative-/goal-selecting ACs from beneficiary-/instrument-/theme-selecting ACs with distinct forms of morphological marking. Of these languages, 42 show a locative-/goal-selecting AM with the form of *-i* or similar. Similar here refers to suffixal forms that show an excrescent consonant at the morpheme boundary between the suffix and the stem, which is written as *-(C)i* or *-Ci*. This is frequently found in languages where some or all word-final consonants were lost historically, which is an areal feature in Sulawesi (Sneddon 1993), and also occurs in many languages of the Barrier Islands (Nothofer 1986), but has no material bearing on the properties of the morpheme or its origins.

There are two cases in the sample where the form of the locative-/goal-selecting suffix differs from the prevailing pattern. Balinese has the form *-in*, which is also found in the Ngenó-Ngené variety of Sasak (Austin 2001), but not the Menó-Mené, Meriaq-Meriku, and Ampenan varieties. The origin of this form is unclear. It is also found in Betawi and Colloquial Jakarta Indonesian, where it marks both locative-/goal-selecting applicatives, and beneficiary-/instrument-/theme-selecting applicatives. Ikranagara (1975) has suggested that *-in* in Betawi was borrowed from Balinese. Because both Balinese and Ngenó-Ngené divide applicative functions in the same manner across *-in* and the other form of AM *-ang*, and because Menó-Mené and Meriaq-Meriku have neither *-in* nor the ability to form locative-/goal-selecting applicatives, I conclude that *-in* and *-ang* were both present in Proto-BSS, and was lost in Menó-Mené, Meriaq-Meriku, and Ampenan Sasak. Ampenan Sasak additionally shows extension of *-an* to take on previous functions of *-in*, though this is observed with only a few bases. In Sundanese, the locative-/goal-selecting applicative is always marked with the verbal suffix *-an*.

Additionally, there are four languages of the sample in which the shape of the locative-/goal-selecting applicative co-varies with TAM category. Partial paradigms for these are presented in Table 6.9. As shown in the table, in Standard Javanese locative ACs are marked by *-i* in basic indicative verb forms, and the propositive mood in AV, but *-an* in non-volitional verb forms and the propositive mood in PV (Oglobin 2005). Tengger Javanese is very similar, though in PV the propositive mood has been lost as a distinct construction (Connors 2008). In Toba Batak, the locative-/goal-selecting AC is marked with the suffix *-an* in non-volitional verb forms, promissory mood, and completive participials bearing *-in-*, but *-i* elsewhere (Nababan 1981; van der Tuuk 1971 [1864-1867]). In Bobongko, *-an* is used to mark the locative-/goal-selecting construction in PV realis forms, with *-i* used in all other forms (Mead 2001). As discussed in §6.1, in PAn and PMP, LV was marked with *-an* in indicative mood, *-ay* in optative mood, and *-i* in imperative mood, and similar alternations are observed in present day Totoli and Southern Sama for LV. Thus, in these four languages, pivot-neutral AMs marking the selection of locative/goal roles show evidence of being derived from earlier LV morphology, with shift towards *-i* found especially in indicative forms, and remnants of earlier *-an* still found sporadically in certain TAM categories.

Furthermore, in terms of function, there is a good deal of overlap between LV and pivot-neutral locative-/goal-selecting applicatives. Both are frequently used with postural verb roots,

Table 6.8: Locative/goal-selecting applicative morphemes

Gen. Gp.	Language	AM	Role of Applied Phrase						Other Functions		
			LOC	GOAL	REC	STIM	CONT	ADDR	CAUS	PLUR	INTENS
NWS-BI	Alas Batak	-i	✓	✓		✓					
NWS-BI	Karo Batak	-i	✓	✓		✓			✓	✓	
NWS-BI	Toba Batak	-i/-an	✓	✓		✓			✓	✓	
NWS-BI	Gayo	-i	✓	✓		✓				✓	
NWS-BI	Nias	-(C)i	✓	✓		✓			✓		
ENG	Enggano	-(C)i	✓			✓					
NAS	Nasal	-i	✓	✓	✓	✓			✓	✓	✓
LAM	Lampung Api	-i	✓	✓		✓		✓	✓	✓	✓
T-T	Dampelas	-i	✓	✓	✓	✓					
T-T	Pendau	-i	✓	✓	✓	✓		✓			✓
T-T	Tajio	-i	✓	✓	✓	✓		✓	✓		
T-T	Totoli*	-i		✓	✓						
K-P	Ledo Kaili	-i	✓	✓		✓	✓		✓		
K-P	Moma	-i	✓	✓				✓	✓		
K-P	Uma	-i	✓	✓				✓	✓		
K-P	Behoa*	-i	✓	✓				✓			
S-B	Balantak	-i	✓	✓	✓	✓		✓	✓	✓	
S-B	Bobongko	-i/-an	✓	✓			✓				
B-T	Mori Bawah	-Ci		✓		✓			✓	✓	✓
B-T	Mori Bawak	-Cari	✓								
B-T	Moronene	-Ci		✓							
M-B	Muna	-Ci	✓	✓		✓				✓	
M-B	Busoa	-Ci	✓	✓		✓					
M-B	Tukang Besi	-(VC)i	✓	✓		✓	✓				✓
W-W	Wotu	-i	✓	✓				✓	✓		
W-W	Wolio	-i	✓	✓				✓	✓		
W-W	Laiyolo	-i	✓	✓					✓		
SSUL	Embaloh	-i		✓				✓			
SSUL	Bugis	-Ci	✓	✓		✓					
SSUL	Coastal Konjo	-i	✓	✓		✓				✓	✓
SSUL	Makasar	-i	✓	✓		✓				✓	
SSUL	Duri	-i	✓	✓			✓				
SSUL	Bambam	-i	✓	✓		✓	✓	✓			
SSUL	Seko Padang	-i	✓						✓		
MAL	S. Barisan Mal.	-i	✓	✓	✓	✓	✓	✓	✓	✓	✓
MAL	Jambi Malay	-i	✓			✓	✓		✓	✓	
MAL	Brunei	-i	✓	✓	✓				✓		
MAL	Kendayan	-i?	✓	✓	✓	✓		✓	✓		
MAL	Indonesian	-i	✓	✓	✓	✓	✓	✓	✓	✓	✓
SUN	Sundanese	-an	✓	✓	✓	✓	✓	✓	✓	✓	✓
JAV	Javanese	-i/-an	✓	✓	✓			✓	✓	✓	✓
JAV	Tengger Jav.	-i/-an	✓	✓	✓		✓		✓	✓	
JAV	Suriname Jav.	-i	✓	✓	✓				✓		
MAD	Madura	-e	✓	✓	✓			✓	✓		
BSS	Balinese	-in	✓	✓	✓	✓	✓	✓	✓		

* For these languages, the genetic classification listed is disputed.

Table 6.9: Morphological marking for locative applicatives and TAM in selected languages

PAn (LV)		Std. Javanese	Tengger Jav.	Toba Batak	Bobongko
Indic., neutral		Indicative	Indicative	Indicative	Unrealized
*-an	AV	<i>N- -i</i>	<i>N- -i</i>	<i>mang-/mar-/ma- -i</i>	<i>mon- -i</i>
	PV	<i>Ø-/di- -i</i>	<i>Ø-/di- -i</i>	<i>Ø-/di- -i</i>	<i>ku-/o- -i</i>
Indic., perf.		PV, archaic	PV, archaic	Compl. participial	Realized
*-in- -an	AV	(no form)	(no form)	<i>-um- -i</i>	<i>non- -i</i>
	PV	(no data)	(no data)	<i>ni- -an</i>	<i>-in- -an</i>
		Non-volitional	Non-volitional	Non-volitional	
	PV	<i>ka- -an</i>	<i>kə- -an</i>	<i>ha- -an</i>	
Imper./Neg.		Imper./Irr.	Imper./Irr.	Imperative	Imperative
*-i	AV	<i>-an-a</i>	(no data)	(no form)	<i>pon- -i</i>
	PV	<i>-an-a</i>	(no data)	<i>-i</i>	<i>-i</i>
Opt./Hort.		Propositive	Propositive	Promissory	
*-ay	AV	<i>N- -i</i>	<i>N- -i</i>	(no form)	
	PV	<i>-an-é</i>	(no form)	<i>-an</i>	

e.g. ‘sit’, ‘sleep’, with a location role as the applied phrase, e.g. ‘to sit on a bench’, ‘to sleep on a mat’. Pivot-neutral locative-/goal-selecting applicatives are also frequently found with verbs expressing directional or caused motion, e.g. ‘to throw rocks at s.t./s.o.’ and goal applied phrases. While the selection of goal roles as pivot is not commonly reported in the languages of the sample with Philippine-type voice systems, Chen (2017: 167–169) reports that goal roles (and some that I would label recipients) may be selected as the pivot in LV constructions in a number of Philippine-type languages, especially with ditransitive base verbs, but also with transitive and intransitive base verbs.

To date, it has been proposed that the pivot-neutral applicative suffix *-i*, either (i) derives from LV morphology (Sirk 1996: 194–195; Wolff 1996; Ross 2002: 52–55), or (ii) derives from the capture of reflex of the PAn and PMP general locative preposition *i, which is found widely in the Austronesian language family (Starosta, Pawley & Reid 1982: 155).

Based on the survey data, the first proposal is likely to be true in West Nusantara, but the second cannot be ruled out. In languages where the form of the locative-/goal-selecting applicative marker is *-an* in remnant TAM categories and *-i* in others (Toba Batak, Bobongko, Standard Javanese, Tengger Javanese), this alternation is strong evidence that these constructions derive from earlier LV constructions. In these cases, *-i* has been extended outside of imperative mood into other categories, and is almost always observed when the locative-selecting applicative co-occurs with AV prefixes, which was not possible in earlier stages of PAn/PMP and represents a realignment of Philippine-type voice morphology into a two-way symmetrical system with pivot-neutral applicatives. Even though these remnants of earlier LV *-an are observed in just a small number of languages, they are fairly broadly distributed geographically and genetically. These cases also show that it is possible that in languages with just *-i* marking or *-an* marking for pivot-neutral locative-/goal-selecting constructions, these AMs also reflect earlier LV morphology. Functionally, the pivot-neutral locative-/goal-selecting applicatives show many meanings that are commonly observed for LV, but perhaps more frequent and common usage with directed and caused motion events. All of these can be seen as consistent with the locative preposition *i, which has a broad range of uses including marking of spatial direction or path in addition to

location in space or time. If grammaticalization of a postverbal locative preposition *i* did occur, convergence of this with reflexes of the LV verbal suffix *-i and extended uses of this suffix in place of earlier *-an* may explain some differences in meaning between LV and the newer pivot-neutral constructions. But capture of the preposition *i alone, would not explain cases of remnant *-an* suffixes or the Sundanese locative-/goal-selecting AM *-an*.

6.2.3 Forms and development of pivot-neutral applicative morphemes selecting beneficiaries, instruments, and themes

In the large majority of West Nusantara languages, the pivot-neutral locative-/goal-selecting applicative contrasts with an AM that marks the selection of a beneficiary, instrument, and/or theme role as the applied phrase. The pattern of association between selection of these roles with one distinct form of applicative marking is again a pervasive and highly stable pattern in West Nusantara. In this section, I examine the forms of the morphemes that mark constructions of this type and implications for our understanding of their historical development. In parallel to the locative-/goal-selecting constructions described above, I argue that this set of constructions derives from earlier Philippine-type CV constructions (pivot-selecting applicatives) and that the morphological marking for the present-day constructions shows a complex pattern of shift and replacement of the original voice morphology for CV. Therefore, though the pivot-neutral AMs cannot be traced to a single reconstructed form of source morphology, the functional pattern seen in these constructions may be explained by inheritance, but does not appear to be consistent with a pattern of independent innovations.

Table 6.10 shows pivot-neutral AMs for which the primary applicative function is to select beneficiary, instrument and/or theme roles as the applied phrase. In the survey data, I identified 49 AMs of this type in 45 languages. The Enggano applicative suffix *-(C)a?a* is somewhat unusual because it may select either an instrument role or a goal role as the applied phrase, while in all other cases, locative and goal roles are excluded in ACs marked by these affixes. Otherwise, AMs that mark selection of beneficiaries, instruments, and/or themes, may be associated with selection a number of other semantic roles as the applied phrase, however these patterns of association are less frequently observed in the survey data. Applied phrases with the semantic role of recipient (e.g. ‘to give to s.o.’) are attested for a little less than one-third of such AMs (14 of 49), however it was not always possible to distinguish clearly recipients from beneficiaries in these languages. Meanwhile, the selection of applied phrases with the role of stimulus (e.g. ‘to be afraid of’, ‘to listen to’) and content (e.g. ‘to tell s.t.’, ‘to speak about s.t.’) are each found to be associated with about a third of these AMs (each in 16 of 49). Addressee roles (e.g. ‘to talk to s.o.’) may be selected as the applied phrase with a smaller proportion of this type of AMs (10 of 49), as are circumstantial roles (e.g. ‘to die because of’, ‘to go for a purpose’) (8 of 48). In terms of non-applicative functions, for about half of these AMs a causative function is attested (25 of 49). In only a few cases, do we find pluractional aspectual and intensive functions for these AMs, and these functions are not included in the table (see discussion in §6.4.3).⁴

⁴In Totoli, the benefactive/instrumental AM is *-an*, and iterative aspect is also indicated by means of a verbal suffix *-an* (Himmelmann & Riesberg 2013), however given that *-an* is also associated with iterative aspect when it serves as a locative-/goal-selecting AM, as with Sundanese *-an*, perhaps this should be considered a separate morpheme in Totoli.

Table 6.10: Beneficiary/instrument/theme-selecting applicative morphemes

Gen. Gp.	Language	AM	Role of Applied Phrase								Other
			BEN	INST	THM	REC	STIM	CONT	ADDR	CIRC	CAUS
NWS-BI	Alas Batak	-ken	✓	✓	✓	✓		✓			✓
NWS-BI	Karo Batak	-ken		✓	✓		✓				✓
NWS-BI	Toba Batak	-hon	✓	✓	✓		✓			✓	✓
NWS-BI	Gayo	-(n)en			✓		✓				✓
NWS-BI	Nias	'-ö			✓		✓	✓	✓		✓
ENG	Enggano	-(C)aʔa		✓							✓
NAS	Nasal	-kun	✓	✓	✓	✓	✓				✓
LAM	Lampung Api	-ko	✓		✓		✓	✓			✓
T-T	Dampelas	-a'o	✓		✓	✓					✓
T-T	Pendau	-a'	✓	✓	✓	✓					✓
T-T	Tajio	-ao	✓	✓							✓
T-T	Totoli	-an	✓	✓							✓
K-P	Ledo Kaili	-aka		✓	✓						✓
K-P	Ledo Kaili	-ka	✓								
K-P	Moma	-aka					✓				
K-P	Moma	-ka	✓	✓		✓					✓
K-P	Uma	-ki	✓	✓		✓		✓	✓	✓	
K-P	Behoa*	-á	✓					✓	✓		✓
S-B	Balantak	-kon	✓	✓		✓	✓	✓	✓	✓	✓
B-T	Bobongko	-akon	✓								
B-T	Mori Bawah	-ako	✓	✓						✓	
B-T	Mori Bawah	-Cako			✓						✓
B-T	Moronene	-ako	✓	✓		✓		✓	✓		
B-T	Tolaki	-Cako	✓	✓		✓		✓	✓	✓	
M-B	Muna	-ghoo	✓	✓		✓				✓	✓
M-B	Busoa	-ho	✓	✓	✓					✓	
M-B	Tukang Besi	-ako	✓	✓	✓			✓	✓		
W-W	Wotu	-a	✓					✓			
W-W	Laiyolo	-aka			✓		✓				
W-W	Laiyolo	-ka	✓			✓					
W-W	Wolio	-aka	✓	✓	✓	✓	✓	✓	✓		
SSUL	Embaloh	-ang	✓		✓	✓	✓				
SSUL	Bugis	-Ceng	✓	✓		✓				✓	✓
SSUL	Coastal Konjo	-ang	✓								
SSUL	Makassar	-ang	✓	✓	✓						
SSUL	Duri	-an	✓	✓	✓			✓			
SSUL	Bambam	-am	✓	✓				✓			
SSUL	Seko Padang	-ing		✓	✓						
MAL	S. Barisan Mal.	-ka	✓	✓	✓		✓	✓			✓
MAL	Jambi Malay	-kan/-an	✓		✓		✓	✓			✓
MAL	Brunei	-kan	✓		✓			✓			✓
MAL	Kendayan	-an	✓		✓			✓			
MAL	Indonesian	-kan	✓	✓	✓		✓	✓			✓
MAD	Madurese	-agi	✓	✓	✓			✓			✓
SUN	Sundanese	-keun	✓	✓	✓		✓	✓			✓
JAV	Javanese	-aké	✓	✓	✓			✓			✓
JAV	Tengger Jav.	-en/-na	✓		✓						✓
JAV	Suriname Jav.	-aké	✓	✓	✓						✓
BSS	Balinese	-ang	✓	✓	✓		✓				✓

* For these languages, the genetic classification listed is disputed.

In West Nusantara, the marker of beneficiary-/instrument-/theme-selecting pivot-neutral applicatives is almost exclusively suffixal. The shapes of these AMs may be divided into two main types, (i) those which show a vowel + nasal sequence, which I will call -AN suffixes, and (ii) those which show an initial *-k* or *-ak* sequence, e.g. *-kan* or *-aka*, which I will call -K and -AK suffixes, respectively, following Sirk (1996). However, as will be discussed below, the shapes of the -K and -AK suffixes show some curious variance. In addition, some Malayic languages show an alternation between an -AN suffix and a -K suffix.

6.2.3.1 Benefactive/instrumental applicative morphemes of the -AN type

Languages that mark the selection a beneficiary, instrument, and/or theme role as the applied phrase with an -AN suffix are in the minority in the survey data. Nonetheless, this type of form is still widely distributed across West Nusantara languages, even more so, if we include discussion of AMs with with the form *-an* in Sama-Bajau, which have become generalized to also mark selection of locative/goal roles in some members of the subgroup (e.g. West Coast Bajau). These forms are shown in Table 6.11. Due to the stable pattern of association between selection of the beneficiary, instrument, and theme semantic roles as the applied phrase with a single form, and the shapes of these forms, for the most part these morphemes are clearly derived from earlier PAn **-an*, that marked CV in imperative and negative clauses.

Table 6.11: Pivot-neutral applicative morphemes of the -AN type

Larger Subgroup	Source	Branch	Forms
South Sulawesi	<i>*-an</i>	Northern Makassaric	Duri <i>-an</i> ; Bambam <i>-am</i> Coastal Konjo, Makassar <i>-ang</i>
		Tamanic-Bugis	Embaloh <i>-ang</i> ; Bugis <i>-Cang</i>
		Seko	Seko Padang <i>-ing</i> [†]
		Badaic	Behoa* <i>-á</i> [-a]
Tomini-Tolitoli	<i>*-an</i>	Tolitoli	Totoli* <i>-an</i>
Bali-Sasak-Sumbawa	<i>*-an</i>	Bali	Balinese <i>-ang</i>
		Sasak-Sumbawa	Ngené-Ngené <i>-ang</i> , Ampenan <i>-an</i> , Menó-Mené <i>-angk/-at ~ -ant/-am/-an</i>
Malayic	<i>*-an</i>	Kendayan Malay	Kendayan <i>-AN</i> [-(?)a(t)n] Jambi, Minangkabau <i>-an</i> (~ <i>-kan</i>)
Sama-Bajaw	<i>*-an</i>	Borneo Coast Bajau Inner Sulu Sama Western Sulu Sama	West Coast Bajau, Indonesian Bajau <i>-an</i> Cent. Sama <i>-an</i> ; (Southern Sama <i>-an</i> cv) Pangutaran Sama <i>-an</i>
Javanese	<i>*-(?)ən</i> (?)	—	Std. Jav. <i>-na</i> , <i>-né</i> (~ <i>-aké/-akən</i>); Tengger <i>-na ~ -ən</i>
Sumatran	<i>*-ən</i> (?)	NW Sumatra	Gayo <i>-(n)ən</i>

* For these languages, the genetic classification listed is disputed. † These forms show unexpected sound correspondences.

Proto South Sulawesi **-an* is reconstructed as the beneficiary-selecting marker by Mills (1975: 192–196) and almost all South Sulawesi languages show regular reflexes of this form.⁵ As dis-

⁵Seko Padang *-ing* does not show regular correspondences consistent with **-ang* (Mills 1975). However, the

cussed above, the form *-an should also be reconstructed to Proto Bali-Sasak-Sumbawa. In Meno-Mené Sasak, this suffix has fused with enclitic person indexes to form portmanteau suffixes denoting selection of a beneficiary applied argument and person indexing for this argument, e.g. *-angk* < *-an + =*k* 1SG, *-am* < *an + =*m* 2SG (Austin 2001). Proto-Malayic *-an is reconstructed as an applicative suffix by Adelaar (2006: 78–79). It is seen in Kendayan (Salako) and most Minangkabau dialects (with others showing *-kan*), while in some Jambi varieties there is unpredictable variation between *-an* and *-kan* (Adelaar 1984: 406; Yanti 2010: 598). Generally the *-an* suffix is thought to be older than *-kan*, which cannot be reconstructed for Proto-Malayic and appears to have replaced the AM *-an in many Malayic languages. In Sama-Bajaw, *-an is reconstructed as the indicative CV suffix by Pallesen (1985) and has been reanalyzed as a generalized applicative marker (see §5.8.4.2 above).

Though they are included in Table 6.11, Standard Javanese *-na*, *-né* and Tengger *-na*, which mark beneficiary-/instrument-/theme-selecting ACs in certain non-indicative TAM categories, may not reflect earlier CV *-an directly. The paradigm for these forms in the non-formal register are shown in Table 6.12 alongside that for PAn PV and CV. The Javanese morphemes are analyzed as a shortened suffixal form *-n* + imperative/irrealis *-a* or propositive *-é* (Oglobin 2005: 600). Adelaar (2011) traces this shortened suffix to a Pre-Javanese transitivizing suffix *-(?)ən*, which I note could have itself been a replacement for earlier CV *-an in imperative/negative **an-i* and optative/hortative **an-ay*. Gayo also has a verbal suffix *-(n)en* that shows similarities with the proposed transitive suffix **-(?)ən* in Javanese, though Gayo *-(n)en* is much more productive as a causative marker than an applicative marker and its origin is even less certain.

Table 6.12: Benefactive/instrumental applicatives and TAM in Javanese

PAn (PV)	PAn (CV)		Std. Javanese	Tengger Jav.
Indic., neutral	Indic., neutral		Indicative	Indicative
*-ən	*Si-/Sa-	AV	<i>N- -aké</i>	<i>N- -ən</i>
		PV	<i>Ø-/di- -aké</i>	<i>Ø-/di- -ən</i>
Indic., perf.	Indic., perf.		PV, archaic	PV, archaic
*-in- -Ø	*-in- -an	PV	<i>-in- -aké</i>	(no data)
			Non-volitional	Non-volitional
		PV	<i>ka- -aké</i>	<i>kə- -Ø</i>
Imper./Neg.	Imper./Neg.		Imper./Irr.	Imper./Irr.
STEM	*-an-i	AV	<i>N- -n-a</i>	<i>N- -ən</i>
		PV	<i>-n-a</i>	<i>-na</i>
Opt./Hort.	Opt./Hort.		Propositive	Propositive
*-a	*-an-ay	AV	<i>N- -aké</i>	<i>N- -na</i>
		PV	<i>-n-é</i>	(no form)

6.2.3.2 Benefactive/instrumental applicative morphemes of the -K and -AK types

Languages that mark the selection a beneficiary, instrument, and/or theme role as the applied phrase with a -K or -AK suffix are quite numerous in the survey data. Such forms show a very

Badaic languages of Central Sulawesi probably form an exclusive subgroup under South Sulawesi together with Seko, rather than belonging to Kaili-Pamona (Zobel 2020). In Behoa (Badaic), the beneficiary-selecting AM is *-á*, which is consistent with **-ang*, as the backed-vowel *á* (IPA [ɑ]) reflects PMP **a* followed by a deleted final back consonant or nasal (Truong in prep).

Table 6.13: Applicative morphemes of the -K and -AK types

Larger Subgroup	Source	Branch	Forms
Sumatran	*-(a)kən	NW Sumatra	Karo <i>-kən</i> ; Alas <i>-akən</i> ; Toba Batak <i>-hon</i>
		Barrier Islands	Mentawai <i>-ake</i> ; Nias -'ö [-ʔʏ]
		Nasal*	Nasal <i>-kun</i>
		Lampungic*	Lampung Api <i>-ko</i> [†]
		Enggano*	Enggano <i>-aʔa</i> [†]
Malayic	(-kan)	Malay	Brunei, Std. Indonesian <i>kan</i> ; Jambi, Minangkabau <i>-kan</i> (~ <i>-an</i>); S. Barisan <i>-ka</i>
Javanese	(akən)	—	Old Jav., Std. Jav. <i>-akən</i> ; Std. Jav. <i>-aké</i> [†] (~ <i>-na</i> , <i>-ne</i>); Suriname Jav. <i>-ké</i> [†]
Madura		—	Madurese <i>-agi</i> [†] [-ak ^h ɛ], Kangean <i>-aghən</i> [†] [-ak ^h ən], Mad. Bawean <i>-akən</i>
Sundanese	*-kən	—	Sundanese, Badui <i>-keun</i> [-kɤn]
Tomini	*ako	—	Dampelas <i>-a'o</i> , Pendau <i>-a'</i> , Tajio <i>-ao</i>
Eastern Celebic	*ako(n)	Saluan-Banggai	Balantak <i>-ako</i> [†] , Bobongko <i>-akon</i>
		Bungku-Tolaki	Mori Bawah, Moronene <i>-ako</i> ; Mori Bawah, Tolaki <i>-(C)ako</i>
		Muna-Buton	Muna <i>-ghoo</i> [†] [-ʋoo]; Busoa <i>-ho</i> [†] ; Tukang Besi <i>-ako</i> ; Cia-Cia <i>-(ʔ)aso</i>
Kaili-Wolio	*aka	Kaili-Pamona	Kaili-Ledo, Moma <i>-(C)aka</i> , <i>-ka</i> ; Uma <i>-ki</i> (< <i>-ka</i> + <i>-i</i>)
		Wotu-Wolio	Laiyolo, Wolio <i>-aka</i> ; Laiyolo <i>-ka</i> ; Wotu <i>-a</i>

[†] These forms show unexpected sound correspondences/shapes.

broad distribution in West Nusantara, being found among the non-Malayic languages of Sumatra and the Barrier Islands, the Celebic languages of Sulawesi, Sundanese, Javanese, Madurese, and many Malayic languages.⁶ Table 6.13 shows applicative markers of this type in languages of West Nusantara. Unlike the -AN suffixes, reconstructing and/or identifying earlier source morphology for these sets of suffixes is problematic. While there is quite a bit of apparent similarity between the forms listed, when the data are examined from a historical-comparative perspective, a number of puzzling trends appear.

First, in Malayic and Javanese, two subgroups with well-established reconstructions and written records dating back several centuries, none of the suffixes of this type can be reconstructed to the proto-language. Proto-Malayic did not have a suffix *-kan* (Adelaar 1984), though the preposition *akan* may date back to Old Malay. The suffix *-kan* also alternates with *-an* in some Malay languages including Minangkabau, where the variation is dialectal (Adelaar 1984: 406), and Jambi Malay, where the variation is both lexical and dialectal (Yanti 2010: 598–600). As for Javanese, the high register or *krama* form of the beneficiary-instrument-theme-selecting AM is *-akən*. However, *-akən* cannot be reconstructed to Proto-Javanese, even though there is evidence for a morpheme *akən* in Old Javanese (Adelaar 2011). It is also not clear why the corresponding neutral

⁶-K and -AK type suffixes do not appear to be well-attested in Borneo outside of Malay languages that are more recent arrivals, e.g. Banjar *-akan* and Brunei Malay *-kan*. However, this is difficult to say definitively due to the general sparsity of languages with pivot-neutral applicatives in Borneo.

register form of *-akən* is *-aké* (or *-ké*) as this does not follow known patterns for transformation of *ngoko* forms into *krama* forms.

Second, the appearance of final *-n* in the *-K* and *-AK* suffixes is inconsistent. In the non-Malayic languages of Sumatra and the Barrier Islands proposed to belong to an exclusive subgroup Sumatran (see McDonnell and Billings, forthcoming), final *-n* is present for the Batak languages and Nasal and is lost as expected in Mentawai and Nias, but is lost unexpectedly in Enggano and Lampung. The Enggano form also shows an unexpected second vowel (o is the expected reflex of *ə). Meanwhile, in Malayic languages, the form of the suffix is *-ka* in Besemah (South Barisan Malay) (McDonnell 2016), but *-kan*, *-an*, and *-a* are also attested in south Sumatran Malay varieties, as well as many phonetic variants of these (Anderbeck & McDowell 2020: 91–101). The suffix *-ka* is also attested in Iban, where *-ka* has been proposed to derive from recent capture of the allative preposition *ka* (Adelaar 1984: 409). In most languages of Sulawesi, final consonants have been lost or reduced (Sneddon 1993). In Saluan-Banggai, which does retain final nasals, however, we see final *-n* in this suffix in Bobongko but not Balantak. There is also variance in the occurrence of the initial *-a* of the suffix, but this is less problematic, and may result from phonological processes resulting in reduction or deletion of the suffix vowel at the morpheme boundary with the stem.

Third, while most forms presented in Table 6.13 show a first consonant that is consistent with earlier *k, a few have problematic correspondences for this consonant. Muna *-ghoo* has initial *gh* (a phonemic uvular trill /R/), which is a regular reflex of PMP *q but not *k (van den Berg 1991). Likewise, the corresponding Busoa applicative suffix *-ho* is problematic, as Busoa /h/ reflects PMP *q or *s but not *k. Madurese *-aghi* /-ak^hε/ is not clearly expected from **-akən*, as **-n* is retained in Madurese, and *g is the source of *gh* according to Nothofer (1975).

Fourth, while some diachronic accounts of these suffixes have treated them as captured prepositions, as with Malay *akan* and Javanese *akən* as mentioned above, prepositions of the requisite shapes are not widely attested. Capture of allative *ka* would not explain a large number of forms showing initial *-a*. Aside from these cases, and *Tukang Besi ako*, there are very few languages in which a preposition with a consistent form **akən* or **akan* is attested, nor are such prepositions found in other Austronesian languages of West Nusantara, the Philippines, or Taiwan.

Given these observations, there is insufficient evidence to demonstrate that suffixes of this type are derived from a common inherited morphological source. Crucially, though, together with the reflexes of the *-an* type, we do see a stable pattern of association between constructional form and meaning. That is, the pattern whereby the selection of beneficiaries, instruments, and themes—alongside circumstantial roles like reason and purpose, and stimuli of experiential verbs—is associated with one morphological marker is extremely stable among languages with pivot-neutral applicatives in West Nusantara. This is observed to be a stable pattern of constructional association as far back as PAn in the form of pivot-selecting CV constructions, and thus dates back far before the breakup of PMP (and observed splits of CV into separate IV and BV constructions in the Philippines and Northeast Borneo). This pattern is also cross-linguistically unusual, as discussed in §6.2.1.4 above, and is unlikely to result from independent innovations in various genetic groupings below PMP. The observed distribution also is not likely to be explained by a simple pattern of borrowing, given that the shapes of the AMs are themselves quite varied, and their broad geographic distribution does not support the inference that these constructions were primarily spread through language contact.

Therefore, I argue that the most likely explanation for the pivot-neutral ACs showing selection

of beneficiary, theme, and/or instrument applied phrases in West Nusantara is that they generally developed from an inherited CV construction. In languages with -AN type AMs marking these constructions, the source morphology is PAn CV *-an (imperative/negative mood), but in languages showing the -K and -AK type suffixes, earlier CV morphology has been replaced. This replacement must have taken place after the breakup of PMP, and sometimes much later. Such replacement would have had to occur multiple times, that is, in various branches under PMP. While the source morphology for the replacement AMs is not always clear, where historical evidence is available, as for Malayic and Javanese, the picture is complex, and sometimes involves only partial replacement and/or successive replacements at different diachronic stages. Further historical-comparative research is necessary to more fully ascertain the stages in the diachronic development of the forms of these AMs seen in other languages of West Nusantara.

6.3 Irregular morphological marking of ACs

There are seven languages of the sample for which some irregular pattern of morphological marking is attested on pivot-neutral ACs, aside from paradigmatic allomorphy conditioned by a TAM category. These cases are described in brief below.

- In Bobongko, the prefix *poN-* is required in PV but not AV with both locative-/goal-selecting ACs and beneficiary-selecting ACs (see §5.9.4.3).
- In Pendau, stem-former prefixes appear irregularly on locative-selecting ACs marked with *-i* and instrument-selecting ACs marked with *-a'*. Such constructions are only possible in PV (see §5.9.4.4).
- In Tajio, stem-former prefixes appear on locative-selecting ACs marked with *-i* (Mayani 2013: 181–184). These constructions appear in both AV and PV with intransitive bases, but with transitive bases, curiously, neither the stem-former nor the suffix are required to take a location as a core argument in AV.
- In Uma, the suffix *-i* marks locative/goal-selecting constructions. Martens (1988b: 194–195) reports that the prefix *po-* or *poN-* is used with the suffix *-i* in constructions with the meaning ‘to use as a place for X’. Martens analyzes the prefixes as transitivity markers.
- In Bambam, the prefix *paC-* is required on intransitive bases in beneficiary-selecting ACs marked with *-am* and locative/goal-selecting ACs marked with *-i*. Campbell (1989: 101–104) analyzes this prefix as a transitivity marker.
- In Makasar, the prefix *pa-* is found on a number of theme- and instrument-selecting ACs with *-ang* (Jukes 2020: 312–117). With certain motion verbs, *pa-* appears on the verb in locative-selecting ACs marked with *-i* (Jukes 2020: 309–311). In other constructions, *pa-* is a causative marker in Makasar.
- In Merina Malagasy, in certain CV constructions, the verb must be marked with both a prefix and the suffix *-an* (Pearson 2001: 40). The shape of the prefix varies and is lexically determined (see also §6.1.4 above).

I note that the shapes of prefixes which are observed to occur irregularly on ACs in languages of the sample shows close resemblance to prefixal components found regularly in LV and CV voice morphology in other languages of West Nusantara, see §6.1.3 and §6.1.4 above. This suggests that examples of the sort presented in this section with pivot-neutral ACs might occur as remnants of earlier LV and CV pivot-selecting constructions. It is also possible that the verbal prefixes seen on pivot-neutral ACs may more directly reflect earlier PMP verbal prefixes *paR- ‘distributive’, *paN- ‘durative’, and sometimes *pa- ‘causative’. I also note that the meanings of these prefixes resemble many of the non-applicative functions or constructional meanings found for pivot-neutral applicatives in languages of West Nusantara, which are described in the following section.

6.4 Polyfunctionality of West Nusantara applicative morphemes

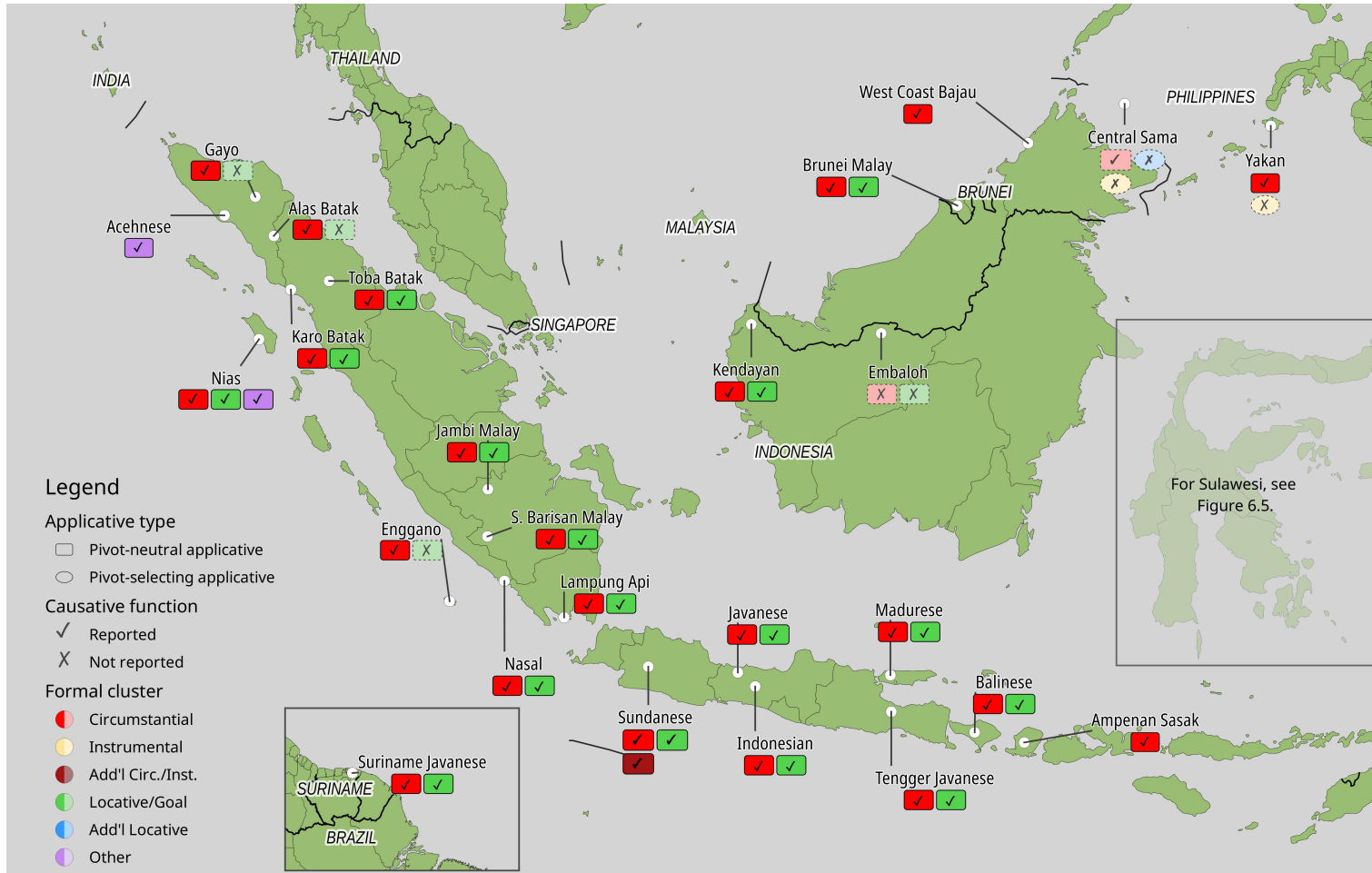
In West Nusantara languages, AMs that mark the verb in pivot-neutral ACs have many non-applicative functions, and as such are highly polyfunctional. The occurrence of some of these non-applicative functions was mentioned briefly in previous sections, and they include forming causative constructions (sometimes in combination with other morphological marking), indicating aspectual meanings, indicating greater intensity, marking highly lexicalized changes in semantic meaning, and forming comparative constructions. These are discussed in turn below in greater detail.

6.4.1 Causative functions of applicative morphemes

In a causative construction, an A argument with the role of causer or instigator is introduced to the argument structure of clause (Zúñiga & Kittilä 2019: 15–17). Causative functions are highly prevalent for pivot-neutral AMs found in the languages of the sample. For languages with pivot-neutral applicatives, in the large majority (40 of 50 languages), at least one AM also is reported to mark causative constructions.

As shown in Figure 6.4, in most of West Nusantara, the norm is for both the AM selecting beneficiary, instrument, and/or theme applied phrases and the AM selecting locative and goal applied phrases (if present) to have causative functions. In Sulawesi, however, as shown in Figure 6.5, causative functions appear to be less prevalent, being reported with at least one AM in about two-thirds of languages (17 of 25) and less than half of the total number of pivot-neutral AMs (23 of 56). In addition, when they are attested, causative functions of AMs in Sulawesi may be of low productivity. For example, in *Tukang Besi*, the applicative suffix *-(VC)i* is reported with a causative meaning in a single example, *ma-i* ‘to bring s.t.’ cf. *mai* ‘to come’ (Donohue 1999: 243). In the sample, there are only ten languages with pivot-neutral applicatives for which causative functions are not reported for any attested AM; eight of these are spoken in Sulawesi (Bambam, Bobongko, Busoa, Duri, Coastal Konjo, Moronene, Muna, and Tolaki), and one additional language, Embaloh, belongs to the South Sulawesi subgroup but is spoken in Borneo. The final language of this type is Central Sama. These numbers do not include additional languages in which an AM-marked construction is causative only in irregular cases when a causative prefix is also marked on the verb (see §6.4.2 below).

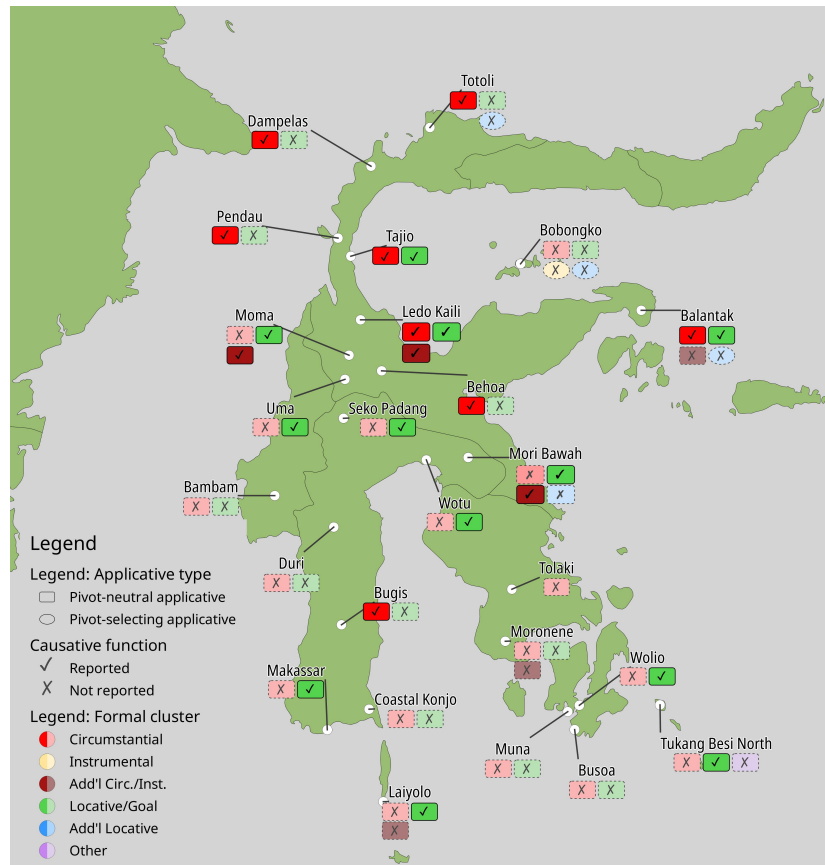
Figure 6.4: Causative functions of applicative morphemes in West Nusantara



Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

In languages of West Nusantara with pivot-selecting applicatives, the morphological marking associated with these constructions (Philippine-type LV and CV constructions) are not generally observed to form causative constructions. In these languages an instigating causer may typically be introduced to the argument structure of verb by means of one of more distinct causative morphemes. The most common type of these is a prefix that appears to derive from PMP *pa-: *pa-*, *po-* or *p-* in Sabahan languages, *pa-* or *pe-* in Sama-Bajau languages, and *(m)amp-* in Merina Malagasy, which appears to be composed of the AV prefix *maN-* plus *p-*. Such causative prefixes generally may co-occur across multiple voice constructions in the language, including some pivot-selecting constructions (see discussion in §6.5.1).

Figure 6.5: Causative functions of applicative morphemes in Sulawesi



Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

In general, designated causative prefixes derived from PMP *pa- are also extremely common in Sulawesi, being found in all 25 languages of the sample spoken in Sulawesi, all of which have pivot-neutral applicatives. This may account for the lower prevalence and less productive use of pivot-neutral AMs as causative markers in many of these languages. Outside of Sulawesi, of the remaining languages of West Nusantara with pivot neutral applicatives, only about half (13 of 25) show a causative prefix derived from *pa-.

While I attempted to assess the productivity of causative prefixes based on source material for the survey, it is difficult to make reliable comparisons across languages of the sample here because of large disparities in the level of descriptive detail in source material and available quantity of

examples, textual material, and lexical resources. Still, in a good number of languages of the sample with little to no productive use of causative prefixes, it is the pivot-neutral applicative affixes that appear to fill this gap through extremely productive and frequent use as morphological causative markers, as in Lampung Api, Sundanese, Javanese, Balinese, Jambi Malay, South Barisan Malay, and Nasal, among others.

6.4.2 Use in irregular causative constructions

In the previous section, I discussed the many languages of the sample for which morphological marking for causative constructions may consist solely of affixation of the verb with an AM. However, in some of these languages, we observe irregular occurrence of a pivot-neutral AM on bases bearing a distinct causative prefix. In such cases, both affixes must be present for the causative meaning to be exhibited with certain lexical bases. This is also observed in some languages in which the relevant pivot-neutral AMs do not show causative functions on their own.

This type of irregular occurrence of AMs on verbs in causative constructions is seen in about a third of the languages of the sample with pivot-neutral applicatives (16 of 50), including six languages in which the relevant AM does not independently function as a causative marker.

For example, in Muna, the suffix *-Ci* does not generally function as a causative marker; morphological causative constructions are marked with the prefix *fo-* as in the example in (131). But with a handful of verbs, in causative constructions, the verb is necessarily marked with both *fo-* and the locative/goal-selecting applicative suffix *-Ci*, as in (132).⁷

(131) Muna, causative construction marked with *fo-*

a. *ao-ndawu*
1S.RLS-fall
'I fall'

b. *ae-fo-ndawu piso*
1S.RLS-CAUS-fall knife
'I drop a knife'

(van den Berg 2013: 64)

(132) Muna, causative construction marked with *fo-*

a. *A-foroghu oe karindi.*
1S.RLS-drink water cold
'I am drinking cold water.'

(van den Berg 2013: 64)

b. *Ai-no no-fo-foroghu-ti-ane oe-no*
younger.sibling-3S.POSS 3S.RLS-CAUS-drink-LOC.APPL-BEN.APPL:3S water-3SG.POSS
kalembungo.
young.coconut

'He gave his younger sister coconut juice to drink'

(“*foroghu*” van den Berg & Marafad 2016)

⁷The verb in (132b) shows a special type of AM-marking with an indirect object suffix, see discussion in §6.5.3.

Likewise in Ledo Kaili, the verbal prefix *po-* is a morphological causative marker. For the verb *nang-inda* ‘to borrow s.t. (AV)’, however, the causative verb stem is *nom-popa-inda-ka* ‘to loan or lend out s.t.’ which is formed with both *po-* (in its variant form *popa-*) and the theme-selecting AM *-ka*.

Additional examples from Karo Batak include *pe-sirang-ken* ‘to cause to separate’ cf. *sirang* ‘to separate’ and *pe-pinem-ken* ‘to make drink, suckle, breastfeed’ cf. *pinem* ‘to drink’ (Woollams 1996: 52). Note that Karo Batak *-ken* forms causative verb stems on its own as in *keriken* ‘to deplete, use up’ cf. *keri* ‘depleted’, *kabangken* ‘to make fly, blow away’ cf. *kabang* ‘to fly’, and this is found with a much broader set of bases (Woollams 1996: 57–58).

The required appearance of the causative morpheme together with the AM in certain causative constructions suggests one possible pathway by which applicative suffixes in West Nusantara may have developed causative functions. Perhaps previously both applicative suffixes and causative prefixes were marked on certain lexical verbs, alongside verbs with only one or the other. As the causative prefixes subsequently became less productive the stems with only the applicative suffixes became preferred, especially in cases where the verb expressed both the participation of a instigating causer, and a constructional meaning consistent with an AC, such as motion of a theme, and use of an instrument to affect a change of state, among others.

6.4.3 Aspectual and intensive meanings of applicative morphemes

In many languages of West Nusantara, aspectual meanings are associated with AMs. These aspectual meanings may be grouped under the cover term *pluractional aspect* and include iterative aspect (i.e. repeated action), habitual aspect (i.e. characteristic action), and pluractional action (i.e. an action performed by multiple actors and/or on multiple undergoers) (see Truong & McDonnell 2022; Wood 2007; Mattioli 2019). Durative aspectual meanings are also sometimes observed, i.e. an event or state extends over some period of time. For examples of such aspectual meanings in Sundanese, see §2.10.2.

In addition to aspectual meanings, AMs in West Nusantara languages may be associated with an intensive meaning, such that greater intensity, thoroughness, exertion, or force is indicated for an action, event, or state when the AM marks the verb. An example is given from Coastal Konjo in (133) below. In source material for the survey, aspectual and intensive meanings were not found for pivot-selecting AMs; for this reason, only languages with pivot-neutral applicatives are included in the discussion that follows (50 total languages in the sample).

(133) Coastal Konjo, Intensive meaning of *-i*

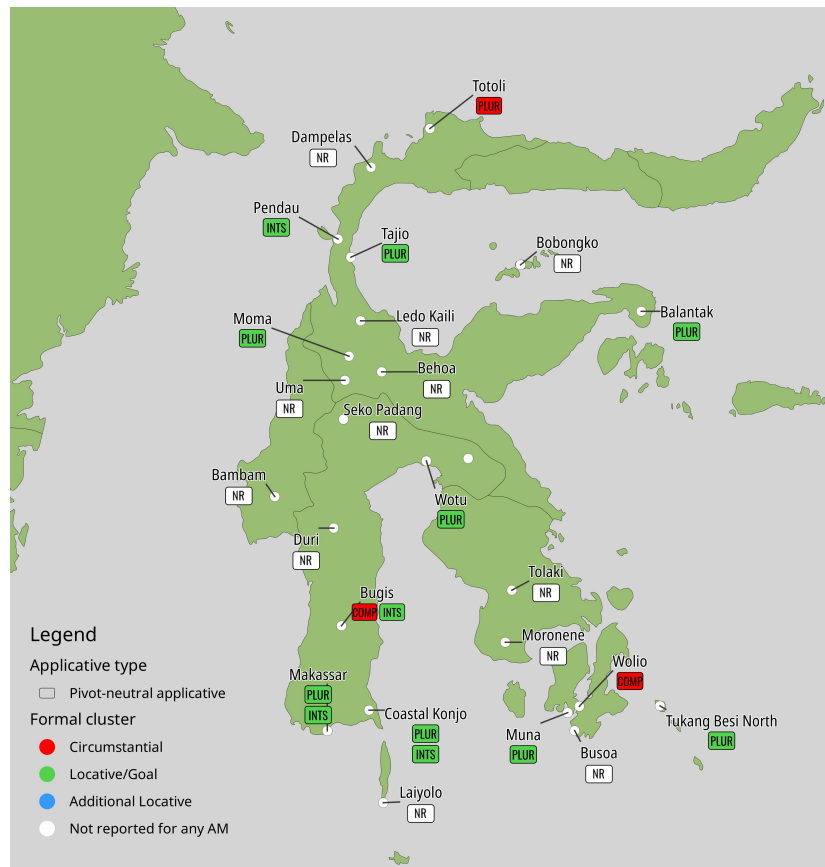
- a. *Kupakaháji'i.*
ku-paka-haji'-i
 1.ERG-CAUS-good-3.ABS
 ‘I fix it (make it good).’
- b. *Kupakahajíkii.*
ku-paka-haji'-i-i
 1.ERG-CAUS-good-INTENS-3.ABS
 ‘I repair it thoroughly.’

(B. Friberg 1991: 117)

As shown in Figures 6.7 and 6.6, pluractional aspectual meanings are quite common for pivot-neutral AMs in West Nusantara, being found in a fairly large portion of the languages of the sample (20 of 50). These are broadly distributed geographically, although not universally reported. Pluractional aspectual meanings are most commonly associated with locative/goal-selecting AMs, being found with such forms in 17 languages (out of 50 total) in the sample. In just a handful of languages (3 of 50; Yakan, Totoli, and Mori Bawah), pluractional aspect is reported as a meaning of a benefactive/instrumental or partially generalized benefactive AM.

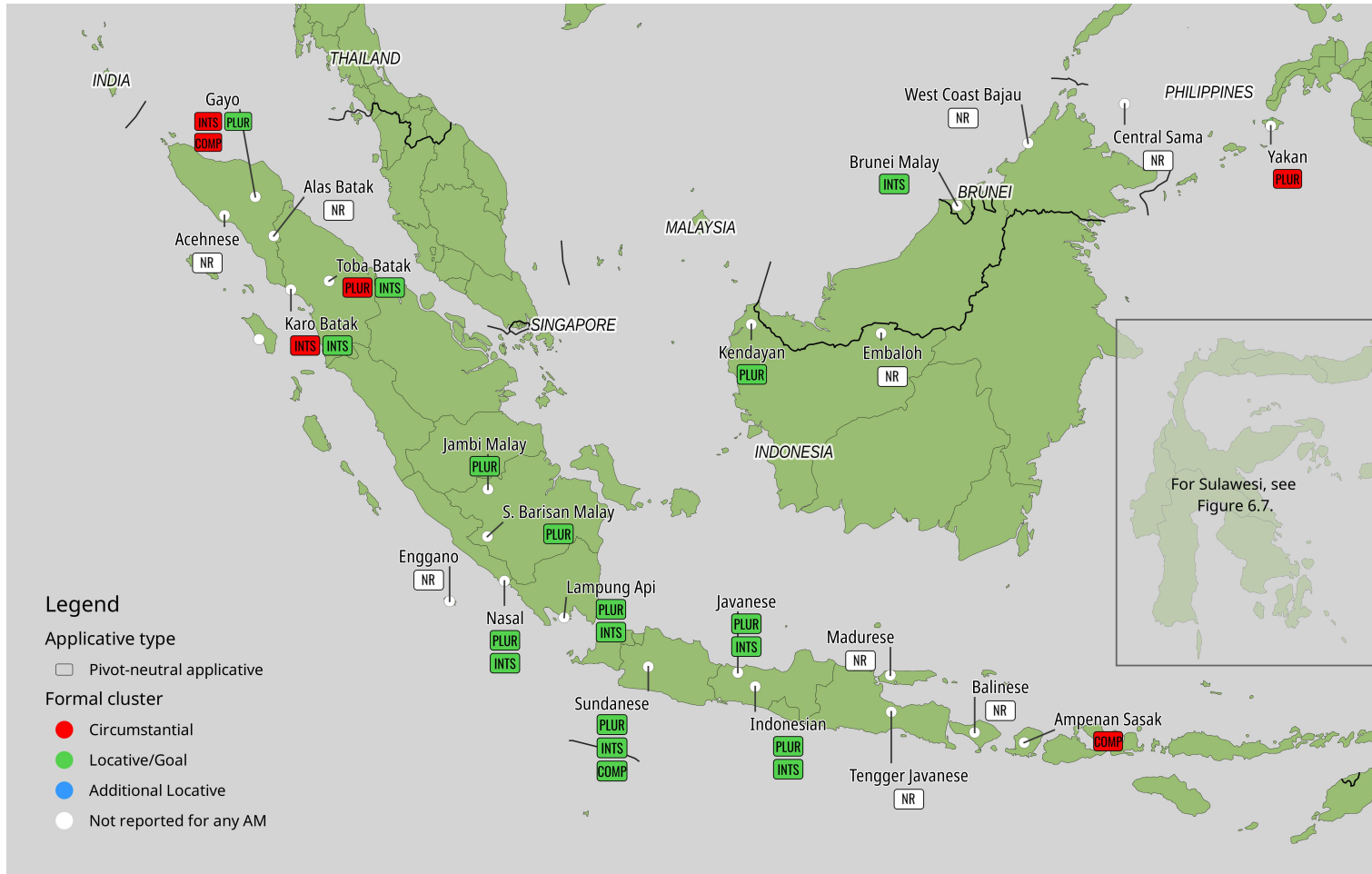
Intensive meanings are less commonly reported for AMs in the survey data than are aspectual meanings, being found in about one quarter of the languages of the sample (13 of 50, or 26%). Again, intensive meanings are most commonly associated with locative/goal-selecting AMs, being found with such forms in ten languages of the sample. Intensive meanings are uncommon with beneficiary-/instrument-/theme-selecting AMs, being found in just three languages (3 of 50): Nasal, Gayo, and Mori Bawah. With Nasal *-kun* and Gayo *-(n)en* the intensive meaning primarily applies to perception events and indicates active or careful perception. Mori Bawah *-Cako* can indicate an intense or haphazard manner of action as in *me'iniako* 'to hold oneself tightly or strongly' and *morawuosako* 'to scatter in a rough manner, strew' (Mead 2005: 703).

Figure 6.6: Aspectual, intensive, and comparative functions of applicative morphemes in Sulawesi



Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

Figure 6.7: Aspectual, intensive, and comparative functions of applicative morphemes in West Nusantara



Includes geodata from Natural Earth (public domain) and language data from Hammarström et al. (2022).

6.4.4 Comparative and other degree constructions

In a relatively small number of languages of the sample, a comparative degree construction is formed by morphological marking of the verb with an AM (see §2.10.4 for examples in Sundanese). This occurs in six languages of the sample, which are shown in Table 6.14. Despite being uncommon, this function is fairly widely distributed geographically and across genetic groupings in West Nusantara. However, it is difficult to draw conclusions about the origin of this function, which is not found with pivot-selecting applicative markers.

In one additional language, Toba Batak, comparative constructions may be formed with the suffix *-an*, e.g. *balgá* ‘big’, *balgaán* ‘bigger’; *daó* ‘far’, *daoán* ‘farther’ (Nababan 1981: 71–72). In Toba Batak, the locative/goal-selecting applicative has the form *-an* only with non-volitional PV verbal forms marked with *ha-*, and completive participial PV verbs marked with *ni-*, while *-i* is found with AV verbs and simple indicative PV verbs. Thus it is not clear that the comparative suffix *-an* and the locative applicative suffix *-an* should be considered the same morpheme.

The fact that the comparative degree function is found with both benefactive/instrumental AMs and locative/goal-selecting AMs in the survey data, but only the latter when the associated form is *-an*, suggests that the connection between the comparative function and the AM, may have come about through the convergence of two separate morphemes with similar shapes, at least in some cases.

Table 6.14: Applicative morphemes with a comparative degree function

Language	AM	Role of AppP	CAUS	COMP
Gayo	-(n)en	THM	Y	Y
Bugis	-(C)eng	BEN, INST	N	Y
Mori Bawah	-aka	BEN, INST	N	Y
Wolio	-aka	BEN, THM	N	Y
Sundanese	-an	LOC, GOAL	Y	Y
Ampenan Sasak	-an	BEN, THM, LOC, GOAL	Y	Y

In a few languages of the sample, we see the AM in constructions expressing causative meanings in which causee comes to be characterized by a greater degree of the base property, e.g. ‘to make more X’. These constructions may be marked by both the AM in addition to a distinct causative morpheme. Some examples include Karo Batak *pebiringken* ‘to make blacker’, cf. *pebiring* ‘to blacken’ and *biring* ‘black’; and *peganjangken* ‘to put up even higher’ cf. *peganjang* ‘to put up high’ and *ganjang* ‘high’ (Woollams 1996: 62). Verb stems like these are also found in Karo Batak with the applicative suffix *-i*, sometimes with the same bases, e.g. *peganjangi* ‘to make higher’; *pekitiki* ‘to make smaller’, cf. *kitik* ‘small’ (Woollams 1996: 63). Similar examples also appear with *-i* in Wotu (Mead 2013: 33).

6.5 Syntactic properties of applicative constructions

6.5.1 Co-occurrence with other major voice constructions

As described in §4.5.2 above, for each language in the sample, the co-occurrence of ACs with other major voice constructions (e.g. AV, PV, passive, antipassive) was investigated. This was primarily included as a diagnostic for pivot-selecting vs. pivot-neutral applicatives. In this section, summary findings are reported for this pattern.

There are 59 languages in the sample with applicatives of any type. In nine of these languages, the AC(s) strictly do not co-occur with other basic transitive constructions (AV, PV, or each other). These languages represent Philippine-type voice systems in which the only ACs are essentially voice constructions in which a peripheral semantic role is selected as the syntactic pivot (LV, CV). In addition, we generally do not see any co-occurrence of LV and CV constructions with intransitive constructions. Some authors do report that it is possible for certain nonactor voice constructions—including pivot-selecting applicatives—to co-occur with distinct morphologically marked causative constructions (see Kroeger 2005: 422–424; Dillon 1994: 123–124; Gault 1999: 20–22), though this is not always addressed in the source material for the survey.

Some examples of morphologically marked causative constructions in co-occurrence with pivot-selecting applicatives are given in (134) and (135) below. Note that compatibility of causative constructions with pivot-selecting applicatives vary by language, for example, in Tatana, only AV, PV, and IV co-occur with causative constructions marked with causative prefixes.

(134) Kimaragang, Causative constructions

- a. *Pa-akon-on kuh poh i Jaiwan tu wilton*
CAUS-eat-PV 1SG.GEN yet NOM J. because hungry
'I'll have **Jaiwan** eat something (i.e. give him something to eat) first, because he's hungry.' (Causative + PV)
- b. *Nunuh ot i-pa-akan nuh do tanak dot s<um>usu poh?*
what NOM IV-CAUS-eat 2SG.GEN GEN child REL <AV>milk yet
'**What** will you feed a child that is still nursing?' (Causative + IV)
(Kroeger 2005: 423)

(135) Sama Bangingi', Causative constructions

- a. *Pa-bowa-na ma aku surat na*
CAUS-PV.bring-3SG.GEN OBL 1SG.ABS letter 3SG.GEN
'He will send **his** letter with me' ('He will cause his letter to be carried by me.') (PV)
(Gault 1999: 20)
- b. *Pa-bowa-han-ta kaa surat*
CAUS-bring-BV-1PL.INCL.GEN 2SG.ABS letter
'I'll send **you** a letter.' ('I'll cause you to be brought a letter.') (BV) (Gault 1999: 21)

In the remaining 50 languages of the sample, at least one pivot-neutral AC is attested in which the applied phrase does not obligatorily map to the relevant privileged syntactic argument (e.g. pivot, subject, absolutive). These results will be discussed below according to type of voice system.

Five languages of the sample show pivot-neutral applicatives alongside a Philippine-type voice system (productive or marginal). In four of these, it appears that the pivot-neutral applicative(s) do not combine with all transitive voices. In Totoli, neither of the two pivot-neutral applicative formations (one goal-selecting and one beneficiary-/instrument-/theme-selecting) co-occur with LV, even though the LV construction is otherwise very productive (Himmelmann & Riesberg 2013). In Central Sama and Yakan, similarly pivot-neutral ACs marked with *-an* do not co-occur with IV or LV (Brainard & Behrens 2002; Townsend 2017). In Bobongko, we see some co-occurrence of IV with locative/goal ACs marked with *-i*, and possible co-occurrence of benefactive ACs marked with *-akon* with LV, but not vice versa (Mead 2001). Only in Balantak, do we see co-occurrence of pivot-neutral applicative across all transitive voice values, AV, PV, and LV, and this is observed with all three AMs in the language (van den Berg & Busenitz 2012). These examples have been discussed in more detail above in §5.8.4.2 and §5.9.4.

There are 30 languages of the sample that show two-way symmetrical voice systems (non-marginal). Of these, almost all (28 of 30) show free co-occurrence of pivot-neutral applicatives across major transitive voice constructions, i.e. AV, PV and any additional passive or antipassive constructions.

In Pendau, the instrument-selecting ACs marked with *-a'* and the locative-selecting ACs marked with *-i* do not co-occur with AV, while beneficiary-selecting AC with *-a'* and goal-selecting ACs with *-i* do co-occur with AV (Quick 2007). As discussed in §5.9.4, the restricted ACs probably represent remnant LV and IV constructions, and it is possible that Pendau should be considered a marginal Philippine-type voice system, though it was not coded as such in the survey.

In one two-way symmetrical voice language, Laiyolo, it was not possible to determine whether the pivot-neutral applicative suffixes co-occur with the AV construction based on the source material. Applicative suffixes *-i* and *-aka* in Laiyolo are attested to co-occur with the zero-marked (PV) transitive construction and the passive construction marked with *ri-* (Belding, Laidig & Maingak 2001).

Eight languages of the sample show marginal two-way symmetrical voice systems and pivot-neutral applicatives, while seven show asymmetrical voice systems and pivot-neutral applicatives. These will be discussed together because of overlap in the types of major voice constructions represented.

Four of these languages (Acehnese, Nias, Enggano, and Wolio) show only one basic transitive voice construction, at least in main clauses, and a passive or possible passive construction. In Wolio and Nias, there are remnants of an alternation between A-oriented and P-oriented transitive clauses found in restricted clauses types (i.e. relative clauses in Nias marked with *aN-* vs. *ni-*, participial clauses in Wolio marked with *mo-* vs. *i-*). Across both types of marking in these restricted clauses, AM-marking may co-occur.

Nine languages of the sample show one zero-marked P-oriented basic transitive construction, a passive construction, and one or more constructions marked with a nasal prefix that show a restricted distribution. These include three Bungku-Tolaki languages (Moronene, Mori Bawah, and Tolaki) and six South Sulawesi languages (Embaloh, Bugis, Coastal Konjo, Makassar, Bambam, Seko Padang). The constructions marked with the nasal prefix may be labelled actor focus, semi-

transitive, or antipassive, and typically show lower semantic transitivity and lower specificity or definiteness of P. In almost all cases, AM-marking in these nine languages is found to co-occur across all major voice constructions available. In Mori Bawah, however, while the AMs *-Ci*, *-Cari*, and *-Cako* co-occur with all major voice constructions, *-ako* does not co-occur with the antipassive construction marked with *poN-*, at least with some functions. I find it surprising that ACs in almost all cases, do co-occur with the less semantically transitive constructions, including antipassive construction, as the properties of these appear to be at odds with identified discursive functions of applicatives, which include higher topicality and referentiality of the applied phrase (see e.g. Peterson 2007: 120–121).

Finally, two languages of the sample (Muna, Busoa) also show two distinctive alternations in main clauses that co-vary with definiteness of P, but neither of these are P-oriented, instead showing a privileged subject relation. In addition to an agentless passive construction, Muna and Busoa show an alternation between A-oriented and P-oriented participial clauses with transitive verbs. ACs in the languages appear to generally co-occur across all of the aforementioned constructions, though in the case of Busoa, examples with *-Ci* in participials are limited.

6.5.2 Syntactic properties of the applied phrase in pivot-selecting ACs

As described in §4.5.3 and 4.5.4, for languages with ACs, patterns of indexing of the applied phrase and syntactic properties of the applied phrase (e.g. syntactic privilege, access to relativization) were investigated in the survey. In this section, I present results for these patterns for pivot-selecting constructions, while the following section presents results for pivot-neutral constructions.

It is important to note that information about indexing and syntactic properties of the applied phrase were not clearly reported in many descriptive accounts used as source material in the survey. In the absence of descriptive comment, coding values for the survey were determined on the basis of example sentences and other textual material available. Even so, it was not possible to determine some of the information included in the survey questionnaires for a number of languages of the sample, and this will be noted below.

In languages of the sample with Philippine-type voice alternations, of which there are 14 in the sample, the indexing and syntactic properties of the applied phrase in LV and CV constructions are generally described as equivalent to those of the pivot argument in other basic constructions, i.e., A in AV and P in PV. This pattern is as expected.

In Philippine-type languages of the Sabahan group, certain clitic pronominal forms may be used to index clausal arguments on the verb or verbal complex. The distribution and properties of these clitic pronouns or “second position clitics” are described for Kimaragang (Kroeger 2005) and Tatana (Dillon 1994). In these languages only non-pivot A arguments and pivot arguments may be expressed by means of a second position clitic, though the pronominal sets from which these are drawn are distinct. The set for the former is usually labeled *genitive* while the set for the latter is usually labeled *nominative*. For Keningau Murut, Timugon Murut, and Tombonuo, some details are unclear in the source material, but the sets of pronouns and their use in argument indexing appears to be similar. For all five Philippine-type Sabahan languages of the sample, the pronominal index used for the pivot argument in AV and PV may also be used for the applied phrase in LV and CV constructions.

In terms of syntactic properties, the applied phrase in LV and CV constructions in these languages is obligatorily the pivot argument. In Kimaragang, Tatana, and Keningau Murut, behavioral properties of the applied phrase and other pivot arguments is more clear. In these languages, it is reported that the applied phrase generally shows syntactic behavior similar to A in AV and P in PV. The pivot argument—including the applied phrase in LV and CV—may be fronted, relativized, and/or clefted (Dillon 1994: 30–33; 71–72; Cohen 1999: 7–8; Kroeger 2005: 405, 412) while non-pivot arguments do not have access to these operations. Some examples from Tatana are given in (136). Information on syntactic behavior of the applied phrase and other pivot arguments was not available on Timugon Murut and Tombonuo from the source material, including evidence of access to relativization.

(136) Tatana, Relativization of applied phrase

- a. *Ulun gii [p<in>a-taak-an mu do duit]*
 person REL <PST>-CAUS-give-RV 2S.GEN DAT money
naka-panau do goi' i.
 AV.NVOL-walk ADVZ late.afternoon PN.NOM
 ‘The man [who was given money by you] left yesterday.’ (RV)
- b. *Isai do kadai [p<in>am-(b)ali-an mu do akanon diti]?*
 who DAT shop <PST>LV-buy-LV 2SG.GEN DAT food PROX
 ‘Who owns the shop [where the food was bought by you]?’ (LV) (Dillon 1994: 32)

In the the remaining Philippine-type languages of the sample, there does not appear to be use of bound indexing for non-A arguments; bound indexes are not used for P of transitive BCs, nor for the applied phrase in ACs.

6.5.3 Syntactic properties of the applied phrase in pivot-neutral ACs

In this section, I report results of the survey for syntactic properties of the applied phrase in pivot-neutral constructions, as compiled for the 50 languages of the sample showing this type of AC. Because the properties of the applied phrase surveyed are often closely correlated, I begin with a general overview the results for the surveyed patterns together, then give more details for particular languages that showing unexpected patterns.

6.5.3.1 General overview of results

The survey data show that languages with pivot-neutral applicatives in West Nusantara in large part show expected patterns of indexing, mapping to a privileged syntactic relation, and access to relativization for the applied phrase.

Results on use of bound-indexing for the applied phrase are shown in Table 6.15. About one-third of the languages (19 of 50) generally show no bound-indexing of P on the verbal complex in transitive clauses, and no bound-indexing of S in passive constructions. For these languages, as expected, there is no bound-indexing of applied phrases in ACs. In more than half of the languages of the sample (30 of 50), bound-indexing of P on the verbal complex is observed, at least for some person and number categories, and in some clause types. Of these, the large majority

(21 of 30) show bound-indexing of the applied phrase in ACs in the same manner characteristic of P in BCs. However, in a small number of cases (6 of 30), there are mixed patterns of indexing for applied phrases in ACs, some of which diverge from expected patterns. These cases include Muna, Busoa, *Tukang Besi*, *Mori Bawah*, and *Tolaki*. For three languages (3 of 30), patterns of indexing could not be determined across all AMs: *Alas Batak*, *Wolio*, and *Seko Padang*. For one language (*Enggano*) it could not be determined whether there is bound indexing of S in passive constructions.

Table 6.15: Bound-indexing of the applied phrase

Indexing pattern	No. lgs.	%	No. lgs.	%
Bound-indexing of P	30	60		
Like indexing of AppP			21	44
Mixed indexing of AppP			6	12
Undet. indexing of AppP			3	6
No bound-indexing of P	19	38		
Undetermined	1	2		
TOTAL	50	100		

With respect to mapping of the applied phrase to a privileged syntactic relation, each language was assigned a coding value in a manner consistent with its type of voice and alignment system. Generally, this pattern was evaluated based on the P-oriented constructions in a given language, such as PV constructions, zero-marked P-oriented basic transitive clauses, passive constructions, and P-oriented relative clauses or participial phrases. In such constructions, there is evidence of syntactic privilege for the P argument, which may be a pivot, a subject, or an absolutive.

Table 6.16: Mapping of the applied phrase to the privileged syntactic argument

Mapping pattern	No. lgs.	%
Yes, in P-oriented const.	40	80
Mixed, in P-oriented const.	8	16
Not applicable	1	2
Undetermined	1	2
TOTAL	50	100

Results on mapping of the applied phrase to the privileged syntactic argument in pivot-neutral ACs are shown in Table 6.16. Most languages (38 of 50) show evidence that the applied phrase may map to the privileged syntactic argument across P-oriented clause types, and across the inventory of AMs. In a smaller number of languages (8 of 50), the applied phrase in certain types of pivot-neutral ACs may not map to the privileged syntactic relation in P-oriented constructions, though in other types of ACs, they may. These include the five languages with mixed patterns of indexing for the applied phrase mentioned above, plus *Balantak*, *Bobongko*, and *Makasar*. In one language (*Acehnese*), there is no evidence of a privileged syntactic argument (Durie 1988),

so this pattern is not applicable. In one language (Moronene), it was not possible to determine the pattern based on the source material.

With respect to access to relativization, this behavioral property tends to correlate with mapping to the privileged syntactic argument, though there is some variance across languages of the sample as to which grammatical relations may be the head noun in a relative clause. In some languages participial clauses or gerunds appear to form headless relative clauses, and these structures were used as evidence of access to relativization. Coding for this pattern was evaluated on the basis of descriptive comment and examples in textual material, or only the latter if the author did not explicitly mention this possibility for the applied phrase in ACs.

Table 6.17: Access to relativization for the applied phrase

Access pattern	No. lgs.	%
Yes, generally	33	66
Mixed, by type of AC	5	10
Undetermined	12	24
TOTAL	50	100

Results on access to relativization for the applied phrase in pivot-neutral ACs are shown in Table 6.17. In most languages (33 of 50) there was evidence that the applied phrase can head a relative clause, in at least one P-oriented clause type. For five languages, there was evidence that the applied phrase in at least some types of ACs cannot head a relative clause. These languages are: Balantak, Bobongko, Mori Bawah, Tolaki, and Tukang Besi. For a good number of languages of the sample (12 of 50), this pattern could not be determined due to lack of descriptive comment and lack of sufficient examples. In most of these, the applied phrase does appear to map to the type of grammatical relation which may head a relative clause, but there are simply no clear examples in the source material of AM-marked verbs in relative clauses.

While most languages with pivot-neutral applicatives in West Nusantara show expected patterns of coding and behavioral properties, it must be noted that evidence of this for all types of peripheral roles that may be selected as the applied phrase was often not available. Also, I did not attempt to evaluate the syntactic properties of applied phrases and companion phrases in A-oriented constructions, as this is not necessarily apparent from coding, and behavioral properties for such are rarely, if ever, addressed in the source material. Furthermore, non-canonical ACs may exist along canonical ACs in these languages, as will be discussed in §6.5.4. Thus, the results may overstate apparent similarities between P arguments in BCs and applied phrases in ACs in some languages of the sample.

In the remainder of this section I will present the patterns for indexing, mapping to a privileged syntactic argument, and access to relativization in languages that show unexpected or mixed patterns across types of ACs. The eight languages in which these are found are concentrated in Sulawesi and include Balantak and Bobongko (Saluan-Banggai); Muna, Busoa, and Tukang Besi (Muna-Buton); Mori Bawah and Tolaki (Bungku-Tolaki); and Makasar (South Sulawesi).

6.5.3.2 Unexpected patterns of indexing

Six languages of the sample show mixed patterns of indexing for the applied phrase. These patterns show animacy effects as beneficiaries are more likely to be indexed, and to be indexed using special bound forms, than are instruments, themes, and circumstantial roles (e.g. reason, purpose). Mixed patterns of indexing are also observed more commonly in ditransitive ACs compared to monotransitive ACs.

In Muna, Busoa, Mori Bawah, and Tolaki, the applied phrase may show special forms for pronominal indexes used in ACs marked with the benefactive/instrumental AM. In these languages, the P argument, when definite, is usually indexed on the verb as an enclitic or suffixal pronominal form in transitive BCs.

In Bungku-Tolaki languages, the forms of the special indexes are analyzed as arising from phonological coalescence of the applicative suffix *-ako* with the absolutive enclitics used to index P in BCs (Mead 1998: 207–212; cf. Edwards 2012: 53–54). Such forms are shown for the Mori Bawah in Table 6.18. In BCs, the verb stems either takes no bound-indexing or indexing with the absolutive suffixes in the left column. In certain ACs, we either see *-ako* and no following index on the verb, while in others we see one of the fused forms in right column. The special indexes in Mori Bawah and their corresponding forms in Tolaki, are regularly used to index beneficiary applied phrases, but only rarely with instrument or reason applied phrases.

Table 6.18: Fusion of Mori Bawah *-ako* with absolutive pronouns (Mead 2005: 704)

	Absolutive	<i>ako</i> + absolutive
1SG	aku	akune
2SG	ko	akomu
3SG	o	akono
1PL.INCL	kita	akita
1PL.EXCL	kami	akami
2PL	komiu	komiu
3PL	ira	ako'ira

In Muna and Busoa, the use of the special bound-indexes is similar with some complications. A comparison of Muna pronominal suffixes is given in Table 6.19.⁸ Van den Berg (2013: 68) calls the forms used for indexing of P in BCs “direct object suffixes,” and the forms used for indexing beneficiary and recipient applied phrases “indirect object suffixes,” as Muna has a subject relation (S and A) and an object relation (P) in basic main clauses. As in Mori Bawah, in BCs the verb either bears no bound-index for the object, or a direct object suffix from the left column. In certain ACs, the verb either bears the suffix *-ghoo* and no object index, or a form from the right column.⁹

⁸Note that there are no first person dual or plural inclusive pronominal suffixes used to index clausal arguments; free pronouns must be used instead.

⁹As shown in the table, the Muna direct object and indirect object suffixes do not differ in form for all person and number categories. But when they do, it does not appear that these forms result from fusion of *-ghoo* with the object suffixes. Instead, the shapes of the indirect object suffixes are consistent with the fusion of an earlier suffix with the form *-an* preceding the object suffix. If so, this *-an*, would be a likely reflex of PAn and PMP CV **-an*, which was replaced by *-ghoo*. In Busoa, the pronominal suffixes are somewhat different from those in Muna, and

Table 6.19: Muna pronominal suffixes for direct and indirect objects

	Direct object	Indirect Object
1SG	-kanau	-kanau
2SG	-ko	-angko
2SG.POL	-kaeta	-kaeta
3.SG	-e	-ane
1DU.INCL	—	—
1PL.INCL	—	—
1PL.EXCL	-kasami	-kasami
2	-ko-omu	-angko-omu
2PL.POL	-kaeta-amu	-kaeta-amu
3PL	-da	-anda

In Muna and Busoa, the indirect object suffixes are used to index beneficiary and instrument applied phrases. It does not appear that the indirect object suffixes are used to index reason applied phrases, which are also found in ACs marked with *-ghoo* in Muna and *-ho* in Busoa.

In all four languages that show special forms of indexes, the typical indexing for P and the special indexing for the applied phrase may stack on the verb. However the conditions governing the type of suffixes and mapping used when both the applied phrase and the companion phrase are eligible to be indexed on the verb are complex and differ across the languages. There do however, appear to be some animacy effects observed. In Muna and Tolaki, the instrument may only be indexed when it is a third person referent and in that case, the companion phrase may be indexed with the special form, while the applied phrase is indexed with the regular object/absolute suffix. In general, then it appears that the use of the special forms of indexes in ACs in these four languages shows animacy effects. The special forms are commonly used for indexing of beneficiaries and recipients, which are typically animate, sometimes used for certain types of patient companion phrases, which are possibly animate or inanimate, but only rarely used with instruments and reason roles, which are typically inanimate.

Tukang Besi and Makasar show only one form of bound-indexing for nonactor arguments whether in BCs or ACs. In both languages, the applied phrase is indexed on the verb in this manner in some types of ACs but not others.

In Tukang Besi, the applicative suffix *-ako* marks ACs in which the applied phrase is a beneficiary, instrument, theme, reason, or purpose role. All of these except themes and purpose phrases may be indexed on the verb in ACs with the normal object suffixes (Donohue 1999: 232–242). Interestingly, it appears that the companion phrase is never indexed on the verbs in ACs marked with *-ako*; even if the applied phrase is not indexed or cannot be indexed. Comitative applied phrases in ACs marked with *-ngkene* and locative/goal applied phrases in ACs marked with *-i* may be indexed on the verb, and again the companion phrase is these types of ACs is

also show dialectal variation and stem-conditioned allophony (van den Berg 2020), but in at least in the third person singular, where the direct object suffix is *-e* and the indirect object suffix is *-ene* (and variant *-ane*) this proposal might be supported. In any case, this constitutes some evidence that Muna *-ghoo* and Busoa *-ho* may be newer forms of AM-marking, and this might explain why the shapes of the suffixes show unexpected sound correspondences as mentioned in §6.2.3 above.

never indexed.

In Makasar, patterns of indexing for the applied phrase varies in ACs marked with the applicative suffix *-ang*. In Makasar the general pattern in BCs is that, when P is definite, it is indexed on the verb in a zero-marked transitive clause using an absolutive enclitic. If P is not definite, the zero-marked transitive construction typically cannot be used, only a semi-transitive or antipassive construction in which the same participant is not indexed. In ACs, the pattern of indexing observed is influenced by the transitivity of the base. With intransitive bases, the applied phrase, which is usually a theme, is always indexed on the verb. With monotransitive bases, the beneficiary applied phrase is always indexed on the verb, and the clause is not sensitive to the definiteness of the companion phrase. Instruments appear to follow the same pattern, though the examples are not always clear. With ditransitive bases, which only occur with lexical verb meaning ‘give’, the recipient is always indexed on the verb, and the theme companion phrase is not. These ACs with ‘give’ however, are sensitive to the definiteness of the theme companion phrase, and cannot be used with indefinite themes.

6.5.3.3 Unexpected patterns of syntactic behavior

Eight languages of the sample show unexpected syntactic behavior for the applied phrase. In general in locative- and goal-selecting ACs the applied phrase is allowed to be the privileged syntactic argument in P-oriented constructions. In ACs selecting beneficiaries, instruments, themes, and other circumstantial roles as the applied phrase, the patterns are more mixed. Interestingly, indexing of the applied phrase does not necessarily covary with access to syntactic operations.

In Balantak and Bobongko, the syntactic properties of the applied phrase vary by AM. The applied phrase in locative- or goal-selecting AC with *-i* (variant *-an* in Bobongko) pattern much like P in BCs and can become the pivot in a PV construction. Accordingly, they are eligible to be the head of a relative clause in Balantak, and the same would be likely hold for Bobongko though evidence of such is not available (Mead 2001).

The applied phrase in ACs marked with *-ako* in Balantak and *-akon* in Bobongko, however, may not hold the pivot relation in PV. In Balantak the applied phrase in ACs marked with *-ako* are of various roles—including beneficiary, instrument, theme, and reason—but all alike are never realized as the pivot in PV. In Bobongko only beneficiary applied phrases are observed with *-akon*, and these also may not hold the pivot relation in PV (but see also §5.9.4 on ACs marked with *-akonan* in Bobongko). Accordingly, these types of roles are not eligible to head a relative clause in Balantak, and the same would likely hold for Bobongko.

There is also a type of AC marked with *-ii* in Balantak, which selects a beneficiary as the applied phrase. In such ACs, only the beneficiary may be the pivot in P. The companion phrase (theme or patient) cannot be the pivot in PV, as evidenced by word order and the patterning of articles with NP arguments in such clauses (van den Berg & Busenitz 2012: 108–109). This is in keeping with the proposal that *-ii* arose as a fusion of the locative/goal suffix *-i* plus a following person name marker *i*.

In Tukang Besi, most types of applied phrases may be the subject of a passive construction, and are eligible to head a P-oriented relative clause marked with the prefix *i-* on the verb. Only reason and purpose applied phrases may not be the subject of passive construction, nor head a P-oriented relative clause. The companion phrase typically does not show object properties in ACs, except in locative/goal-selecting ACs, where it may head a P-oriented relative clause, but

not map to subject in a passive construction. See also Donohue (1999) for extensive discussion of other syntactic properties of the applied phrase and companion phrase across roles and types of AM-marking.

In Makasar, Jukes (2020) argues that indexing of an argument with absolutive enclitics or ergative proclitics does not necessarily correlate with grammatical relations. So while some types of applied phrase are indexed with absolutive enclitics, this does not necessarily equate to status as syntactically privileged. With respect to relativization, applied phrases may head a relative clause (Jukes 2020: 229).

Finally, in the languages that show special indexes for the applied phrase in some ACs, the pattern of indexing does not necessarily covary with access to syntactic operations. In Muna and Tolaki, though beneficiaries are indexed with the special suffixes and instruments are usually not, both types of applied phrase may be the subject of a passive participial clause. (There is no evidence by which to compare reason applied phrases in these two languages.) In Mori Bawah, however, only beneficiaries may be the subject of a passive construction in ACs marked with *-ako* with transitive bases. Instrument and reason applied phrases may not be the subject in passive constructions, unless the base is intransitive. The patterns for Busoa are not clear. In these languages, the passive participials may be used like a headless relative clause, so passivization and relativization are not separate indicators of access to syntactic operations.

6.5.3.4 Summary and discussion

As shown in this section, the survey data indicates that, in most languages of the sample with pivot-neutral applicatives, the applied argument in ACs shows coding and syntactic properties similar to that of P in BCs, at least for the type of indicators which can be assessed based on the source material. However, there are eight languages of the sample, all located in Sulawesi, which show unexpected patterns in the syntactic properties of applied phrases.

Generally it appears that in locative-/goal-selecting ACs marked with *-i* the applied phrase shows coding and syntactic behavior like P in BCs. For beneficiary-/instrument-/theme-selecting ACs, the patterns tend to show the influence of animacy, as well as sensitivity to the transitivity of the base. With transitive bases, unexpected patterns in syntactic properties are more commonly observed, since both the applied phrase and a companion phrase may be arguments of the clause, and there is, it appears, some type of competition between them for syntactic properties. The influence of animacy is also observed, with beneficiaries being most likely to show indexing and access to syntactic operations, as compared to instruments and themes, with reason and purpose applied phrases being even less likely to do so. Thus in cases where both an applied phrase and a companion phrase appear to be clausal arguments in an AC, the animate participant is more likely to show special forms of indexing, or to be indexed to the exclusion of the other, and more likely to show access to syntactic operations.

The patterns discussed in this section, however, have been coded and described primarily based on AM-marked clauses that resemble canonical ACs in structure. This is because these are the type of data in which comparison can most readily be made between properties of the applied phrase in ACs and P in BCs. In the following section, I discussed non-canonical ACs that exist alongside more canonical ACs in the languages of the sample.

6.5.4 Alternations in the realization of the applied phrase

In this section, in order to provide greater descriptive detail, I present various types of non-canonical ACs that exist alongside canonical ACs in the same languages with respect to status of the applied phrase. It is important to note that non-canonical structures for ACs are often mentioned in passing by authors in source material, or sometimes are not explicitly mentioned, though included textual examples may happen to show more than one possible structure for ACs. The data presented here must thus be considered non-exhaustive; nevertheless, the examples shown below point to the range of diversity observed. Based on these data, it appears that alternations in the realization of the applied phrase (i.e. coding as core or oblique) may be influenced by voice construction (AV, PV), word order, and whether the companion phrase is realized or unrealized.

6.5.4.1 Possessor beneficiaries

In Sundanese, the beneficiary in an AC marked with *pang-* *-keun* may be realized as an unmarked NP, but are also frequently instead expressed as the possessor of the companion phrase, which is a patient or theme participant. This appears to be especially common in AV structures.

- (137) Sundanese, possessor beneficiary
Abi m-(p)ang-ngumbah-keun piring Ema.
1SG AV-BEN.APPL-wash-BEN.APPL dish mother
'I washed mother's dishes for (her).' (CT1-003)

In the Sundanese examples, it is possible that the possessor NP bears is not formally mapped to the beneficiary role in argument structure, but instead is interpreted as the beneficiary through pragmatic inference. This, however, is not necessarily the case in other languages.

For instance, Balantak show a similar type of beneficiary-selecting AC marked with *-kon*. In example (138) below, the theme companion phrase is *doi'-ku* and include the possessive suffix *-ku* 'our'. This sentence does not necessarily mean that the item to be acquired belongs to 'us', though it must mean that the item is acquired 'for us'. Unlike the Sundanese case, the possessor beneficiary is obligatory in Balantak *-kon* marked benefactive ACs; the beneficiary may not be realized as an unmarked NP or as a PP, and may not be unrealized. This construction is noted to occur with transitive bases in Balantak. In other types of *-kon* marked ACs, overt realization of the applied phrase is not obligatory, and van den Berg & Busenitz (2012) note that overt realization of both the applied phrase and the companion phrase is very rare in AV in such ACs. This suggests that the realization of the applied beneficiary as a possessor is related to avoidance of a sequence of two adjacent unmarked NPs after the verb.

- (138) Balantak, possessor beneficiary
Koo-si a mang-ala-kon doi'-ku.
2SG-SEQ ART AV-get-BEN.APPL money-1SG
'You are the one who will get money for us' (van den Berg & Busenitz 2012: 101)

6.5.4.2 Beneficiary/recipient applied phrases with oblique marking

In a number of cases, alternations are observed in the coding of the beneficiary or recipient participant in AM-marked constructions. Examples have already been presented for Sundanese ACs in §2.7 above, showing that in some clauses marked with *pang-* *-keun*, the beneficiary is realized as a core argument and coded as an unmarked NP core argument, while in others, the beneficiary is coded as an oblique PP with *keur* ‘for’ (or a possessor phrase as mentioned above). Likewise, examples for Indonesian were presented in §3.3 above, showing that in some clauses marked with *-kan*, the beneficiary is realized as a core argument and coded as an unmarked NP, while in others, the beneficiary is coded as an oblique PP with *untuk* ‘for’ (see also Vamarasi 1999: 76). See also the discussion of Bobongko beneficiary applied phrases showing oblique marking in §5.9.4.3.

In Pendau, the suffix *-a* marks benefactive and instrumental ACs, among other functions. Quick (2007: 292–293) reports alternations observed in the coding of the beneficiary/recipient participant when the verb *bagi* ‘to give’ is suffixed with *-a*. The beneficiary may either be realized in absolutive case, which is typical for P arguments, or marked with the preposition *sono* ‘with’ which is typical for oblique comitative phrases.

(139) Pendau, Alternations in coding of the recipient

- a. *A’u mombagi-a’ oo bulaan.*
a’u M-pong-bagi-a’ ’oo bulaan
 1SG.ABS AV.IRR-SF-give-BEN.APPL 2SG.ABS gold
 ‘I will give you gold.’ (AC, absolutive beneficiary)
- b. *Ula uo nombagia’ doi’ sono langkai uo.*
ula ’uo N-pong-bagi-a’ doi’ sono langkai ’uo
 snake yonder AV.RLS-SF-give-BEN.APPL money with male yonder
 ‘That snake gave the money to the man’. (AC, PP beneficiary) (Quick 2007: 292)

In Bugis, Hanson (2003: 196–197) reports that alternations in the coding of beneficiary/recipient in ACs marked with *-Ceng* are possible with some lexical verbs but not others. Some verbs allow coding of the beneficiary as a prepositional phrase, as in (140) with the verb *nasu-ng* ‘cook for’. Other verbs with *-Ceng* require the beneficiary to be either coded as an unmarked NP (with indexing of the companion phrase on the verb) or indexed on the verb (with the companion phrase coded as an unmarked NP), see §7.5.

(140) Bugis, Oblique marked beneficiary

- U-nasu-ng-ngi nanre ku Ali.*
 1.ERG-cook-BEN.APPL-3.ABS rice PREP A.
 ‘I cooked rice for Ali.’ (Hanson 2003: 196)

6.5.4.3 Locative applied phrases with oblique marking

In Pendau, locative applied phrases in ACs marked with the applicative suffix *-i* may be realized as unmarked NP, but are also observed to be encoded with oblique marking. In the example below,

this is observed with the locative proclitic *ri=*. This *ri=* represent the usual marking for locative obliques in BCs.

(141) Pendau, Locative applied phrase with oblique marking

Ribongkarongo'u niponyoputi'u.

ri=bongkarong='u ni-pong-soput-i='u

LOC=hut=1SG.GEN IV.RLS-SF-shoot-LOC.APPL=1SG.GEN

'I shot (it) at/beside my hut.'

(Quick 2007: 300)

The Pendau example looks similar to certain clauses bearing LV marking in languages with pivot-selecting applicatives (see 5.9.4.4 on the Pendau locative-selecting applicative as a remnant of LV). An example is given from Kimaragang below in LV. The pivot argument (i.e., applied phrase) normally bears nominative case marking in Kimaragang, including locative pivots in LV. But in this example, while the verb is marked with *-on* indicating LV, there is no nominative argument, only a dative marked phrase expressing a location, *sid tana* 'on the earth'. Kroeger (2005: 408) considers this phrase "an adverbial element" which has been fronted. It is not clear whether *sid tana* should be considered an applied phrase in this example.

(142) Kimaragang, Oblique marked locative phrase in LV

Sid tana yah n-odop-on

DAT earth 1PL.EXCL.GEN PST-sleep-LV

'We slept on the ground (after the house burned down).'

(Kroeger 2005: 408)

Another similar example is found in Keningau Murut, as shown in (143). The pivot is Keningau is usually a bare NP for common nouns or marked with *i* for personal names. Here the verb is in LV but the location phrase is marked with the locative noun phrase marker *tio'* rather than bare. It is possible that this example does not consist of a single verbal clause; even so, it closely resembles the Pendau structure analyzed as a oblique marked locative applied phrase.

(143) Keningau Murut, Oblique marked locative phrase in LV

Man-amaal-an ku du baloy kuy tio' Patikang

NPST-build-LV 1SG.NPIV.A CN.OBL house 1SG.POSS.DIST LOC Patikang

'The place I will build my house is there in Patikang'

(Cohen 1999: 42)

6.5.4.4 Goal/recipient applied phrases with oblique marking

In Javanese, the suffix *-i* is a pivot-neutral AM that may select a location, goal, or recipient applied phrase, among others. With this AM and verbal bases expressing transfer events, Bintoro (1980: 309) reports alternations in the coding of the recipient applied phrase in conversational Javanese as shown in the examples below. Other authors do not report that preposition-marking is possible for the recipient in *-i* marked clauses (Hemmings 2013; Vander Klok & Evans 2022), though see below on similar alternations in Madurese.¹⁰

¹⁰There are a number of discrepancies between the account given by Bintoro (1980) for verbs of transfer with the suffix *-i* and those presented by other authors (Hemmings 2013; Vander Klok & Evans 2022), including the fact

(144) Javanese, Alternations in coding of the recipient

- a. *Bapak ng-(k)irim-i Jono.*
father AV-send-LOC.APPL J.
'Father sent Jono something' (*dhuwit* 'money' may be contextually understood). (AC)
- b. *Bapak ng-(k)irim-i marang Jono.*
father AV-send-LOC.APPL to J.
'Father sent (something) to Jono' (*dhuwit* 'money' may be contextually understood). (AC)
- c. *Bapak ng-(k)irim-i dhuwit marang Jono.*
father AV-send-LOC.APPL money to J.
'Father sent money to Jono'. (AC) (Bintoro 1980: 309)

In Yakan, some examples of apparent alternations in the coding of recipient applied phrases occur only with specific lexical verbs. In Yakan, most transitive verbs can appear in bare form in P-oriented transitive verbal clause. Some lexical roots, however, cannot be used in this manner. The root *teppad* 'throw', is one such root; it must either bear marking with the prefix *mag-* in A-oriented clauses, or marking with *-an* in P-oriented clauses. In the former case, *mag-teppad* always take a theme as its complement. In the latter case, two possible structures are found as shown in (145). Brainard & Behrens (2002) analyze clauses like (145a) as a type of BC, in which the verbal suffix functions to form a transitive verb stem, but does not select a peripheral role as core. Here a theme participant maps to P, just as with *mag-teppad*, and the recipient/goal participant is coded with oblique phrase marking.¹¹ They analyze clauses like (145b) as a type of AC, in which *-an* is an AM that selects a recipient/goal as the applied phrase. Here, both the recipient/goal and theme are coded as core arguments, mapping to R and T respectively. This analysis is fairly well-supported given other examples of clauses with *-an* suffixed on the verb in Yakan (see also discussion of "vacuous application" of *-an* in West Coast Bajau in Miller 2007: 292–293). That notwithstanding, the pair of clauses in (145) closely resembles the alternation in (144) above from Javanese.

(145) Yakan, Alternations in coding of the recipient

- a. *Teppad-an-ne tolang-in pī si asuhin.*
throw-SF-ERG.3S bone-DEF DIR OBL dog-DEF
'She threw the bone to the dog.' (BC)
- b. *Teppad-an-ne asuh-in tolang.*
throw-GOAL.APPL-ERG.3S dog-DEF bone
'She threw the dog a bone'. (AC) (Brainard & Behrens 2002: 155)

that Bintoro shows examples in which the verbs *ng-(k)irim-i* 'send to', *marang-i* 'give to', *mènèh-i* 'give to', and *ng-layang-i* 'write to' may take a theme as the P argument. It is possible that such examples have differing levels of acceptability across speakers, and this may be related to differences in dialect, speech style, or other sociolinguistic factors, and/or properties of specific lexical bases.

¹¹I have glossed the suffix *-an* in this function as a stem-former, in keeping with the terminology used elsewhere in this study. Brainard & Behrens (2002: 104) call it a "classifier" because whether the suffix is required is lexically-determined, thus differentiating two classes of verb.

6.5.4.5 Instrument applied phrases with oblique marking

In Madurese, the verb *tambu* ‘hit’ may be used in an instrumental AC marked with the suffix *-agi*. In AV, the relationship between the AC and BC is fairly straightforward and both are mono-transitive. In the BC, the patient is the non-subject core argument and coded as an unmarked NP, while the instrument is coded as an oblique PP. In the AC, the verb is marked with *-agi*, the mapping is reversed and the instrument applied phrase is coded as an unmarked NP, and the patient companion phrase is coded as an oblique PP.

(146) Madurese, Instrument-selecting AC in AV

- a. *Ale’ n-(t)ambu’ burus bi’ bato.*
younger.sibling AV-hit dog with rock
‘Little Brother hit the dog with rocks.’ (BC in AV)
- b. *Ale’ n-(t)ambu’-agi bato dha’ burus.*
younger.sibling AV-hit-INST.APPL rock to dog
‘Little Brother hit the dog with rocks.’ (AC in AV)

(Davies 2010: 309)

(147) Madurese, Instrument-selecting AC in PV

- a. *Burus-sa e-tambu (bi’) bato bi’ Ale’.*
dog-DEF PV-hit (by) rock by younger.sibling
‘Little Brother hit the dog with rocks.’ (BC in PV)
- b. *Batu e-tambu’-agi (dha) burus bi’ Ale’.*
Rock PV-hit-INST.APPL (to) dog by younger.sibling
‘Little Brother hit the dog with rocks.’ (AC in PV)

(Davies 2010: 309)

In PV clauses, however, an alternation is observed in the structure of clauses with *tambu* ‘hit’. Such clauses appear to represent either monotransitive or ditransitive structures as indicated by the parentheses surrounding the preposition marking the postverbal NP in the example below. Note that, if the postverbal argument does not immediately follow the verb, but instead the agent phrase intervenes between the two, the preposition is obligatory. In this example, alternation in the coding of phrases expressing the instrument and the patient is not necessarily the result of the applicative operation, as it is observed in both the BC and AC. In addition, it is not clear in this example whether the prepositional marking is a reliable indicator of oblique syntactic status¹²

6.6 Summary of findings and implications

In this chapter, I have presented results from the typological survey related to properties of applicative systems in West Nusantara languages, and properties of ACs in those systems. Some major findings of this chapter are as follows.

¹²In a good number of languages of West Nusantara, the non-pivot actor in a PV clause is also observed to be optionally marked with a preposition when adjacent to the verb, but obligatorily marked with the preposition when not. In some cases such actor phrases show properties of core clausal arguments. This suggests that the status of the patient and instrument phrases in the Madurese examples discussed here is not clear-cut.

With respect to the distribution of forms of applicative morphology and their functions, a predominant pattern for pivot-neutral AMs is observed whereby one form of AM-marking is used for locative/goal-selecting ACs, and another is used for beneficiary-/instrument-/theme-selecting ACs. For the locative/goal-selecting ACs, based on this pattern and observed allomorphy in the relevant AMs a number of languages, I have argued that ACs of this type are derived from earlier LV constructions marked with *-i/*-an. For the beneficiary-/instrument-/theme-selecting ACs, I have argued that ACs of this type are most likely derived from earlier CV constructions marked with *-an, though in many languages replacement of the original suffix with a -K or -AK suffix has taken place.

With respect to non-applicative functions of AMs, the distributional patterns tend to be influenced by geographic location. Sulawesi languages in particular, show less prevalence and lower productivity of causative functions for AMs, than are observed in other parts of West Nusantara. The pluractional function is very widespread, and is strongly associated with locative/goal-selecting AMs, though also found in one case (Totoli) with benefactive/instrumental *-an*. The comparative function is sparsely attested, but also fairly broadly distributed across West Nusantara languages. It is associated with the form *-an*, whether it be a locative-goal-selecting AM, or a partially generalized benefactive-selecting AM, raising the possibility that the association of this function with the relevant AMs originated via some merger of two formerly distinct morphemes.

With respect to syntactic properties of the applied phrase in ACs, most languages of the sample show coding and behavior similar to that of P, though this similarity may be overstated due to the type and quantity of data available. However, some languages of Sulawesi show varying coding and behavior for the applied phrase, in patterns that are influenced by transitivity of the base and animacy of clausal arguments. Non-canonical realizations for the applied phrase are also observed across languages of the sample, though it is not possible to make an exhaustive description of these due to limitations of the survey and the available data.

Taken together, a major implication of these findings is that functional patterns as well as formal patterns are important in understanding the origin of the pivot-neutral applicatives in languages of West Nusantara. Besides this, the results indicate some ways that languages of western Indonesia in the area centered around Java and Madura show some similarities that are not reflected elsewhere, especially in Sulawesi, where the range of properties found for ACs and AMs reflect a fair amount of diversity. Such diversity is perhaps still understated, due to the fact that detailed description is not available for many languages of Sulawesi, and the accounts given here primarily feature descriptive details from eight languages described almost entirely by three authors who are areal experts. Therefore, there is still much work to do in documenting and understanding the diversity of applicative systems of West Nusantara languages, though the patterns outlined here may stand as a guide, albeit incomplete, to the possible range of this diversity.

Part III: Form and Meaning in West Nusantara Applicative Constructions

Chapter 7

A functional typology of applicative constructions in languages of West Nusantara

7.1 Introduction

In this chapter, I present a functional typology of pivot-neutral applicative constructions (ACs) and other constructions marked with applicative morphemes (AMs), based on a sample of 24 West Nusantara languages.¹ In these languages, ACs are marked by one of a small number of AMs that while serving many non-applicative functions are clearly separate from morphology signalling symmetrical voice alternations (see §3.4 on distinguishing symmetrical voice and applicatives). I focus on systems showing pivot-neutral applicatives, because AMs in these systems show more polyfunctionality, and ACs show more diverse syntactic properties, compared to systems with only pivot-selecting applicatives. Applicative systems showing pivot-neutral ACs are also well-represented in West Nusantara, while for pivot-selecting ACs, languages outside of West Nusantara make up the great part of diversity of applicative systems, particularly Formosan and Philippine languages.

To explore the range of diversity and commonality found across systems showing pivot-neutral ACs in West Nusantara, in this chapter, I describe AM-marked clausal constructions in these languages according to properties of their fixed form and associated meaning. As discussed in §1.4.1, the fixed form of a construction includes the form of the morphological marking on the verb, and syntactic properties of the clause, such as the number, structural position, and coding of clausal arguments. The meaning of a construction here refers to semantic properties of clause, such as the semantic roles that map to various positions in the clausal structure, and other elements of semantic meaning that are generalizable across bases that may fill the verbal slot in

¹An earlier version of this chapter was published as McDonnell & Truong (2024). However, it has been significantly expanded and revised with the addition of 16 languages to the original eight included in the sample, and incorporation of the constructional framework and typological categories used in Parts I and II of this study. I would like to recognize five individuals who shared their expertise on some of the languages in this chapter: Khairunnisa for Sasak, Dewi Setiani for Sundanese, Hendi Feriza for Besemah (South Barisan Malay), and Wawan Sahrozi and Johan Safri for Nasal.

these constructions.

The results show that beneficiary-selecting ACs show different properties than theme- and instrument-selecting ACs even though they are marked with the same AMs. The former almost always show maximally ditransitive structures, while the latter typically show monotransitive structures with remapping of the companion phrase. Locative- and goal-selecting ACs to a lesser extent also show some differences in properties. Causative functions are also attested for AMs, but not equally across all languages of the sample. Still, in all but three languages of the sample, at least one AM can also mark causative constructions. Finally, I show that the selection of independent voice or valency alternations plus applicatives modulates access of constituents expressing peripheral roles to positions which are syntactically privileged and prominent in discourse, and this is seen both in languages with symmetrical voice systems and those which have developed asymmetrical voice systems.

This chapter is organized as follows. §7.2 introduces the language sample. §7.3 presents a brief overview of the basic morphosyntax of these languages. §7.4 gives a general introduction to applicative affixes in languages of the sample. The next five sections describe ACs in these languages according to the semantic role of the applied phrase. These include beneficiaries and recipients (§7.5), instruments and themes (§7.6), goals and locations (§7.7), circumstantial roles and comitatives (§7.8), and other types of participants (§7.9). §7.10 describes aspectual and intensive meanings and other semantic effects marked by AMs. §7.11 describes causative AM-marked constructions in languages of the sample. In §7.12, the relationship between applicatives and other major voice (or valency) alternations is described. The chapter concludes with a summary of findings in §7.13.

7.2 Sampling of languages

A total of 24 languages are included in the sample for this chapter, as listed in Table 7.1.² The languages included were selected out of the 50 total languages with pivot-neutral applicatives in the larger sample for the typological survey (see §4.2). To be eligible for selection, either detailed descriptive accounts with robust numbers of examples of ACs or fairly comprehensive lexical resources with example sentences for a language were required. For each genetic grouping used in the typological survey, I attempted to include two languages from different primary branches, unless the grouping is comprised of more than 50 total languages, in which case, up to three languages were selected.

For a number of genetic groupings, no languages with adequate resources were found. These include Lampung, Enggano, and Wotu-Wolio. For Javanese languages only Standard Javanese [jav] was included, due to lack of examples of ACs for Tengger and Suriname Javanese. Acehnese, the sole Chamic language with applicatives, was excluded because its sole applicative morpheme *peu-* is primarily a causative marker and very few lexical verbs take applicative meanings when

²In Table 7.1, I use lower-level subgroups for genetic affiliation as outlined in §4.2 for the typological survey. Thus Nasal is treated as a separate genetic grouping from Northwest Sumatra-Barrier Islands (Toba Batak and Nias), even though these likely subgroup together under the proposed genetic affiliation Sumatran (Billings & McDonnell forthcoming). Sumatran and some other higher-level subgroups were mentioned in §6.2.3 in reference to reconstruction of specific AM forms.

Table 7.1: Language sample used for the functional typology

No.	Language	Gen. Grp.	Branch	Location	Sources
1	Toba Batak	NWS-BI	Batak	Sumatra	(van der Tuuk 1971 [1864-1867])
2	Nias	NWS-BI	Barrier Isl.	Barrier Isl.	(Brown 2001)
3	Nasal	NAS	none	Sumatra	(McDonnell fieldnotes)
4	W. Coast Bajau	GRB	Sama-Bajau	Borneo	(Miller 2007)
5	Yakan	GRB	Sama-Bajau	S. Philip.	(Brainard & Behrens 2002)
6	Kendayan	MAL	none	Borneo	(Adelaar 2005b)
7	S. Barisan Mal.	MAL	Malay	Sumatra	(McDonnell 2016)
8	Std. Indonesian	MAL	Malay	(wide use)	(Sneddon et al. 2010)
9	Sundanese	SUN	none	Java	(Truong fieldnotes)
10	Javanese	JAV	none	Java	(Oglobin 2005; Hemmings 2013; Vander Klok & Evans 2022)
11	Madurese	MAD	none	Madura	(Davies 2010)
12	Balinese	BSS	Bali	Lsr. Sundas	(Arka 2003; Artawa 1998)
13	Sasak	BSS	Sasak-Sumb.	Lsr. Sundas	(Khairunnisa & McDonnell in prep)
14	Pendau	T-T	Tomini	Sulawesi	(Quick 2007)
15	Ledo Kaili	K-P	Northern	Sulawesi	(D. Evans 2003)
16	Behoa	K-P*	Badaic	Sulawesi	(Shore 2016)
17	Balantak	S-B	Eastern	Sulawesi	(van den Berg & Busenitz 2012)
18	Mori Bawah	B-T	Eastern	Sulawesi	(Mead 1998)
19	Tolaki	B-T	Western	Sulawesi	(Edwards 2012)
20	Muna	M-B	Nuclear	Sulawesi	(van den Berg 2013)
21	Tukang Besi	M-B	Tukangbesi- Bonerate	Sulawesi	(Donohue 1999)
22	Bugis	SSUL	Bugis	Sulawesi	(Hanson 2003; Sirk 1983)
23	Makasar	SSUL	Makassar	Sulawesi	(Jukes 2020)
24	Duri	SSUL	Northern	Sulawesi	(Valkama 1993)

prefixed with *peu-*. For the Greater Barito linkage, only the Sama-Bajau languages have pivot-neutral applicatives; two are included from separate branches of the Sama-Bajau subgroup.

7.3 Basic morphosyntax

In terms of morphological typology, the languages of West Nusantara included in the sample range from moderately agglutinative (e.g. Pendau in which a single stem may show several prefixes and suffixes) to relatively isolating (e.g. Ampenan Sasak in which a given stem takes only up to one prefix and one suffix). These languages show little to no case-marking. As mentioned in §1.4.2 earlier, however, an important type of syntactic coding in these languages is the selection of pronominal forms indexing clausal arguments from particular sets, which co-varies with grammatical relations. Languages of the sample that show more use of case-marking include Nias, which uses morphophonological changes on nouns (nominal mutation) to indicate grammatical relations (Brown 2001), and Yakan and Tukang Besi, which make use of preposed case marking particles on NPs (Brainard & Behrens 2002; Donohue 1999). In most languages of the sample, however, lexical NP arguments are unmarked when core, and marked with a preposition when oblique. In the languages of the sample, arguments of the verb are often unrealized when their reference is recoverable in the discourse.

A summary of some relevant properties of voice and case marking systems for the languages of the sample is given in Table 7.2 below. In the remainder of this section, these properties will be discussed in further detail, including illustrative examples.

In the languages of the sample, intransitive predicates may be unmarked, or they may show marking with a number of different affixes according to their semantic properties. These include affixes indicating stative meaning, and dynamic, reciprocal, or non-volitional action, among others. In languages with two-way symmetrical and Philippine-type voice systems, the verb in a transitive clause is typically marked for voice. An example is given for Toba Batak in (148), showing that AV is marked with the prefix *maN-*, while PV is marked with the prefix *di-*. Ledo Kaili is another language that has distinctive affixes for each transitive voice, as shown in (149).

(148) Toba Batak, Voice alternations

- a. *Man-jaha buku guru i.*
AV-read book teacher DEF
'The teacher read a book.' (AV)
- b. *Di-jaha guru buku i.*
PV-read teacher book DEF
'A teacher read the book.' (PV)

(Schachter 1984: 127–128)

Table 7.2: Voice and case marking in languages of the sample

Language	Voice	Case marking	Major transitive alternations
Balantak	Philippine-type	limited (pronominals)	AV / PV / LV
Behoa	two-way symmetrical	limited (pronominals)	AV / PV
Duri	two-way symmetrical	limited (pronominals)	AV / PV
Javanese	two-way symmetrical	limited (pronominals)	AV / PV
Ledo Kaili	two-way symmetrical	limited (pronominals)	AV / PV
Madurese	two-way symmetrical	limited (pronominals)	AV / PV
Nasal	two-way symmetrical	limited (pronominals)	AV / PV
Pendau	two-way symmetrical	limited (pronominals)	AV / PV
Std Indonesian	two-way symmetrical	limited (pronominals)	AV / PV
Toba Batak	two-way symmetrical	limited (pronominals)	AV / PV
Ampenan Sasak	two-way symmetrical	none	AV / PV
Balinese	two-way symmetrical	none	AV / PV
Kendayan	two-way symmetrical	none	AV / PV
S. Barisan Malay	two-way symmetrical	none	AV / PV
Sundanese	two-way symmetrical	none	AV / PV
W. Coast Bajau	two-way symmetrical	none	AV / PV
Tukang Besi	two-way symmetrical	Y (preposed particles)	AV / PV
Yakan	Philippine-type (marginal)	Y (preposed particles)	Transitive / semi-transitive / IV
Bugis	two-way (marginal)	limited (pronominals)	Transitive / semi-transitive
Makasar	two-way (marginal)	limited (pronominals)	Transitive / semi-transitive
Mori Bawah	asymmetrical	limited (pronominals)	Transitive / semi-transitive
Muna	asymmetrical	limited (pronominals)	Transitive / semi-transitive
Tolaki	asymmetrical	limited (pronominals)	Transitive / semi-transitive
Nias	two way (marginal)	Y (mutation)	None in main clauses

(149) Ledo Kaili, Voice alternations

- a. *Soso nang-(k)ande loka.*
gecko AV-eat banana
'The gecko eats a banana.' (AV)
- b. *Loka ni-kande nu soso.*
banana PV-eat by gecko
'The banana is eaten by a gecko.' (PV)

(D. Evans 2003: 495, English translation & glosses added)

In both Toba Batak and Ledo Kaili, if A represents certain first person or second person pronominal categories, it may be realized as a pronominal prefix or proclitic that replaces the PV prefix. This is shown in (150) for Ledo Kaili, which only shows this type of actor indexing prefix in irrealis mode clauses.

(150) Ledo Kaili, Irrealis PV with bound-indexing for A

- Ku-kande tapia ngena.*
1SG.IRR.PV-eat mango later
'Later, I will eat a mango.' (PV)

(D. Evans 2003: 496, English translation & glosses added)

In a good number of languages with two-way symmetrical systems and marginal two-way systems, the P-oriented transitive construction is zero-marked. That is, it is distinguished by lack of voice morphology in comparison to A-oriented constructions, which are generally marked with one or more verbal prefixes. An example of this is given for West Coast Bajau in (151). Here the AV construction is marked with the prefix *N-*, while the PV construction is zero-marked. Note also the special form of the non-pivot A when pronominal. Non-pivot A in PV is encoded as an enclitic pronoun, i.e. =*ku* 'I', in (151), whereas all other core arguments are coded as unmarked independent pronouns and are not otherwise morphologically distinguished for case.

(151) West Coast Bajau

- a. *Aku boi m-(b)oo iyo pitu*
1SG CMPL AV-bring 3SG to.here
'I brought him/her here.' (AV)
- b. *Boi boo=**ku** iyo pitu*
CMPL [PV]bring=1SG.NPIV.A 3SG to.here
'I brought him/her here.' (PV)

(Miller 2007: 140)

Other languages with a zero-marked PV construction and a prefix-marked AV construction are Duri and Behoa. In Balantak, only agented realis PV is zero-marked, otherwise, PV verbs are marked with voice affixes. Kendayan shows prefixal marking for AV and optional marking of PV with a proclitic *di=* (see §5.8.2 on similar constructions in neighboring languages).

In some symmetrical voice languages, the non-pivot A has special coding when it is a lexical NP. Sundanese marks AV and PV constructions with prefixes *ng-* and *di-*, respectively, as in (152).

The A argument in PV is most commonly marked by a preposition *ku* ‘by’, as in (152b), though *ku* may be omitted when the actor immediately follows the verb (Kurniawan 2013). Kaili Ledo, as shown in (149b) above, exhibits similar marking of the non-pivot A with *nu* ‘by’.

(152) Sundanese, Voice alternations

- a. *Asep m-(b)euli baju.*
A. AV-buy clothes
‘Asep bought clothes.’
- b. *Baju di-beuli ku Asep.*
clothes PV-buy by A.
‘Asep bought clothes.’

(FM4-050)

In the Nasal examples in (153), PV is unmarked when A is first or second person, as in (153a), and optionally marked with a dedicated prefix when A is third person, as in (153b). This pattern is also found in South Barisan Malay.

(153) Nasal, PV marking

- a. *lahan ni kak khadu kam=suah.*
field that PFV finish 1PL.EXCL.NPIV=[PV]burn
‘We already burned the field.’
- b. *lahan ni kak khadu (di-)suah=nyo.*
field that PFV finish PV-burn=3SG
‘He already burned the field.’

(McDonnell fieldnotes)

Pendau shares similarities to Nasal, but also shows additional complexity in its system of verbal morphology, see §5.9.4.4. Pendau distinguishes realis and irrealis mode, which is reflected in two sets of voice prefixes. Like Nasal, Pendau also shows distinct forms of pronominal clitics used to index first and second person non-pivot A arguments in certain PV constructions. These appear in preverbal position, in which case no dedicated voice prefix is observed on the verb (see Quick 2007: 374-375).

Among the languages of the sample, Ampenan Sasak and Tukang Besi made be said to show the least verbal morphology marking symmetrical voice categories. As mentioned in §5.6, Ampenan Sasak, still shows a contrast between A-oriented and P-oriented constructions, but the predicate is typically unmarked; with the former constructions only optionally bearing the AV prefix *N-*. This constitutes a two-way symmetrical diathesis system, with the alternation between A-oriented and P-oriented clauses signaled through a combination of word order and coding of A arguments, including the use of clitic pronominal forms (see Khairunnisa 2022 for more details).

In Tukang Besi, similarly, there are two types of transitive clauses, neither of which show morphological marking on the verb. One type of transitive clause shows indexing of P on the verb, and the other does not. However in relative clauses, morphologically-marked distinctions akin to PV and AV do appear on the verb, and this co-varies with access to syntactic operations for the argument marked with *na* ‘nominative’ in the clause. Therefore, Tukang Besi is treated as a two-way symmetrical system by Donohue (2002), and also likewise in this study.

Some languages of the sample that show zero-marked transitive constructions in alternation with prefix-marked constructions have been described differently from the characterization above. One such language is Makasar (South Sulawesi) which has an unmarked, basic transitive clause alongside several marked constructions (Jukes 2020). In the zero-marked basic transitive clause, there is no required verbal morphology signalling voice or valency, but A and P arguments are indexed with pronominal clitics, as in (154).

- (154) Makasar, Transitive construction
ku=kanre=i taipa=nu
 1=eat=3 mango=2FAM.POSS
 ‘I eat your mangoes.’ (Jukes 2020: 257)

In other constructions, the verb in Makasar may be affixed with one of a number of verbal prefixes that are likely cognate with voice prefixes in the other languages of the sample. However, Jukes (2013) analyzes these as *valency-signaling* prefixes. The verbal prefixes *aN(N)-* and *aN-* are analogous to AV. Of these two, only *aN(N)-* triggers nasal substitution of the first consonant of the root, while *aN-* does not. Jukes analyzes *aN(N)-* as a marker of a so-called semi-transitive clause. Unlike the zero-marked basic transitive clause, in clauses marked with *aN(N)-*, P is not indexed on the verb with an enclitic pronoun, and P may not be definite. As example is given below in (155).

- (155) Makasar, Semi-transitive construction
angng-(k)anre=a' taipa
 STR-eat=1 mango
 ‘I eat mangoes’ (Jukes 2020: 257)

Jukes (2020) analyzes the verbal prefix *aN-* as the marker of a type of clause he calls Actor Focus. An example is shown in (156). In this construction, the A argument must occur in the preverbal position but is not indexed on the verb as a proclitic, while P is indexed as a pronominal enclitic. See also §6.5.3 and Jukes (2020: 240-242, 330-333) for further discussion of patterns of argument indexing in Makasar and grammatical relations.

- (156) Makasar, Actor-focus construction
kongkong am-buno=i_i miong=ku_i
 dog AF-kill=3 cat=1.POSS
 ‘a dog killed my cat.’ (Jukes 2020: 269)

Bugis, another South Sulawesi language, is analyzed by D. Laskowske (2016) as a two-way symmetrical voice system. Under such an analysis, the zero-marked basic transitive clause is considered a PV construction, while the semi-transitive and so-called Actor Focus constructions are AV constructions, albeit showing lower semantic transitivity than that of PV. Yakan (Sama-Bajau), shows similar alternations that likewise may be treated as AV and PV constructions, or a zero-marked transitive P-oriented construction and a semi-transitive or antipassive construction marked with *mag-* (plus less productive pivot-selecting ACs). For my purposes in this chapter, it is not necessary to definitively categorize languages by type of voice system or declare some

of these to be ‘truly’ symmetrical. Still, I seek to demonstrate that while the languages in the sample have diverse systems for marking transitive constructions, there are often clear parallels in the types of voice and valency alternations observed, and clear similarities in the ways that applicatives function within these systems.

Languages of the sample with asymmetrical voice systems include Muna, Mori Bawah, Tolaki, and Nias. Nias does not show verbal alternations of the type we have been discussing in main clauses, while the three Sulawesi languages in this set do. These languages show an alternation in the form of morphological marking or argument indexing on the verb depending on the definiteness of the P argument, which has been called “definiteness shift” (van den Berg 1995). However, unlike symmetrical voice languages, there is little evidence from syntactic behavioral properties by which one such alternation may be considered A-oriented and the other P-oriented. An example is given in (157) below from Muna, where for realis clauses, the third person singular A index *ae-* is used with an indefinite P argument, while the corresponding A index *a-* is used with a definite P argument. Note that the choice of *a-* or *ae-* class indexes does not co-vary with access to syntactic operations for types of core arguments; only S or A may head a relative clause with an *a-* class subordinate verb form (a participial marked with *-um-* *-no*) and likewise only S or A may head a relative clause with an *ae-* class subordinate verb form (a participial marked with *me-* *-no*) (van den Berg 2013: 232–234).

(157) Muna, Definiteness shift

- a. *Ae-alo-mo kapulu.*
3SG.RLS-take-PERF machete
‘I took a machete.’
- b. *A-ala-mo kapulu-ku.*
3SG.RLS-take-PERF machete-1SG.POSS
‘I took my machete.’

(van den Berg 1995: 169)

In Mori Bawah and Tolaki, in transitive clauses with definite P arguments, the verb is zero-marked, and P is obligatorily indexed on the verb. In corresponding clauses with indefinite P arguments, the verb is prefixed with *poN-* in Mori Bawah (*po-* in Tolaki) which is glossed as antipassive. In neither case is P syntactically privileged. As ‘antipassive’ constructions in Mori Bawah and Tolaki, and *ae-* class indexed clauses in Muna are less semantically transitive than zero-marked or *a-* class indexed clauses, these three languages also show similarities to Makasar and Bugis.

In addition to transitive alternations, many languages of the sample have a ‘true passive’ construction with the exception of Pendau, Toba Batak, and possibly Yakan. In some of these, the passive construction is marked with a dedicated affix, as with West Coast Bajau (*-in-*), Amepenani Sasak (*te-*), and Makasar (*ni-*). In other languages, as in Nasal and Sundanese, the passive construction may be marked with the PV prefix, however, see Chen & McDonnell (2019) on difficulties in distinguishing PV and passive constructions when both are marked with the same verbal morphology).

Finally, some languages of the sample, most notably Pendau, make use of a number of *stem-former* prefixes. These prefixes do not express semantic content of their own, but instead function to form an augmented stem which is grammatically required for further affixation with voice pre-

fixes, applicative suffixes, and other derivational morphology (see Quick 2007: 99–108 and discussion in §5.9.4.4). Stem-former prefixes in the languages of the sample typically have shapes like *pong-*, *po-*, *pe-*, and *popo-*, alongside allomorphs showing vowel harmony. Such prefixes are also found in Behoa and Kaili Ledo, though their use in these languages is perhaps less pervasive than in Pendau. Some prefixes with the same shape have been treated as transitivizers or transitive markers by some authors, and some are also treated as causative prefixes (see §7.11). Stem-former prefixes are commonly observed in Austronesian languages of the Philippines and Sulawesi (see e.g. Himmelmann & Wolff 1999).

7.4 Applicative morphology

Each language of the sample has between one and four AMs that mark ACs, which are listed in Table 7.3. In four languages—West Coast Bajau, Yakan, Tolaki, and Ampenan Sasak—there is only one AM. Most languages of the sample (14 of 24) have two AMs, one marking selection of beneficiaries, instruments, and/or themes as the applied phrase, and the other marking selection of locations and goals as the applied phrase. For applied phrases with other semantic roles such as stimulus or content, however, there is no such specialization across AMs (see §7.9). Five languages of the sample show three AMs—Nias, Sundanese, Kaili Ledo, Balantak, and Tukang Besi—and one language of the sample—Mori Bawah—shows four. In such languages the additional AMs can be unique innovations, e.g. Tukang Besi comitative *-ngkene*, may arise from an apparent split of an AM, e.g. Mori Bawah *-Cako* and *-ako*, or can represent use of primarily causative prefix with applicative functions, i.e. Nias *fa-* (see discussion in §6.2.1.3).

The role expressed by the applied phrase in ACs in these languages typically may be expressed as an oblique in a corresponding BC with the same verbal stem without AM-marking. Such obliques are most commonly encoded as a prepositional phrase. Thus, ACs in these languages are predominantly *optional* applicatives (see discussion in §3.1.3.1). There are a few cases, however, in which there is no (monoclausal) equivalent BC, which will be noted below in the relevant sections. Throughout the following sections, I also make reference to the semantic participant expressed by the P argument in a BC, and its syntactic realization in the corresponding AC. As discussed in §3.1.3.4, I refer to this participant as the *companion phrase* because it generally may be expressed alongside the applied phrase in the AC, and may or may not be encoded as a core clausal argument.

7.5 Beneficiaries and recipients

7.5.1 Distribution and general properties of beneficiary-selecting ACs

All but one languages in the sample (Nias) has an applicative construction in which a beneficiary role is selected to map to the applied phrase. In many cases, this same construction is also found with a recipient role expressed as the applied phrase, though beneficiary and recipients are not consistently distinguished in descriptive source material.

The languages of the sample vary in terms productivity of the beneficiary-selecting applicative and compatibility with various bases. In all the languages of the sample, benefactive applica-

Table 7.3: Applicative morphology by semantic role of the applied phrase

Language		Single Form				
		BEN	INST	THM	LOC	GOAL
W. Coast Bajau	<i>-an</i>	✓	✓	✓	✓	✓
Yakan	<i>-an</i>	✓		✓		
Sasak	<i>-an</i>	✓	✓		✓	✓
Tolaki	<i>-Cako</i>	✓	✓			

Language		Form 1			Form 2		
		BEN	INST	THM	LOC	GOAL	
Toba Batak	<i>-hon</i>	✓	✓	✓	<i>-i/an</i>	✓	✓
Nasal	<i>-kun</i>	✓	✓	✓	<i>-i</i>	✓	✓
Kendayan	<i>-an</i>	✓		✓	<i>-i</i>	✓	✓
S. Barisan Mal.	<i>-ka</i>	✓	✓	✓	<i>-i</i>	✓	✓
Std. Indonesian	<i>-kan</i>	✓	✓	✓	<i>-i</i>	✓	✓
Javanese	<i>-aké</i>	✓	✓	✓	<i>-i</i>	✓	✓
Madurese	<i>-agi</i>	✓	✓	✓	<i>-e</i>	✓	✓
Balinese	<i>-ang</i>	✓	✓	✓	<i>-i</i>	✓	✓
Pendau	<i>-a'</i>	✓	✓	✓	<i>-i</i>	✓	✓
Behoa	<i>-á</i>	✓			<i>-i</i>	✓	✓
Muna	<i>-ghoo</i>	✓	✓		<i>-i</i>	✓	✓
Bugis	<i>-Ceng</i>	✓	✓	✓	<i>-i</i>	✓	✓
Makasar	<i>-ang</i>	✓	✓	✓	<i>-i</i>	✓	✓
Duri	<i>-an</i>	✓	✓	✓	<i>-i</i>	✓	✓

Language		Form 1			Form 2			Form 3
		BEN	INST	THM	LOC	GOAL		
Nias	<i>-'ö</i>			✓	<i>-(C)i</i>	✓	✓	<i>fa-</i> THM, GOAL
Sundanese	<i>-keun</i>	✓	✓	✓	<i>-an</i>	✓	✓	<i>pang- -keun</i> BEN
Kaili Ledo	<i>-ka</i>	✓			<i>-i</i>	✓	✓	<i>-aka</i> THM, INST
Balantak	<i>-kun</i>	✓	✓	✓	<i>-i</i>	✓	✓	<i>-ii</i> BEN
Tukang Besi	<i>-ako</i>	✓	✓	✓	<i>-(VC)i</i>	✓	✓	<i>-ngkene</i> COM

Language		Form 1			Form 2			Forms 3 & 4
		BEN	INST	THM	LOC	GOAL		
Mori Bawah	<i>-ako</i>	✓	✓		<i>-(C)i</i>	✓	✓	<i>-Cako</i> <i>-Cari</i> THM GOAL, STIM

tives may attach to at least a subset of transitive bases, including roots with the meanings that express transfer, acquisition, and conveyance, e.g. ‘give’, ‘buy’, and ‘bring’, and acts of creation, e.g. ‘make’, ‘build’. In about half of the languages of the sample, beneficiary-selecting applicatives also attach to transitive roots expressing processes, such as ‘wash’, ‘grind’, and ‘pound’. In only a few languages are beneficiary-selecting applicatives attested with bases expressing sensory or cognitive events, e.g. Sundanese *pang-ambeu-keun* ‘to smell for s.o.’, Sasak *dengah-an* ‘to hear for s.o.’, or with intransitive bases expressing activities, e.g. Muna *ne-lagu-lagu-ghoo* ‘to sing to/for s.o.’, Behoa *mam-pe-kakae-á* ‘to pray for s.o.’.

Example (158) shows an beneficiary-selecting AC from Nasal with the suffix *-kun*.

(158) Nasal, Beneficiary-selecting AC with *-kun*

a. *Azma ny-(s)anik buwak gin anak=nyo.*

A. AV-make snack for child=3SG
 ‘Azma made snacks for her children.’ (BC)

b. *Azma ny-(s)anik-kun anak=nyo buwak.*

A. AV-make-BEN.APPL child=3SG snack
 ‘Azma made her children snacks.’ (AC)

(McDonnell fieldnotes)

In the BC in (158a), the beneficiary is expressed as a PP, as is typical of most languages of the sample. In Makasar and Salako, the expression of a beneficiary with a PP in this manner is only possible with an apparently borrowed preposition *untu’/untuk* ‘for’ from Indonesian. Jukes (2020: 315) reports that only younger speaker of Makasar use this borrowed preposition. Expressing the beneficiary as a PP is also possible in Ampenan Sasak with a preposition *umaq*, but this is apparently rare and found only in elicited examples (Khairunnisa & McDonnell in prep).

Sundanese shows two different benefactive applicatives: *-keun* and *pang- -keun*. The prior is much less productive as a benefactive applicative. It is restricted to a small number of the transitive bases, and takes a beneficiary applied phrase, as in (159), repeated from (56) in §2.8. The suffix *-keun* may also act as a causative, as in (160).

(159) Sundanese, Beneficiary-selecting AC with *-keun*

a. *Udi m-(b)uka panto keur kuring.*

U. AV-open door for 1S
 ‘Udi opens the door for me.’

b. *Udi m-(b)uka-keun kuring panto*

U. AV-(b)uka-BEN.APPL 1S door
 ‘Udi opens the door for me.’

(Hanafi 1997: 23)

(160) Sundanese, Causative construction with *-keun*

a. *Jandela peupeus.*

window break
 ‘The window breaks / is broken.’ (BC)

- b. *Abi m-(p)eupeus-keun jandela.*
 1SG AV-break-CAUS window
 ‘I broke the window.’ (Causative) (CT1-017)

As discussed in §2.7, Sundanese beneficiary-selecting ACs with *pang-* *-keun* have a substitutive benefactive meaning. They are also much more productive than beneficiary-selecting ACs with *-keun*. Benefactive ACs with *pang-* *-keun* are found with a wide range of intransitive and transitive bases. When *pang-* *-keun* attaches to an intransitive base, such as *peupeus* ‘break’ in (160a), the AC is *not* monotransitive as we might expect. Instead as shown in (161), the AC is ditransitive and has both a causative meaning, with A expressing an instigating causer, and an applicative meaning, with one non-A core argument expressing the beneficiary and the other expressing a patient or theme role.

- (161) Sundanese, Beneficiary-selecting *pang-* *-keun*
Euis di-pang-meupeus-keun kalapa ku abi.
 E. PV-BEN.APPL-break-BEN.APPL coconut by 1SG
 ‘I broke open a coconut for Euis.’ (AC) (CT1-017)

Note that I do not analyze the prefixal element of the circumfix *pang-* *-keun* as a beneficiary-selecting AM in its own right. Prefixal *pang-* is not observed to occur on its own as a benefactive AM, and with transitive bases *pang-* *-keun* does not necessary show a causative meaning, see example (168) below. (See also §6.2.1.3 on possible origins of this circumfix.)

Balantak also shows two AMs that form beneficiary-selecting ACs: the suffixes *-kon* and *-ii*. ACs formed with these two suffixes show different structural properties. As shown in (162), benefactives ACs marked with *-kon* are always monotransitive. The companion phrase (patient) is encoded as a core argument (P), while the beneficiary applied phrase is encoded as a possessor phrase modifying the companion phrase as in *gala-ni Goris*, where *-ni* indicates a following personal possessor NP (see also §6.5.4).

- (162) Balantak, Beneficiary-selecting AC with *-kon*
- a. *Sina-gku man-taring gala boni Goris.*
 mother-1SG AV.IRR-cook vegetables for.PA G.
 ‘My mother is cooking vegetables for Doris.’
- b. *Sina-gku man-taring-kon gala-ni Goris.*
 mother-1SG AV.IRR-cook-BEN.APPL vegetables-3SG.PA G.
 ‘My mother is cooking vegetables for Doris.’ (van den Berg & Busenitz 2012: 100)

In contrast, beneficiary-selecting ACs marked with *-ii* are ditransitive, and both the companion phrase and the beneficiary applied phrase show are encoded as unmarked core arguments. In such ACs, only the beneficiary may become the subject in PV, as in (163) (see van den Berg & Busenitz 2012: 107).

(163) Balantak, Beneficiary-selecting AC with *-ii*

Tama-ngku ni-wawau-ii-mo wala'on.
 father-2SG PV.RLS-make-BEN.APPL-PERF boiled.water
 'My father has already been made a hot drink.' (van den Berg & Busenitz 2012: 108)

Aside from the Balantak benefactive ACs with *-kon*, in all languages of the sample, beneficiary-selecting ACs may be ditransitive. While only a few descriptions (i.e. Quick 2007, Miller 2007, Jukes 2013) discuss the core/oblique status of the beneficiary applied phrase and patient/theme companion phrase, in the languages of the sample, generally both may be considered core since they may appear unmarked (i.e. not in a PP) when expressed as a lexical NP. This applies in all 23 languages of the sample in which this construction occurs. Note that while benefactive ACs are maximally ditransitive in these language, more than one structure is sometimes possible, and not all of these are ditransitive (see §7.5.3).

7.5.2 Beneficiary-selecting ACs with special properties

Beneficiary arguments are indexed differently than P arguments in BCs in three languages of the sample, all spoken in Sulawesi: Muna, Mori Bawah, and Tolaki. In these languages, there is a special form of indexing on the verb for beneficiaries and recipients. These forms are also used uncommonly with other types of peripheral participants (see §6.5.3.2 for more on the special indexing suffixes in these languages). These morphemes are analyzed as fused forms of an applicative suffix and suffixal indexes for objects or absolutive arguments, though there are some complications, and they may show the same shape as the object/absolutive suffixes in some person and number categories.

Examples from Muna are shown in (164) with the verb *gholi* 'buy'. In the BC in (164a), the verb is not suffixed. The theme immediately follows the verb, and is encoded as an unmarked NP *bhadhu* 'shirt', while the beneficiary, if expressed, is encoded as a PP consisting of a preposition followed by a free pronoun or NP, in this case *so insaidi* 'for us'. In a benefactive AC with an NP beneficiary, as shown in (164b), the verb is suffixed with *-ghoo*, and the beneficiary is expressed as an unmarked NP, *ina-ku* 'my mother'. The beneficiary applied phrase precedes the theme which is also an unmarked NP, *o pae* 'rice'.³ Alternately, as shown in (164c), when the beneficiary is pronominal, a special form indexing the beneficiary appears on the verb, in this case *-angko* 'for you', again followed by an unmarked NP expressing the theme. The suffix *-ghoo* does not co-occur with 'indirect object' argument indexes like *-angko*; they are in complementary distribution.

(164) Muna, Beneficiary-selecting ACs with *ae-* class prefixes

- a. *ama-mani ne-gholi bhadhu so insaidi*
 father-1PL.EXCL.POSS 3SG.RLS-buy shirt for 1PL.EXCL
 'Our father bought a shirt for us.' (BC) (van den Berg 2013: 82)
- b. *ae-gholi-ghoo ina-ku o pae*
 1SG.RLS-buy-BEN.APPL mother-1SG.POSS ART rice
 'I buy rice for my mother.' (AC, NP beneficiary) (van den Berg 2013: 176)

³The article *o* is not used with NPs immediately following the verb in Muna (van den Berg 2013: 102).

- c. *ae-gholi-angko* *pae*
 1SG.RLS-buy-BEN.APPL:2SG.IO rice
 ‘I buy rice for you.’ (AC, pronominal beneficiary) (van den Berg 2013: 70)

In the examples above, *gholi* is used with *ae-* class actor indexing. In BCs the *ae-* class prefixes are always used with an indefinite P argument, which is not indexed on the verb with a direct object suffix. This is shown in (164a) above where the theme (‘rice’) is indefinite and not indexed. Interesting, in beneficiary-selecting ACs in Muna, the companion phrase in the AC still obeys the definiteness constraint; if *ae-* class marking is used, the theme must be indefinite and is not indexed on the verb, though the beneficiary may be definite and is indexed of the verb with indirect object suffixes when pronominal.

In BCs, if *gholi* ‘buy’ is used with *a-* class actor indexing, P must be definite and is indexed on the verb with a direct object suffix when pronominal. This is shown in (165a), where the theme argument is indexed on the end of the verb stem with *-e* ‘it’. In beneficiary-selecting ACs with *a-* class indexing of A, both the beneficiary and the theme are indexed on the verb when pronominal, as in (165b). The beneficiary applied phrase is indexed using the indirect object suffix form, and the theme companion phrase is indexed using the direct object suffix.

(165) Muna, Beneficiary-selecting ACs with *a-* class prefixes

- a. *a-gholi-e* *so ihintu*
 1SG.RLS-buy-3SG for 2SG
 ‘I bought it for you’ (AC, pronominal beneficiary) (van den Berg 2013: 143)
- b. *a-gh<um>oli-angko-e*
 1SG-<IRR>buy-BEN.APPL:2SG.IO-3SG
 ‘I will buy it for you (AC, both objects pronominal) (van den Berg 2013: 71)

In Tolaki, beneficiary and recipient roles are always indexed on the verb in ACs, regardless of whether they are pronominal or expressed as a full NP. Examples are shown in (166). A special set of pronominal indexes (glossed as dative) is used, and these forms are distinct from the absolutive forms used to index P in BCs, at least in some person categories.

(166) Tolaki, Indexing of beneficiary applied phrases

- a. *Kuposusuanggee* *banggonannggu.*
Ku-po-susuaN-kee *banggona-nggu*
 1SG.NOM-INDEF.P-sing-BEN.APPL:3.DAT friend-1SG.POSS
 ‘I sang for my friend.’ (AC)
- b. *Ku-tidu-’i-ko’o.*
 1SG.NOM-punch-3.ABS-BEN.APPL:2.DAT
 ‘I’ll punch him for you/ I’ll get him (back) for you.’ (AC) (Edwards 2012: 54)

In Mori Bawah, beneficiary-selecting constructions are marked on the verb with the AM *-ako*. Beneficiary (or recipient) roles are indexed on the verb in applicative constructions. The index follows the suffix *-ako*, and has fused with it, except in third person forms. When a beneficiary or recipient is indexed on an AM-marked verb, normal indexing of the patient/theme on the verb

lapses, and the companion phrase is not indexed. This is shown in the example below with the verb *ala* ‘take’.

(167) Mori Bawah, Beneficiary-selecting AC

- a. *i-potae bange andio: tewala kanatuu, Puu-puu,*
 3SG-say monkey this when like.that Pigeon
io bou-mu-mo koa ku-'ala-o
 CN fish-2SG.POSS-PFV just 1SG.NOM-take-3SG.ABS
 ‘The monkey said: “In that case, I’ll take your fish.’ (BC) (Esser 2011: 104)
- b. *kuri'a-no ala-akita balu-balu-mu ka to-kita-o*
 say-3SG.POSS get-APPL:1PL.INCL RDP-goods-2SG.POSS and 1PL.INCL.NOM-see-3SG.ABS
 He says, ‘Come, fetch your wares for us, that we may see them.’ (AC)
 (Esser 2011: 322)

7.5.3 Properties of the applied phrase and interactions with voice/valency

Across languages of the sample some variance is observed in the interaction of beneficiary-selecting ACs with symmetrical voice and other valency alternations. In general, the languages of the sample showing symmetrical voice allow benefactive applicatives in AV. However, in Toba Batak, it is reportedly rare for beneficiary-selecting ACs specifically to be used with AV (van der Tuuk 1971 [1864-1867]: 105). PV forms are favored instead, and these are obligatorily marked with the prefix *pa-* in addition to the applicative suffix *-hon*. Sundanese has a general restriction on ditransitive constructions, which must occur in PV for the large majority of verbal bases. In the example in (168) below, the beneficiary is the pivot argument (R) and the theme, *balanjaan* ‘shopping purchases’, is an additional non-A core argument (T).

(168) Sundanese, Possessor beneficiary

- a. *Icih ny-(c)okot duit.*
 I. AV-take money
 ‘Icih took the money.’ (BC) (FM4-050)
- b. *Indung di-pang-nyokot-keun balanja-an ku Udi.*
 mother PV-BEN.APPL-take-BEN.APPL shopping-NMLZ by U.
 ‘Udi took the shopping purchases (in the house) for mother.’ (AC) (CT1-025)

When Sundanese beneficiary ACs occur in AV, they most often show monotransitive clausal structure with the companion phrase (theme or patient) selected as the P argument. The beneficiary participant, if overtly mentioned, is most commonly expressed as the possessor of the companion phrase, which is shown in (169) below. In other cases, the beneficiary is also commonly unrealized when understood from context or may be overtly expressed as a PP (see examples and discussion in §2.7, analogous Nasal examples in (180)–(181) below).

(173) West Coast Bajau, Beneficiary-selecting AC in PV and passive

- a. *Boi sembali emma'=ku kambing e ta' Saiman.*
 CMPL [PV]slaughter father=1SG.NPIV goat DEM LOC S.
 'My father slaughtered the goat for Saiman.' (BC)
- b. *Boi sembali-an emma'=ku Saiman kambing tu.*
 CMPL [PV]slaughter-BEN.APPL father=1SG.NPIV S. goat DEM
 'My father slaughtered (for) Saiman the goat.' (AC)
- c. *Saiman boi s<in>embali-an kambing le' emma'=ku.*
 S. CMPL <PASS>slaughter-BEN.APPL goat LOC father=1SG.POSS
 '(For) Saiman was slaughtered a goat by my father.' (AC) (Miller 2007: 278-280)

In Nasal and West Coast Bajau, the patient companion phrase does not appear to be a possible pivot argument in either PV or passive constructions (see Miller 2007: 280 for discussion of West Coast Bajau). In Pendau, Quick (2007) reports that either the applied phrase or the patient companion phrase may occur in the preverbal position, as in (174). For such examples, the beneficiary applied phrase *io* 3SG.ABS is the pivot argument in either case.

(174) Pendau, Recipient applied phrase in PV

- a. *Io nipogabua'o'u vea.*
io ni-po-gabu-a'='u vea
 3SG PV.RLS-SF-COOK-BEN.APPL=1SG.GEN raw.rice
 'I cooked rice for him/her.' (AC)
- b. *Vea nipogabua'o'u io.*
vea ni-po-gabu-a'='u io
 raw.rice PV.RLS-SF-COOK-BEN.APPL=1SG.GEN 3SG
 'I cooked rice for him/her.' (Passive AC) (Quick 2007: 292)

Even some languages which have distinct patterns of indexing of the applied recipient or beneficiary with 'dative' or 'indirect object' indexes, the recipient or beneficiary may be the privileged syntactic argument of a P-oriented construction. This is reported for Muna and Mori Bawah in the sample, and an example is given from Muna below showing a passive participial phrase marked with *ne-*.

(175) Muna, Beneficiary-AC, passive participle

- aini-ha-e-mo robhine ne-owa-ghoo-ku se-tuwu bheta*
 this-PRED-3SG.POSS-PERF woman PASS.PART-bring-BEN.APPL-1SG.NPIV.A one-CL sarong
 'this is the woman to whom I have taken a sarong' (van den Berg 2013: 234)

In Tolaki, it appears that the dative marked beneficiary may not be the privileged syntactic argument of a passive construction (see Edwards 2012). In Balantak, the same is true of applied beneficiaries in ACs marked with *-kon*; only the patient/theme companion phrase may become the pivot of a PV clause (van den Berg & Busenitz 2012: 97). However, in beneficiary-selecting

ACs marked with *-ii* in Balantak, as mentioned above, the beneficiary is the pivot in PV, not the companion phrase.

In Makasar, both arguments in zero-marked transitive constructions are indexed on the verb, as in the BC in (176a). Regardless of whether a verb bears an AM or not, a maximum of two arguments may be indexed on the verb. In ditransitive benefactive ACs, the A argument and the beneficiary applied phrase are indexed on the verb (Jukes 2020: 314-315), while the companion phrase is not. This is shown in the AC in (176b), where the verb is suffixed with the AM *-ang*, and the beneficiary applied phrase is indexed on the verb with the 2nd person enclitic *=ko*, while the theme companion phrase is not indexed.

(176) Makasar, Indexing of beneficiary applied phrase

a. *ku=balli=i baju=a*
 1=buy=3 shirt=DEF
 ‘I bought the shirt.’ (BC)

b. *ku=balli-ang=ko baju.*
 1=buy-BEN=2FAM shirt
 ‘I bought you a shirt.’ (AC)

(Jukes 2020: 314-315)

Likewise, in *Tukang Besi*, in one of the two transitive constructions, P is indexed on the verb. In such constructions, only the beneficiary, not the companion phrase, may be indexed on the verb in a beneficiary-selecting AC marked with the AM *-ako* (Donohue 1999).

As mentioned above, in *Mori Bawah*, only the beneficiary may be indexed on the verb in a beneficiary-selecting AC in transitive constructions, while in *Tolaki* and *Muna*, both beneficiary and patient or theme may be indexed on the verb concurrently in transitive constructions, provided that the companion phrase is definite. In constructions that are less semantically transitive or ‘antipassive’ constructions, the companion phrase is indefinite in these languages, but the beneficiary may be definite and can be indexed.

(177) *Mori Bawah*, Beneficiary-selecting AC in antipassive

Aku mon-tena mo-wawa-akontu inahu eu.
 1SG.FUT PART:APASS-command PART:APASS-bring-BEN.APPL:2PL.ABS vegetable spinach
 ‘I will send someone to bring you spinach.’ (Mead 2005: 704)

According to Hanson (2003: 198–201), in *Bugis*, either the beneficiary applied phrase or the patient/theme companion phrase but not both, may be indexed on the verb in a zero-marked transitive construction, with the beneficiary being preferred. The choice appears to be conditioned by pragmatic considerations. Examples are given in (178) and (179). This pattern is not reported for any other languages of the sample. Hanson (2003: 201) also notes that beneficiary-selecting ACs marked with *-Ceng* do not occur if both the beneficiary and the patient/theme are indefinite.

(178) *Bugis*, Indexing of companion phrase in benefactive AC

a. *Magai deqna nanre ri bola-e?*
 why NEG rice LOC house-DEF
 ‘Why is there no rice in the house?’

- b. *Ani na-nasu-ng-ngi iko.*
 A. 3.ERG-COOK-BEN.APPL-3.ABS 2SG
 ‘Ani (has already) cooked it for you.’ (Hanson 2003: 198)

(179) Bugis, Indexing of beneficiary applied phrase in benefactive AC

- a. *Pakkogai iko m-anre?*
 how 2SG AV-eat
 ‘How did you eat?’
- b. *Ani na-nasu-k-kaq nanre.*
 A. 3.ERG-COOK-BEN.APPL-1.ABS rice
 ‘Ani cooked me rice.’ (Hanson 2003: 198–199)

As mentioned in §7.3, arguments of the verb are commonly unrealized in the languages of the sample. This appears to apply to the beneficiary applied phrase in most languages of the sample. In Ampenean Sasak, in fact, it is most common for the beneficiary or recipient applied phrase in such ACs to be unrealized (Khairunnisa & McDonnell in prep). In such cases, note that even though the beneficiary applied phrase may be unrealized, a benefactive meaning is still present in the clause. For example, in the Nasal AC in (180), a beneficiary is inferred from context. In Sundanese, this non-realization of the applied beneficiary is also possible. Most authors of available source material for other languages of the sample do not report whether this is possible. Donohue (1999: 232–234) however, notes that it is not possible to omit or delete the recipient applied argument as an unspecified referent in *Tukang Besi*.

(180) Nasal, Unrealized beneficiary applied phrase

- Azma ny-(s)anik-kun buwak.*
 A. AV-make-BEN.APPL snack
 ‘Azma made snacks (for them).’ (AC) (McDonnell fieldnotes)

In beneficiary-selecting ACs in Nasal, it is also possible for the beneficiary/recipient applied phrase to occur as an oblique PP even though the verb is suffixed with the benefactive AM *-kun*, as in (181). This pattern is also found in Sundanese, Pendau, and Indonesian (see §6.5.4).

(181) Nasal, Preposition marked beneficiary applied phrase

- Azma ny-(s)anik-kun buwak gin anak=nyo.*
 A. AV-make-BEN.APPL snack for child=3SG
 ‘Azma made snacks for her child.’ (AC) (McDonnell fieldnotes)

For beneficiary-selecting ACs in Nasal, it is thus possible for the beneficiary applied phrase to be expressed as a PP with an AM on the verb, as in (181), or without an AM, as in (158a). Crucially though, if the beneficiary argument is expressed as an unmarked NP, the AM *-kun* must be suffixed on the verb, as in (158b). Furthermore, if *-kun* is suffixed on the verb and the beneficiary of recipient applied phrase is unrealized, the AC must be interpreted as benefactive, as in (180); that is to say, the existence of a beneficiary participant that is distinct from the referent

of the A argument and that of the companion phrase is inferrable from clauses like (180). These properties together constitute good evidence that -kun indeed functions to form an AC in all such cases, despite the fact that the beneficiary may be encoded as a PP or be unrealized. The same appears to hold for Sundanese, Pendau, and Indonesian.

7.6 Instruments and themes

In this section, I present ACs in which an instrument or theme is selected as the applied phrase. As noted in the discussion of instrument- and theme-selecting applicatives in Sundanese (§2.5), it is not always possible to distinguish clearly instruments from themes, as instruments are often directed into motion, and thus may also be classified as themes. In the discussion that follows, an entity is treated as a instrument when it appears to be manipulated to some effect or purpose by an agent, rather than merely moved or directed into motion. Instrument-selecting and theme-selecting applicative constructions are taken up in turn below.

7.6.1 Instrument-selecting ACs

Of the 24 languages of the sample, 20 have an AC that selects an instrument applied phrase, while four do not: Nias, Behoa, Ampenan Sasak, and Kendayan. In almost all languages in which instrumental ACs are attested, the AM used is the same as the one that selects a beneficiary/recipient applied phrase.⁵ In Kaili-Ledo, the suffix *-(C)aka* marks instrumental ACs, while *-ka* marks benefactive ACs. The base verb in an instrument-selecting AC is commonly transitive. Base verbs meaning ‘to hit’, ‘to strike’, ‘to chop’, and ‘to buy’ are frequently found, and also attested are ‘to stab’, ‘to shoot’, ‘to sew’, ‘to scoop up’, ‘to dig’, ‘to pay for’, ‘to write’, ‘to make’, and ‘to kill’. In a few cases, intransitive base verbs are also used, such as ‘to go’, and ‘to return home’, which are found in *Tukang Besi* and *Muna* in instrumental ACs with meanings such as ‘to go by means of (vehicle)’.

Two examples of instrumental ACs are given below. The Sundanese example in (182) is repeated from (40) in §2.5.2.

(182) Sundanese, Instrument-selecting AC

- a. *Udi ny-(c)oel sambel maké témpé.*
 U. AV-scoop chili.sauce AV.use soybean.cake
 ‘U. scooped up chili sauce using (a piece of) soybean cake.’ (BC)
- b. *Udi ny-(c)oel-keun témpé kana sambel.*
 U. AV-scoop-INST.APPL soybean.cake onto chili.sauce
 ‘U. used (a piece of) soybean cake to scoop up chili sauce.’ (AC) (CT1-020)

⁵In Sundanese, the *-keun* suffix marks instrumental ACs as well as benefactive ACs with certain verbs, while *pang-* *-keun* exclusively marks benefactive ACs. In Balantak, *-kon* marks both instrumental and benefactive ACs, while *-ii* exclusively marks benefactive ACs.

(183) Toba Batak, Instrument-selecting AC

- a. *mangombak tanggurung ni horbo dohot indalu*
AV.strike back of buffalo with rice.pestle
'to strike the back of a buffalo with a rice pounder' (BC)
- b. *mangombak-kon indalu tu tanggurung ni horbo*
AV.strike-INST.APPL rice.pestle onto back of buffalo
'to strike a rice pounder on the back of a buffalo' (AC)

(van der Tuuk 1971 [1864-1867]: 103, English translation & glosses added)

As shown in the above examples, in an instrument-selecting AC, the companion phrase is commonly a patient or goal (i.e., an endpoint of directed motion). In BCs, this patient/goal participant is the P argument. In a BC, the instrument role is realized as an oblique PP or adverbial phrase, if it is expressed. Instruments may also be expressed as the complement of the verb meaning 'to use, to wear', e.g. Sundanese *maké*. In some languages, this verb appears to be in the process of grammaticalization to a preposition.

In two languages of the sample, Ledo Kaili and Duri, though instrumental ACs are attested, there are very few textual examples, and the syntactic properties of these constructions are unclear.

In about half of the remaining languages that show this type of construction, it appears that instrumental ACs are exclusively monotransitive. These include Toba Batak, Nasal, Balantak, South Barisan Malay, Indonesian, Balinese, Javanese, and West Coast Bajau. As seen in the AC examples from Sundanese and Toba Batak above, the instrument applied phrase is coded as an unmarked core argument and the companion phrase is expressed as an oblique PP rather than an unmarked NP. Following Zúñiga & Kittilä (2019), this type of construction is referred to as a *remapping* AC, as we observe a difference in the coding of the companion phrase compared to the BC. Sundanese follows the same pattern with a single exception; when *-keun* is suffixed to *beuli* 'buy', the AC is ditransitive.

The rest of the languages of the sample allow ditransitive instrumental ACs. These are primary found in languages of Sulawesi, and include Pendau, Tukang Besi, Bugis, Makasar, Mori Bawah, Tolaki, and Muna, as well as Yakan, which is spoken in the southern Philippines. However, note that even closely related languages may differ in the maximal transitivity of instrumental ACs, for example, Yakan allows such constructions to be ditransitive, while West Coast Bajau does not. In Madurese, instrumental ACs in AV are consistently monotransitive, but it appears that such constructions in PV may be optionally ditransitive (see Davies 2010: 309, and discussion in §6.5.4).

In Pendau, instrument-selecting ACs are consistently ditransitive and are only found in PV (see §5.9.4.4. The instrument applied phrase and the patient or goal companion phrase are both analyzed as core arguments of the verb. An example is given in (184).

(184) Pendau, Instrument-selecting AC

- a. *Pae rosunung nijimo nuuram.*
pae ro-sunung nijimo nu=uram
rice PV.IRR-burn 3PL.GEN INST=medicine
'They burned (or smoked) the rice with medicine (for medicinal purposes).' (BC)

- b. *Uram roposununua' nijimo pae.*
uram ro-po-sunung-a' nijimo pae
 medicine PV.IRR-SF.INST-burn-INST.APPL 3PL.GEN rice
 'They burned (or smoked) the rice with medicine (for medicinal purposes).' (AC)
 (Quick 2007: 297–298)

In the BC in (184a), the patient argument *pae* 'rice' is the pivot argument in PV. The instrument *uram* 'medicine' is not a core argument, but marked with *nu=*, the genitive case marker which is found on oblique instruments as well as possessors. In the AC in (184b), the instrument applied phrase is the pivot and the patient companion phrase is expressed as a postverbal core argument. The companion phrase is realized as an unmarked NP, which is evidence that it is not an oblique phrase.

In Makasar, instrument-selecting ACs are also ditransitive (see Jukes 2020: 315–316). This is shown in (185).

- (185) Makasar, Instrument-selecting AC
anjo selek=a na=buno-ang=i bali=a
 that kris=DEF 3=kill-INST.APPL=3 enemy=DEF
 'He killed the enemy with that kris. (AC)' (Jukes 2020: 316)

In Muna, Tolaki, and Mori Bawah, instrumental ACs also appear to be ditransitive, as both the instrument and patient may be coded as unmarked NPs. In these languages, with respect to indexing and access to syntactic operations, there are some differences in the behavior of applied instruments as compared to applied beneficiaries.

Unlike applied beneficiaries, in Mori Bawah, the applied instrument is generally not indexed on the verb, only the patient may be indexed, when it is definite. In Mori Bawah, the applied instrument also may not be the subject of a passive construction, but only the patient.

- (186) Mori Bawah, Instrument-selecting AC
Lauro andio te'ingka ku='oho-akomiu.
 rattan PROX near.future 1SG.NOM=bind-INST.APPL:2P.ABS
 'In a moment I will bind you with this rattan.' (Mead 2005: 704)

In Muna, the instrument is usually not indexed on the verb, but there are a few exceptions. For example, the instrument and the patient may both be indexed if both are pronominal and the instrument is a third person referent; also in constructions with non-canonical word order, it appears that a fronted instrument may be indexed on the verb (van den Berg 2013: 180–181). However, the applied instrument may be the subject of a passive construction in Muna, just like the applied beneficiary. An example of indexing both the beneficiary and the patient is given in (187).⁶

⁶This sentence has other possible interpretations, and may also be interpreted as a benefactive AC, e.g. 'I will wash it for them.' (see van den Berg 2013: 180).

(187) Muna, Instrument-selecting AC

a-k<um>adiu-ande-e

<IRR>1SG-wash-APPL:3PL.IO-3SG.OBJ

‘I will wash them with it.’

(van den Berg 2013: 180)

In Tolaki, the patterns are less clear. It appears that indexing of the instrument on the verb is generally acceptable only if a definite patient is also indexed on the verb, and the instrument is a third person referent. However, the applied instrument may only be the subject of a passive construction without any kind of indexing on the verb (Edwards 2012: 86).

In all three of these languages, when a definite patient is indexed on the verb in an instrumental AC, instead of the basic form of the applicative suffix, e.g. *-ako* or *-ghoo*, a dative or indirect object pronominal index appears on the verb. This supports the analysis that these special indexes are functionally equivalent to the applicative suffix plus a following pronominal index for an object or absolutive argument. Definiteness of the instrument does not trigger a shift in verbal morphology, though definiteness of the patient does.

In languages with instrumental ACs, it is not reported whether the instrument applied phrase has a special pragmatic status. However, in Sundanese, we find that the instrumental AC is preferred over the BC if the instrument represents salient or unexpected information; this is especially the case in PV, where the instrument applied phrase is the pivot.

7.6.2 Theme-selecting ACs

Of the 24 languages of the sample, 23 show theme-selecting ACs. In such constructions, the applied argument is a semantic theme role, i.e. an entity which changes location in a directed motion event. In almost all cases, the applicative morpheme that marks theme-selecting ACs also marks instrument- and beneficiary-selecting ACs. Nias, which has neither beneficiary- nor instrument-selecting ACs, is an exception.⁷ Theme applied phrases are found in only a limited number of examples in Nias, Yakan, Kendayan, Ampenan Sasak, Behoa, and Duri (however, see the discussion below in §7.9, where the product of a bodily function verb may be considered a theme). Tolaki does not appear to show theme-selecting ACs.

As mentioned above, instrument applied phrases in these languages often share semantic similarities with themes, as instruments used for chopping, hitting, and similar events are directed into motion by an agent (see also Kroeger 2007). This is not necessarily true for all bases, however; with base verbs like ‘buy’, the instrument applied phrase in an AC (typically some type of currency) is not necessarily in motion. This is even more the case for instrument-selecting ACs that allow a mechanical instrument to be the applied phrase, e.g. Muna *tampoli-ghoo* ‘to sew with (e.g. hand, sewing machine)’.

In the languages of the sample, we find various types of constructions in which AM-marking coincides with a theme applied phrase. These are discussed in turn below according to properties of the base verb.

First, theme applied phrases are found in ACs with transitive bases that describe an event of directed motion, e.g. ‘to pelt/throw (at)’, ‘to spray (at)’, ‘to shoot (at)’. In corresponding BCs, the

⁷In Nias, there is one construction that may be considered a theme-selecting applicative, which is formed with the prefix *fa-*, usually a causative marker, and the verbs meaning ‘to pelt’ and ‘to throw’, see Brown (2001: 228–232).

P argument is a goal, while the theme is coded as an oblique phrase if expressed. ACs of this type select a theme as the applied phrase, show monotransitive clausal structure, and are considered remapping, with the companion phrase expressed as an oblique PP. In the example from West Coast Bajau in (188), an instrument-selecting AC is shown with the verb *seput* ‘spray’. In the BC in (188a), the P argument is the semantic goal, *using e* ‘the cat’, and no theme argument is expressed. In the AC in (188b), the theme applied phrase, *dalit* ‘venom’, is encoded as an unmarked NP and represents the pivot (a core argument). The goal companion phrase *ta’ using e* ‘at the cat’, is expressed as an oblique PP marked with the locative preposition *ta’*.

(188) West Coast Bajau, Theme-selecting AC

a. *Using e ai Ø-seput soo dilaw.*
 cat DEM PFV PV-spray snake yesterday
 ‘A snake sprayed the cat (with venom) yesterday.’ (BC)

b. *Ai Ø-seput-an soo dalit ta’ using e*
 PFV PV-spray-THM.APPL snake venom LOC cat DEM
 ‘A snake sprayed venom at the cat.’ (AC)

(Miller 2007: 290)

A similar alternation is described in Toba Batak by Van der Tuuk (1971 [1864-1867]: 104) for the verbs *mamodil* ‘to shoot with a gun (AV)’ and *mangultop* ‘to shoot with a blowpipe (AV)’. With these verbs, the goal in a BC (i.e. the target of the shot) is the P argument. When the same verb root is marked with the AM suffix *-hon*, a theme applied phrase may be selected. Thus, the P argument of “*mamodilhon* is the bullet or that which acts as such, as, for example, *inal* [‘wooden rod for shooting at birds’]”, and the P argument of “*mangultoppon* is the arrow (*nakkat*)” (van der Tuuk 1971 [1864-1867]: 104). Alternately, the applied phrase with these same verbs when suffixed with *-hon* may be the instrument, i.e. a blowpipe and gun, respectively. The companion phrase in these ACs is the goal, which is encoded as a PP, if expressed.

Second, theme applied phrases are found in ACs with base verbs that describe an act of locomotion. Such ACs are found in Toba Batak, West Coast Bajau, Pendau, Mori Bawah, Makasar, and Bugis. In Mori Bawah, these are formed with the applicative suffix *-Cako*, while beneficiary- and instrument-selecting constructions are formed with *-ako*, though the two are almost certainly related historically. In addition, in a number of languages besides those mentioned above, similar constructions are attested with the verb meaning ‘to run, to flee’, though not necessarily with other locomotion verbs. Examples of theme-selecting constructions with ‘run’ were found in Yakan, South Barisan Malay, Indonesian, Javanese, Balinese, Ampenan Sasak, and Behoa, among others. In theme-selecting ACs with locomotion base verbs, the theme shows semantic similarities to a comitative participant or causand. The base verb in such ACs may be intransitive or transitive.

With bases representing intransitive locomotion verbs, a theme-selecting AC is monotransitive, with the theme applied phrase showing core encoding. For example, in Toba Batak, the verb meaning ‘fly’ is *habang*. When marked with the AM *-hon* and the AV prefix *maN-*, the resulting stem *makkabakkon* means ‘to fly away with (s.t.)’, and it selects as a core argument the entity that is flown with (van der Tuuk 1971 [1864-1867]: 1977). An example from Behoa is shown in (189) with the verb *langka* ‘to run’. In such examples, the A argument in the AC is an instigating causer, thus this type of construction may be considered both causative and applicative.

(189) Behoa, Theme-selecting AC

- a. *Ahung=ku' me-langka mai pe=mo' na-nóntóhi.*
 dog=1S.POSS AV-run come just=1SG.ABS.CMPL 3S.ERG-head.for
 'My dog ran straight away towards me.' (BC)
- b. *Mam-pe-langka-ha=ko' áná'=ku' nodo=he' to ra-tawani.*
 AV-TR-RUN-THM.APPL=2SG.ABS children=1SG.POSS like=3PL.ABS REL 3PL.ERG-capture
 'You ran off with my children as if they were captives.' (AC)
 ("langka", Shore 2016, glosses added)

With bases representing transitive locomotion verbs, theme-selecting ACs are generally mono-transitive and considered remapping, with the theme applied phrase showing core encoding and the companion phrase (a goal or path) expressed as an oblique PP. Examples are given from West Coast Bajau in (190) and (191).

(190) West Coast Bajau, Theme-selecting AC

- a. *Aku boi n-(s)embet Azizy engko' surat.*
 1SG CMPL AV-chase A. with letter
 'I chased Azizy with the letter.'
- b. *Aku boi n-(s)embet-en surat e ta' Azizy.*
 1SG CMPL AV-chase-THM.APPL letter DEM LOC A.
 'I rushed the letter to Azizy.' (Miller 2007: 238)

(191) West Coast Bajau, Theme-selecting AC

- a. *Ai Ø-keta Pirik suang e.*
 PERF PV-CROSS P. river DEM
 'Pirik crossed the river.'
- b. *Ai Ø-keta-an Pirik using e pe dembila' suang.*
 PERF PV-CROSS-THM.APPL P. cat DEM to.there across river
 'Pirik carried the cat across the river.' (Miller 2007: 237)

In (190b) the applied phrase, *surat e* 'the letter', again is semantically similar to a comitative participant or causand. The sentence in (190b) could be paraphrased as 'I chased Azizy **with the letter**' or 'I caused **the letter** to chase Azizy.' In (191b), the applied phrase, *using e* 'the cat', is likewise semantically similar to a comitative participant or causand. In the event described in (191b), the agent participant crosses the river while carrying the cat, thus causing the cat to also cross the river.

Third, theme-selecting AMs are also found on base verbs that normally select a theme as the P argument without any applicative marking. In the AM-marked clauses with such verbs, there is no change in the semantic role selected as P compared to corresponding BCs, however there is usually a semantic emphasis on the act of directed motion. This type of theme-selecting construction is found with certain verbs of transfer in Sundanese, West Coast Bajau, Toba Batak, and Makasar, among others, e.g. 'to send/send to', 'to give', 'to push'. For instance, with the

Sundanese verb *surung* ‘push’, both the unsuffixed form and the form marked with *-keun* take a theme as P. However, the verb bearing the suffix *-keun* is only used when the event of pushing described is an act of directed motion towards a goal. Consider the example in (192).

(192) Sundanese, Directed motion construction with *-keun*

- a. *Keur abi ulin ka lapang, aya Pak Haji keur ny-(s)urung roda.*
 while 1SG play to field, exist Sir Haji while AV-push cart
 ‘When I was playing at the field, Pak Haji was there pushing his cart.’ (unmarked)
- b. *Udi tos ny-(s)urung-keun mobil ka imah.*
 U. already AV-push-THM.APPL car to house
 ‘Udi is done pushing the car to the house (i.e. it is now at the house).’ (marked)

(CT1-027)

In (192a), the verb *nyurung* shows no applicative marking, and is used in this context because the event of pushing described does not direct the theme towards an identifiable and salient endpoint.⁸ In (192b), the same verb is used with the AM *-keun*. In this context, *nyurungkeun* is used because the event describes an act of directed motion. When the completive marker *tos* is used with this verb, the clause entails that the theme has already undergone the change in location described by the goal phrase. van der Tuuk (1971 [1864-1867]: 104) discusses a similar use of the theme-selecting AM *-hon* with the Toba Batak verb *tongos* ‘to send’. This verb may take a theme (e.g. *sorat* ‘letter’) as a core argument without any AM-marking, but the form marked with *-hon* is preferred if directed motion is emphasized, and when a recipient or goal “is either stated or in the mind of the speaker.”

Verbs of bodily function are also found in AM-marked constructions that may be considered theme-selecting ACs (see §7.9 below).

7.7 Goals and locations

7.7.1 Distribution and general properties

The AM that selects goals and locations, in the majority of languages differs from those that select the semantic roles discussed thus far (i.e. instruments, themes, beneficiaries, recipients). Exceptions are West Coast Bajau, Yakan, and Ampenan Sasak, which each have a single applicative suffix *-an*. In Mori Bawah, locative- and goal-selecting ACs are marked with both the suffix *-Ci* and the suffix *-Cari*. Locative/goal-selecting ACs are found in almost all the languages of the sample, however, Tolaki does not appear that have this type of AC.

Locative- and goal-selecting ACs typically occur on intransitive and transitive bases in the languages of the sample. In Makasar, the locative/goal-selecting AM *-i* appears to be compatible only with intransitive bases, though Jukes (2020) does not address this point explicitly. Common intransitive bases found in locative/goal-selecting ACs include activity verbs (e.g. ‘swim’, ‘jump’, ‘run’), posture verbs (e.g. ‘sit’, ‘stand’, ‘sleep’), and bodily function verbs (e.g. ‘cough’, ‘vomit’). In Muna, only a small number of transitive bases may be suffixed with *-Ci*, and the resulting verbs

⁸The term *Pak Haji* refers to a distinguished man who has completed a religious pilgrimage.

often have idiosyncratic meanings rather than meanings consistent with a locative/goal-selecting applicative (van den Berg 2013: 292-293). For languages other than Makasar and Muna, transitive verbs commonly used as a base in such constructions include ‘put’, ‘pour’, ‘plant’, ‘teach’, and ‘pay’. Consider the examples from West Coast Bajau in (193)–(194).⁹

(193) West Coast Bajau, Locative-selecting AC

a. *Sesok e ai pe-rekot ta' jing.*
house.lizard DEM PFV INTR-stick LOC zinc
‘The house lizard has stuck to the zinc.’ (BC)

b. *Ai rekot-on sesok jing e.*
PFV [PV]stick-LOC.APPL house.lizard zinc DEM
‘The house lizard has stuck to the zinc.’ (AC)

(Miller 2007: 283)

(194) West Coast Bajau, Goal-selecting AC

a. *Ai enna'(-an)=ni gula' diam kupi'.*
PFV [PV]place-LOC.APPL=3SG.NPIV sugar inside coffee
‘(S)he put sugar in the coffee.’ (BC)

b. *Ai enna'-an=ni kupi' e gula'.*
PFV [PV]place-LOC.APPL=3SG.NPIV coffee DEM sugar
‘(S)he put sugar in the coffee.’ (AC)

(Miller 2007: 285, slightly modified)

In these examples, the locative role in (193) and the goal role in (194), are expressed in oblique PPs in BCs. In ACs marked with *-an*, these roles are selected as the applied phrase and encoded as unmarked core arguments. In West Coast Bajau, locative- and goal-selecting ACs are consistently valency-increasing. In Nasal, an increase in transitivity is observed when the applicative suffix *-i* attaches to intransitive bases and a limited number of transitive bases. Compare the examples of the transitive base *ajakh* ‘teach’ in (195) to the transitive base *takhuk* ‘plant’ in (196).

(195) Nasal, Goal-selecting AC, valency-increasing with ‘teach’

a. *yo agi ng-ajakh baso Nasal khan anak=ku.*
3SG PROG AV-teach language Nasal with child=1SG.POSS
‘I am teaching Nasal to my child.’ (BC)

b. *yo agi ng-ajakh-i anak=ku baso Nasal.*
3SG PROG AV-teach-LOC.APPL child=1SG.POSS language Nasal
‘I am teaching my child Nasal.’ (AC)

(McDonnell fieldnotes)

⁹For example (194), Miller (2007: 283) states that the suffix *-an* applies “vacuously” and is optional in the BC without any change in the argument or oblique (see §7.10 for discussion). Thus, in (194a), it is possible to use the *-an* suffix without any change in argument structure. However, in (194b), the ditransitive construction requires the applicative suffix *-an*.

- (196) Nasal, Locative-selecting AC, remapping with ‘plant’
- a. *be-bibai-an n-(t)akhuk jagung di sawah.*
 DISTR-woman-DISTR AV-plant corn LOC rice.paddy
 ‘The women planted corn in the rice paddy.’ (BC)
- b. *be-bibai-an n-(t)akhuk-i sawah khan jagung.*
 DISTR-woman-DISTR AV-plant-LOC.APPL rice.paddy with corn
 ‘The women planted corn in the rice paddy.’ (AC) (McDonnell fieldnotes)

In the BC in (195a), the information that is transmitted (i.e. what is being taught) is expressed as the P argument, while the goal/recipient phrase (i.e. the person being taught) is expressed as an oblique PP. In the AC in (196b), both the goal and the companion phrase are unmarked, showing that the AC is ditransitive. In the BC in (196a), the theme is expressed as an unmarked core argument (P) and the location phrase is encoded as an oblique PP. In the AC in (196b), the location is expressed as an unmarked core argument while the theme companion phrase is expressed as an oblique PP. Thus we see that with some transitive bases, locative/goal-selecting ACs in Nasal are valency-increasing, and with others, they are remapping.

Locative- and goal-selecting ACs in Sundanese, Indonesian, South Barisan Malay, Javanese, Balantak, and Yakan are similar to those in Nasal. Some transitive bases show ditransitive structures in such ACs, while others are considered remapping. Languages like Bugis, Balinese, Ampenan Sasak, Mori Bawah, and Madurese, appear to behave like West Coast Bajau in consistently allowing ditransitive locative/goal-selecting ACs with transitive bases. However, it is not clear how transitive bases pattern when suffixed with a locative-/goal-selecting AM in some languages of the sample, including Salako, Toba Batak, Kaili Ledo, Tukang Besi, and Behoa.

7.7.2 Locative/goal-selecting ACs with special properties

Like West Coast Bajau, in Pendau all goal-selecting ACs show an increase in the value of syntactic transitivity compared to corresponding BCs. However, the Pendau goal-selecting AM *-i* is only found to increase the transitivity of transitive bases when it co-occurs with a stem-former prefix, which was mentioned in §5.9.4.4. Goal-selecting ACs like that shown in (197) require both the stem-former and the AM *-i* to be used on the verb.

- (197) Pendau, Locative applied phrase with stem-former
- a. *bau uo ni-alap ni=kai ri=payangan.*
 fish yonder PV.RLS-take PN=grandfather LOC=boat
 ‘The grandfather took the fish in the boat.’ (BC)
- b. *payangan ni-pong-alap-i ni=kai bau uo*
 boat PV.RLS-SF.TR-take-LOC.APPL PN=grandfather fish yonder
 ‘The grandfather took the fish in the boat.’ (AC) (Quick 2007: 301)

Another exceptional example is found in Pendau where the applied phrase in locative-selecting ACs may alternately be expressed as a PP (rather than an unmarked NP). According to Quick (2007: 300), the locative applied phrase can even be the pivot when it is a PP, which is shown

by its pre-verbal position in (198). See also §6.5.4 for examples of LV clauses that appear to take similar oblique location pivots.

(198) Pendau, Locative applied phrase in PP

Ribongkarongo'u niponyoputi'u.

ri=bongkarong='u ni-pong-soput-i='u

LOC=hut=1SG.GEN PV.RLS-SF-shoot-LOC.APPL=1SG.GEN

'I shot (it) at/beside my hut.' (AC)

(Quick 2007: 300)

7.7.3 Addressees and recipients marked like goals

In some languages, the applicative morpheme that marks locative/goal-selecting ACs is also observed in constructions where the applied phrase is an animate entity rather than a location per se. In such cases, the applied phrase may have the role of addressee or recipient in the event described by the AC marked with this suffix.

7.7.3.1 Addressee-selecting ACs

Addressee-selecting applicative constructions are found with same AM used in locative-/goal-selecting applicatives in 15 languages of the sample with more than one applicative morpheme, and in three languages with only one applicative morpheme (Yakan, West Coast Bajau, and Ampanan Sasak). An example of an addressee-selecting AC is given in (199) below from West Coast Bajau.

(199) West Coast Bajau, Addressee applied phrase

a. *"Buat-in do' aku bue' susu, too' bana kelong=ku tu,"*
make-PV.IMP EMPH 1SG water milk dry very throat=1SG.POSS DEM

Ø-bara'=ni m-aku.

PV-tell=3SG.NPIV LOC-1SG

'"Make me some milk, I am very thirsty," she said to me.' (BC)

b. *Bila teko me-ruma' bara-an=ni emma'=ni uun jomo*
when arrive LOC-house PV-tell-LOC.APPL=3SG.NPIV father=3SG.POSS exist person

mu' lawa' bana.

there beautiful very

'When she arrived home, she told her father that there was a very handsome man there.' (AC) (Miller 2007: 286)

The base in this type of addressee-selecting AC is generally a communication verb, or verb of speaking. Some examples include Nias *fa-hede-si* 'to greet s.o.' cf. *fe-hede* 'to say a greeting', Indonesian *men-cerita-i* 'to tell to s.o.', c.f. *ber-cerita*, 'to tell a story', and Muna *podea-ghi* 'to shout at s.o.' cf. *podea* 'to shout'.

In some instances, when a communicative verb bears the AM, this stem has a specific, lexicalized meaning in addition to indicating that the applied phrase is an addressee. For example,

Indonesian *meng-(k)ata-i* means ‘to scold s.o.’ or ‘to speak badly of s.o.’ (cf. *ber-kata* ‘to say’). In Balantak, *mam-bantil-i* means ‘to advise, instruct, admonish or reprimand s.o.’, or ‘to invite (s.o.) to an event’ cf. (*mem-bantil-kon*, ‘to inform, tell s.o.’). The Sundanese verb *omong-an* (c.f. *omong* ‘to speak’) similarly means ‘to scold or advise s.o.’.

However, note that addressee-selecting ACs can also be marked with the AM that marks selection of beneficiaries, instruments, and themes as the applied phrase, to the exclusion of the AM that marks selection of locative and goal roles. Some examples include, Behoa *mo-uli’á* ‘to speak to s.o./s.t.’ (c.f. *mang-uli* ‘to say’), Duri *kua-n* ‘to say to s.o.’ (c.f. *kua* ‘to say’). Addressee-selecting ACs of this type are less commonly observed than those formed with the locative-/goal-selecting applicative suffix, though of course, in languages with only one AM, it is not necessarily possible to make such a distinction.

7.7.3.2 Recipient-selecting ACs

In some languages with multiple AMs, the locative-/goal-selecting AM is found on the verb in constructions where the applied phrase is a recipient. This can occur with verbs that take either a location applied phrase (inanimate goal endpoint) or a recipient applied phrase (animate recipient endpoint), as with the verbs meaning ‘to send’ or ‘to carry’ in many languages. In other cases, it occurs with verbs that do not generally show an inanimate goal as a possible core argument, such as the verb meaning ‘to sell’, ‘to serve’, or ‘to give’. Some examples are shown below.

(200) Javanese, Recipient-selecting AC with *-i*

a. *aku nguwéh-aké buku menyang Laura*
 1SG AV.give-THM.APPL book to L.
 ‘I gave a book to Laura’ (Theme-selecting AC)

b. *aku nguwéh-i Laura buku*
 1SG AV.give-LOC.APPL L. book
 ‘I gave Laura a book.’ (Recipient-selecting AC)

(Hemmings 2013: 173)

(201) Indonesian, Recipient-selecting AC with *-i*

a. *Aayah meng-(k)irim(-kan) uang kepada saya*
 father AV-send-THM.APPL money to 1SG
 ‘I gave a book to Laura’ (Theme-selecting AC)

b. *Ayah meng-(k)irim-i saya uang*
 father AV-send-LOC.APPL 1SG money
 ‘Father sent me money.’ (Recipient-selecting AC)

(Arka 1993: 94–94)

(202) Balantak, Recipient-selecting AC with *-i*

a. *...kasi ni-tarop-kon a wala'on i-ya'a na sawe.'*
 then PV.RLS-serve-THM.APPL ART boiled.water DEIC-DEM LOC guest
 ‘Then the hot drinks were served to the guests.’ (Theme-selecting AC)

- b. *Tempo man-tarop-i suo' men u'uru i-ya'a tia wala'on*
 time AV-serve-LOC.APPL family REL new DEIC-DEM with boiled.water
 'When they serve the bride and the groom (lit. the new family) a hot drink' (Recipient-selecting AC) (van den Berg & Busenitz 2012: 107)

Note that in these cases we often observe an alternation between the theme-selecting AM and the goal-selecting AM. Thus it appears that in these cases, a recipient applied phrase is treated like a goal. This type of recipient-selecting AC is fairly limited in the languages of the sample and usually is found only with a small number of base verbs. In comparison, constructions marked with beneficiary-selecting AMs are much more commonly observed with recipient applied phrases; such ACs are found across a much larger number of languages and with a much larger set of base verbs. As mentioned above, in many languages of the sample, recipient-beneficiaries are the most common type of role selected as applied phrase in examples of benefactive ACs by far.

7.8 Circumstantial and comitative roles

In a relatively small number of languages of the sample, the applied phrase in an AC may be a circumstantial role, such as a reason, or purpose. Such constructions are found in Toba Batak, Balantak, Mori Bawah, Tolaki, Muna, and Tukang Besi, as well as in a limited number of examples in Pendau. With the exception of Toba Batak, all of these are languages of Sulawesi. In such languages, the base verb in such a construction may typically be either intransitive or transitive. Some examples are given below.

- (203) Muna, Circumstantial-selecting AC

dadi a-laga ta-sendai na-mate-ghoo ka-gharo
 so ART-ant just-little 3S.IRR-die-CIRC.APPL NMLZ-hungry
 'so in a little while Ant will die of hunger.' (Reason-selecting AC)
 (van den Berg 2013: 183)

- (204) Mori Bawah, Circumstantial-selecting AC

Aku h<um>uku-akomu ponako-mu
 1SG.FUT <PART>punish-CIRC.APPL:2SG.ABS steal-2SG.POSS
 'I will punish you on account of your thievery' (Reason-selecting AC)
 (Mead 2005: 704)

- (205) Tukang Besi, Circumstantial-selecting AC

No-lea-ako te langke-a-no te kaitela
 3.RLS-load-CIRC.APPL CORE sail-NMLZ-3.POSS CORE corn
 'They loaded the corn for the voyage.' (Purpose-selecting AC) (Donohue 1999: 240)

As shown in these examples, the applied phrase in circumstantial-selecting ACs may be realized with core encoding. With transitive bases, it appears that the AC may be ditransitive. In

Mori Bawah, Mead (2005: 704) specifies that it is the companion phrase, i.e. patient or other role that is the object of the BC, which is indexed on the verb in the AC, rather than the applied phrase. Donohue (1999: 238–242) notes that reason and purpose applied phrases in ACs show restricted access to syntactic operations compared to P arguments in base clauses, or other types of applied phrases. Neither may head an object relative clause, and purpose applied phrases are never indexed on the verb. In other languages of the sample details about the argument structure and patterns of indexing for circumstantial-selecting ACs are not clear.

The only language of the sample with a productive comitative AC is *Tukang Besi*, in which such constructions are marked with the applicative suffix *-ngkene* (Donohue 1999). In *Tukang Besi*, a comitative applied phrase “is an equal and voluntary participant in the action indicated by the verb, but is viewed by speaker as of secondary importance” when compared to the A argument. Comitative ACs are found with intransitive and transitive bases in *Tukang Besi*. In the latter type, the applied phrase may not become the pivot of a passive clause, though it may in the former. Some examples of comitative ACs are given below. As seen in (206a), a related verb *kene* ‘to accompany’ can be used in a serial verb construction with a similar meaning to a comitative AC, as in (206b).

(206) *Tukang Besi*, Comitative AC

a. *No-wila no-kene ta ina-no*
 3.RLS-go 3.RLS-accompany CORE mother.3.POSS

‘She went, accompanying her mother.’ (Serial verb construction)

b. *No-wila-ngkene te ina-no.*
 3.RLS-go-COM.APPL CORE mother-3.POSS

‘She went with her mother.’ (Comitative AC)

(Donohue 1999: 200)

(207) *Tukang Besi*, Comitative AC

No-homoru-ngkene te kene-no te wurai na ompu-su.
 3.RLS-weave-COM.APPL CORE friend-3SG.POSS CORE sarang NOM grandparent-1SG.POSS
 ‘My grandmother wove a sarong with her friend.’ (Donohue 1999: 229)

Besides such constructions in *Tukang Besi*, we observe some clausal constructions marked with AMs that select P arguments which may be viewed as accompaniments or companions. However, these typically occur only with a very limited set of bases. Examples include *Ledo Kaili ra-kande-ka* ‘to be eaten with (s.t.), of side dishes’ from *kande* ‘to eat’ (D. Evans 2003: 72) and the similar applicative verb *fumaa-ghoo* ‘to eat with’ in *Muna* (van den Berg 2013: 123).

Otherwise applied phrases that appear to have a role of accompaniment or companion in ACs are mostly limited to bases expressing locomotion, e.g. ‘to run off with’, ‘to cross over with’, as discussed for theme-selecting ACs §7.6.2 above. I have treated these as theme roles though they are sometimes calls companions; note that in such constructions the applied phrase does not generally represent a voluntary participant in the clausal event, unlike the applied phrase in *Tukang Besi* *-ngkene* marked comitative ACs.

7.9 Other applied phrases

When AMs attach to intransitive bases, they often select—or licence—a P argument that takes on various semantic roles. Such transitivizing ACs are found in all languages of the sample. In many cases, there is no clear monoclausal BC equivalent. To express the same semantic role in a BC, a subordinate clause or parallel clause often must be used.

7.9.1 Content-selecting ACs

With bases that describe acts of speaking or cognition, an AM commonly selects a content applied phrase that is realized as the P argument. The referent of the applied phrase may be a topic, proposition, or reported speech. An example is given from Sundanese with the verb *carios* ‘talk’ in (208) below, repeated from (60) in §2.9.

(208) Sundanese, Content-selecting AC

a. *Abi ny-(c)arios ka mama, “Ma, abi hoyong miliarian damel.”*
1SG AV-talk to mother, mom 1SG want AV.look.PLUR work
‘I said to my mother, “Ma, I want to look for work.”’ (BC)

b. *Hayang urang kempel ny-(c)arios-keun pa-damel-an.*
let 1PL gather AV.talk-CONT.APPL NMLZ-work-NMLZ
‘Let’s meet up and talk about the job.’ (AC)

(FM4-027)

Other examples with verbs of communication include Toba Batak *mang-hata-hon* ‘to talk about’ from *hata* ‘word, talk’ (see Schachter 1984: 103), Kaili Ledo *nang-oe-ka* or *nang-koe-ki* ‘to boast about s.t.’ from *nang-oe* ‘to boast’ (D. Evans 2003), Muna *po-kamunti-ghoo* ‘to whisper about’ from *po-kamunti* ‘to whisper (to one another)’, and Mori Bawah *pesikeno-ako* ‘to ask about s.t.’ from *pesikeno* ‘to pose a question’. With verbs of cognition, content-selecting ACs are found with Ampenan Sasak *pikir-an* ‘to think about’ from *pikir* ‘to think’ (Khairunnisa & McDonnell in prep) and Sundanese *ng-impi-keun* ‘to dream about, aspire for’ from *ng-impi*, ‘to dream, dream of’.

A number of languages also show a content-selecting AC with the base meaning ‘to tell (a story)’. This is found in Salako, where *ba-curitâ* means ‘to tell (intr.)’ but *ny-(c)uritâ-?âtn* means ‘to tell (s.t.)’, and also in Sundanese, i.e. *ny-(c)arita* ‘to tell a story’ cf. *ny-(c)arita-keun* ‘to tell (s.t.), to tell about (s.t.)’. Similar applicative verbs are found in Indonesian, South Barisan Malay, Javanese, Kaili Ledo, Ampenan Sasak, and Makasar, among others.

7.9.2 Stimulus-selecting ACs

With verbs describing emotional states or responses, AMs commonly select a stimulus (or target) applied phrase that is realized as the P argument. These ACs are found with at least one AM, in all 24 languages of the sample. In corresponding BCs in these languages, the stimulus may be realized as an oblique PP, but in some cases there is no monoclausal equivalent to the AC. Examples from Salako (209) and Ampenan Sasak (210) are given below.

(209) Kendayan, Stimulus-selecting AC

- a. *Berà sidi ià ka Ne? Kulup.*
angry very 3 LOC PN K.
'He was extremely angry at Kulup.' (BC) (Adelaar 2005b: 92)
- b. *Tarutama bapa?=e karas sidi m-(b)era-i? ià...*
especially father-3.POSS hard very AV-angry-APPL 3
'Especially his father got very angry at him...' (AC) (Adelaar 2005b: 86)

(210) Ampenan Sasak, Stimulus-selecting AC

- a. *Ie takut.*
3.SG afraid
'(S)he is afraid.' (BC)
- b. *Ie takut-an berarak.*
3.SG afraid-APPL spider
'(S)he is afraid of spiders.' (AC) (Khairunnisa & McDonnell in prep)

In Toba Batak, in stimulus-selecting ACs of this type, base verbs expressing emotional states take the fossilized prefix *ha-* in addition to the AM suffix *-i*. For example, root *tahut* 'afraid' shows the base verb form *ma-tahut* 'to be afraid' and the AM-marked verb form *mak-ka-tahut-i* 'to be afraid of (s.t.)' (van der Tuuk 1971 [1864-1867]: 134). In Sundanese, the prefix *CVng-* (partial reduplication) is found on some stimulus-selecting ACs. This prefix also indicates greater intensity, as in (211) below, repeated from (69).

(211) Sundanese, Stimulus-selecting AC

- a. *Mariam ceurik lantaran indung=na maot.*
M. cry because mother=3SG.POSS die
'Mariam cried because her mother died.' (BC)
- b. *Mariam ny-(c)eung-ceurik-an indung=na.*
M. AV-RDP-cry-APPL mother=3s.POSS
'Mariam cried intensely about her mother.' (AC) (CT1-006, based on Hanafi 1997: 22)

AM-marking is also found on stimulus-selecting ACs with intransitive perception verbs as bases. In Toba Batak, the intransitive verb *marnangi* means 'to have ears, to be able to hear', while the same root marked with *-hon* in AV is *manangikan* meaning 'to hear (s.t.), to listen to (s.t.)' (van der Tuuk 1971 [1864-1867]: 101). Van der Tuuk (1971 [1864-1867]) writes that the applied phrase of *manangihon* "is something to/for which one listens in order to catch it, either a distant sound, or a word towards which one directs one's hearing" (101). See §7.10 below for discussion of AM-marking used with an intensifying effect on transitive base verbs of perception.

7.9.3 ACs with verbs of bodily function verbs

In many, but not all, languages of the sample, an AM marks selection of a P argument with intransitive base verbs describing bodily processes. Examples include Muna *no-hoda-ghoo* ‘to cough up (s.t.)’ from *hoda* ‘to cough’, Nias *uta-’ö* ‘to throw up (s.t.)’ from *m-uta* ‘to vomit’, and Makassar *na-ta’-me-áng=i cera* ‘he is pissing blood’ cf. *at-ta’-mea=i* ‘he is urinating’. In such cases the applied phrase may be considered a type of theme, though they are not always categorized as such in descriptive accounts. Jukes (2020: 312–314) for example, considers the applied phrase in such examples to be a type of “inherent patient”. Examples of this type are also found in Toba Batak, Kaili Ledo, and Javanese. An example of an intransitive BC and a transitive AC with a bodily function verb is given below.

(212) Sundanese, AC with bodily function verb

a. *Icih utah.*

I. vomit

‘Icih vomited.’ (BC)

b. *Icih ng-utah-keun udang=na*

I. AV-vomit-THM.APPL shrimp=DEF

‘Icih vomited up the shrimp.’ (AC)

(FM4-046)

7.10 Aspect, intensity and other semantic effects

In the languages of the sample, AM-marking is often observed to be associated with a semantic difference in clausal meaning, such as a change in aspect, intensity, manner, or semantic characteristics of the P argument (see §6.4.3).

The AMs that mark locative- and goal-selecting ACs are also commonly found to indicate repeated, iterative, habitual or pluractional aspect, and this may occur with or without a change in argument structure. This was discussed for the Sundanese AM *-an* in §2.10.2. Aspectual functions of this type are also found with Toba Batak *-i*, Salako *-i?*, South Barisan Malay *-i*, Indonesian *-i*, Javanese *-i*, Nasal *-i*, Balantak *-i*, Mori Bawah *-Ci*, Muna *-Ci*, and Yakan *-an*. Makasar *-i* is also sometimes observed with similar effects, but this function is “most likely no longer productive” (Jukes 2020: 306). An example of a pluractional construction is given from Muna below with the locative/goal AM *-Ci*.¹⁰

(213) Muna, pluractional aspect with *-Ci*

a. *Sau hae ne-ala-mu itua?*

wood what PASS.PART-take-2SG that-EMPH

‘What kind of wood did you take?’

¹⁰The form *kaeta* in example (213) is used in informal imperatives and means something like ‘for us (incl.)’ (van den Berg 2013: 72).

- b. *Ala-hi kaeta sau itu, owa-hi kaeta na ini.*
 take-PLUR 2.POL wood that, bring-PLUR 2.POL to here
 ‘Fetch all that wood over there, bring it here.’

(“*ala*”, van den Berg & Marafad 2016, glosses added)

In a number of languages, AMs also function to indicate intensive or careful action. This occurs with transitive base verbs of perception as in Salako *nanang-an* ‘to watch, to look at (AV)’ cf. *nanang* ‘to see (AV)’ (Adelaar 2005b). Similar construction are also found in Indonesian, e.g. *me-lihat-i* ‘to scrutinise, look at intently’ from *me-lihat* ‘to see’, Kaile Ledo, e.g. *ne-talinga-ni* ‘to listen to s.t. on purpose, eavesdrop’ from *talinga* ‘ear’, cf. *nang-epe* ‘to hear s.t.’ (D. Evans 2003: 44, 228).

In other cases, the intensive function of AMs indicates application of greater than usual force. For instance, Toba Batak *pasak* means ‘beat’, and *pasak-kon* may mean ‘beat with s.t.’ (instrumental applicative function) or alternately, ‘do beat s.t.’ (intensive/emphatic function) (Nababan 1981: 70). In Mori Bawah, Mead (2005: 703) notes that *-Cako* can indicate that an action “is performed in a more intense or haphazard manner (without any change in valency of the predicate)”. Interestingly, *-Cako* also indicates that an action “is performed by a large number of people” (702) with some bases, but this appears to be fairly limited in productivity. In *Tukang Besi*, the applicative suffix *-(VC)i* indicates “forceful application” (Donohue 1999: 243), as in *pepe-ki* ‘slap forcefully’, and *busu-ki* ‘punch with forward fist’. An example from Javanese is given below with the AM *-i*, which can indicate intensity as well as intentional action.

(214) Javanese, Greater intensity with *-i*

- a. *Charlotte ng-rusah lawag*
 C. AV-break door
 ‘Charlotte broke the door’
- b. *Charlotte ng-rusak-i lawang*
 C. AV-break-INTENS door
 ‘Charlotte destroyed the door’

(Hemmings 2013: 171)

In Sundanese, the theme-selecting AM *-keun* is also observed to have semantic effects including greater individuation or specificity of P. An example is shown below contrasting the base verb *melak* and the AM-marked verb *melakkeun*, which both mean ‘to plant’ and both select a theme role as the P argument. In (215a), the unmarked verb *melak* is used when the planting of rice is described as a general activity. In (215b), the AM-marked verb *melakkeun* is used in when the clause describes a planting event with a more individuated referent for P. In this case, P refers to rice seeds or seedlings which have been prepared by the farmer ahead of time, as already mentioned in first clause of (215b), preceding the use of *melakkeun*.

(215) Sundanese, Higher individuation of P

- a. *M-(p)elak paré ayeuna mah di sawah, di-sebut=na ny-(s)awah*
 AV-plant rice now PRT in rice.field PV-call=3SG AV-rice.field
 ‘Now planting rice in a paddy is called *nyawah* (making paddies).’

- b. *Saméméh m-pacul ilaharna patani sok sa-sadia-an binih paré heula.*
 before AV-hoe usually farmer go.ahead RDP-ready-CAUS seed rice first
M-(p)elak-keun binih paré (gabah) di-sebut=na tebar.
 AV-plant-THM.APPL seed rice rice.grain PV-call=3SG make.rice.seedling
 ‘Before tilling the ground usually the farmer prepares rice seeds first. Planting the
 rice seeds (grains of rice) is called *tebar* (making rice seedlings).’
 (CT1-030, based on Kustian n.d.)

Specificity effects are also reported for the Javanese applicative suffix *-i* by Vander Klok & Evans (2022), though it appears that these effects apply only to a limited set of bases.

(216) Javanese, Higher specificity of P

- a. *Slamet n-(t)ulis (buku anyer / *buku-ku sing anyer).*
 S. AV-write book new / book-1SG REL new
 ‘Slamet wrote a new book / *my new book.’
- b. *Slamet n-(t)ulis (buku-ku sing anyer / *buku anyer)*
 S. AV-write-LOC.APPL (book-1SG REL new / book new
nganggo tulis-an warna-warna.
 using write-NMLZ RDP-color
 ‘Slamet wrote (my new book / *a new book) with colored writing.’
 (Vander Klok & Evans 2022, citing Sudaryanto 1991:61-62)

In Makasar, the verb *sare* ‘give’ is a ditransitive verb that selects three core arguments. As shown in (217a), when *sare* is used without any AM-marking, the theme is always indefinite, is not indexed on the verb, and is obligatory to mention. When the AM *-ang* is affixed to *sare*, as shown in (217b), the selection of semantic roles and indexing of arguments shows no change. There is, however, a semantic difference in that the theme must be definite in the AM-marked clause. In addition, the theme argument may be unexpressed when the verb is marked with *-ang*, which represents a change in its syntactic properties.

(217) Makasar, Change in definiteness of the theme

- a. *La=ku=sare=ko doe’*
 FUT=1SG=give=2SG money
 ‘I’ll give you some money.’
- b. *La=ku=saré-ang=ko doek=ku*
 FUT=1SG=give-THM.APPL=2SG money=1SG.POSS
 ‘I’ll give you my money.’ (Jukes 2020: 254)

Certain clauses marked with the AM *-an* show similar semantic effects in West Coast Bajau. The AM *-an* is found on many monotransitive and ditransitive verbs in AV without a change in argument structure compared to unsuffixed forms. Miller (2007: 293) described these alternations by stating that “when the *-an*₁ suffix does occur, a specific/referential argument and/or a particular event is involved.” This is similar to the use of Yakan *-an* as an obligatory suffix in the

zero-marked PV constructions with certain bases; this constructions must be used for transitive clauses with definite P arguments. In corresponding semi-transitive or ‘antipassive’ constructions marked with *mag-*, P is indefinite and *-an* is not required (though with some lexical verbs it may be optionally present).

Finally, AM-marking is frequently observed to be associated with lexicalized changes in verbal meaning. As mentioned in §2.5.1, the Sundanese verb *béré* means ‘give’, while *béré-keun* means ‘to hand over’. In Balantak, *mim-bibit* means ‘to carry in the hand’, while *mim-bibit-i* means ‘to attach/make a carrying strap/rope on s.t.’ (van den Berg & Busenitz 2012: 105). For further examples and discussion of the semantic functions of AMs in West Nusantara languages see Truong & McDonnell (2022).

7.11 Causative AM-marked constructions

The use of AMs in forming causative constructions is quite common in languages of the sample, but not universally observed. In some languages of the sample, AMs are the primary markers for morphological causative constructions. This is the case in three Malayic languages of the sample, as well as Nasal, Javanese, Sundanese, Balinese, and Ampenan Sasak. In Nias, the prefix *fa-* should be considered primarily a causative morpheme, though it is found on a few constructions with applicative meanings, as described above.

In other languages, including Toba Batak, Madurese, Yakan, West Coast Bajau, and all 11 languages of Sulawesi represented in the sample, there is a separate causative prefix with the form *pa-*, *paka-* or similar that forms morphological causative constructions. In 13 of the 15 such languages, these prefixes are the primary markers of causative constructions. In Madurese, both *pa-* and the AMs *-agi* and *-e* are productively used to form causative constructions. In Balantak, the AMs *-i* and *-kon* and causative prefixes *po-* and *pa-* are found to form causative constructions across bases according to patterns which are “lexically specific and somewhat idiosyncratic” (van den Berg & Busenitz 2012: 93–94). Still, even in languages where the causative prefixes are the primary markers of morphological causative constructions, AM-marked constructions in which an instigating causer is introduced as the A argument may still be observed with certain lexical bases. Altogether, in 21 languages of the sample, one or more AMs forms causative constructions. In three languages, however, there is apparently no use of AMs to form causative constructions with any level of productivity. These are Muna, Tolaki, and Duri, all language of Sulawesi with productive causative prefixes.

An example of a causative AM-marking construction with a stative base is given below from Balantak with the locative applicative AM *-i*

(218) Balantak, Causative function of *-i* with stative base

- a. *Anak-ku no-baloki'-mo.*
 child1S RLS-big-PFV
 ‘My child is already big (now).’ (van den Berg & Busenitz 2012: 68)
- b. *Ia mae' na Luwak mam-balaki'-i poto-na*
 3SG go LOC Luwak AV-big-CAUS photo-3SG
 ‘He went to Luwak to enlarge his photo.’ (Causative)
 (van den Berg & Busenitz 2012: 96)

With certain bases in Madura, either the causative prefix *pa-* or an applicative suffix may be used to form a causative construction. This is shown with the AM *-agi* in the example below.

(219) Madurese, Causative constructions

- a. *Mutmainah m-(p)a-sossa ca-kanca-na.*
 M. AV-CAUS-sad RDP-friend-DEF
 ‘Mutmainah made her friends sad.’ (Causative)
- b. *Mutmainah ny-(s)ossa’-agi ca-kanca-na*
 M. AV-sad-CAUS RDP-friend-DEF
 ‘Mutmainah made her friends sad.’ (Causative)

The example below shows a causative AM-marked construction from Ampenan Sasak, with the dynamic intransitive verb *kèlèp*, ‘fly’.

(220) Sasak, Causative function of *-an* with dynamic intransitive base

- a. *Pesawat nó kèlèp.*
 plane DET fly
 ‘The plane flew.’
- b. *Pilòt nó kèlèp-an pesawat.*
 pilot DET fly-CAUS plane
 ‘The pilot flew the plane.’ (Khairunnisa & McDonnell in prep)

Note that in some causative AM-marked constructions formed with transitive bases, it is not the causand that maps to P, but the causee. An example is show below from Javanese with the AM *-aké*. Constructions of this type are also found in Indonesian, Sundanese, and Nasal.

(221) Javanese, Causative function of *-aké* with transitive base

- a. *Bambang gawe dolanan iku kanggo bocah kae.*
 B. make toy that for child that
 ‘Bambang made that toy for that child.’
 (Vander Klok & Evans 2022, citing Suhandono 1994:53)
- b. *Klambi-ne di-gawè-kake dening gêrji.*
 shirt-3s PV-make-CAUS by tailor
 ‘He had his shirt made by a tailor.’ (Horne 1974: 193, glosses added)

In languages with productive causative morphemes distinct from AMs, the two types of morphemes often may combine on the same verb. However, in certain languages, such as Pendau, it does not appear that causative prefixes freely combine with applicative suffixes.¹¹

In Toba Batak, when certain transitive base verbs are marked with the prefix *pa-* and the suffix *-hon*, it does appear that *-hon* functions to selects a theme applied phrase, while *pa-* introduces an instigating causer argument. Examples include *pa-pahat-ton* ‘to give (s.t.) to animals to eat’ from

¹¹However, see Quick (2007: 284) for one example in which *pa-* is analyzed as both a causative marker and a stem forming prefix when used with the applicative suffix *-a*’.

pahan ‘to eat’ and *pa-djudjuk-kon* ‘to give (s.t.) to someone to carry on the head’ from *djudjung* ‘to carry on one’s head’ (van der Tuuk 1971 [1864-1867]: 130). In certain other irregular causative verb forms, however, note that *-hon* is also required to be marked on stems formed with *pa-* in AV only (see discussion of similar cases in §6.4.2). These irregular causative stems are found with intransitive or stative base verbs. With such verbs, there is no clear selection of a peripheral role to map to a core argument, and thus no clear applicative function. For example, *jóngjong* means ‘to stand’, while *pa-jóngjong* means ‘to make to stand (up), PV’, but *pa-jónggok-kon* means ‘to make to stand (up), AV’ (Nababan 1981: 103).

In Makasar, causative *pa-* combines with both the benefactive/instrumental AM suffix *-ang* and the locative/goals AM suffix *-i*. In some cases marking with the AM results in the selection of an additional definite argument, as with the verb *kanre* ‘eat’, which is shown in (222) below.

(222) Makasar, Causative *pa-* + AM *-ang*

- a. *ku=pa-kanre=i bembe*
 1=CAUS-eat=3 goat
 ‘I made/let him eat goat (meat).’ (Jukes 2020: 290)
- b. *ku=pa-kanre-ang=i bembe=a*
 1=CAUS-eat-BEN.APPL=3 goat=DEF
 ‘I made/let him eat the goat.’ (Jukes 2020: 296)

With other base verbs in Makasar, the functions of a co-occurring causative prefix and applicative suffix are not easy to distinguish clearly (see Jukes 2020: 295-297). A number of verbs, for example, bear both the prefix *pi-* and one or both of the applicative suffixes. The prefix *pi-* can function as a causative marker, and it also has other semantic functions. This prefix “derives forms with meanings like ‘(examine/inspect/listen) carefully or intently’” (Jukes 2020: 299) and in such cases may be found with or without concurrent AM-marking on the verb.

In some languages, the selection of a stimulus applied phrase with a base verb of perception or emotional state requires both an applicative suffix and another prefix. This was noted earlier for Toba Batak *ha-*, a fossilized prefix of unclear function, and Sundanese *CVng-*, an intensifier or simulfactive marker (see also the discussion of Makasar *pi-* in experiencer-oriented constructions above).

For languages of the sample in which the same AM shows both causative and applicative functions, the distribution of such functions across bases may be influenced by syntactic and semantic properties of the base lexeme, though certainly idiosyncratic and lexicalized patterns are also often reported. AM-marking on stative base verbs and most intransitive dynamic base verbs often results in only the introduction of an instigating causer argument. AM-marking with transitive bases, and intransitive bases of certain semantic subclasses (e.g. speech, perception, emotion as described above), AM-marking commonly is associated with the selection of a peripheral role to map to a non-A argument in the clause (i.e., an applicative function). Additionally, on certain lexical bases, marking with a single AM may be associated with both the introduction of an instigating causer and the selection of an applied phrase that is not a causand. An example is given from Sundanese in (223).

(223) Sundanese, Portmanteau use of *-an*

a. *Cai hujan ng-(k)ucur=na ka solokan.*
 water rain AV-flow=3SG to gutter
 ‘Rain water flows to the gutter.’

b. *Icih ng-(k)ucur-keun cai kana gelas.*
 I. AV-flow-CAUS.THM.APPL water into glass.
 ‘Icih poured water into a glass.’

c. *Gelas di-kucur-an cai ku Icih.*
 glass PV-flow-CAUS.LOC.APPL water by I.
 ‘The glass had water poured into it by Icih.’

(CT1-023)

In example (223a), the base verb *kucur* ‘flow’ is shown to be intransitive without AM-marking. It bears the AV prefix *ng-* but no applicative affix, and takes a single S argument, *cai hujan* ‘rain water’. In example (223b), the AM *-keun* forms a causative construction with *kucur*. Here the verb *ng-(k)ucur-keun* takes an A argument that is an instigating causer, *Icih* ‘personal name’, and a P argument that is a causand, *cai* ‘water’. However, in example (223c), when the AM *-an* occurs on the base verb *kucur*, the resulting clause shows both a causative and applicative meaning. The verb selects a goal applied phrase *gelas* ‘drinking cup’, which maps to the pivot argument in PV, while the causand *cai* ‘water’ is realized as a second non-A core argument in postverbal position.

This type of construction is also possible with transitive bases, in which case the selection of a peripheral role as a non-A core argument may be more unexpected. Consider the following alternation from Sundanese with the verb *injeum* ‘to borrow’. In the BC in (225a), the A argument refers to the agent, that is, the borrower, while the P argument refers to the entity that is transferred in possession, the theme, or thing borrowed. In the AC in (225a), the A argument now refers to an instigating causer, who is not the borrower, but the entity who allows the borrowing to occur, i.e. the lender. In this case it must be a third-party lender, who is not the owner of the thing borrowed. Contra expectation, we do not see the borrower selected as the P argument in the AC in (224b), instead the theme remains the P argument (and represents the applied phrase), and the borrower is expressed in an oblique PP, *ka Icih* ‘to Icih’. This AC is similar in structure to the causative example in (221) above from Javanese with *-aké*.

(224) Sundanese, Causative + Theme-selecting construction

a. *Icih ng-injeum duit ka abi.*
 I. AV-borrow money to me
 ‘Icih borrowed money from me.’ (BC)

(CT1-003)

b. *Duit abi di-injeum-keun ka Icih ku Ema.*
 money 1SG PV-borrow-CAUS.THM.APPL to I. by mother.
 ‘My money was lent out to Icih by Mother.’ (AC)

(CT1-003)

Similar examples are found with recipient applied phrases, as shown in (225).

(225) Javanese, Causative + Recipient-selecting construction

a. *Buku-ne tak silih.*
book-3.POSS 1SG borrow
'I borrowed his book.' (BC)

b. *Bocah=e di-silih-i sarung.*
child=3 PV-borrow-CAUS.LOC.APPL sarung
'They lent the child a sarong.' (AC)

(Horne 1974: 554, glosses added)

However, such examples are not necessarily limited to verbs that express events of caused motion or transfer. Other verbs may acquire the meaning of caused motion or transfer when an applicative affix associated with that constructional meaning is affixed to a base. In the example below, the BC in (226a) describes a perception event in which there is no sense of movement of a theme. In the AC marked with *-keun* in (226b), however, the event described is both a caused perception event and a caused motion event. The A argument is understood as an instigating causer rather than a perceiver, and the applied phrase *kaos kaki Udi* 'Udi's socks' is both a theme and a stimulus.

(226) Sundanese, Causative + Theme-selecting construction

a. *Abi ng-ambeu kue haneut.*
1SG AV-smell cookie warm
'I smell warm cookies.' (BC)

b. *Kaos kaki Udi di-ambeu-keun ka adi=na.*
clothing foot U. PV-smell-CAUS.THM.APPL to younger.sibling=3.POSS
'Udi forced his younger sibling to smell his socks (by moving them close to his/her face).' (AC) (CT1-015)

These examples underscore that AM-marked constructions in these languages may have constructional meanings that are both causative and applicative. This is found even in languages with a separate causative morpheme, e.g. Balantak *bolos-ii* 'to lend s.o. s.t.', from *bolos* 'to borrow s.t.', while Balantak also has the causative prefixes *pa-*, and *po-* (van den Berg & Busenitz 2012). However, it appears that installation of an instigating causer is limited to a very small number of bases with Balantak *-ii*.

7.12 Applicatives and voice

In the languages of the sample, ACs typically freely co-occur across other major voice constructions including AV, PV, and passive voice constructions. As reported in the previous chapter, this is the norm in languages of West Nusantara (see §6.5.1). It is also evident from examples of ACs given above in this chapter, which frequently have been shown in both AV and PV constructions. Even in languages that show reduced semantic transitivity in certain A-oriented constructions, ACs tend to combine freely with these clause types. As mentioned above in §7.3, Makasar is one such language, and has been described to have an asymmetrical voice system. Nonetheless, ACs freely combine with the actor focus prefix *aN-* and the semi-transitive prefix *aN(N)-* (see Jukes

2020: 306). Examples were also given of ACs in the ‘antipassive’ construction in Mori Bawah in (177) and in what may be called semi-transitive clauses with the *ae-* class prefixes in Muna in (164). Likewise, in languages that have a passive construction, ACs generally freely occur with it, as in Ampenan Sasak, which is shown in (227). In the AC in (227b), the verb bears both the passive prefix *te-* and the AM suffix *-an*. The applied phrase, in this case a goal, is selected as the pivot argument when this is the case.

(227) Ampenan Sasak, Passive with goal-selecting AC phrase

- a. *Dengan tólóq babak bajur leq ramuan.*
 people put bark bajur LOC potion
 ‘People put bajur tree bark in the potion.’ (BC)
- b. *Ramuan te-tólóq-an babak bajur siq dengan.*
 potion PASS-put-LOC.APPL bark bajur by people
 ‘The tree bark was added to the potion (lit. made thing) by people.’ (AC)
 (Khairunnisa & McDonnell in prep)

In terms of syntactic restrictions on co-occurrence of ACs with other major voice constructions, as discussed above in §7.5.3 above, certain ditransitive ACs in Sundanese may occur in PV but usually not AV. Likewise, in Pendau, goal ACs marked with *-i* and instrumental ACs marked with *-a* do not occur in AV (see §5.9.4.4). Quick (2007: 304-305) also mentions that a number of roots in Pendau require that the verb be marked with the locative/goal applicative *-i* in PV, as in (228a), whereas marking with the same AM cannot occur in AV, as in (228b). With the same root, omitting the AM *-i* in PV or including it in AV is ungrammatical.

(228) Pendau, Voice restrictions on AC

- a. *Palan roguntuninyo.*
palan ro-guntung-i=nyo
 lamp PV.IRR-light-LOC.APPL=3SG.GEN
 ‘He/she will light the lamp.’
- b. *A’u mo-guntung palan.*
a’u M-pong-guntung palan
 1SG AV.IRR-SF-light lamp.
 ‘I will light the lamp.’ (Quick 2007: 304)

Similar restriction are reported by Miller (2007: 192-193) for West Coast Bajau and Brainard & Behrens (2002) for Yakan, where the AM-marking with *-an* is required on some roots in PV but optional in AV. This type of restrictions may be related to a general tendency for PV constructions to show higher semantic transitivity, and greater individuation or affectedness for the P argument (and thus the applied phrase in ACs). Thus, while ACs do typically co-occur with AV and other A-oriented constructions in the languages of the sample, the use of AM-marking and PV does show a correlation in at least some West Nusantara languages, e.g. Besemah (South Barisan Malay) (McDonnell 2016: 214-215), and this is likely driven by tendencies in discursive usage that have become grammaticalized. See also Donohue (2001), which examines the use of *Tukang Besi* ACs in natural texts and finds an overwhelming tendency for applied phrases to be pivots in discourse.

This and other syntactic restrictions noted in this section are in keeping with the proposal made in Chapter 6, that pivot-neutral ACs are derived from earlier pivot-selecting constructions, in which of course, the peripheral role of location, goal, beneficiary, instrument, or theme, exclusively maps to the pivot relation.

Despite a general lack of syntactic restrictions on voice in ACs across the sample, the combination of voice and applicative morphology allows a participant with a peripheral semantic role, such as a beneficiary, instrument, goal, or location among others to be the pivot (see Davies 2005 for in-depth discussion of these points in Madurese). Further, it is well-known that pivots in symmetrical voice languages are the syntactically privileged argument and thus play an important role in syntactic operations, such as relativization and focus (see e.g. Arka 2003, Riesberg 2014b). In many of the languages of the sample, whether symmetrical or not, such operations are usually restricted to the pivot or subject, and if not, at least to core arguments.

In the languages of West Nusantara, relativization is typically marked by a ‘gap’ in the relative clause, which is co-referential with the head noun. In languages of the sample, the relative clause may be introduced with a linker particle or relativizer, or may be headless, as in Makasar and West Coast Bajau, among others. In asymmetrical voice languages like Nias, Muna, and Mori Bawah, relative clauses, or participial structures that function like headless relative clauses, show A-oriented and P-oriented alternations, which are not observed in main clauses. Thus an applied phrase may be the head noun of a P-oriented relative clause in these languages, provided that it is eligible to be the pivot or subject in P-oriented clause types (see §6.5.3). With the exceptions of Ampenan Sasak and Makasar, the argument that is co-referential with the head noun is restricted to the pivot (or analogous privileged syntactic relation). In Makasar a single argument in the clause may be co-referential with head noun (see below). In Ampenan Sasak, access to relativization is restricted only to core arguments (see Khairunnisa 2022: 84-87).

In Nasal, as is typical for two-way symmetrical voice languages of the sample, voice alternations and applicatives are used together to modulate access of peripheral semantic roles to relativization. As shown in (229a) and (230a), in Nasal, when the verb in the relative clause is in PV, only P may be the head noun. When AM-marking is used together with a PV construction, the applied phrase now maps to the the P argument, and thus may serve as the head noun of the relative clause. In (229b), the verb in the relative clause is marked with the locative/goal-selecting AM *-i*, and accordingly, the goal applied phrase is the head noun. In (229b), the verb in the relative clause is marked with the instrument-selecting AM *-kun*, and accordingly, the instrument applied phrase is the head noun.

(229) Nasal, Instrument-selecting AC in relative clause

- a. *tulis-an* [sai di-tulis anak=ku jenu] kak ku=hapus.
 write-NMLZ REL PV-write child=1SG.POSS earlier PFV 1SG.NPIV=[PV]erase
 ‘I erased the writing that my child wrote (on the wall).’ (BC)
- b. *sisai* [sai di-tulis-i anak=ku jenu] kak ku=bekhesih-kun.
 wall REL PV-write-LOC.APPL child=1SG.POSS earlier PFV 1SG.NPIV=[PV]clean-CAUS
 ‘I cleaned the wall that my child wrote on.’ (AC) (McDonnell fieldnotes)

(230) Nasal, Goal-selecting AC in relative clause

- a. *manuk [sai ku=panggul jenu] lijung.*
chicken REL 1SG.NPIV=[PV]hit earlier flee
'the chicken that I hit earlier ran away.' (BC)
- b. *tungkuk [sai ku=panggul-kun khan manuk jenu] patuh.*
staff REL 1SG.NPIV=[PV]hit-INST.APPL with chicken earlier break
'the staff that I used to hit the chicken broke.' (AC) (McDonnell fieldnotes)

Relative clauses in Makasar do not make use of a relativizer. Instead a relative clause simply follows the head noun, and a clitic =*a* that marks definiteness appears on the right edge of the verb. In zero-marked transitive clauses, generally only the P argument may be the head noun of the relative clause. For A to be eligible to be the head noun of a relative clause, the verb must be marked for an A-oriented type of construction with the actor focus prefix *aN-* or the semi-transitive prefix *aN(N)-* (see Jukes 2020: 228). In Makasar, as in Nasal, a peripheral semantic role is only eligible to be the head noun when the verb in a relative clause is marked with an AM. This is shown in the example in (231). Here the verb is marked with the locative/goal AM *-i*, and accordingly, the locative applied phrase is the head noun. In (232), likewise the verb is marked with the instrumental AM *-ang*, and accordingly, the instrument applied phrase is the head noun. Thus, the use of ACs in combination with voice alternations modulates access of phrases expressing different sets of semantic roles to relativization. In this respect, the combination of P-oriented voice constructions with pivot-neutral applicatives together functions like LV and CV alternations (pivot-selecting applicatives) in Philippine-type languages (see §6.5.2).

(231) Makasar, Goal-selecting AC in relative clause

- sikola [na=mange-i=a agang=ku] baji'=i.*
school 3=go-LOC.APPL=DEF friend=1.POSS good=3
'the school my friend goes to is good.' (AC) (Jukes 2020: 229)

(232) Makasar, Instrument-selecting AC in relative Clause

- sele' [ni-buno-ang=a=i] tarang=i.*
kris PASS-kill-INST.APPL=DEF=3 sharp=3
'the kris he was killed with was sharp.' (AC) (Jukes 2020: 229)

Additionally, certain focus and clefting constructions are also restricted to the pivot argument, or to core arguments more generally, in many languages of West Nusantara. In Makasar, for instance, a peripheral semantic role is only eligible to be focused in clauses where the verb bears AM-marking (Jukes 2020: 228). This is shown in (233). In this type of construction, the argument that occurs in preverbal position receives special pragmatic focus. In canonical word order in Makasar, clausal arguments occur in postverbal position.

(233) Makasar, Focused location applied phrase

- tapper=e' ku=empo-i.*
mat=EC 1=sit-LOC.APPL
'I sit on a mat.' (AC) (Jukes 2020: 311)

In West Coast Bajau, a similar focus construction is observed (Miller 2007: 206-207). Again, the argument in preverbal position receives special pragmatic focus, though in West Coast Bajau this position is restricted specifically only to pivot arguments. As shown in (234a), in PV clauses, the P argument is the pivot and thus can be focused in the preverbal position. Accordingly, when the verb is marked with the AM *-an*, a peripheral semantic role is eligible to be focused. This is shown in (234b), where the beneficiary applied phrase is the pivot and thus may appear in preverbal position, receiving pragmatic focus. Non-pivot arguments cannot be fronted in the same manner. Again, the use of ACs in combination with voice alternations modulates access of phrases expressing different sets of semantic roles to positions of structural and discursive prominence. Just as with relativization, in Philippine-type languages, it is LV and CV constructions that are used to provide access to peripheral semantic roles to fronting and clefting constructions in a similar manner (see §6.5.2).

(234) West Coast Bajau, Focused beneficiary applied phrase

- a. *Telumpa' e boi beli=ni ta' Kuzik.*
 shoes DEM CMPL [PV]buy=3SG.NPIV LOC K.
 'She bought the shoes for Kuzik.' (BC)
- b. *Kuzik boi beli-an=ni telumpa' e dilaw.*
 K CMPL [PV]buy-BEN.APPL=3SG.NPIV shoes DEM yesterday
 'She bought Kuzik the shoes yesterday.' (AC) (Miller 2007: 206)

7.13 Summary of findings

In this chapter, I have presented a functional typology of ACs and other AM-marked constructions based on 24 languages of West Nusantara with pivot-neutral applicatives. AMs in these languages are generally polyfunctional, though this is not observed equally for all functions and across all languages of the sample. In this concluding section, I summarize some key distributional patterns for AMs and AM-marked constructions which are demonstrated in the chapter.

Each of languages in the sample has between one and four AMs; with applicative systems with two distinct AMs being most common. In languages with more than one AM, the predominant pattern is to show one AM that marks beneficiary, recipient, instrument, and theme applied phrases and another AM that marks locative and goal applied phrases, plus possibly addressee and recipient roles. Outliers include West Coast Bajau, Yakan, and Ampenan Sasak which have a single AM that is generalized or partially generalized, and Tolaki, which has a single benefactive/instrumental AM. Six languages of the sample also show more than two AMs.

ACs show differing structural properties depending on the role of the applied argument. ACs with beneficiary applied phrase most commonly occur only with transitive bases. Beneficiary-selecting AC usually show in an increase in the number of core arguments over a corresponding BC, such that these ACs are maximally ditransitive. One exception to this is benefactive ACs marked with Balantak *-kon*, which are maximally monotransitive and valency-preserving with transitive bases.

Instrument-selecting ACs are quite often monotransitive, with the companion phrase (patient or goal) remapped to an oblique phrase. This occurs especially in Malayic languages and

languages of Java and Sumatra. In Pendau and a number of other languages of Sulawesi, instrumental ACs are maximally ditransitive. Most languages with instrument-selecting ACs also have theme-selecting ACs marked with the same AM. The two constructions share semantic similarities in that like themes, most instrument applied phrases in these languages express an entity that is directed into motion. Theme ACs are generally monotransitive in the languages of the sample, with remapping of the companion phrase when the BC is transitive.

Goals, locations, and addressees are typically marked with the same AM. The syntactic properties of these goal ACs is more diverse than both benefactive and theme-selecting ACs. With some exceptions, goal ACs may occur on intransitive and transitive bases, and with transitive bases, may either increase the transitivity or show remapping of the applied phrase and the companion phrase.

When AMs that mark beneficiaries, instruments, goals, and locations attach to intransitive bases, they also select applied phrases expressing various other semantic roles, including content, stimulus, and theme-like products of bodily processes. These constructions result in monotransitive ACs. For many ACs of this type there is no monoclausal BC equivalent.

AM-marked constructions may also show non-applicative functions. Thus some AM-marked constructions do not show selection of an applied phrase but instead have a purely semantic effect such as indicating repeated or pluractional aspect, greater intensity or changes in the properties of non-A arguments, such as individuation, definiteness, or specificity.

Causative functions are also attested for AMs, but not equally across all languages of the sample. In the languages of the sample, eight of the languages show no productive causative morphology distinct from AMs. The other 16 show productive causative prefixes, and in 14 of these, such prefixes are the most productive markers of morphological causative constructions by far. In many such cases, AM-marking and causative prefixation may combine. In all but three languages of the sample, one or both suffixes that mark ACs can also mark causative constructions. For these suffixes, AM-marked constructions may be both causative and applicative, with the AM functioning to select both a causer A argument and a peripheral role as a non-A core argument.

Finally, in West Nusantara languages with symmetrical voice systems, the pivot plays an important role in syntactic operations, and the combination of voice and applicative morphology allows peripheral semantic roles, such as beneficiaries, instruments, goals, and locations to be the pivot. Syntactic operations, such as relativization, in many of the languages in the sample are reserved for the pivot. Analogous patterns are found in many languages with asymmetrical or marginally symmetrical voice systems, which often retain an alternation between P-oriented constructions and A-oriented constructions only in relative clauses or participial structures. In this way the selection of voice alternations plus use of applicatives modulates access of constituents expressing peripheral roles to positions which are privileged syntactically (e.g. pivot, head noun of relative clause) and prominent in discourse (e.g. focused, or highly topical).

Chapter 8

Exploring West Nusantara applicative constructions through the lexicon

8.1 Introduction and rationale

In this chapter, I explore patterns in the observed functions of pivot-neutral applicative morphemes (AMs) when applied to a representative sample of possible bases over the lexicon. These bases will be examined across a sample of nine languages of West Nusantara, and compared in the distribution of the constructional meanings that are observed with AM-marking across them. The goals of the chapter are (i) to investigate the extent to which lexical semantics is predictive of the function of AMs observed and structural properties of AM-marked constructions, (ii) to make a preliminary identification of components of semantic meaning that influence the distribution of functions of AMs across the lexicon, and (iii) to investigate the extent to which these patterns are consistent across languages. The results show that certain lexical meanings are more consistently attracted to particular constructions meanings marked by AMs, which may be used to infer components of semantic meaning that influence these patterns. The results further show that consistency in such patterns is generally not observed equally across the possible bases; instead attraction of bases with certain semantic properties to particular AM-marked functions may be limited to a narrow band of the lexicon, or even just a few highly lexicalized stems. Consistency is also not observed across all languages, such that observed patterns in the distribution of functions of AMs across lexical bases cannot be generalized outside of a small band of western Indonesia. Thus, I conclude that patterns in the observed functions of AMs across the lexicon in well-known cases like Indonesian and Javanese are not predictive of similar patterns in Sulawesi languages and West Nusantara languages spoken on the periphery of the region.

Previous research on applicatives in West Nusantara languages has shown that AMs behave differently with different sets of bases. From such studies it is clear that syntactic properties of base verbs alone, especially verbal subclass (e.g. stative, dynamic intransitive, or transitive) cannot account for the total distribution of applicative, causative, and aspectual functions of AMs across bases. The most often discussed of these cases by far is Indonesian *-kan* marked verbs, and authors have taken different approaches to this problem. Vamarasi (1999) posits that a beneficiary is part of certain lexical verbs' valence (e.g. *beli* 'buy') but not others (e.g. *gendong* 'carry') even while admitting that for the former "their meanings do not seem to demand a Benefactive

NP” (89). Cole & Son (2004) argue that the causative and applicative functions of *-kan* are a unitary syntactic licensing function, the effects of which depend on the thematic structure of the base verb. For this to be tenable, they must also hold that beneficiary roles are licensed in the thematic structure of base verbs compatible with benefactive *-kan* (i.e. semantic structure) but not their syntactic argument structure. Kroeger (2007) invokes lexical semantics frequently in his explanation of the distribution of causative and applicative functions of *-kan*, including the stipulation that “predicates that do not involve motion” are not compatible with the meaning “CAUSE-BECOME-AT” of *-kan*₁ (what I have called theme-selecting ACs), and he also holds that benefactive *-kan*₂ is not available with intransitive bases. Thus it is clear that lexical semantics plays a role in the distribution of functions of AMs across potential lexical bases, but there is disagreement about which ACs are influenced by lexical semantics of the base and little clarity on how these effects might be characterized outside of very limited contexts, i.e. only Indonesian *-kan* marked verbs and to a lesser extent Indonesian *-i* marked verbs (Arka 1993) and Javanese *-i* and *-ake* marked verbs (Hemmings 2013; Vander Klok & Evans 2022).

Meanwhile in Chapter 7, across a sample of 24 West Nusantara languages, I show that particular ACs, e.g. a beneficiary-selecting AC vs. an instrument-selecting AC, are not equally compatible with all types of lexical bases from a given language, and this again is not explained in full by syntactic properties of bases. Semantic properties, e.g. whether the base describes a transfer event, an act of striking, a process performed on material items, or a locomotion activity, also appear to be key correlates of the compatibility of AM-functions with potential bases. In Chapter 6, across a sample of 50 languages, I also show that pivot-neutral AMs in particular are highly polyfunctional, commonly showing causative, aspectual, or intensive meanings (see §6.4). Here again it is evident that these functions are not equally compatible with all potential lexical bases, and may in fact apply to just small subsets of bases. For example, the intensive aspectual meaning of the AM *-(C)i* in *Tukang Besi*, are reported with just a small number of bases, e.g. ‘to punch’, ‘to slap’, ‘to kick’, and ‘to reprimand’ (Donohue 1999: 243).

Thus, it is clear that functions of AMs in these languages are compatible with some lexical bases and not others, and that to some extent this is irrespective of syntactic properties of bases. Moreover, lexical semantics is often invoked or implicated in descriptive and theoretical accounts of these patterns. It follows that lexical semantics plays a role in determining the possible meanings that result when a particular AM combines with a base, and whether it may do so or not. Even so, components of lexical semantics that influence compatibility with the various functions of AMs have not been clearly articulated outside of a few limited contexts, even though the essential problem, i.e. how to predict which functions appear with which bases, is very widespread in West Nusantara.

Therefore, in this chapter, I investigate patterns in the observed functions of AMs across a range of lexical bases with consistent meanings, for a sample of languages of West Nusantara, across which such patterns may be compared. The organization of the remainder of this chapter is as follows. In §8.2, I present the language sample and data sources used. In §8.3, I describe the methods used for compilation of data, including sampling of lexical meanings and coding of data. In §8.4, I give a general overview of results, showing the relative productivity of various AMs and their functions across languages of the sample. In §8.5, I present detailed results for lexical meanings showing fairly consistent patterns of attraction to constructional meaning

across languages.¹ These are presented according to semantic role of the applied phrase including: beneficiaries, themes, instruments, goals, and other locative roles, following by causative constructions and pluractional constructional meanings. In §8.6, the chapter concludes with a summary of findings and discussion of implications.

8.2 Language sample and sources of data

Nine languages of West Nusantara with pivot-neutral applicatives were selected for inclusion in the language sample for this lexical study. These are given in Table 8.1 below. This sample is a subset of the 24 languages used for the functional typology in Chapter 7 (see §7.2). Also listed are the lexical resources used as sources for this portion of the study in addition to descriptive resources already noted in §7.2. For Indonesian, Javanese, and Sundanese, supplementary data from the Leipzig Corpora Collection were also used (see Goldhahn, Eckart & Quasthoff 2012). Because this study is exploratory in nature and nothing of the sort has been previously conducted for western Austronesian languages, I decided to investigate a smaller subset of the 24 languages in more detail.

Table 8.1: Language sample used for the lexical study

No.	Language	Gen. Grp.	Location	Sources
1	Nasal	NWS-BI	Sumatra	McDonnell fieldnotes
2	Yakan	GRB	Philippines	Behrens 2002
3	S. Barisan Malay	MAL	Sumatra	McDonnell fieldnotes
4	Std. Indonesian	MAL	(wide use)	Pusat Bahasa (Indonesia) 2007
5	Sundanese	SUN	Java	Truong fieldnotes
6	Javanese	JAV	Java	Robson & Wibisono 2002; Horne 1974
7	Ampenan Sasak	BSS	Lesser Sundas	McDonnell & Khairunnisa in prep.
8	Balantak	S-B	Sulawesi	Busenitz & Bradbury 2016
9	Muna	M-B	Sulawesi	van den Berg & Marafad 2016

The 24 languages used for the functional typology portion of this study were selected by genetic group and subbranch, according to availability of descriptive, pedagogical, and/or lexical resources. In order to compile the necessary data for this portion of the study, only languages with fairly extensive lexical resources, i.e. dictionaries, and corresponding sets of sentential or clausal examples, either as part of the dictionary, or separately in corpora and descriptive materials, could be included. This significantly restricts the pool of eligible languages, especially for languages with smaller speaker populations and comparatively little previous linguistic research. Ultimately, this results in some imbalance in the sample. Of the 24 languages in the larger sample, nearly half of the languages immediately were excluded, because no lexical resources are available beyond a simple wordlist or lexicon.

I prioritized languages with fairly complete, accessible, and electronically searchable sources of data, and languages for which I have access to original data from fieldnotes, including Sun-

¹In this chapter, I use the term *attraction* in reference to certain lexical meanings that show consistently high compatibility with certain AM-marked constructions across languages. This differs from the use of the term attraction in corpus-based collostructional analysis (see e.g. Stefanowitsch & Gries 2003).

danese data I collected with Dewi Setiani, and data generously shared with me by Bradley McDonnell and Khairunnisa. Six of the languages were already included in a earlier pilot study (Truong & McDonnell 2021). I also sought to avoid exacerbating geographic imbalances in the sample by not oversampling from languages of Sulawesi and Java. In follow up studies, shortcomings of this sample might be remedied by inclusion of one or more languages in the following categories: non-Malayic languages of Sumatra, languages of the Barrier Islands, languages of Borneo, and languages belonging to the South Sulawesi group.

8.3 Methods used for coding and sampling lexical meanings

The Leipzig Valency Questionnaire is used for sampling of lexical meanings (see Appendix B.2). This questionnaire consists of 80 lexical meanings with frames showing a prototypical or targeted clausal argument structure. For each meaning in the sample in each language, one lexical base is selected and coded for various properties as follows. Based on the source material, it is recorded whether the base occurs with each possible AM in the language's inventory. For each AM-marked construction that occurs, information on the syntactic structure of the construction, and the constructional meaning(s) that are observed when base and AM combine is recorded. Each AM-marked construction is then classified by functional type, the semantic role of the applied phrase, and its effect on argument structure. Information on the syntactic structural of corresponding BCs was also recorded.

During data compilation, some meanings included in the Leipzig Valency Questionnaire were found to not be commonly expressed as a single unique lexical item across the languages of the sample. For these meanings substitutions and adjustments were made as listed below. The result was that 71 total meanings were coded separately for bases in the study.

- For BURN, the frame was substituted with “The man burns leaves.”
- BE A HUNTER was omitted, as this was not commonly expressed as one lexeme, and no appropriate substitution was found.
- FEEL COLD was not commonly expressed as one lexeme, so the meaning BE COLD was substituted. Likewise SHOUT was substituted for SHOUT AT.
- For LIVE, bases with the meaning ‘stay’ were sometimes included to better match the frame, “The old people live in town.” For LIKE, bases with the meaning ‘love’ were sometimes substituted to match the frame “The boy likes his new toy.”
- For GO, COME was substituted, as some languages of the sample use a preposition for GO.
- DRESS and WIPE were omitted, due to difficulties compiling bases with similar semantic meanings.
- For BUILD, MAKE was substituted, as most languages of the sample use a morphologically complex form with a base meaning ‘stand’ or ‘arise’ for ‘build’
- The pairs SIT/SIT DOWN, LIVE (STAY)/LEAVE, FEAR/FRIGHTEN, SEE/LOOK AT, DIE/KILL, HIT/BEAT, CARRY/BRING, and SEE/SHOW were coded together, as most languages of the

sample use the same lexical base for both, distinguished by form of verbal morphology. This has the result of reducing the number of transitive base meanings represented, thus PLANT and BORROW were added to compensate for this in part.

For each lexical meaning and each language, one root lexeme was identified, and based on the source material, AMs which may be affixed on the root lexeme were identified. The following data were then compiled for each pair of BC and AC identified.

1. The form of the lexical root.
2. Morphology used on the root in the BC.
3. Valency of the BC
4. Role(s) mapping to core arguments in the BC
5. AM which is affixed on the verb.
6. A brief definition of the AM-marked verb
7. Valency of the AM-marked construction
8. Role(s) mapping to core argument in the AM-marked construction
9. Function of the AM
10. Change in argument structure from BC to AM-marked construction

Sentential examples from published or unpublished sources representing the BC structure and AC structure were also included in notes.

The coding categories used are listed in Table 8.2 below. More than one coding category was assigned when applicable, e.g. an AM-marked construction may be coded as GOAL+PLUR if its structure is consistent with a goal-selecting applicative and it also shows pluractional aspect (while the BC shows neither). The compiled data used for analysis is included in Appendix E. These data were analyzed using R statistical software version 4.2.3 (R Core Team 2023) and the tidyverse and stringr packages (Wickham 2017, 2019).

8.4 Overview of results

In this section, I give an overview of general results from the lexical study including measures of productivity of functions of AMs and specific forms of AMs across languages in the sample.

A summary of the compiled data used for analysis in the study is shown in Table 8.3. Each language is listed, along with its ISO-639-3 code and the number of distinct AMs in its inventory, which ranges from one to three AMs. A total of 71 distinct meanings were used to compile data for the study, but in some languages, the source material did not contain a lexical base which can be used with the structure of the targeted frame for every sampled meaning, resulting in a lower number of total lexical bases. For each language in the sample, between 61 and 71 lexical bases are represented in the compiled data. In each language and for each possible AM, more

Table 8.2: Coding categories used for the lexical study

Functional categories	
APPL	Applicative, selection of peripheral role as core
CAT	Verbalizing with non-verbal base
CAUS	Causative, selection of instigator or effector as S or A)
LEX	Lexicalized change in semantic meaning
INTENS	Intensive meaning
OBLIG	The affix is obligatorily present on a verbal base
OPT	The affix is optionally present with no other observed function
PAT	Selection of patientive P argument not otherwise observed
PLUR	Pluractional aspectual meaning
other	Other than the above functions (e.g., imperative, emphatic)
NA	No applicative affix is attested with this lexical base
Semantic role categories for applicatives	
ADDR	Addressee-selecting (communication events)
BEN	Beneficiary-selecting (transfer and non-transfer events)
COM	Comitative-selecting (i.e., accompanier)
CONT	Content-selecting (cognition and communication events)
GOAL	Goal-selecting (motion events)
INST	Instrumental-selecting applicative (i.e. manipulated entity)
LOC	Location-selecting (i.e. general location or static location)
PATH	Path-selecting (motion events)
PURP	Purpose-selecting (i.e., future intention)
REAS	Reason-selecting (i.e., prior cause)
REC	Recipient-selecting (transfer events)
STIM	Stimulus-selecting (perception and sensory events or states)
TARG	Target-selecting (emotive events or states)
THM	Theme-selecting (motion events and spatial states)
Categories for argument structure	
no-change	Number of arguments and mapping of roles does not differ from the base construction
remapping	Number of non-A core arguments is the same as the base construction, but mapping of roles differs
monotransitive	Maximal number of non-A core arguments is one, representing an increase over the base construction
ditransitive	Maximal number of non-A core arguments is two, representing an increase over the base construction
detransitizing	The maximal number of non-A core arguments represents a decrease compared to the base construction
NA	The base cannot be used as a verbal predicate without an applicative morpheme OR the base is not attested with any applicative morpheme

than one unique AM-marked construction may be recorded, as long as each shows a different function for the AM. It is also possible for a particular AM not to be attested in combination with particular base. For each meaning in given language, between zero and four unique AM-marked constructions are represented in the data, with the total number of AM-marked constructions for each language given in the rightmost column.

Sundanese has the highest total number of unique AM-marked constructions in the data at 144, while Yakan has the least at 44. Four of the nine languages show a much lower number of attested AM-marked constructions: Muna, Balantak, Ampenan Sasak and Yakan, and this cannot be explained in full by the total number of lexical bases represented, as Nasal and South Barisan Malay both have much higher total numbers of AM-marked constructions across the same number of lexical bases or fewer. For Ampenan Sasak and Yakan, which each have only AM, this divergence might be explained by the number of available AMs in their inventories, but for Muna, which has two AMs, and Balantak, which has three, a similar explanation is not tenable.

Table 8.3: Lexical data collected by language

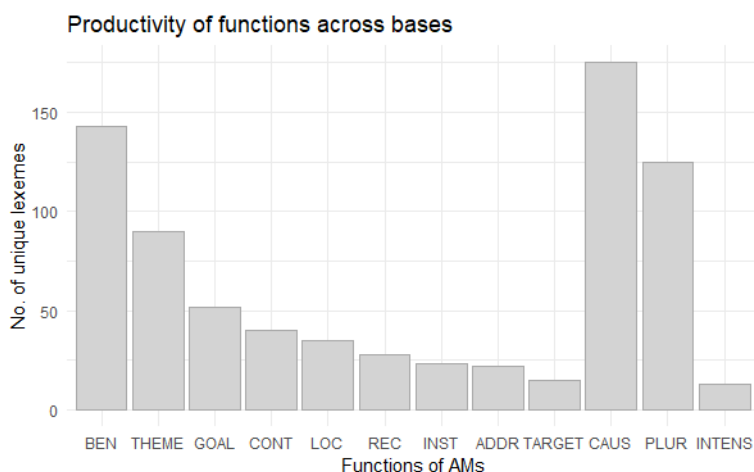
Language	Code	AMs	Bases	Marked const.
Sundanese	sun	3	71	144
Indonesian	ind	2	70	128
Javanese	jav	2	70	126
Nasal	nas	2	61	117
S. Barisan Malay	bes	2	61	96
Ampenan Sasak	sas	1	61	66
Balantak	blz	3	68	55
Muna	mnb	2	65	51
Yakan	yka	1	68	44
TOTAL		17	595	827

Total productivity of the functions of AMs across all languages and all types of AMs is presented in the chart in Figure 8.1. As shown in the chart, the applicative function with the highest total productivity by far is the beneficiary-selecting function, being found with 143 unique base lexemes out of 595 total unique bases. The next most productive applicative function is the theme-selecting function found with 87 unique bases. All other applicative functions are much more limited in productivity across sampled meanings, being found with 50 or fewer total unique bases.

For non-applicative functions of AMs, the causative function is the most productive, being found with 175 unique base lexemes out of 595 total bases. The pluractional aspect function is also highly productive being found with 124 unique lexical bases. Meanwhile, the greater intensity function is the least productive function of all, being found only with 13 unique lexical bases.

To facilitate comparison of data for like AMs across the languages, each AM was assigned to a category based on broad types for pivot-neutral AMs identified in the typological survey, see §6.2. Languages of the sample and the AMs in their inventories by category are shown in Table 8.4. Type I indicates an AM that marks ACs in which a beneficiary, instrument or theme

Figure 8.1: Productivity of functions across all languages and AMs



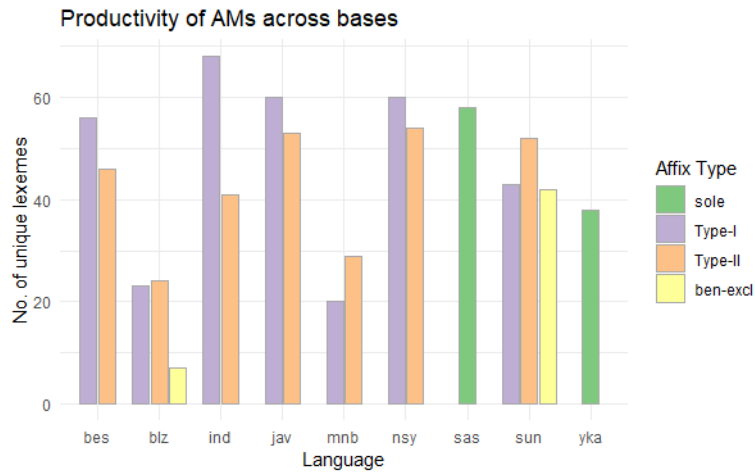
is the applied phrase. Type II indicates an AM that marks ACs in which a locative or goal role is the applied phrase. The category labelled “sole”, indicates a generalized or partially generalized AM that is the sole AM in a given language. The category labelled “Ben-excl” was used for two instances in which an AM has a unique form and is exclusively used in beneficiary-selecting ACs.

Table 8.4: Applicative morphemes by language and functional type

Language	Code	AMs	Type I	Type II	Sole	Ben-excl
Sundanese	sun	3	<i>-keun</i>	<i>-an</i>	—	<i>pang- -keun</i>
Balantak	blz	3	<i>-kon</i>	<i>-i</i>	—	<i>-ii</i>
Indonesian	ind	2	<i>-kan</i>	<i>-i</i>	—	—
Javanese	jav	2	<i>-aké</i>	<i>-i</i>	—	—
Muna	mnb	2	<i>-ghoo</i>	<i>-Ci</i>	—	—
Nasal	nas	2	<i>-kun</i>	<i>-i</i>	—	—
S. Barisan Malay	bes	2	<i>-ka</i>	<i>-i</i>	—	—
Ampenan Sasak	sas	1	—	—	<i>-an</i>	—
Yakan	yka	1	—	—	<i>-an</i>	—

Data showing the total productivity of individual AMs by type and language is presented in the chart in Figure 8.2. The languages of the sample primarily spoken in Java and Sumatra, i.e. South Barisan Malay, Nasal, Javanese, and Sundanese—along with Indonesian—show relatively high productivity of all AMs in their inventories, which each of these being found with at least 40 unique bases across meanings sampled. Sasak *-an* is similarly productive, being found with 58 unique bases, and Yakan *-an* nearly rises to a similar level of productivity, being found with 38 unique bases in the data. In Balantak and Muna, however, the total level of productivity for all AMs is quite a bit lower than any of the previously mentioned languages, with each of the AMs being found with between 20 to 29 unique bases, except the Balantak special benefactive AM *-ii* which is found with only seven unique bases across sampled meanings, representing the

Figure 8.2: Productivity of functions across AMs by type and language



least productive AM in the data by this measure. By comparison, the productivity of the special benefactive AM *pang-* *-keun* in Sundanese is among the most productive AMs, being found with 42 unique lexical bases.

The data presented in this section indicate that there is substantial variance in the overall productivity of functions of AMs, with beneficiary-selecting, causative, and pluractional functions being among the most productive functions of AMs by far. There is also substantial variance in the productivity of individual AMs across languages observed in the data. The Sulawesi languages Muna and Balantak consistently show low productivity of AMs, even taking into account the type and number of AMs found in these languages.

One explanation for the varying levels of productivity of AM may be low usage of AMs in certain languages to form causative constructions. Across AM-marked constructions in the data, the mean number of constructions that may be considered causative is 26.55 per language, and the median is 25. This refers to constructions in which an instigating causer is introduced as the A argument of the AM-marked construction. Languages with overall low productivity of AMs tend to show values much lower than the median for this measure. In the data, the causative function is found in only 16 unique AM-marked constructions in Sasak, 11 in Balantak, 10 in Yakan and 3 in Muna. On the high end, Javanese, Sundanese, and Indonesian shows the most usage of AMs in forming causative constructions, ranging from 45 to 48 unique AM-marked constructions.

The incidence of pluractional meanings is also a contributor to variance in overall levels of productivity. Across AM-marked constructions in the data, the mean number of constructions that show a pluractional meaning of any type (durative, habitual, iterative, plural participants) is 13.89 per language, and the median is 11 per language. Languages with low overall productivity of AMs tend to show values much lower than the median for this measure. In the data, the pluractional meaning is found in only nine unique AM-marked constructions in Yakan and Javanese, and one in Balantak and Sasak. On the high end, the pluractional function is found with 32 AM-marked constructions in Sundanese, 24 in South Barisan Malay, and 23 in Nasal.

In the following section, I will look at more detailed patterns for various applicative functions across the lexical meanings sampled in the data, and will identify lexical meanings that show more consistent patterns of distribution for these functions AM-marked constructions. However,

it must be kept in mind throughout that some languages and some individual AMs show low overall productivity across the data. Thus, a portion of the variance observed in the distribution of applicative functions across lexical meanings is due to this larger pattern, rather than semantic components of meaning that may be ascribed to the lexical bases represented.

8.5 Lexical meanings showing consistent patterns of association

In this section, I identify lexical meanings that show consistent patterns of compatibility with various AM-marked constructions, according to their constructional meanings or functions, or attraction to certain AM-marked functions. To compare the relative compatibility of meanings with a given AM-marked function, for each base meaning attested with the function in any language, the total number of languages in which that function is attested is tallied, and the mean, median, and standard deviation are then calculated for the function across base meanings. To be considered consistently compatible or attracted to an AM-marked function, a cut-off level is calculated as one standard deviation above the mean, rounding down to the nearest integer. Any base meaning that is attested with this AM-marked function in a number of languages equal to or greater than the cut-off level is included for analysis in the subsections that follow.

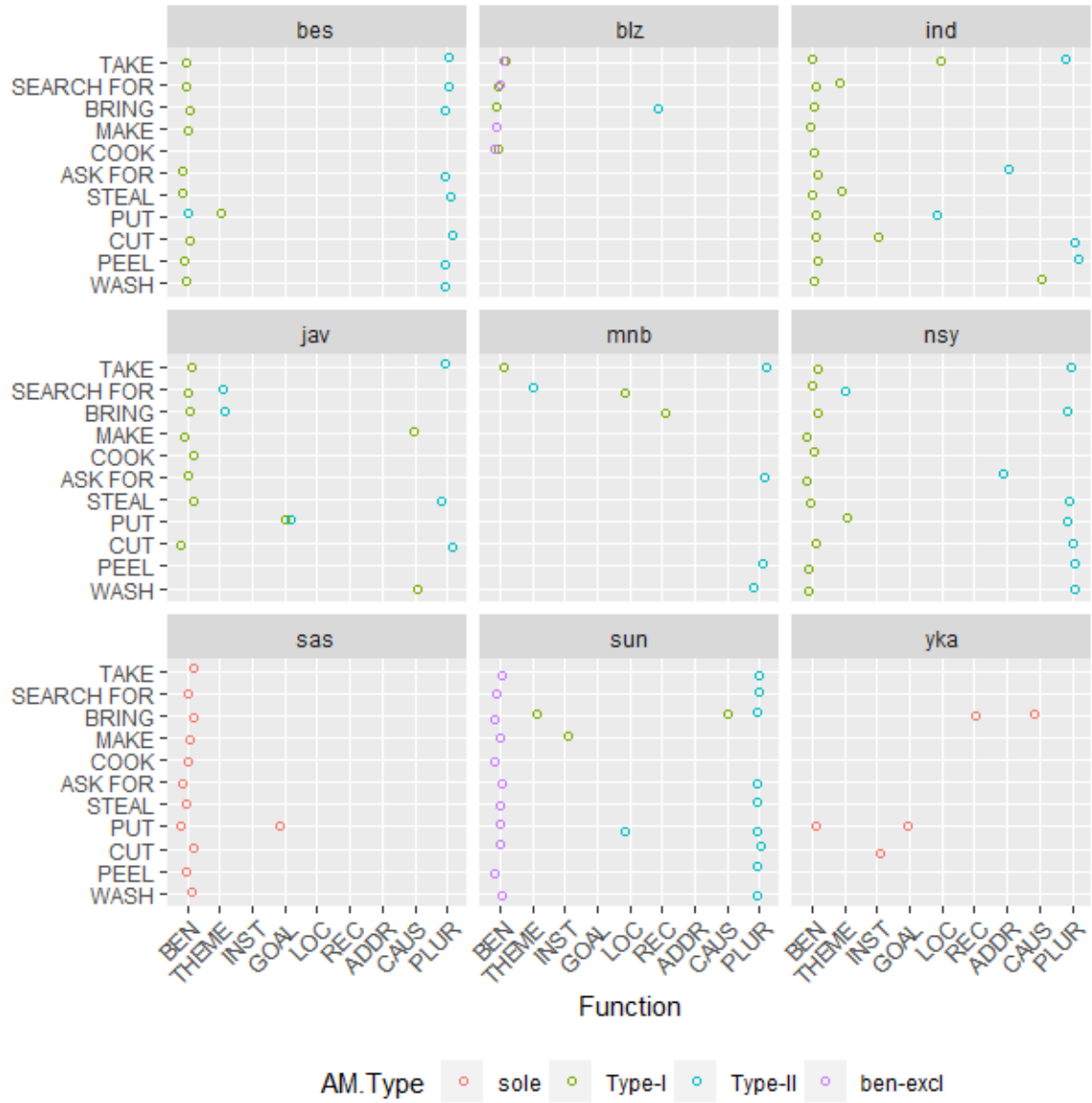
8.5.1 Beneficiary-selecting constructions

In the compiled data, the beneficiary-selecting applicative function of AMs is found with 49 lexical meanings (49 of 71, or 69%) across all languages and AMs. Across these meanings, the mean number of languages showing the beneficiary-selecting function with any AM is 3, the median number of languages is 2 and the standard deviation is 2.16. Meanings which were considered to show consistently moderate to high levels of attraction to the beneficiary-selecting function are those for which this function is attested in at least five languages of the sample. There are 11 such meanings in the data, representing 97 unique lexemes in the data. The patterns of AM-marked constructions for these are represented in the chart in Figure 8.3.

As shown in the chart, across languages of the sample, the most common pattern is for one or more AMs of the Type I, sole, or special benefactive categories to mark a beneficiary-selecting AC with these base meanings, while a Type II AM, if present, indicates pluractional aspect with the same base. In Sasak and Balantak, only the beneficiary-selecting function is typically indicated with these base meanings, while in Muna only pluractional meanings are indicated with moderate consistency, and Yakan, shows neither at any level of consistency.

Lexical bases in this set may be semantically divided into three subsets on the basis of their semantic characteristics. First, bases showing attraction to the beneficiary-selecting function include verbs indicating acquisition or transfer of possession: TAKE, SEARCH FOR, BRING, ASK FOR, and STEAL. The meaning PUT could possibly be related to this set, though it shows a different possible pattern as well, being found with locative or goal-selecting ACs in five of the nine languages. Second, base meanings showing attraction to the beneficiary-selecting function include verbs of creation, in which a created item is fabricated or otherwise made to exist: BUILD and COOK. Finally, base meanings showing attraction to the beneficiary function include verbs

Figure 8.3: Lexical meanings associated with beneficiary-selecting ACs



describing processing of materials or items: CUT, PEEL, and WASH. These are generally on the more moderate side of attraction to the benefactive function, compared to the other two sets.

Together these subsets describe components of meaning for lexical bases which show a relative attraction to ACs with benefactive meanings in these languages. These patterns reinforce that beneficiary-selecting ACs in these languages most commonly indicate the participation of a recipient-beneficiary; this participant accrues benefit by virtue of gaining access to desirable materials or items in a desirable condition. As reported above, there is also fairly high incidence of pluractional meanings observed with Type II AMs with the set of bases showing consistent benefactive meanings. This might result from the focus on tangible materials or items as the undergoer of the eventive action; and the consideration that acquiring or processing multiple of such items is an activity that is commonly relevant in usage. The bases that show attraction to the beneficiary-selecting ACs are almost exclusively monotransitive, though benefactive ACs forms with these bases may be either monotransitive or ditransitive.

8.5.2 Theme-selecting constructions

The theme-selecting applicative function of AMs is found with 24 lexical meanings (24 of 71, or 34%) across all languages and AMs. The mean number of languages showing the theme-selecting function with any AM is 4.17, the median number of languages is 4, and the standard deviation is 2.44. Meanings which were considered to show consistently moderate to high levels of attraction to the beneficiary-selecting function are those for which this function is attested in at least six languages of the sample. There are nine such meanings in the data, representing 81 unique lexical bases of which 53 show at least one possible theme-selecting AC. The patterns of AM-marked constructions for these meanings are represented in the chart in Figure 8.4.

As shown in the chart, the most common pattern is for a Type I or sole AM to mark a theme-selecting AC with these base meanings. With some bases, the same AM may mark a separate causative construction, e.g. Sasak, in which *pelai-an* can make ‘to run (away) with s.t.’ or ‘to make s.t. run quickly’.

These lexical bases may be semantically divided into three subsets on the basis of their semantic characteristics. First, bases showing attraction to the theme-selecting function include lexemes indicating the relative position of a theme or a change in position of a theme: POUR, PUSH, BORROW, FILL, MEET, THROW and TIE. Second, one base meaning showing attraction to the theme-selecting function describes a bodily function, COUGH. Third, the meaning RUN show a high level of consistent attraction to the theme-selecting AC, with clausal meanings like ‘he ran off (with) the goods’. The lexemes representing RUN may mean ‘to run away, to flee’, e.g. Indonesian *lari*, which indicates directional motion, though they are sometimes also found in examples indicating a locomotive activity, e.g. ‘to run fast, to be running’. Verbs composed of roots meaning ‘run’ plus a Type-I AM also appear to be highly lexicalized in a good number of cases, taking on the specific connotation ‘to elope’, ‘to run off with s.o. in order to take them as a spouse’.

Lexical bases in this set that show a theme-selecting AC in the data are split in base valency, with 19 unique bases showing a base valency of one, and 31 showing a base valency of two. For three unique lexical bases found with this function, the base does not function as a verbal predicate without AM-marking. These include *tali* ‘rope’ in Javanese and Sundanese, which is nominal, and Balantak *dudul* ‘push’, which is not used as a verb without AM-marking. Despite

Figure 8.4: Lexical meanings associated with theme-selecting ACs

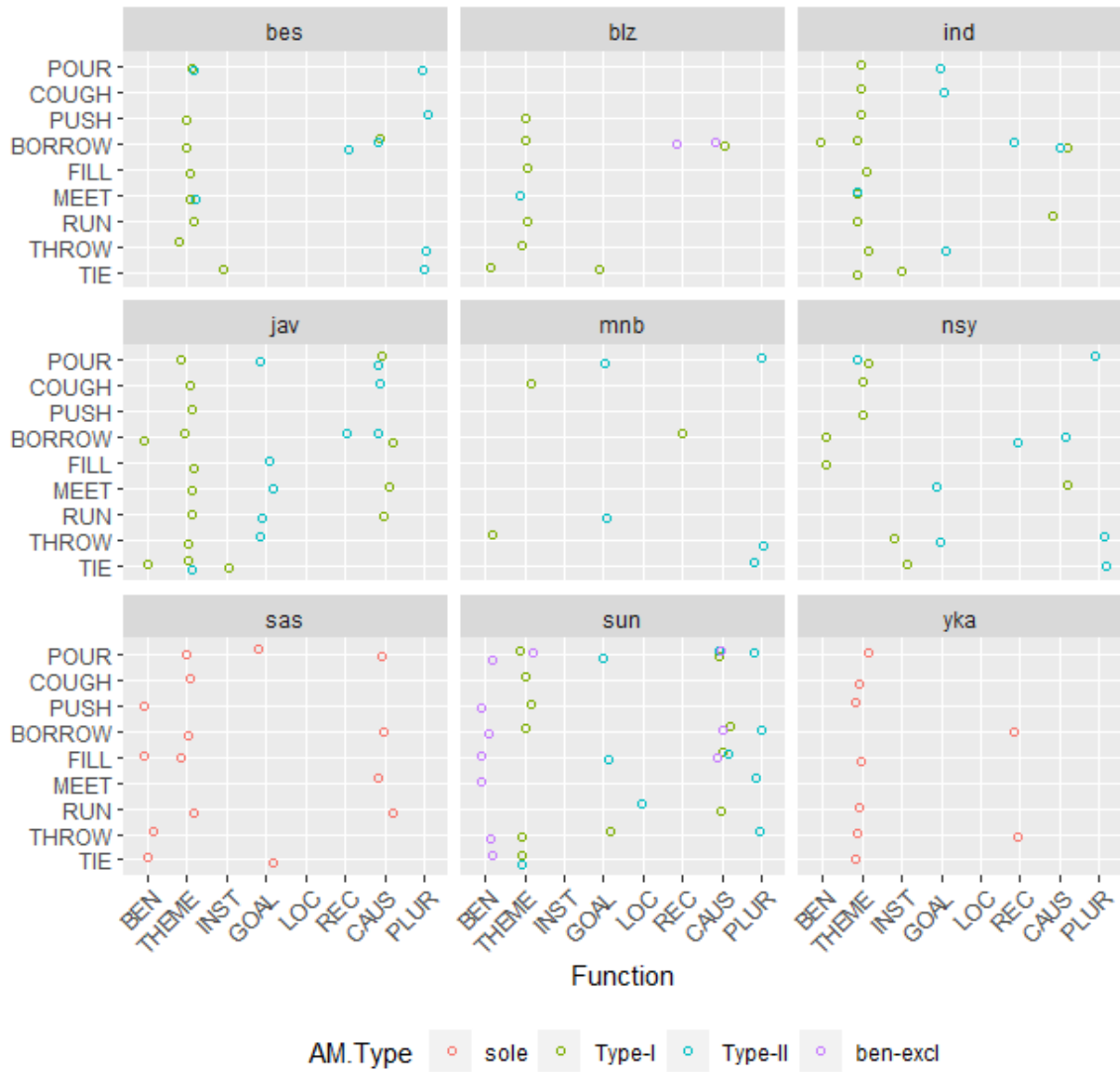
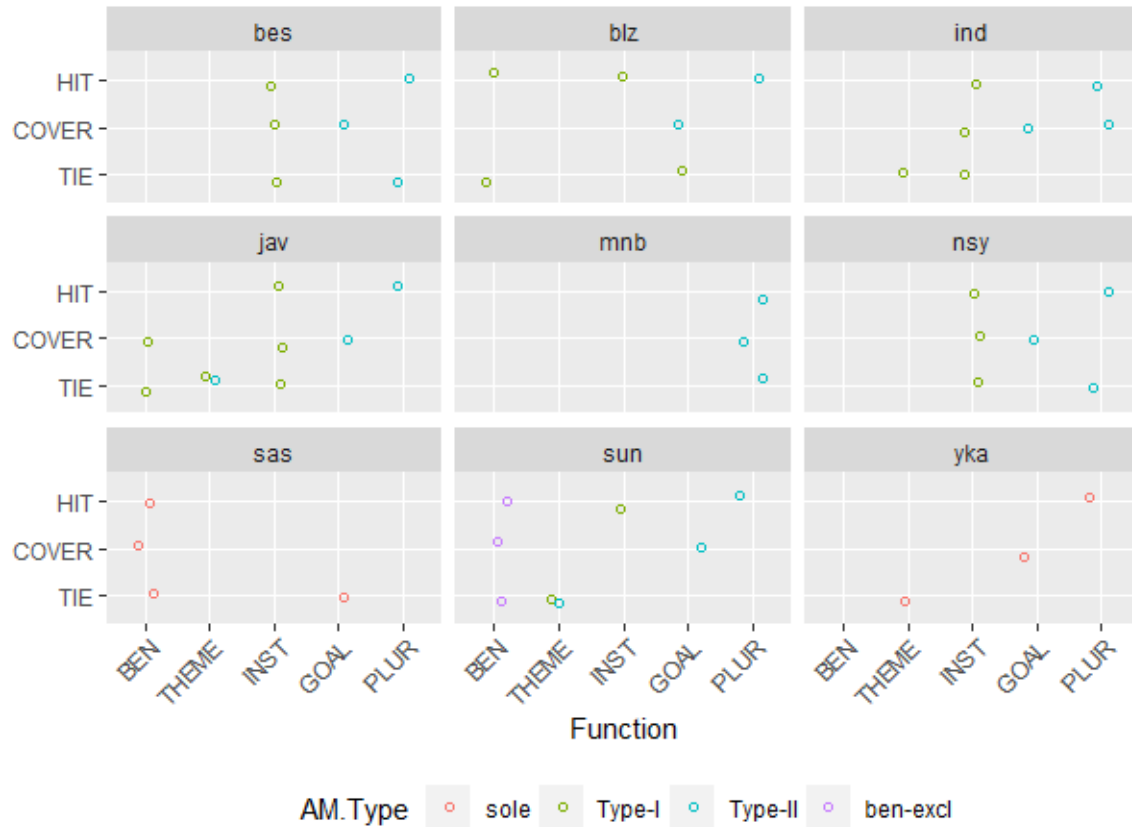


Figure 8.5: Lexical meanings associated with inst-selecting ACs



this split, there is strong consistency in the valency of theme-selecting ACs formed on these bases, 59 of the 62 unique AM-marked theme-selecting constructions for this set of meanings are monotransitive, and just three are ditransitive.

8.5.3 Instrument-selecting ACs

The instrument-selecting applicative function of AMs is among the least productive functions. It is found with ten lexical meanings (24 of 71, or 14% of possible meanings) across all languages and AMs and 23 total unique bases. Across these meanings, the mean number of languages showing the instrument-selecting function with any AM is 2.3, the median number of languages is 1.5, and the standard deviation is 1.77. Meanings which were considered to show consistently moderate level of attraction to the beneficiary-selecting function are those for which this function is attested in at least four languages of the sample. There are only three such meanings in the data, representing 14 unique lexical bases across the nine languages. The patterns of AM-marked constructions for these meanings are represented in the chart in Figure 8.5.

Sasak *-an* and Yakan *-an* do not show an instrumental-selecting function. In Yakan, there is an IV pivot-selecting AC marked with *paN-* but it is not clear how productive this construction is, and it is not included in this study. It is difficult to draw conclusions as to the patterns in the functions of AMs across these bases given the sparsity of data. In the data it observed that Type I

AMs generally mark instrument-selecting ACs, and the bases attracted to this function are most consistent across four languages: South Barisan Malay, Indonesian, Javanese, and Nasal.

Of the three meanings that comprise this set—HIT, COVER, and TIE—examples of the latter two typically show an inherent semantic instrument, e.g. the object that covers something, or the material that binds something. In a few cases, lexical bases representing these meanings are only used as nouns expressing such instruments, e.g. the base *tali* in Sundanese and Javanese is a noun that refers to a rope or other similar material, and these bases cannot be used as verbal predicates without AM-marking. With the meaning HIT, some lexical bases represented in the data do specifically mean to hit something with an object, e.g. Javanese *thuthuk*, but most do not, e.g. Sundanese *teunggeul* and Balantak *bobok*, which describe a general act of striking, including striking acts done with an object or part of the agent's body.

The bases in this set that occur with the theme-selecting function are almost always monotransitive, with 12 of the 14 unique bases showing a base valency of two. The last two bases, Javanese *tali* 'rope', and the Sundanese base with the same form, as mentioned above, do not operate as a verbal predicate without AM-marking.

8.5.4 Goal-selecting ACs

The goal-selecting applicative function of AMs is quite a bit less productive compared to the beneficiary-selecting function and even the theme-selecting function, but much more productive than the instrument-selecting function. It is found with 19 lexical meanings (19 of 71, or 27% of possible meanings) across all languages and AMs and 52 total unique bases. Across these meanings, the mean number of languages showing the goal-selecting function with any AM is 2.84, the median number of languages is 2, and the standard deviation is 2.19. Meanings which were considered to show consistently moderate to high level of attractions to the goal-selecting function are those for which this function is attested in at least four languages of the sample. There are six such meanings in the data, representing 51 unique lexical bases across the nine languages, of which 32 are associated with at least one goal-selecting AC. The patterns of AM-marked constructions for these meanings are represented in the chart in Figure 8.6.

Of the meanings that comprise this set—COME, COVER, RAIN, POUR, PUT, THROW—all six have a semantic component of directional motion. However these bases vary in the type of motion expressed. COME expresses self-motion or simple motion. COVER, POUR, PUT, and THROW express caused motion. RAIN describes an meteorological event and is not a motion verb per se. It does describe an event that entails motion (i.e. of rainwater), which may be why it is attracted to the goal-selecting function of AMs.

The lexical bases in this set that are found in goal-selecting ACs show varying base valency. Four of the 32 unique bases show a base valency of zero (i.e. zero-place RAIN verbs), while 11 unique bases show a base valency of one and 16 show a base valency of two. The last remaining unique base is Balantak *talop* 'cover (n.)' which does not operate as a verbal predicate without AM-marking. Despite this split for base valency, the goal-selecting ACs are remarkably consistent in their valency. 32 of the 34 unique goal-selecting ACs in the data are monotransitive, and just two are ditransitive.

Figure 8.6: Lexical meanings associated with goal-selecting ACs

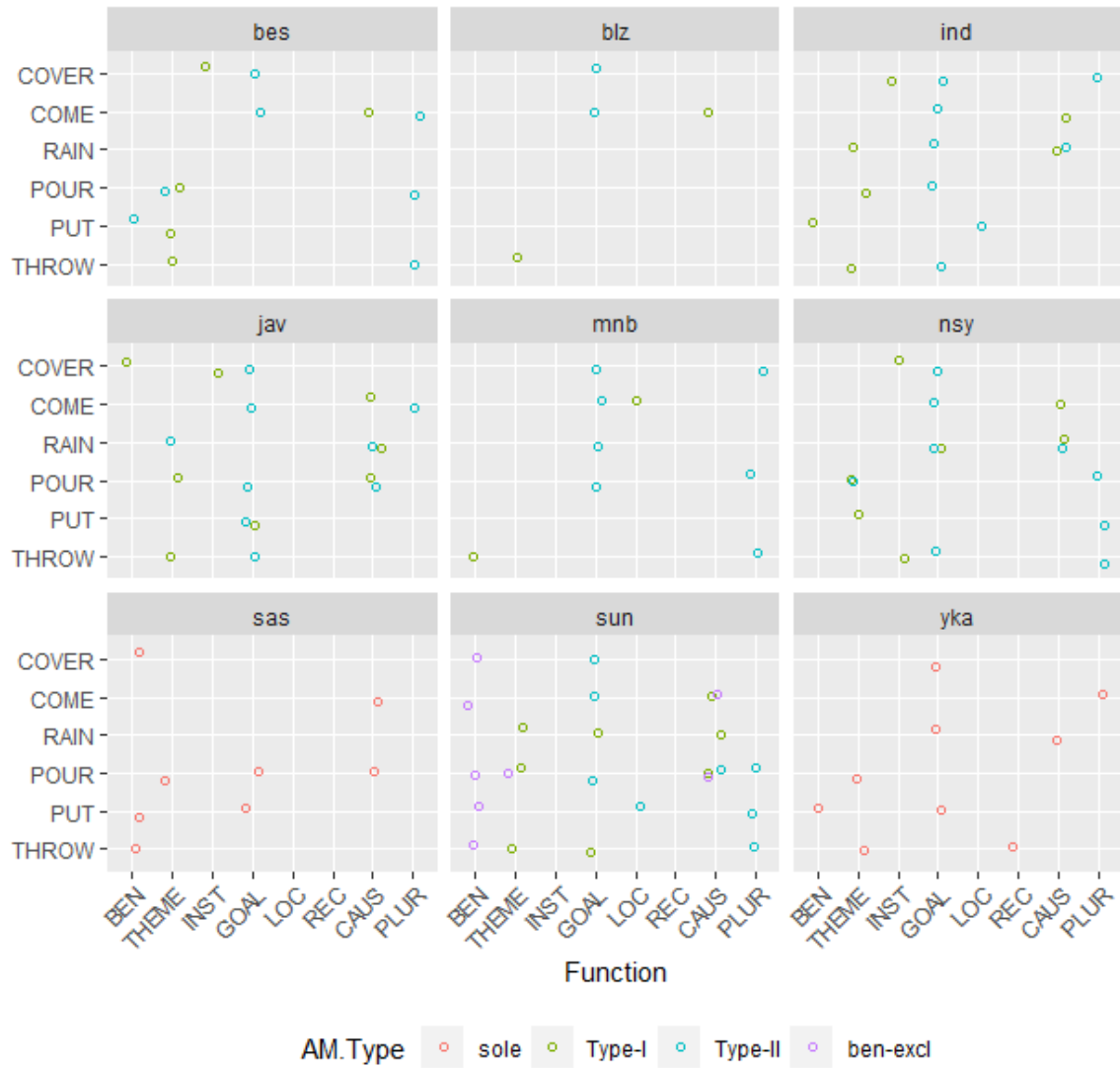
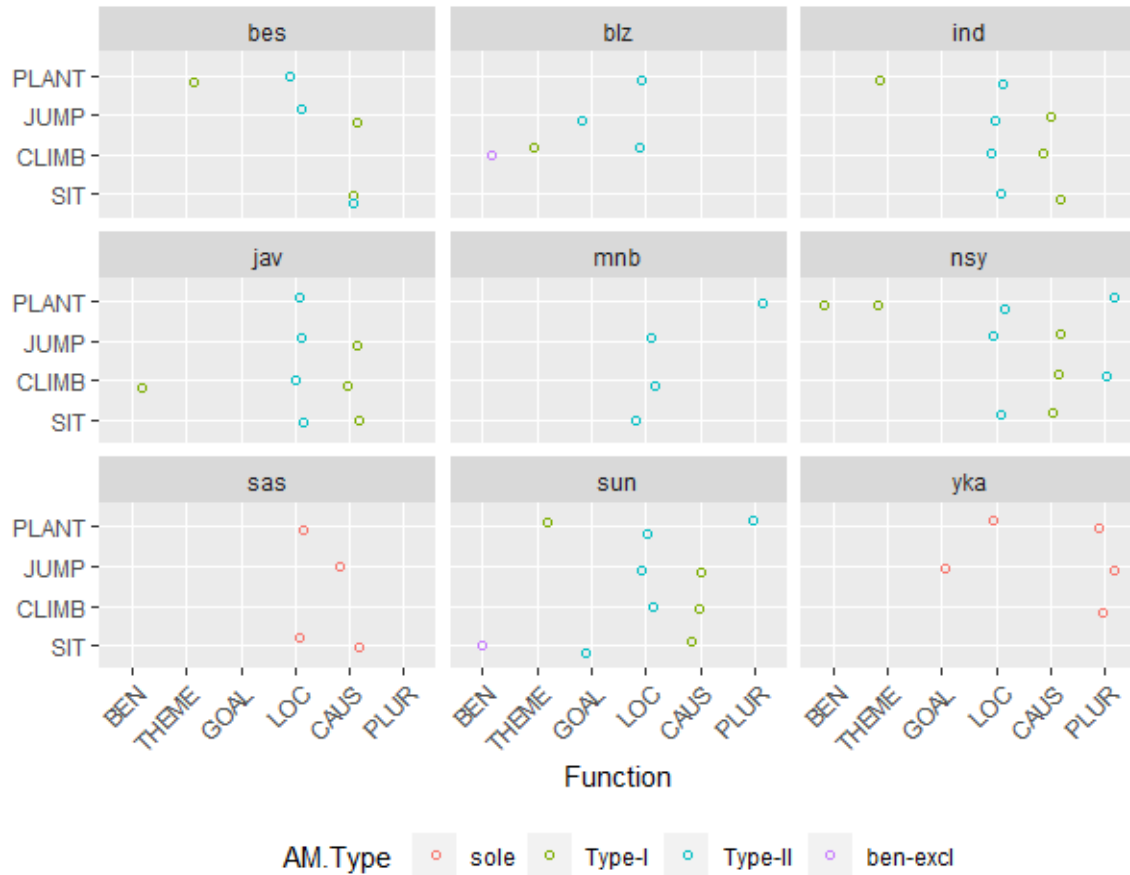


Figure 8.7: Lexical meanings associated with locative-selecting ACs



8.5.5 Locative-selecting ACs

The locative-selecting applicative function of AMs show relatively low productivity, though not as low as the instrument-selecting function. It is found with 11 lexical meanings (11 of 71, or 15% of possible meanings) across all languages and AMs and 35 total unique bases. Across these meanings, the mean number of languages showing the goal-selecting function with any AM is 3.18, the median number of languages is 2, and the standard deviation is 2.52. Meanings which were considered to show consistently moderate level of attraction to the locative-selecting function are those for which this function is attested in at least five languages of the sample. There are four such meanings in the data, representing 34 unique lexical bases across the nine languages, of which 24 are associated with at least one AC that selects a location, source, or path as the applied phrase. The patterns of AM-marked constructions for these meanings are represented in the chart in Figure 8.7.

Sasak shows just two locative-selecting ACs for this set of meanings, and Yakan only one. This is in keeping with the general pattern in these two languages for low productivity of this function. In the languages of Java and Sumatra, and in Indonesian, the primary pattern is for a locative-selecting applicative function to be marked with a Type II AM on these bases, while a causative meaning is marked with the Type I AM on the same base, which is observed a fair

portion of the time. In Balantak and Muna, we also see Type II AMs marking the locative-selecting applicative function, but almost no instance of causative functions for the same bases with AM-marking. Semantically the bases are somewhat varied, including two locomotion verbs (JUMP, CLIMB), one postural verb (SIT), and one verb of placing (PLANT).

The bases that are used with locative-selecting ACs in the data vary by base valency. 15 of the unique bases show intransitive BCs, while the other nine show transitive BCs. Despite this fact, across unique ACs in the data, valency of the AC is largely consistent. Of the 25 unique locative-selecting ACs, 22 are monotransitive and only three are ditransitive.

8.5.6 Causative constructions

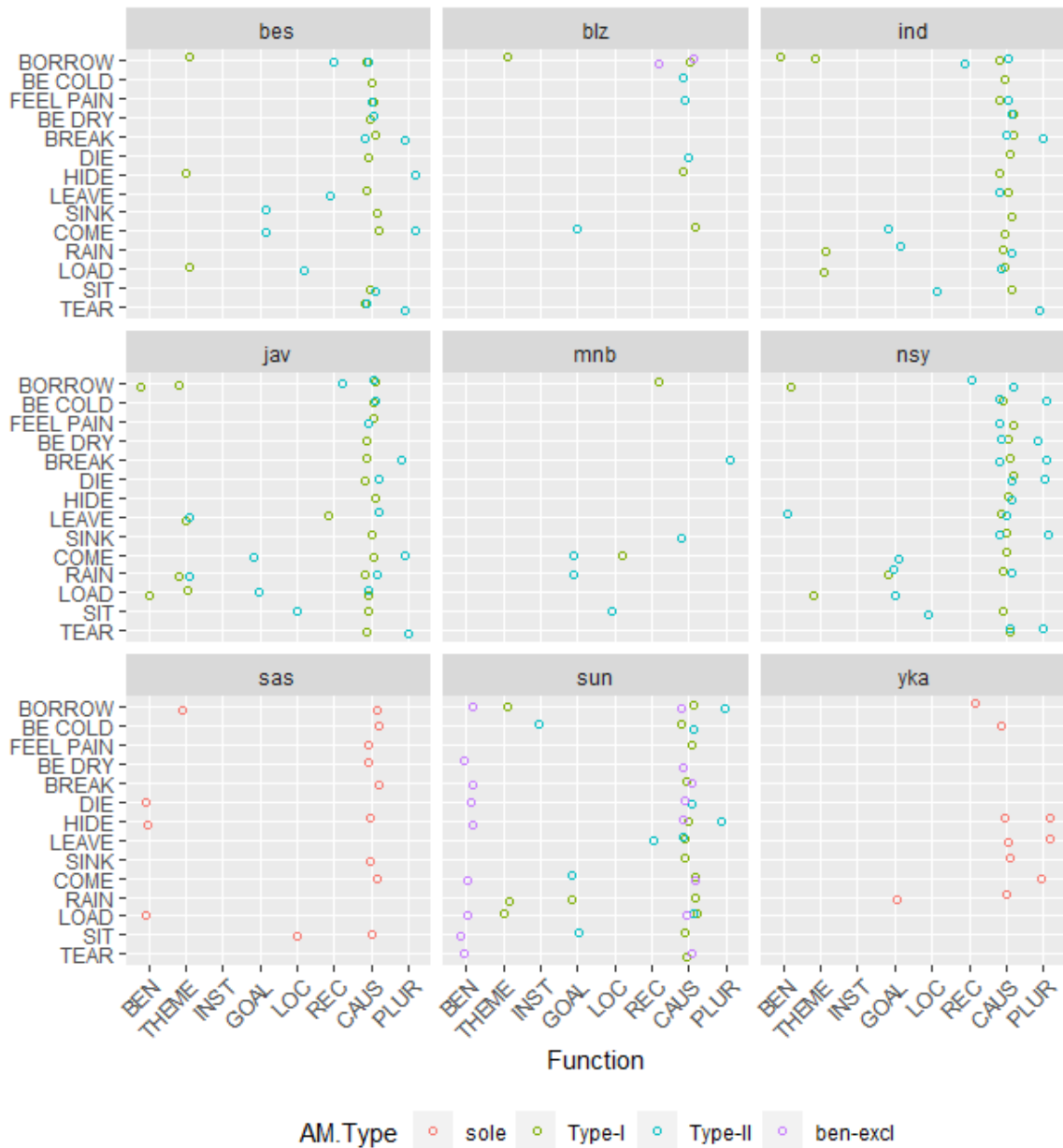
As mentioned earlier, the causative function of AMs shows by far the highest productivity of any one function across the sampled meaning. It is found with 47 lexical meanings (47 of 71, or 66% of possible meanings) across all languages and AMs and 174 total unique bases. Across these meanings, the mean number of languages showing the causative function with any AM is 4.91, the median number of languages is 4, and the standard deviation is 3.19. Meanings which were considered to show consistently moderate to high levels of attraction to the causative-selecting function are those for which this function is attested in at least seven languages of the sample. There are 14 such meanings in the data, representing 120 unique lexical bases across the nine languages, of which 86 are associated with at least one causative AM-marked construction, i.e. a construction in which an instigating causer is introduced as an A argument. The patterns of AM-marked constructions for these meanings are represented in the chart in Figure 8.8. As seen in the chart, causative constructions can be marked by all types of AMs in the data, and this is commonly found for Type I AMs, Type II AMs, and sole AMs, but not as commonly with special benefactive AMs. The patterns observed for this set of bases varies primarily by language. Muna, Yakan and Balantak show only a small number of bases found in causative constructions, with Sasak showing a moderate number of bases with causative meanings, while the remaining five languages—South Barisan Malay, Indonesian, Javanese, Nasal, and Sundanese—being rife with causative AM-marked constructions.

The bases observed with causative AM-marked constructions are mostly intransitive, as reflected in 69 unique bases, while five zero-place bases and 12 monotransitive bases are also found in this set. Semantically, there is a wide variance as well. This set includes bases that expressive stative conditions (e.g. BE COLD, BE DRY), inchoatives (e.g. TEAR, BREAK), and motion verbs (e.g. COME, LEAVE), among others. Causative AM-marked constructions are nonetheless mostly monotransitive, with 111 unique AM-marked causative showing monotransitive structure, while 16 show ditransitive structure, and only three are intransitive.

8.5.7 Pluractional meanings

The pluractional meaning is also very productive across the data, coming behind only the causative function and the beneficiary-selecting function in productivity across unique bases. The pluractional meaning of AMs is found with 125 unique bases representing 48 meanings (48 of 71, 68%). Across such bases, the mean number of languages in which a pluractional meaning is observed is 2.60, while the median number of languages is 2, and the standard deviation is 1.70. Meanings were considered to show consistently moderate to high levels of attraction to the pluractional

Figure 8.8: Lexical meanings associated with causative AM-marked constructions



function in this study include those found with this function in at least four or more languages. This is the case for 14 meanings representing 120 unique lexical bases, of which 67 show a pluractional constructional meaning with at least one AM. Patterns for these bases are shown in the chart in Figure 8.9.

As seen in the chart, the pluractional meaning is not generally found for Sasak *-an*, and only sparsely attested for Yakan *-an*. In other languages the pluractional meaning is quite productive across this set, except in Balantak. The pluractional meaning is associated almost exclusively with Type II suffixes in the languages of the sample.

Semantically these bases include a good number of verbs denoting intentional actions to effect an entity (e.g. HIT, EAT, CUT, BURN), as well as a few bases expressing inchoative events (e.g. BREAK, TEAR), and few expressing caused motion actions (e.g. TAKE, PLANT, STEAL). In terms of valency, the large majority the bases are monotransitive (60 of 67), while the remaining seven are intransitive. Despite this variance in base valency, the pluractional AM-marked constructions are exclusively monotransitive across 67 unique AM-marked pluractional constructions.

8.6 Summary of findings and implications

To conclude the chapter, I outline the major findings of the lexical study described above and discuss implications of these findings, including how they might inform future research on the relationship between lexical semantics and applicatives in West Nusantara languages.

The analysis presented in this chapter demonstrates one way of measuring the distribution of constructional meanings of AMs across potential bases in the lexicon. This distribution is not monolithic, but marked by many kinds of variance observed in the data. The productivity of AMs shows variance across the lexicon, across forms of AMs, across functions of AMs, and across languages. These patterns underscore some observations that descriptive linguists have often grappled with in previous research, especially the observation that the function of a given AM in combination with a given lexical base is not completely predictable from the syntactic properties of the base, such as its base valency, its syntactic subclass (e.g. stative, dynamic intransitive, transitive), or other properties of the structure of the BC, such as the mapping of semantic roles to specific positions in argument structure. Lexicalization of AM-marked stems also clearly plays a role, which is evident with specific combinations of base meaning and AM-marking, e.g. Type I AMs showing the meaning ‘to elope’ with base lexemes meaning ‘to run’ and Type II AMs consistently showing both causative and theme-selecting functions with bases meaning ‘to borrow’, such that the AM-marked stems have meanings like ‘to lend something’.

This complexity notwithstanding, the results presented in this chapter point to some semantic properties of lexical bases that influence the compatibility or attraction of such bases to certain functions or constructional meanings associated with AMs. This is most evident in the data with those functions which have higher overall productivity.

For example, lexemes indicating acquisition or transfer of possession (e.g. TAKE, BRING) show a strong attraction to the beneficiary-selecting function of AMs, as do lexemes that describe the creation or processing of materials (e.g. MAKE, WASH). Such lexemes show an inherent semantic compatibility with the constructional meaning of beneficiary-selecting ACs in these languages, as these ACs are associated with the participation of a recipient-beneficiary, and thus receiving possession, and also are associated with accrual of benefit, and thus activities that are

Figure 8.9: Lexical meanings associated with pluractional AM-marked constructions



useful or productive.

Bases with a different set of semantic properties were shown to have strong compatibility or attraction to theme-selecting ACs in the data. In large part, such bases describe the (stative) position of entity relative some reference point (e.g. TIE, as in ‘I tie a bandage on her arm’, or FILL, as in ‘The crater is filled with water’), or a change in the relative position of an entity (e.g. PUSH, as in ‘He pushed the cart down the street’, or RUN/FLEE, as in ‘They fled from the flood’). Such bases show an inherent semantic compatibility with the constructional meaning of the theme-selecting ACs in these languages, as such AC are associated with caused motion, and these bases all show an entity in space or motion in their semantic meanings. This entity can take different structural positions in BC, including an S argument, as with RUN/FLEE, or a P argument, as with PUSH, or an oblique PP, as with FILL. This shows that it is not the mapping of the targeted role in syntactic structure of the BC that is important, but that such a role is present in what can be called the semantic structure of the base.

This last point is related to the findings in this study regarding structural properties of BCs and AM-marked constructions. For almost every function of AM examined, bases that show attraction to this function are split across those showing intransitive BCs and those showing monotransitive base structures, with a handful of strictly non-verbal bases also observed. Yet the transitivity of the AC is generally not predicted by the transitivity of the base, but is instead nearly fixed, with most AM-marked constructions being almost exclusively monotransitive, except for beneficiary-selecting ACs, which may be ditransitive in some languages. Likewise, the mapping of roles to argument structure in an AC is not strictly predicated on the mapping of roles found in the corresponding BC. So with the theme-selecting ACs mentioned above, the entity that is located or moving in space in the BC always maps to P in theme-selecting ACs, no matter the grammatical relation held by the same referent in the BC. This shows that AM-marked constructions and lexical bases both have identifiable semantic properties that influence their compatibility with one another, though these effects may not be observed to apply equally over the entire lexicon, but instead shown more strongly with a subset of possible bases, i.e. lexicalization is occurring.

Finally, the results of this study give rise to another very important implication. Overlaid on the patterns of compatibility or attraction between lexical bases and AM-marked constructional meanings, is an overarching pattern of varying productivity of those constructional meanings across languages. Thus it is unmistakably clear that patterns in the behavior of AMs and AM-marked constructions cannot be generalized from western Indonesian and Malayic languages to other West Nusantara languages, especially the languages of Sulawesi, and languages located on the extremities of West Nusantara, such as Yakan, which is spoken in the Southern Philippines, and Ampenan Sasak, which is spoken in the Lesser Sundas just outside the boundary dividing West Nusantara and East Nusantara. In Muna and Balantak, for instance, the meanings that AM-marked stems show in Indonesian and Sundanese are usually expressed using other grammatical resources of the language. This includes the use of stems marked with stative and intransitive verbal morphology in contrast to basic transitive voice marking (e.g. AV, PV), and the use of causative prefixes of various kinds. For example, Sundanese uses AV marking to express the state of being full with a material (*ng-eusi*) and Type II AM-marking to express an action causing a container (goal) to be filled (*ng-eusi-an*). In Muna, however, these two meanings can be expressed by use of active participial marking (*mo-pono-no* ‘full’) vs. causative marking (*feka-pono* ‘to fill up s.t.’). In languages with more elaborate systems of verbal morphology, modulation between intransitive and transitive verbal constructions using other types of morphology thus may serve

the functions of AM-marking observed in languages with less elaborate morphological systems, and the presence of other kinds of grammatical resources in a language may preclude the AMs from taking on these functions at all, or showing them with specific lexical bases.

As the lexical study described in this chapter is exploratory in nature, its findings are far from providing an complete account, or even an adequate account, of the relationships between lexical semantics and functions of applicative morphology in West Nusantara languages. The results do however suggest some directions that further research might take. In further research, it could be advantageous to include more transitive bases, and to include more bases representing different semantic classes of verbs. For instance, the instrument-selecting function of ACs is not well represented in the data, perhaps because of a lack of appropriate meanings in the target list for sampling, such as more nouns that can be used as implements, or activities that typically make use of implements. Additional non-Malayic languages of Sumatra and South Sulawesi languages should also be included, in order to verify the extent to which semantic relationships between bases and ACs are attributable to sub-areal patterns. West Nusantara languages that do not show applicatives might also be considered for comparison, along with constructions utilizing other kinds of valency modulating morphology, e.g. causative prefixes, stative and other intransitive marking. Nonetheless, it is clear even from the limited sampling of languages included in this study, that conclusions based on research on well-known languages like Indonesian and Javanese and their applicative systems cannot necessarily be generalized to other West Nusantara languages with applicatives, even if the forms of the applicative morphology found in these languages show historical continuity and functional similarities. For this reason, it is all the more important to include diverse representation of applicative systems in future research on the properties, usage, and development of applicatives.

Chapter 9

Conclusion

In this concluding chapter, I summarize some of the major findings and contributions of this study and discuss implications for future research.

From the start, the goal of the study has been to understand the applicative systems of West Nusantara in typological context, but also on their own terms, in the context of the diachronic and synchronic systems in which they developed and are used. In addition, I have given special attention to broadening the description and cross-linguistic comparison of applicative constructions (ACs) in languages of West Nusantara. Accordingly, the depth and the breadth by which I have described ACs in West Nusantara is a major contribution of this study. In it, I have shown how diversity in the types of data and perspectives considered can shed light on the functions, properties, and development of applicatives in these languages. This has been driven by two factors.

First, the conceptual framework for applicatives that I developed in Part I of this study allows for broad inclusion in the types of constructions that were considered relevant. In Chapters 1 and 3, I outlined a constructional approach whereby ACs are understood as a conventionalized pairing of a fixed form and consistent meaning, and I showed how elements of form and meaning for ACs might be further elaborated. This focus led to consideration of a variety of clausal constructions with functional similarities or formal similarities to canonical applicatives, including non-canonical ACs of various types, and non-applicative AM-marked constructions with diverse functions.

Second, I have sought to emphasize distinctions that are meaningful in the internal linguistic systems of West Nusantara languages. In Chapter 2, I gave a thorough descriptive account of the applicative system of Sundanese, a language which has received relatively little attention in previous research on applicatives. As part of that case study, I showed that ACs in Sundanese cannot be clearly distinguished from causative constructions, pluractional meanings, and category-changing constructions marked by the same morphemes. Furthermore in Sundanese the overlap between applicative and non-applicative functions of AMs is observed across diverse forms of AM-marking and in diverse types of ACs, e.g. goal-selecting ACs with *-an*, theme-selecting ACs with *-keun* and beneficiary-selecting ACs with the unique applicative circumfix *pang-* *-keun*.

In keeping with these two driving factors, in part of Chapter 3, I examined the relationship between applicatives and symmetrical voice in western Austronesian languages. On the one hand, in languages like Balinese, symmetrical voice and applicatives are two separate dimensions

of the verbal system. In languages like Kimaragang, on the other hand, certain symmetrical voice constructions (LV, CV, etc.) can be considered applicatives in cross-linguistic context on the basis of their form and function. But within the internal grammatical system of the individual language, they operate within a single paradigm of alternations that includes constructions that are clearly not applicatives (AV, PV). On this basis I proposed the terms *pivot-neutral* applicative for the Balinese type and *pivot-selecting* applicative for the Kimaragang type, and showed how this would allow for better integration of Philippine-type languages of West Nusantara like Kimaragang—and Formosan and Philippine languages—into a cross-linguistic typology of applicatives.

In Part II of the study, I conducted a large scale typological survey of the languages of West Nusantara, examining data in 85 languages of the region representing extensive geographic diversity and genetic diversity under Malayo-Polynesian. This yielded a breadth and depth of data on applicatives in West Nusantara on a scale far larger than any previous research. Based on the survey, in Chapter 5, I established the basic distributional facts for applicatives in West Nusantara languages. The results include a number of important patterns not previously identified.

First, while the presence of pivot-neutral applicatives are the norm in West Nusantara, there are conspicuous subareal patterns whereby applicatives are missing in Mainland Southeast Asia north of Peninsular Malaysia, and Borneo south of Sabah. This lack of applicatives correlates with other typological changes in the relevant languages, including more reliance on fixed word order to signal grammatical relations, shift to analytic rather than morphological means for modulating voice and valency, and loss of suffixation. In mainland Southeast Asia it is well known that these changes were caused by language contact with non-Austronesian languages, especially members of the Austroasiatic family. In Borneo, there are also indications that these changes have spread through contact, beginning in southwest Borneo in the area where Land Dayak languages are spoken and gradually affecting North Sarawak languages and Greater Barito languages in adjacent areas, as well as some Malayic languages like Iban and Mualang. While Adelaar (2006) previously identified some of these trends in West Kalimantan and southern Central Sarawak, the extent of these patterns is larger than he recognized at the time. Still, the initial trigger for these changes in Borneo is not identified, though one must wonder whether speakers of non-Austronesian languages present in Borneo upon the arrival of Malayo-Polynesian groups did not influence the course of development for the newcomers' languages.

Second, in West Nusantara not only is the presence of pivot-neutral applicatives the norm, pivot-neutral applicatives are associated with diverse typological properties of language, and are found across widely differing systems of alignment and across many types of symmetrical and asymmetrical voice systems. This casts doubt on the association of (pivot-neutral) applicatives with a proposed category of Indonesian-type languages showing two-way symmetrical voice systems and use of special coding for non-pivot actors, among other properties.

Third, instead of association with the emergence of a particular typological profile in West Nusantara, there are indications that the presence of pivot-neutral applicatives here is associated specifically with the decline of Philippine-type voice. Surprisingly, pivot-neutral applicatives and pivot-selecting applicatives (Philippine-type LV and CV constructions) are found to overlap in a number of languages of West Nusantara, including some members of the Sama-Bajau subgroup and some languages spoken in the northern third of Sulawesi representing various microgroups. Particularly in Central Sama, Pendau and Bobongko, we see incomplete stages in the transition from Philippine-type voice to pivot-neutral applicatives, showing how the former might have receded in productivity, while the latter expanded. The extension of applicatives into co-occurrence

with both AV and PV constructions drove reorganization from an integrated Philippine-type voice system (AV, PV, LV, CV) to the separation of symmetrical voice (AV and PV) and pivot-neutral applicatives (locative/goal AM, benefactive/instrument AM) as independent paradigmatic alternations in the verbal system. There are indications that this did not happen all at once; for example, in Pendau the goal- and beneficiary-selecting applicatives co-occur with AV marking, but locative- and instrumental-selecting applicatives are only found in PV, despite overlap in the use of suffixal AM-marking.

In Chapter 6, results of the survey are presented showing patterns in the distribution of forms of AMs, and properties of ACs and other AM-marked constructions. These results show that the predominant pattern in West Nusantara is for one form of AM to be associated with selection of locative and goal applied phrases, and another form of AM to be associated with selection of beneficiary, theme, and instrument applied phrases. This pattern of functional distribution across AMs is pervasive and broadly distributed across the region. It also shows marked similarities to the pattern of functional distribution across the LV and CV constructions in PMP and PAN. Furthermore, the association of benefactive and instrumental ACs with the same AM to the exclusion of locative and goal ACs is cross-linguistically unusual (see Peterson 2007).

Based on evidence from synchronic allomorphic alternations between *-i* and *-an* with locative- and goal-selecting ACs in some languages of West Nusantara, I have argued that AMs marking locative/goal-selecting ACs with either the form *-i* or *-an* are derived from LV morphology. For AMs that select beneficiaries, instruments, and themes, we see more variance in the form of the suffix. In South Sulawesi languages, Balinese-Sasak-Sumbawa, and Sama-Bajau we see exclusive use of a morpheme with the form *-an* or some regular reflex of this. Elsewhere, we observe a variety of forms like *-akən*, *-kan*, and *-ako*, but there are also irregularities in expected sound correspondences for these forms that complicate the picture. In addition, there is some evidence that older forms of benefactive/instrumental AMs have been replaced with newer forms, e.g. Javanese *-ʔən* replaced with *-akən*, and Proto Malayic *-an* replaced with *-kan*. Therefore, I have argued that the most likely explanation is that these constructions are derived from earlier CV constructions, and were marked with a reflex of the PMP imperative/negative CV suffix **-an* in earlier stages, with later replacement of the AM form in many languages of West Nusantara.

In Chapter 6, I also present evidence for a number of subareal patterns in the properties of AC and AMs in West Nusantara. With respect to polyfunctionality, AMs in languages of Sulawesi are less likely to show causative functions than AMs in other western Indonesian languages. With respect to properties of ACs, a number of languages of Sulawesi show unexpected patterns of indexing and syntactic properties for the applied phrase. These patterns appear to be driven by animacy effects whereby beneficiaries tend to be treated differently than instruments, themes, and circumstantial roles like reason and cause. More animate participants are more likely to be indexed on the verb, more likely to show distinctive forms of such indexing, and more likely to show access to syntactic operations and to be allowed to map to subject or pivot in P-oriented clause types (passive, PV).

In Chapter 7, I present a comprehensive functional typology of ACs and other AM-marked constructions based on a sample of 24 West Nusantara languages. ACs are categorized by the semantic role selected as the applied phrase and then described according to the observed range of syntactic and semantic properties of such ACs across the languages. Beneficiary and recipient-selecting ACs again show different properties than theme- and instrument-selecting ACs, even though they are marked with the same AMs. The former almost always show maximally ditran-

sitive structures, while the latter typically show monotransitive structures with remapping of the companion phrase to an oblique relation. Locative- and goal-selecting ACs to a lesser extent also show some differences in properties. Some goal-selecting ACs are maximally ditransitive, but both locative- and goal-selecting ACs show remapping structures in a good number of languages.

In Chapter 8, I explored the relationship between lexical semantics and functions observed for AM-marked constructions with particular sets of lexemes as bases, based on a sample of nine languages of West Nusantara. The results show that some components of lexical semantics do influence attraction of lexical bases to certain AM-marked constructions. For example, verbs that express relative position or movement of an entity are attracted to theme-selecting ACs, while verbs of acquisition and creation, among others, are attracted to beneficiary-selecting ACs. The results also indicate that syntactic properties of a BC do not necessarily predict properties of a AC. Instead, there appears to be a preferred mapping of the peripheral role to structural position in the AC and a maximal transitivity for the AC that is typically fixed.

An important implication of the results in Part III is that the productivity of ACs and other functions of AMs shows a large degree of variance. This variance is observed across the lexicon, across functions of AMs, across forms of AMs, and across languages. Commonalities in these patterns may be observed in certain subareal groupings. For this reason, while certain well-known and often studied languages like Indonesian and Javanese do tend to pattern alike with respect to their applicative systems, these patterns cannot be generalized to languages of Sulawesi and other West Nusantara languages in outlying geographic locations. I have suggested that this is because AMs are in a competition of sorts with other grammatical resources in a given language for functional space. Many functions of AM-marked constructions can be filled by other types of verbal alternations in languages that show the relevant morphological resources. Such resources might include many kinds of valency modulating morphology, whenever two values for valency are contrastively marked; thus stative/intransitive markers, general transitivity affixes, voice markers, and (non-applicative) causative morphology may participate. And while I did not investigate this systematically in this study, there is no doubt that certain functions of AMs observed in this study are filled by analytic clausal constructions in some languages of West Nusantara, and more broadly in the Austronesian language family as well. These also represent grammatical resources of the languages in question, and include serial verb constructions, clause chaining, and use of auxiliaries, among others.

In terms of further research, there are several directions implicated by this study. Regarding diachronic development, more historical-comparative research is needed, particularly on the forms of pivot-neutral AMs that mark beneficiary, instrument, and theme applied phrases. This research needs to take into account many more possible cognate forms and detailed analysis of sound correspondences than were possible to consider in this study. Research on languages of Borneo may also be of particular interest. Few Borneo languages in Indonesian Kalimantan and parts of Central and North Sarawak could be included in this study due to lack of basic documentation and grammatical description. Patterns of language contact in Borneo are also particularly complex but may be of considerable value in explanation for why Borneo languages show typological changes associated with lack of applicatives.

In terms of studying applicatives in synchronic use, the influence of lexical semantics is a promising direction for further research. Here it would be of value to investigate the extent to which the subareal patterns shown for Sulawesi versus other western Indonesian languages hold.

Inclusion of more non-Malayo languages of Sumatra and South Sulawesi languages would aid in such assessment. Finally, adopting a constructional approach to applicatives of course implicates further studies of applicatives in usage, especially usage in natural discourse and larger scale corpus studies from which frequency patterns might be investigated. The high degree of lexicalization observed for applicatives in these languages makes such study all the more important. At present, it is difficult to conduct these types of research because nearly all corpus resources for West Nusantara languages are concentrated in a very small number of languages, including Indonesian and Javanese. Even smaller documentary linguistic records, collections of textual material, and lexical resources are lacking for the large majority of languages of West Nusantara. For this reason, descriptive linguistic research based on field study and language documentation must go hand in hand with deeper studies of particular features of language in this region, including applicatives.

Lastly, study of applicatives in natural speech and interactive speech events is needed. This study has touched on some functions of applicatives related to semantic transitivity and information structure, like use of ACs to indicate higher specificity or affectedness, higher topicality, and given status in the discourse context for the referent of the applied phrase. And as mentioned in Chapter 7, in many languages using an applicative clausal structure allows the applied phrase to appear in fronted or clefted positions that signal focused information. These types of functions can only be reliably identified and adequately described through study of natural connected speech. However, at present, the only research of this type for West Nusantara is Donohue's (2001) study of *Tukang Besi* applicatives in narrative texts. Interactional and pragmatic uses of applicatives have also received relatively little attention in the literature. But several authors mention marking of the predicate with a benefactive AM as a means of expressing a request or polite imperative, and a number of speakers with whom I have worked find that benefactive imperatives are among the most natural examples of ACs in daily usage, especially with the beneficiary unrealized and left open to interpretation via inference. This could explain how the PAn/PMP imperative CV suffix *-an came to be the general marker of pivot-neutral benefactive/instrumental ACs. Moreover in languages that lost morphologically-marked mood distinctions, frequent use of benefactive ACs in requests could subsequently lead to grammaticalization of the benefactive AM as a general imperative marker. Therefore, I consider new research based on natural connected speech and interactional speech events to be of great importance for further investigation of the synchronic functions of applicatives with implications for their diachronic development.

Appendices and Back Matter

Appendix A: Language sample for typological survey

Table A.1 lists languages included the sample for the typological survey described in Chapter 4. The following genetic groups are noted by abbreviation in the table: BSS = Bali-Sasak-Sumbawa, B-T = Bungku-Tolaki, CHA = Chamic, ENG = Enggano, GRB = Greater Barito, JAV = Javanese, K-P = Kaili-Pamona, LAM = Lampungic, LND = Land Dayak, MAL = Malayic, M-B = Muna-Buton, M-K = Melanau-Kajang, MAD = Madurese, NAS = Nasal, NSAR = North Sarawak, NWS-BI = Northwest Sumatra-Barrier Islands, REJ = Rejang, S-B = Saluan-Banggai, SAB = Sabahan, SSUL = South Sulawesi, SUN = Sundanese, T-T = Tomini-Tolitoli, W-W = Wotu-Wolio. Languages with a disputed genetic affiliation are marked with an asterisk (*). The designation of genetic groups and their primary branches in the table is given according to Eberhard, Simons & Fennig (2021). Regions are listed according to major island within West Nusantara. Mainland Southeast Asia (MSEA) here refers to the continental portion of Southeast Asia north of Penninsular Malaysia. Madura, which is separated from the island of Java by the narrow Madura Strait, is included in the region designated as Java.

Table A.1: Languages included in sample for typological survey

Language	Code	Group	Prim. branch	Region	Sources
Batak Alas-Kluet	btz, bata1292	NWS-BI	Batakic	Sumatra	Soravia 2007
Batak Karo	btx, bata1293	NWS-BI	Batakic	Sumatra	Woollams 1996
Batak Toba	bbc, bata1289	NWS-BI	Batakic	Sumatra	Nababan 1981; Percival 1981; Schachter 1984; van der Tuuk 1971 [1864-1867]
Gayo	gay, gayo1244	NWS-BI	Gayo	Sumatra	Eades 2005
Nias	nia, nias1242	NWS-BI	Nuc. Barrier Isl.	Barrier Isl.	Brown 2001, 2005
Enggano	eno, engg1245	ENG	NA	Barrier Isl.	Crowley n.d.; Edwards 2015
Nasal	nsy, nasa1239	NAS	NA	Sumatra	McDonnell fieldnotes
Dampelas	dms, damp1237	T-T	Tomini	Sulawesi	Moro 2010
Pendau	ums, pend1242	T-T	Tomini	Sulawesi	Quick 2007
Tajio	tdj, taji1246	T-T	Tomini	Sulawesi	Mayani 2013
Totoli	twe, toto1304	T-T	Tolitoli	Sulawesi	Himmelman & Riesberg 2013; Riesberg 2014b
Kaili, Da'a	kzf, daak1235	K-P	Northern	Sulawesi	Barr 1988a, 1988b
Kaili, Ledo	lew, ledo1238	K-P	Northern	Sulawesi	D. Evans 1996, 2003
Moma	myl, moma1242	K-P	Northern	Sulawesi	Adriani & Esser 1939
Uma	ppk, umaa1242	K-P	Southern	Sulawesi	Martens 1988a, 1988b
Behoa	bep, beso1237	K-P*	Southern	Sulawesi	Shore 2016
Balantak	blz, bala1315	SLB	Eastern	Sulawesi	van den Berg & Busenitz 2012
Bobongko	bgb, bobo1255	SLB	Western	Sulawesi	Mead 2001
Mori Bawah	xmz, mori1268	B-T	Eastern	Sulawesi	Esser 2011; Mead 1998, 2005
Moronene	mqn, moro1287	B-T	Eastern	Sulawesi	Andersen 2013; Andersen & Anderson 2005; Mead 1998
Tolaki	lbw, tola1247	B-T	Western	Sulawesi	Edwards 2012
Busoa	bup, buso1238	M-B	Nuc. Muna-Buton	Sulawesi	van den Berg 2020
Muna	mnb, muna1247	M-B	Nuc. Muna-Buton	Sulawesi	van den Berg 2013
Tukang Besi North	khc, tuka1248	M-B	Tukangbesi-Bonerate	Sulawesi	Donohue 1999, 2001
Wotu	wtw, wotu1240	W-W	Wotu	Sulawesi	Mead 2013
Laiyolo	lji, laiy1246	W-W	Kalao	Sulawesi	Belding, Laidig & Maingak 2001
Wolio	wlo, woli1241	W-W	Wolio-Kamaru	Sulawesi	Anceaux 1952
Embaloh	emb, emba1238	SSUL	Bugis	Borneo	Adelaar 1995
Bugis	bug, bugi1244	SSUL	Bugis	Sulawesi	Hanson 2003; D. Laskowske 2016; Sirk 1983
Konjo, Coastal	kjc, coas1295	SSUL	Makassar	Sulawesi	B. Friberg 1991, T. Friberg 1995
Makassar	mak, maka1311	SSUL	Makassar	Sulawesi	Jukes 2020
Duri	mvp, duri1242	SSUL	Northern	Sulawesi	Valkama 1993
Bambam	ptu, bamb1270	SSUL	Northern	Sulawesi	Campbell 1989
Seko Padang	skx, seko1243	SSUL	Seko	Sulawesi	T. Laskowske 2006; Laskowske & Arif 2000; Payne & Laskowske 1997

Table A.1: Languages included in sample for typological survey (cont.)

Language	Code	Group	Prim. branch	Region	Sources
Cham, Eastern	cjm, east2563	CHA	Coastal	MSEA	Thurgood 2005
Bih	ibh, biha1246	CHA	Highlands	MSEA	Nguyen 2013
Tsat	huq, tsat1238	CHA	Highlands	MSEA	Thurgood, Thurgood & Fengxiang 2014
Acehnese	ace, achi1257	CHA	Acehnese	Sumatra	Durie 1985
Malay, Patani	mfa, patt1249	MAL	Malay	MSEA	Tadmor 1995
Kerinci	kvr, keri1250	MAL	Malay	Sumatra	Ernanda 2017
Besemah	pse, cent2053	MAL	Malay	Sumatra	McDonnell 2016
Malay, Jambi	jax, jamb1236	MAL	Malay	Sumatra	Yanti 2010
Brunei	kxd, brun1242	MAL	Malay	Borneo	Clynes 2001
Bazaar Malay	zlm, mala1478	MAL	Malay	Singapore	Aye 2005
Kendayan	knx, kend1254	MAL	Kendayan	Borneo	Adelaar 2002, 2005b
Mualang	mtd, mual1241	MAL	Ibanic	Borneo	Tjia 2007
Indonesian	ind, indo1316	MAL	Malay	Other	Sneddon et al. 2010
Urak Lawoi'	urk, urak1238	MAL	Urak Lawoi'	MSEA	Hogan 1988, 1999
Bali	ban, bali1278	BSS	Bali	Lesser Sundas	Arka 2003
Ampenan Sasak	sas, sasa1249	BSS	Sasak-Sumbawa	Lesser Sundas	Khairunnisa 2022
Sumbawa	smw, sumb1241	BSS	Sasak-Sumbawa	Lesser Sundas	Austin 2001; Shiohara & Arka forthcoming
Madura	mad, nucl1460	MAD	NA	Java	Davies 2010
Sunda	sun, sund1252	SUN	NA	Java	Hardjadibrata 1985; Kurniawan 2013; personal field notes
Javanese	jav, java1254	JAV	NA	Java	Hemmings 2013; Oglobin 2005
Tengger	tes, teng1272	JAV	NA	Java	Connors 2008
Javanese, Suriname	jvn, cari1276	JAV	NA	Americas	Villierius 2019
Rejang	rej, reja1240	REJ	NA	Sumatra	McGinn 1982
Lampung Api	ljp, lamp1243	LAM	NA	Sumatra	Walker 1976
Malagasy, Merina	plt, plat1254	GBR	East	Africa	Pearson 2001; Rasoloson & Rubino 2005
Ma'anyan	mhy, maan1238	GBR	East	Borneo	Gudai 1985
Paku	pku, paku1239	GBR	East	Borneo	Diedrich 2018
Yakan	yka, yaka1277	GBR	Sama-Bajaw	Philippines	Brainard & Behrens 2002
Bajau, West Coast	bdr, west2560	GBR	Sama-Bajaw	Borneo	Miller 2007
Sama, Central	sml, cent2092	GBR	Sama-Bajaw	Philippines	James 2017
Sama, Southern	ssb, sout2918	GBR	Sama-Bajaw	Philippines	Akamine 2003
Ngaju	nij, ngaj1237	GBR	West	Borneo	Harde land 1858
Benyadu'	byd, beny1237	LND	Benyadu'	Borneo	Sommerlot 2020
Bakati'	bei, beka1241	LND	Bakati'	Borneo	Sudarsono 2002
Matéq	xem, kemb1249	LND	Southern	Borneo	Connell 2013
Ribun	rir, ribu1241	LND	Southern	Borneo	Sommerlot 2020

Table A.1: Languages included in sample for typological survey (cont.)

Language	Code	Group	Prim. branch	Region	Sources
Melanau, Central	mel, cent2101	M-K	Melanau	Borneo	R. A. Blust 1988; Chou 1999
Berawan, West	zbw, west2564	NSAR	Berawan-Lower Baram	Borneo	B. Clayre 1997
Belait	beg, bela1260	NSAR	Berawan-Lower Baram	Borneo	Clynes 2005
Kelabit	kzi, kela1258	NSAR	Dayic	Borneo	Hemmings 2016
Lundayeh	lnd, lund1271	NSAR	Dayic	Borneo	Mortensen 2021
Kayan, Baram	kys, bara1370	NSAR*	Kayan-Kenyah	Borneo	Omar 1983
Penan, Eastern	pez, east2485	NSAR	Kayan-Kenyah	Borneo	Sercombe 2006
Punan Tubu	puj, puna1266	NSAR	Punan Tubu	Borneo	Soriente 2013
Kimaragang	kqr, kima1244	SAB	Dusunic	Borneo	Kroeger 2005
Ida'an	dbj, idaa1241	SAB	Dusunic	Borneo	Goudswaard 2005
Murut, Keningau	kxi, keni1249	SAB	Murutic	Borneo	Cohen 1999
Murut, Timugon	tih, timu1262	SAB	Murutic	Borneo	Prentice 1969, 1995
Murut, Serudung	srk, seru1246	SAB	Murutic	Borneo	Townsend 2017
Tombonuo	txa, tomb1244	SAB	Paitanic	Borneo	Levinsohn & King 1991; King 1984, 1991; King 1991
Tatana	txx, tata1257	SAB	Tatana	Borneo	Dillon 1994

Appendix B: Questionnaires used in the typological survey

B.1 Questionnaire A

Part I: General properties of the language

Name of language:

ISO-639-3 language code:

Glottolog language code:

Genetic group:

[Bali-Sasak-Sumbawa; Bungku-Tolaki; Chamic; Enggano; Greater Barito; Javanese; Kaili-Pamona; Lampungic; Land Dayak; Madurese; Malayic; Melanau-Kajang; Muna-Buton; Nasal; North Sarawak; NW Sumatra-Barrier Islands; Rejang; Sabahan; Saluan-Banggai; South Sulawesi; Sundanese; Tomini-Tolitoli; Wotu-Wolio]

Location:

[Africa; Americas; Barrier Islands; Borneo; Java & Madura; Lesser Sundas; Malacca Straits; Other MSEA; Peninsular Malaysia; Philippines; Sulawesi; Sumatra]

Major division of genetic group:

Part II: Structural properties of the language

Pattern 1: Word order

In basic transitive clauses (A-oriented) is the predominant word order Verb-P?

- **Y:** source indicates predominant word order is Verb-P, NP expressing A generally does not intervene
- **N:** source indicates predominant word order other than Verb-P
- **free:** there is no predominant word order

Pattern 2: Morphological alignment

With respect to morphological marking of core arguments in basic intransitive and transitive clauses, how do S, A, and P pattern?

- **accusative:** S and A are marked alike, to the exclusion of P
- **ergative:** S and P are marked alike, to the exclusion of A
- **neutral:** S, A, and P are marked alike (and distinct from non-core arguments)
- **split-S:** S is marked like A when agentive and P when patientive (see Mithun & Chafe 1999: 578)
- **pivot-non-pivot:** S shares distinct marking with A in A-oriented transitive constructions, and P in P-oriented transitive constructions
- **mixed-NPIV.A:** A of P-oriented transitive constructions shows special marking, otherwise S, A, and P are marked alike.
- **mixed (other):** marking of S, A, and P otherwise varies according to construction

Pattern 3: Syntactic alignment

With respect to syntactic properties of core arguments in basic intransitive and transitive clauses, what is the primary patterning of S, A, and P?

- **accusative:** S and A share special syntactic properties to the exclusion of P
- **ergative:** S and P share special syntactic properties to the exclusion of A
- **neutral:** S, A, and P alike share special syntactic properties distinct from those of non-core arguments
- **split-S:** special syntactic processes generally apply to equally to A and S when agentive, and equally to P and S when patientive
- **pivot-non-pivot:** S shares special syntactic properties with A in A-oriented transitive constructions, and P in P-oriented transitive constructions
- **mixed:** no one primary patterning can be said to hold for syntactic properties of S, A, and P across basic intransitive and transitive clauses

Pattern 4: Symmetrical voice

What is the pattern of voice and/or diathetical alternations found in basic transitive clauses in the language?

- **Philippine-type:** Three or more basic transitive constructions are observed, distinguished by mapping of semantic role to grammatical relation. These include at least two distinct nonactor oriented constructions.
- **marginal-Philippine-type:** Three or more transitive constructions are observed, distinguished by mapping of semantic role to grammatical relation. However, one or more of these is non-productive or substantially restricted in distribution.

- **two-way-symmetrical:** Two basic transitive constructions are observed, one A-oriented, and one P-oriented.
- **marginal-two-way-symmetrical:** Two transitive constructions are observed, one A-oriented, and one P-oriented. However, one of these is non-productive or substantially restricted in distribution.
- **asymmetrical:** No alternations in basic transitive constructions are observed which are distinguished by mapping of semantic role to grammatical relation.

Pattern 5: Morphological case-marking

Is morphological case-marking used in the language?

- **Y:** Morphological elements (or morphological processes) that attach to nouns or NP constituents are generally used to indicate grammatical or spatial relations.
- **Limited:** Case-marking morphological elements are found to attach only to a small subclass of nominals, or case is marked morphologically by means of sets of pronominal elements, but is not marked on nouns generally.
- **N:** No morphological elements (or processes) are found that attach to nouns or NP constituents and indicate grammatical or spatial relations.

Pattern 6: Order of noun and possessor

What is the order of possessed noun and possessor in the language?

- **N-Poss:** The possessed noun generally precedes the possessor.
- **Poss-N:** The possessor generally precedes the possessed noun.
- **mixed:** Both orders are observed.

Pattern 7a: Presence of morphological causative construction

Does the language have a morphological causative construction?

- **Y:** The language has a morphological causative construction.
- **N:** The language does not have a morphological causative construction.

Pattern 7b: Source of morphological causative marker

Which of the following are sources of morphological causative markers in the language?

- **appl:** Causative morphology shares the same form with an applicative morpheme in the language.
- ***pa-:** Causative morphology in the language apparently derives from PMP *pa- 'causative'.
- **other:** Causative morphology in the language apparently derives from some other source.

- **NA:** The pattern is not applicable because the language has no morphological causative construction.

Pattern 7c: Productivity of reflex of *pa-

*What is the productivity of morphological causatives derived from *pa-?*

- **high:** productive with a large number of lexical roots across multiple syntactic categories (transitive verbs, intransitive dynamic verbs, stative verbs, nouns, etc.)
- **medium:** productive with a large number of lexical roots that primarily belong to one syntactic category
- **low:** only attested with a limited number of lexical roots.
- **NA:** There is no morphological causative derived from *pa-.

Part III: Properties of the applicative system

Pattern 8a: Presence of applicatives

Does the language have applicative constructions? An applicative is defined as a clausal construction in which overt morphological marking on the verbal complex coincides with the selection of a non-agent, non-patient semantic role to map to a core argument in the clause.

- **Y:** Applicative constructions as defined above are identifiable in the language based on the source material.
- **N:** Applicative constructions as defined above are not identifiable in the language based on the source material.

Pattern 8b: Co-occurrence with other constructions

Does applicative marking in the language freely co-occur with the basic transitive constructions of the language and the passive (if applicable)?

- **Y:** At least one of the applicative morphemes in the language co-occurs with the major diathetical alternations of the language, including all the basic transitive constructions and any passive construction.
- **N-Partial:** None of the applicative morphemes in the language co-occurs with all the basic transitive constructions of the language and any passive constructions. However, at least one applicative morpheme co-occurs with more than one such construction, under certain conditions.
- **N:** The applicative morphemes in the language do not co-occur with other basic transitive constructions of the language or the passive construction.
- **no-appl:** There are no applicatives, so this pattern is not applicable.

Pattern 9: Indexing of the applied phrase

In applicative constructions, does the applied phrase show person-indexing on the verb in the manner generally characteristic of P in monotransitive clauses?

- **Y:** The applied phrase generally shows person indexing on the verb in the same manner as other P arguments in the language.
- **N:** The applied phrase generally does not show person indexing on the verb, even though other P arguments in the language do show such indexing.
- **mixed:** Under certain conditions but not others, the applied phrase shows indexing of person and number features on the verb in the same manner as other P arguments in the language are indexed.
- **no-index:** The pattern is not applicable because the language does not generally make use of person indexing for P arguments.
- **no-appl:** The pattern is not applicable because the language does not have applicatives.

Pattern 10: Syntactic privilege and the applied phrase

Does the applied phrase show evidence of status as a syntactically privileged argument across basic clause types which co-occur with applicative constructions?

- **Y-obl:** Yes, there is evidence that the applied phrase always holds a syntactically privileged relation across basic clause types with which the applicative construction co-occurs.
- **Y-opt:** There is evidence that the applied phrase generally holds a syntactically privileged relation in P-oriented basic clause types.
- **mixed:** There is evidence that the applied phrase may hold a syntactically privileged relation in basic clause types; this co-varies according to the form of the applicative morpheme and/or semantic role of the applied phrase.
- **N:** The applied phrase generally does not appear to hold a syntactically privileged relation in the clause, or does so only in marked, infrequent constructions.
- **no-appl:** The pattern is not applicable because the language does not have applicatives.

Pattern 11: Access to relativization

Does the applied phrase have access to relativization?

- **Y:** Yes, the applied phrase may generally be the head noun of a relative clause, subject to the same conditions as other P arguments.
 - **N:** No, the applied phrase generally may not be the head noun of a relative clause.
 - **mixed:** The applied phrase sometimes may be the head noun of a relative clause, but this is subject to additional conditions not observed for other P arguments.
 - **no-appl:** This pattern is not applicable because the language does not have applicatives.
-

B.2 Questionnaire B

Questionnaire B: Properties of applicative morphemes

Pattern A: Source morphology

What is the apparent source morphology from which the applicative morpheme is derived?

- ***akən**: The applicative morpheme in question appears to derive from *akən ‘on, upon; instrumental preposition’
- ***i**: The applicative morpheme in question appears to derive from PMP *i ‘locative preposition’
- ***-an/*-ən**: The applicative morpheme in question appears to derive from PMP voice morphology *-an or *-ən.
- ***pa-**: The applicative morpheme in questions appears to derive from PMP *pa- or a combination of *pa- and another prefix, e.g. *paka-, popa-, etc.
- **other**: The applicative morpheme in question appears to derive from another known historical source.
- **undetermined**: The historical source of the applicative marked in question is not clear.

Pattern B: Source morphology type

What is the category of the historical source morphology from which the applicative morpheme is derived?

- **adpos**: The source morphology was an adposition.
- **verb**: The source morphology was a verb.
- **caus**: The source morphology was a causative marker.
- **case-mrkr**: The source morphology was a case marker.
- **voi-mrkr**: The source morphology was a voice marker.
- **other**: The source morphology is known to have a different category or function than that listed above.
- **undetermined**: The category or function of the source morphology is not known.

Pattern C: Semantic roles of the applied phrases

What are the semantic roles of the applied phrase in constructions bearing the applicative morpheme?

- **BEN**: Beneficiary, i.e., a participant who accrues a benefit through an event or state of affairs.
- **REC**: Recipient, i.e., an entity that receives possession (physical or otherwise) of an entity

- **LOC:** Static locative, i.e. the static or generalized location of a state or event
- **GOAL:** Goal, i.e., the end point of an entity that changes location in a motion event.
- **THM:** Theme, i.e., an entity undergoes a change of location or is located in space
- **INST:** Instrument, i.e., an inanimate entity manipulated to some effect in a event
- **STIM:** Stimulus, i.e., the object of an act of perception, stimulus of an emotional response
- **CONT:** Content, i.e., the content of an act of speaking or cognition.
- **CIRC:** Circumstantial, i.e., the reason or purpose for a state or event.
- **COM:** Comitative, i.e., a participant that accompanies an actor or mover.
- **ADDR:** Addressee, i.e., a participant that is the intended receiver of some communication

Pattern D: Co-occurrent marking with other morphology

Must the applicative morpheme co-occur together with another morpheme on the verbal complex in certain applicative constructions?

- **Y-SF:** Yes, the applicative morpheme appears together with a “stem-former” in certain applicative constructions. The stem-former does not have a semantic meaning of its own, but it may have a syntactic or morphological function.
- **Y-CAUS:** Yes, the applicative morpheme appears together with a causative morpheme in certain applicative constructions. The causative morpheme has an independent causative function in other constructions.
- **Y-other:** Yes, the applicative morpheme appears together with some other morphological marking on the verbal complex that is not generally found in basic transitive constructions.
- **N:** No, the applicative morpheme generally does not co-occur with other morphological marking on the verb necessary to form the applicative constructions.

Pattern E: Other functions

Which of the following non-applicative functions are also marked by the applicative morpheme?

- **CAUS:** The applicative morpheme also forms causative constructions, in which an instigating causer participant is introduced and selected to map to A.
- **CAUS-combo:** The applicative morpheme co-occurs with another morphological causative on the verbal complex in certain (non-applicative) causative constructions.
- **PLUR:** The applicative morpheme indicates pluractional aspect, e.g. durative, habitual, iterative, multiple actors or undergoers.
- **INTENS:** The applicative morpheme indicates greater intensity, e.g. greater volitionality, greater application of force, etc.

- **COMP:** The applicative morpheme indicates comparative degree of a gradable quality.
 - **sem-change:** The applicative morpheme is associated with some other semantic change in the meaning of a verb, e.g. 'discard' cf. 'throw', 'elope' cf. 'run'.
 - **none:** No non-applicative functions are indicated from the source material.
-

Appendix C: Database for typological survey

The data compiled for the typological survey described in Part II is given below. Table C.1 contains the data collected by language using Questionnaire A. Table C.2 contains the data collected by form of applicative morpheme (AM) using Questionnaire B.

Abbreviations for genetic affiliation: BSS = Bali-Sasak-Sumbawa, B-T = Bungku-Tolaki, CHA = Chamic, ENG = Enggano, GRB = Greater Barito, JAV = Javanese, K-P = Kaili-Pamona, LAM = Lampungic, LND = Land Dayak, MAL = Malayic, M-B = Muna-Buton, M-K = Melanau-Kajang, MAD = Madurese, NAS = Nasal, NSAR = North Sarawak, NWS-BI = Northwest Sumatra-Barrier Islands, REJ = Rejang, S-B = Saluan-Banggai, SAB = Sabahan, SSUL = South Sulawesi, SUN = Sundanese, T-T = Tomini-Tolitoli, W-W = Wotu-Wolio. Languages with a disputed genetic affiliation are marked with an asterisk (*).

Undetermined values are indicated as [und] or [undet]. For other coding values, please refer to the relevant questionnaire in Appendix B.

Table C.1: Survey data for Questionnaire A

Language	Glotto	ISO	Gen. grp.	Location	Maj. Div.	1 WrOrd	notes	2 MrpAlg	3 SynAlg	4 Voice	5 Case	6 N-Poss	7a MrpCau	7b SrcCau	7c CauPrd	8a Appl	8b Co-oc	9 Indx	10 SynPrv	11 AccRcl
Alas Batak	bata1292	btz	NWS-BI	Sumatra	Sumatra	y	AVP	mixed-npiv.A	und	2-wy sym	lim	N-Poss	y	*pa-	med	y	y	und	y-opt	y
Koro Batak	bata1293	btx	NWS-BI	Sumatra	Sumatra	y	AVP	piv-npiv	piv-npiv	2-wy sym	lim	N-Poss	y	*pa-	high	y	y	y	y-opt	y
Toba Batak	bata1289	bbc	NWS-BI	Sumatra	Sumatra	y	VPA	mixed-npiv.A	piv-npiv	2-wy sym	lim	N-Poss	y	*pa-	high	y	y	y	y-opt	y
Gayo	gayo1244	gay	NWS-BI	Sumatra	Sumatra	y	AVP	piv-npiv	piv-npiv	2-wy sym	lim	N-Poss	y	*pa-	med	y	y	y	y-opt	y
Nias	nias1242	nia	NWS-BI	Barr. Isl.	Barr. Isl.	y	VPA	mixed (other)	accus	mrg 2-wy sym	y	N-Poss	y	*pa-, appl	low	y	y	no-indx	y-opt	y
Enggano	engg1245	eno	ENG	Barr. Isl.	Barr. Isl.	y	AVP	neutral	accus	asymmet	y	N-Poss	y	*pa-, appl	med	y	y	und	y-opt	und
Nasal	nasa1239	nsy	NAS	Sumatra	na	y	AVP	piv-npiv	piv-npiv	2-wy sym	lim	N-Poss	y	*pa-, appl	und	y	y	y	y-opt	y
Dampelas	damp1237	dms	T-T	Sulawesi	Tomini	y	AVP	mixed-npiv.A	piv-npiv	2-wy sym	n	N-Poss	y	*pa-	und	y	y	no-indx	y-opt	y
Pendau	pend1242	ums	T-T	Sulawesi	Tomini	y	AVP	mixed-npiv.A	piv-npiv	2-wy sym	lim	N-Poss	y	*pa-	high	y	n-part	no-indx	mixed	y
Tajjo	pend1242	tdj	T-T	Sulawesi	Tomini	y	AVP/VPA	mixed-npiv.A	piv-npiv	2-wy sym	lim	N-Poss	y	*pa-	high	y	y	no-indx	y-opt	y
Totali*	toto1304	twe	T-T	Sulawesi	Tolitoli	y	AVP/VPA	mixed-npiv.A	piv-npiv	Phil-tyt	n	N-Poss	y	*pa-	med	y	n-part	no-indx	mixed	y
Da'a Kaili	daak1235	kzf	K-P	Sulawesi	na	y	AVP	mixed-npiv.A	und	2-wy sym	lim	N-Poss	y	*pa-	high	n	na	no-appl	no-appl	no-appl
Ledo Kaili	ledo1238	lew	K-P	Sulawesi	na	y	AVP/VPA	mixed-npiv.A	piv-npiv	2-wy sym	lim	N-Poss	y	*pa-	high	y	y	no-indx	y-opt	y
Moma	moma1242	myl	K-P	Sulawesi	na	y	AVP	mixed-npiv.A	und	2-wy sym	lim	N-Poss	y	*pa-	high	y	y	no-indx	y-opt	und
Uma	umaa1242	ppk	K-P	Sulawesi	na	y	AVP	ergative	piv-npiv	2-wy sym	lim	N-Poss	y	*pa-	high	y	y	no-indx	y-opt	y
Behoa*	beso1237	bep	K-P	Sulawesi	na	y	AVP	piv-npiv	piv-npiv	2-wy sym	lim	N-Poss	y	*pa-	high	y	y	y	y-opt	y
Balantak	bala1315	blz	S-B	Sulawesi	na	y	AVP	mixed-npiv.A	piv-npiv	Phil-tyt	lim	N-Poss	y	*pa-	low	y	y	no-indx	mixed	mixed
Bobongko	bobo1255	bgb	S-B	Sulawesi	na	und	und	mixed-npiv.A	piv-npiv	mrg Phil-tyt	lim	N-Poss	y	*pa-	und	y	y	no-indx	mixed	mixed
Mori Bawah	mori1268	xmz	B-T	Sulawesi	na	n	VAP/VPA	ergative	accus	asymmet	lim	N-Poss	y	*pa-, appl	high	y	n-part	mixed	mixed	mixed
Moronene	moro1287	mqn	B-T	Sulawesi	na	n	VAP/VPA	mixed (other)	accus	asymmet	n	N-Poss	y	*pa-	und	y	y	und	und	und
Tolaki	tola1247	lbw	B-T	Sulawesi	na	und	und	mixed-other	accus	asymmet	lim	N-Poss	y	*pa-	med	y	y	mixed	mixed	mixed
Buloso	buso1238	bup	M-B	Sulawesi	na	y	AVP	accusative	accus	asymmet	lim	N-Poss	y	*pa-	med	y	y	mixed	mixed	y
Muna	muna1247	mnb	M-B	Sulawesi	na	y	AVP	accusative	accus	asymmet	lim	N-Poss	y	*pa-, appl	high	y	y	mixed	mixed	y
Tukang Besi	tuka1248	khc	M-B	Sulawesi	na	y	VPA	mixed (other)	piv-npiv	2-wy sym	y	N-Poss	y	*pa-, appl	high	y	y	mixed	mixed	mixed
Wotu	wotu1240	wtw	W-W	Sulawesi	na	y	AVP/VPA	mixed (other)	piv-npiv	2-wy sym	lim	N-Poss	y	*pa- + appl	und	y	y	y	y-opt	y
Laiyolo	laiy1246	lji	W-W	Sulawesi	na	y	AVP	accusative	piv-npiv	2-wy sym	lim	N-Poss	y	*pa-, appl	med	y	n-part	y	y-opt	y
Wolio	woli1241	wlo	W-W	Sulawesi	na	n	VAP	accusative	und	mrg 2-wy sym	lim	N-Poss	y	*pa-, appl	med	y	y	und	y-opt	und
Embaloh	emba1238	emb	SSUL	Borneo	Borneo	n	VAP	ergative	ergative	mrg 2-wy sym	lim	N-Poss	y	*pa-	und	y	y	y	y-opt	und
Bugis	bugi1244	bug	SSUL	Sulawesi	Sulawesi	y	AVP/VPA	ergative	piv-npiv	mrg 2-wy sym	lim	N-Poss	y	*pa- + appl	high	y	y	y	y-opt	y
Coastal Konjo	coas1295	kjc	SSUL	Sulawesi	Sulawesi	n	VAP	ergative	accus	mrg 2-wy sym	lim	N-Poss	y	*pa-	und	y	y	y	y-opt	und
Makassar	maka1311	mak	SSUL	Sulawesi	Sulawesi	y	VPA	ergative	neutral	mrg 2-wy sym	lim	N-Poss	y	*pa-, appl	high	y	y	mixed	mixed	y
Duri	duri1242	mvp	SSUL	Sulawesi	Sulawesi	y	AVP	ergative	piv-npiv	2-wy sym	lim	N-Poss	y	*pa-	med	y	y	y	y-opt	und
Bambam	bamb1270	ptu	SSUL	Sulawesi	Sulawesi	y	AVP	ergative	neutral	mrg 2-wy sym	lim	N-Poss	y	*pa-, appl	high	y	y	y	y-opt	y
Seko Padang	seko1243	skx	SSUL	Sulawesi	Sulawesi	und	und	ergative	ergative	mrg 2-wy sym	lim	N-Poss	y	*pa-, appl	und	y	y	und	und	und
Eastern Cham	east2563	cjm	CHA	MSEA	MSEA	y	AVP	neutral	accus	asymmet	n	N-Poss	y	*pa-, appl	low	n	na	no-appl	no-appl	no-appl
Bih	bihai1246	ibh	CHA	MSEA	MSEA	y	AVP	neutral	accus	asymmet	n	N-Poss	y	*pa-, appl	low	n	na	no-appl	no-appl	no-appl
Acehnese	achi1257	ace	CHA	Sumatra	Sumatra	free	free	split-S	split-S	asymmet	lim	N-Poss	y	*pa-	high	y	y	y	NA	und
Tsat	tsat1238	huq	CHA	MSEA	MSEA	y	AVP	neutral	accus	asymmet	lim	mixed	n	na	na	n	na	no-appl	no-appl	no-appl
Pattani Malay	patt1249	mfa	MAL	MSEA	MSEA	y	AVP	mixed-npiv.A	und	2-wy sym	lim	N-Poss	y	other	low	n	na	no-appl	no-appl	no-appl
Kerinci	keri1250	kvr	MAL	Sumatra	Sumatra	y	AVP	mixed (other)	und	2-wy sym	y	N-Poss	y	*pa-, appl	med	n	na	no-appl	no-appl	no-appl
S. Barisan Mal.	cent2053	pse	MAL	Sumatra	Sumatra	y	AVP/VPA	neutral	piv-npiv	2-wy sym	n	N-Poss	y	appl	na	y	y	y	y-opt	y
Jambi Malay	jamb1236	jax	MAL	Sumatra	Sumatra	y	AVP	neutral	mixed	2-wy sym	n	N-Poss	y	appl	na	y	y	no-indx	y-opt	y
Brunei	brun1242	kxd	MAL	Borneo	Borneo	n	VAP/AVP	mixed (other)	piv-npiv	2-wy sym	lim	N-Poss	y	appl	na	y	y	y	y-opt	y
Sing. Baz. Mal.	mala1479	zlm	MAL	Other	other	y	AVP	neutral	accus	asymmet	n	mixed	n	other	na	n	na	no-appl	no-appl	no-appl
Kendayan	kend1254	knx	MAL	Borneo	Borneo	y	AVP	mixed-npiv.A	und	2-wy sym	n	N-Poss	y	appl	na	y	y	no-indx	y-opt	und
Mualang	muai1241	mtd	MAL	Borneo	Borneo	y	AVP	neutral	piv-npiv	2-wy sym	n	N-Poss	y	*pa-	low	n	na	no-appl	no-appl	no-appl
Indonesian	indo1316	ind	MAL	Other	other	y	AVP	mixed-npiv.A	piv-npiv	2-wy sym	lim	N-Poss	y	appl	na	y	y	y	y-opt	y
Urak Lawoi'	urak1238	urk	MAL	MSEA	other	y	AVP	neutral	und	asymmet	n	N-Poss	y	other	na	n	na	no-appl	no-appl	no-appl
Bali	bali1278	ban	BSS	Lsr Sundas	na	y	AVP	mixed-npiv.A	piv-npiv	2-wy sym	n	N-Poss	y	appl	low	y	y	no-indx	y-opt	y
Sasak	sasa1249	sas	BSS	Lsr Sundas	na	y	AVP	mixed (other)	mixed	2-wy sym	n	N-Poss	y	appl	na	y	y	no-indx	y-opt	und
Sumbawa	sumb1241	smw	BSS	Lsr Sundas	na	und	und	accusative	und	asymmet	lim	N-Poss	y	other	low	n	na	no-appl	no-appl	no-appl
Madura	nuc1460	mad	MAD	Java	na	y	AVP	mixed-npiv.A	piv-npiv	2-wy sym	lim	N-Poss	y	*pa-, appl	med	y	y	no-indx	y-opt	y
Sunda	sund1252	sun	SUN	Java	na	y	AVP	mixed-npiv.A	piv-npiv	2-wy sym	n	N-Poss	y	appl	na	y	y	y	y-opt	y
Javanese	jawa1254	jav	JAV	Java	Java	y	AVP	mixed-npiv.A	piv-npiv	2-wy sym	lim	N-Poss	y	*pa- + appl	na	y	y	no-indx	y-opt	y
Tengger	teng1272	tes	JAV	Java	Java	y	AVP/VPA	neutral	und	2-wy sym	lim	N-Poss	y	appl	na	y	y	y	y-opt	und
Suriname Jav.	cari1276	jvn	JAV	Americas	other	y	AVP	mixed-npiv.A	piv-npiv	2-wy sym	lim	N-Poss	y	appl	na	y	y	y	y-opt	y
Rejang	reja1240	rej	REJ	Sumatra	na	y	AVP	mixed-npiv.A	piv-npiv	2-wy sym	lim	N-Poss	y	other	med	n	na	no-appl	no-appl	no-appl
Lampung Api	lamp1243	ljp	LAM	Sumatra	na	y	AVP	mixed-npiv.A	piv-npiv	2-wy sym	lim	N-Poss	y	appl	na	y	y	no-indx	y-opt	y
Merina Malag.	plat1254	plt	GRB	Africa	na	y	VPA	piv-npiv	piv-npiv	Phil-tyt	lim	N-Poss	y	*pa-	high	y	n	y	y-obl	y
Ma'anyan	maan1238	mhy	GRB	Borneo	na	y	AVP	mixed-npiv.A	piv-npiv	2-wy sym	lim	N-Poss	y	other	na	n	na	no-appl	no-appl	no-appl
Paku	paku1239	pku	GRB	Borneo	na	y	AVP	mixed-npiv.A	piv-npiv	2-wy sym	lim	N-Poss	y	other	high	n	na	no-appl	no-appl	no-appl
Yakan	yaka1277	yka	GRB	Philippines	other	n	VAP/VPA	piv-npiv	piv-npiv	mrg Phil-tyt	y	N-Poss	y	*pa-	med	y	n-part	no-indx	y-opt	y
W. C. Bajau	west2560	bdr	GRB	Borneo	na	y	AVP	piv-npiv	piv-npiv	2-wy sym	n	N-Poss	y	*pa-	high	y	y	no-indx	y-opt	y

Table C.1: Survey data for Questionnaire A (cont.)

Language	Glotto	ISO	Gen. grp.	Location	Maj. Div.	1 WrdsOrd	notes	2 MrpAlg	3 SynAlg	4 Voice	5 Case	6 N-Poss	7a MrpCau	7b SrcCau	7c CauPrd	8a Appl	8b Co-oc	9 Indx	10 SynPrv	11 AccRel
Central Sama	cent2092	sml	GRB	Philippines	other	y	VPA	mixed (other)	piv-npiv	mrg Phil-typ	lim	N-Poss y	*pa-	med	y	n-part	y	y-opt	y	
South. Sama	sout2918	ssb	GRB	Philippines	other	y	VPA	mixed-npiv.A	piv-npiv	Phil-typ	lim	N-Poss y	*pa-	high	y	n	y	y-obl	y	
Ngaju	ngaj1237	nij	GRB	Borneo	Borneo	y	AVP	piv-npiv	piv-npiv	2-wy sym	n	N-Poss y	*pa-	med	n	na	no-appl	no-appl	no-appl	
Benyadu'	beny1237	byd	LND	Borneo	na	y	AVP	piv-npiv	mixed	2-wy sym	lim	N-Poss und	und	und	n	na	no-appl	no-appl	no-appl	
Bakati'	beka1241	bei	LND	Borneo	na	y	AVP	mixed-npiv.A	piv-npiv	2-wy sym	n	N-Poss y	other	med	n	na	no-appl	no-appl	no-appl	
Mateq	kemb1249	xem	LND	Borneo	na	y	AVP	neutral	piv-npiv	2-wy sym	n	N-Poss y	other	low	n	na	no-appl	no-appl	no-appl	
Ribu	ribu1241	rir	LND	Borneo	na	y	AVP	mixed-npiv.A	piv-npiv	2-wy sym	n	N-Poss und	und	und	n	na	no-appl	no-appl	no-appl	
Cent. Melanau	cent2101	mel	M-K	Borneo	na	y	AVP	mixed-npiv.A	und	2-wy sym	lim	N-Poss y	*pa-	med	n	na	no-appl	no-appl	no-appl	
W. Berawan	west2564	zbw	NSAR	Borneo	na	y	AVP	mixed-npiv.A	mixed	2-wy sym	lim	N-Poss y	*pa-	med	n	na	no-appl	no-appl	no-appl	
Belait	bela1260	beg	NSAR	Borneo	na	y	AVP/VPA	mixed-npiv.A	piv-npiv	2-wy sym	lim	N-Poss y	*pa-	med	n	na	no-appl	no-appl	no-appl	
Kelabit	kela1258	kzi	NSAR	Borneo	na	y	AVP	neutral	piv-npiv	Phil-typ	n	N-Poss y	*pa-	low	y	n	no-indx	y-obl	y	
Lundayeh	lund1271	lnd	NSAR	Borneo	na	y	AVP/VPA	piv-npiv	piv-npiv	mrg Phil-typ	lim	N-Poss y	other	med	y	n	no-indx	y-obl	y	
Baram Kayan*	bara1370	kys	NSAR	Borneo	na	y	AVP	piv-npiv	und	2-wy sym	lim	N-Poss y	*pa-	und	n	na	no-appl	no-appl	no-appl	
E. Penan	east2485	pez	NSAR	Borneo	na	y	AVP	mixed-npiv.A	piv-npiv	2-wy sym	lim	N-Poss y	*pa-	und	n	na	no-appl	no-appl	no-appl	
Punan Tubu	puna1266	puj	NSAR	Borneo	na	y	AVP	piv-npiv	piv-npiv	2-wy sym	lim	N-Poss y	*pa-	und	n	na	no-appl	no-appl	no-appl	
Kimaragang	kima1244	kqr	SAB	Borneo	na	n	VAP/VPA	piv-npiv	piv-npiv	Phil-typ	y	N-Poss y	*pa-	high	y	n	no-indx	y-obl	y	
Ida'an	idaa1241	dbj	SAB	Borneo	na	y	AVP	piv-npiv	piv-npiv	2-wy sym	lim	N-Poss y	*pa-	high	n	na	no-appl	no-appl	no-appl	
Keningau Mur.	keni1249	kxi	SAB	Borneo	na	n	VAP	piv-npiv	und	Phil-typ	y	N-Poss y	*pa-	high	y	n	no-indx	y-obl	und	
Timugon Mur.	timu1262	tih	SAB	Borneo	na	n	VAP/VPA	piv-npiv	piv-npiv	Phil-typ	y	N-Poss y	*pa-	high	y	n	y	y-obl	no-appl	
Serudung Mur.	seru1246	srk	SAB	Borneo	na	y	VPA	mixed-npiv.A	piv-npiv	2-wy sym	lim	N-Poss y	*pa-	und	n	na	no-appl	no-appl	no-appl	
Tombonuo	tomb1244	txa	SAB	Borneo	na	n	VAP	piv-npiv	und	Phil-typ	y	N-Poss y	*pa-	high	y	n	und	y-obl	und	
Tatana	tata1257	txx	SAB	Borneo	na	y	VPA	piv-npiv	piv-npiv	Phil-typ	y	N-Poss y	*pa-	high	y	n	y	y-obl	y	

Table C.2: Survey data for Questionnaire B

Language	Glotto	ISO	Gen. Grp.	AM Form	Type	A SrcMph	B SrcTyp	C1 BEN	C2 REC	C3 LOC	C4 GOAL	C5 THM	C6 INST	C7 STIM	C8 CONT	C9 CIRC	C10 COM	C11 ADDR	D CoMrp	E1 CAUS	E2 CAUS-c	E3 PLUR	E4 INTENS	E5 COMP	E6 sem-chg
Alas Batak	bata1292	btz	NWS-BI	ken	I	*akən	adpos	✓	✓	–	✓	–	✓	–	✓	–	–	–	–	–	–	–	–	–	–
Alas Batak	bata1292	btz	NWS-BI	-i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Karo Batak	bata1293	btx	NWS-BI	ken	I	*akən	adpos	–	–	–	–	✓	✓	–	–	–	–	–	–	–	–	–	–	–	–
Karo Batak	bata1293	btx	NWS-BI	-i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Toba Batak	bata1289	bbc	NWS-BI	hon	I	*akən	adpos	✓	–	–	–	✓	✓	–	–	–	–	–	Y-SF	✓	–	–	–	–	–
Toba Batak	bata1289	bbc	NWS-BI	-i/-an	II	*-an/*-ən	voi-mrkr	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Gayo	gayo1244	gay	NWS-BI	-i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Gayo	gayo1244	gay	NWS-BI	(n)en	I	*-an/*-ən	voi-mrkr	–	–	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–
Nias	nias1242	nia	NWS-BI	'ō	I	*akən	adpos	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Nias	nias1242	nia	NWS-BI	-(C)i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	Y-CAUS	✓	–	–	–	–
Nias	nias1242	nia	NWS-BI	fa-	other	*pa-	caus	–	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Enggano	engg1245	eno	ENG	-(C)aʔa	I	*akən	adpos	–	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Enggano	engg1245	eno	ENG	-(C)i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Nasal	nasal1239	nsy	NAS	-kun	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Nasal	nasal1239	nsy	NAS	-i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Dampelas	damp1237	dms	T-T	a'o	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Dampelas	damp1237	dms	T-T	-i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Pendau	pend1242	ums	T-T	a'	I	*akən	adpos	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Pendau	pend1242	ums	T-T	-i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Tajio	taji1246	tdj	T-T	ao	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Tajio	taji1246	tdj	T-T	-i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Totoli*	toto1304	twe	T-T	-i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Totoli*	toto1304	twe	T-T	-an	I	*-an/*-ən	voi-mrkr	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Ledo Kaili	ledo1238	lew	K-P	-aka	I	*akən	adpos	–	–	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–
Ledo Kaili	ledo1238	lew	K-P	-i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Ledo Kaili	ledo1238	lew	K-P	-ka	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Moma	moma1242	myl	K-P	-aka	I	*akən	adpos	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Moma	moma1242	myl	K-P	-i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Moma	moma1242	myl	K-P	-ka	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Uma	umaa1242	ppk	K-P	-i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Uma	umaa1242	ppk	K-P	-ki	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Behoa*	beso1237	bep	K-P	-á	I	*-an/*-ən	voi-mrkr	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Behoa*	beso1237	bep	K-P	-i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Balantak	bala1315	blz	S-B	-kon	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Balantak	bala1315	blz	S-B	-i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Balantak	bala1315	blz	S-B	-ii	other	other	other	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Bobongko	bobo1255	bgb	B-T	-akon	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Bobongko	bobo1255	bgb	B-T	-i/-an	II	other	voi-mrkr	–	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Bobongko	bobo1255	bgb	B-T	poN-	CV	other	other	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Bobongko	bobo1255	bgb	B-T	poN--an	LV	other	other	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Mori Bawah	mori1268	xmz	B-T	-ako	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Mori Bawah	mori1268	xmz	B-T	-Ci	II	*i	undet	–	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Mori Bawah	mori1268	xmz	B-T	-Cari	II	other	undet	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Mori Bawah	mori1268	xmz	B-T	-Cako	I	*akən	adpos	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Moronene	moro1287	mqn	B-T	-ako	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Moronene	moro1287	mqn	B-T	-Ci	II	*i	undet	–	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Tolaki	tola1247	lbw	B-T	-Cako	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Muna	muna1247	mnb	M-B	-ghoo	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Muna	muna1247	mnb	M-B	-Ci	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Busoa	buso1238	bup	M-B	-ho	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Busoa	buso1238	bup	M-B	-Ci	II	*i	undet	–	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Tukang Besi	tuka1248	khc	M-B	-ako	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Tukang Besi	tuka1248	khc	M-B	-(VC)i	II	*i	undet	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Tukang Besi	tuka1248	khc	M-B	-ngkene	other	other	other	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Wotu	wotu1240	wtw	W-W	-a	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Wotu	wotu1240	wtw	W-W	-i	II	*i	undet	–	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Laiyolo	laiy1246	lji	W-W	-aka	I	*akən	adpos	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Laiyolo	laiy1246	lji	W-W	-ka	I	*akən	undet	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Laiyolo	laiy1246	lji	W-W	-i	II	*i	undet	–	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Wolio	woli1241	wlo	W-W	-aka	I	*akən	adpos	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Wolio	woli1241	wlo	W-W	-i	II	*i	undet	–	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Embaloh	emba1238	emb	SSUL	-ang	I	*-an/*-ən	voi-mrkr	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Embaloh	emba1238	emb	SSUL	-i	II	*i	undet	–	–	–	✓	–	–	–	–	–	–	–	–	–	–	–	–	–	–

Table C.2: Survey data for Questionnaire B (cont.)

Language	Glotto	ISO	Gen. Grp.	AM Form	Type	A	B	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	D	E1	E2	E3	E4	E5	E6
						SrcMph	SrcTyp	BEN	REC	LOC	GOAL	THM	INST	STIM	CONT	CIRC	COM	ADDR	CoMprp	CAUS	CAUS-c	PLUR	INTENS	COMP	sem-chg
Bugis	bugi1244	bug	SSUL	-Ceng	I	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Bugis	bugi1244	bug	SSUL	-Ci	II	*i	undet	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Coastal Konjo	coas1295	kjc	SSUL	-ang	I	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Coastal Konjo	coas1295	kjc	SSUL	-i	II	*i	undet	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Makassar	maka1311	mak	SSUL	-ang	I	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Y-CAUS	✓	✓	✓	✓	✓
Makassar	maka1311	mak	SSUL	-i	II	*i	undet	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Y-CAUS	✓	✓	✓	✓	✓
Duri	duri1242	mvp	SSUL	-an	I	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Duri	duri1242	mvp	SSUL	-i	I	*i	undet	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Bambam	bamb1270	ptu	SSUL	-am	I	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Y-SF	✓	✓	✓	✓	✓
Bambam	bamb1270	ptu	SSUL	-i	II	*i	undet	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Y-SF	✓	✓	✓	✓	✓
Seko Padang	seko1243	skx	SSUL	-ing	I	*akən	adpos	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Seko Padang	seko1243	skx	SSUL	-i	II	*i	undet	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Acehnese	achi1257	ace	CHA	peu-	other	*pa-	caus	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
S. Barisan Mal.	cent2053	pse	MAL	-ka	I	*akən	adpos	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
S. Barisan Mal.	cent2053	pse	MAL	-i	II	*i	adpos	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Jambi Malay	jamb1236	jax	MAL	-kan/-an	I	*akən	adpos	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Jambi Malay	jamb1236	jax	MAL	-i	II	*i	undet	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Brunei	brun1242	kxd	MAL	-kan	I	*akən	adpos	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Brunei	brun1242	kxd	MAL	-i	II	*i	undet	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Kendayan	kend1254	knx	MAL	-an	I	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Kendayan	kend1254	knx	MAL	-i?	II	*i	undet	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Indonesian	indo1316	ind	MAL	-kan	I	*akən	adpos	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Indonesian	indo1316	ind	MAL	-i	II	*i	undet	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Bali	bali1278	ban	BSS	-ang	I	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Bali	bali1278	ban	BSS	-in	II	*i	undet	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sasak	sasa1249	sas	BSS	-an	I	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Madura	nucl1460	mad	MAD	-agi	I	*akən	adpos	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Madura	nucl1460	mad	MAD	-e	II	*i	undet	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sunda	sund1252	sun	SUN	-keun	I	*akən	adpos	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sunda	sund1252	sun	SUN	-an	I	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sunda	sund1252	sun	SUN	pang--keun	other	other	other	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Javanese	java1254	jav	JAV	-aké	I	*akən	adpos	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Javanese	java1254	jav	JAV	-i/-an	II	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Tengger	teng1272	tes	JAV	-en/-na	I	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Tengger	teng1272	tes	JAV	-i/-an	I	other	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Suriname Jav.	cari1276	jvn	JAV	-aké	I	*akən	adpos	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Suriname Jav.	cari1276	jvn	JAV	-i	II	*i	undet	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lampung Api	lamp1243	ljp	LAM	-ko	I	*akən	adpos	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lampung Api	lamp1243	ljp	LAM	-i	II	*i	undet	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Merina Malag.	plat1254	plt	GRB	-a-	IV	other	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Merina Malag.	plat1254	plt	GRB	-an	BV	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Merina Malag.	plat1254	plt	GRB	PFX--an	CV	other	other	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Y-SF	✓	✓	✓	✓	✓
Yakan	yaka1277	yka	GRB	-an	other	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Yakan	yaka1277	yka	GRB	paN-	IV	other	other	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
W. C. Bajau	west2560	bdr	GRB	-an	other	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Central Sama	cent2092	sml	GRB	-an	other	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Central Sama	cent2092	sml	GRB	-an	IV	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Central Sama	cent2092	sml	GRB	paN--an	LV	other	other	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
South. Sama	sout2918	ssb	GRB	paN--an/paN--in	LV	other	other	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
South. Sama	sout2918	ssb	GRB	-in--an/-an	BV	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
South. Sama	sout2918	ssb	GRB	pan-	IV	other	other	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Kelabit	kelai258	kzi	NSAR	peN-/peneN-	IV	other	other	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lundayeh	lund1271	lnd	NSAR	piN-	IV	other	other	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Kimaragang	kima1244	kqr	SAB	-an	BV	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Kimaragang	kima1244	kqr	SAB	-i	IV	other	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Kimaragang	kima1244	kqr	SAB	-on	LV	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Keningau Mur.	keni1249	kxi	SAB	-in/-an/-i'	BV	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Keningau Mur.	keni1249	kxi	SAB	pa--on/pina-/pa--o'	IV	other	other	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Keningau Mur.	keni1249	kxi	SAB	m/n/pang--an	LV	other	other	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Timugon Mur.	timu1262	tih	SAB	-an/-in/-i?	BV	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Timugon Mur.	timu1262	tih	SAB	CV-	IV	other	other	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Timugon Mur.	timu1262	tih	SAB	-an	LV	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Tombonuo	tomb1244	txa	SAB	-i-	IV	other	voi-mrkr	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table C.2: Survey data for Questionnaire B (cont.)

Language	Glotto	ISO	Gen. Grp.	AM Form	Type	A	B	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	D	E1	E2	E3	E4	E5	E6
						SrcMph	SrcTyp	BEN	REC	LOC	GOAL	THM	INST	STIM	CONT	CIRC	COM	ADDR	CoMrp	CAUS	CAUS-c	PLUR	INTENS	COMP	sem-chg
Tombonuo	tomb1244	txa	SAB	-an/-i	BV	*-an/*-ən	voi-mrkr	✓	✓	-	-	-	-	-	-	-	-	✓	-	-	-	-	-	-	-
Tatana	tata1257	txx	SAB	i-	IV	other	voi-mrkr	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-
Tatana	tata1257	txx	SAB	poN-	IV	other	other	-	-	-	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-
Tatana	tata1257	txx	SAB	-an/-i	BV	*-an/*-ən	voi-mrkr	✓	✓	✓	✓	-	-	-	✓	-	-	-	-	-	-	-	-	-	-
Tatana	tata1257	txx	SAB	poN--an	LV	other	other	-	-	✓	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Appendix D: Sampling of meanings for the lexical study

For the case study in Chapter 2 and the lexical study in Chapter 8, a core sample of verbal stems is based on the list of meanings given in the Leipzig Valency Project Questionnaire (Hartmann, Haspelmath & Taylor 2013). This list was created by the members of the Leipzig Valency Classes Project (Andrej Malchukov, Bernard Comrie, Iren Hartmann, Martin Haspelmath, Bradley Taylor and Søren Wichmann) and is intended as a representative sample of the verbal lexicon. The seventy core verbal meanings and ten additional meanings are listed here for reference, for a total of eighty verbal meanings. Some substitutions were made as appropriate for the West Nusantara cultural context, as described in Chapter 8.

No.	Meaning label	Typical context
1	RAIN	It rained yesterday.
2	BE DRY	The ground is dry.
3	BURN	The house is burning.
4	SINK	The boat sank.
5	ROLL	The ball is rolling.
6	BE A HUNTER	This man is a hunter.
7	BE HUNGRY	The baby is hungry.
8	BE SAD	The little girl was sad.
9	DIE	The snake died.
10	FEEL COLD	I'm cold.
11	FEEL PAIN	My arm is hurting./I'm feeling pain in my arm.
12	SCREAM	The man screamed.
13	LAUGH	The little girl laughed.
14	PLAY	The child is playing.
15	LIVE	The old people live in town.
16	LEAVE	The boy left the village.
17	GO	The woman went to the market.
18	SING	The boy sang (a song).
19	JUMP	The girl jumped.
20	SIT DOWN	The children sat down on the bench.
21	SIT	The children sat on the floor.

22	RUN	The horse is running.
23	CLIMB	The men climbed (up) the tree.
24	COUGH	The old man coughed.
25	BLINK	I blinked (my eyes).
26	SHAVE	The man shaved his beard/cut his hair.
27	DRESS	The mother dressed her daughter.
28	WASH	The mother washed the baby.
29	EAT	The boy ate the fruit.
30	HELP	I helped the boys.
31	FOLLOW	The boys followed the girls.
32	MEET	The men met the boys.
33	HUG	The mother hugged her little boy.
34	SEARCH FOR	The men searched for the women.
35	THINK	The girl thought about her grandmother yesterday.
36	KNOW	The girl knew the boy.
37	LIKE	The boy liked his new toy.
38	FEAR	The man feared the bear.
39	FRIGHTEN	The bear frightened the man.
40	SMELL	The bear smelled the boy.
41	LOOK AT	The boy looked at the girl.
42	SEE	The man saw the bear.
43	TALK	The girl talked to the boy about her dog.
44	ASK FOR	The boy asked his parents for money.
45	SHOUT AT	The woman shouted at the children.
46	TELL	The girl told the boy a funny story.
47	SAY	They said "no" to me.
48	NAME	The parents called the baby Anna.
49	BUILD	The men built a house out of wood.
50	BREAK	The boy broke the window with a stone.
51	KILL	The man killed his enemy with a club.
52	BEAT	The boy beat the snake with a stick.
53	HIT	The boy hit the snake with a stick.
54	TOUCH	The boy touched the snake with a stick.
55	CUT	The woman cut the bread with a sharp knife.
56	TAKE	The man took the money from his friend.
57	TEAR	The girl tore the page from the book.
58	PEEL	The boy peeled the bark off the stick.
59	HIDE	The boy hid the frog from his mother.
60	SHOW	The girls showed pictures to the teacher.
61	GIVE	We gave the books to the children.
62	SEND	The girl sent flowers to her grandmother.
63	CARRY	The men carried the boxes to the market.
64	THROW	The boy threw the ball into the window.
65	TIE	The man tied the horse with a rope to the tree.
66	PUT	I put the cup onto the table.

67	POUR	The man poured water into the glass.
68	COVER	The woman covered the boy with a blanket.
69	FILL	The girl filled the glass with water.
70	LOAD	The farmer loaded hay onto the truck./The farmer loaded the truck with hay.
<hr/>		
71	BRING	The girl brought flowers to me.
72	PUSH	The boy pushed the girl (into the water).
73	DIG	The woman is digging for potatoes.
74	WIPE	The women wiped dirt off the table.
75	STEAL	The thief stole money from the old lady.
76	GRIND	The women ground the seeds (with mortar and pestle).
77	HEAR	The boy heard the bear.
78	TEACH	The old lady taught the girl a song.
79	COOK	The women cooked the meat.
80	BOIL	The water is boiling.
<hr/>		

Appendix E: Database for lexical study

The table below contains the data used in the lexical study described in Chapter 8.

Abbreviations used in general coding: af = affixed verb or AM-marked construction, bs = base, cat = grammatical category, chg = change, cls = clause, detr = detransitive, ditr = ditransitive, fos = fossilized, ISO = ISO-639-3 code for language, mntr = monotransitive, remp = remapping, syn = syntactic, str = clause structure, vl = valency.

Abbreviations and values used in coding of non-applicative functions for the AM-marked construction: CAT = verbalizing with non-verbal base, CAUS = causative, selection of instigator or effector as S/A, LEX lexicalized change in semantic meaning, INTENS = intensive meaning, OBLIG = affix obligatorily present on a verbal base, OBLIG.PV = affix obligatorily present in PV, OPT = affix is optionally present with no other observed function, PAT = selecting of patientive P argument otherwise not observed, PLUR = pluractional aspectual meaning, other = other function (e.g. imperative, emphatic), na = no AM is attested with this lexical base.

Abbreviations for semantic roles used in coding of argument structure and applicative functions: actv = activity, addr = addressee, ben = beneficiary, csnd = causand, cgnr = cognizer, com = comitative, comr = communicator, cnsr = consumer, cnsd = consumed, cont = content, crtr = creator, crtn = creation, dum = dummy, emtr = emoter, expr = experiencer, istg = instigator, inst = instrument, loc = location, matl = material, movr = mover, obsr = observer, pat = patient, prcr = perceiver, prfr = performer, prfc = performance, purp = purpose, reas = reason, src = source, stim = stimulus, targ = targ, thm = theme.

Table E.1: Lexical data by base meaning

ID	Meaning	ISO	Base	BsMorph	BsCat	BsVI	Affix	AfVI	Function	SynChg	AfMeaning	BsStr	AfStr
1	RAIN	blz	usan	bare	unclear	0	-i/-ii/-kon	na	na	na	na		
1	RAIN	ind	hujan	bare	unclear	0	-i	2	CAUS+GOAL	mnr	shower (s.t./s.o) w/ s.t.		istg, goal, thm=PP
1	RAIN	ind	hujan	bare	unclear	0	-kan	2	CAUS+THM	mnr	rain down (s.t.) to s.o.		istg, thm, goal=PP
1	RAIN	jav	udan	bare	unclear	0	-aké	2	CAUS+THM	mnr	rain down (s.t.) on (s.t./s.o.)		istg, thm, goal=PP
1	RAIN	jav	udan	bare	unclear	0	-i	2	CAUS+THM	mnr	rain down (s.t.) on s.t.		istg, thm, goal=PP
1	RAIN	mnb	ghuse	no-	noun	1	-i	2	GOAL	mnr	rain on (s.t.)	dum	dum, goal
1	RAIN	mnb	ghuse	no-	noun	1	-ghoo	na	na	na	na	dum	
1	RAIN	nsy	hujan	bare	verb	0	-kun	2	CAUS+GOAL	mnr	water (s.t.), let be rained on		istg, goal
1	RAIN	nsy	hujan	bare	verb	0	-i	2	CAUS+GOAL	mnr	rain down on (s.o.) using s.t.		istg, goal, thm=PP
1	RAIN	sun	hujan	bare	unclear	0	-keun	2	CAUS+GOAL	mnr	let (s.t.) to be rained on		istg, goal
1	RAIN	sun	hujan	bare	unclear	0	-keun	2	CAUS+THM	mnr	rain down (s.t.) to s.o.		istg, thm, goal=PP
1	RAIN	sun	hujan	bare	unclear	0	-an/pang- -keun	na	na	na	na		
1	RAIN	yka	ulan	bare	verb	0	-an	2	CAUS+GOAL	mnr	put (s.t.) in the rain		istg, goal
2	BE DRY	bes	keghing	bare	verb	1	-i	2	CAUS	mnr	dry (s.t.) a little more	pat	istg, pat
2	BE DRY	bes	keghing	bare	verb	1	-ka	2	CAUS	mnr	dry (s.t.) a little more	pat	istg, pat
2	BE DRY	blz	kangkung	bare	verb	1	-i/-ii/-kon	na	na	na	na	pat	
2	BE DRY	ind	kering	bare	verb	1	-i	2	CAUS	mnr	leave (s.t.) to dry	pat	istg, pat
2	BE DRY	ind	kering	bare	verb	1	-kan	2	CAUS	mnr	make (s.t.) dry	pat	istg, pat
2	BE DRY	jav	garing	bare	verb	1	-aké	2	CAUS	mnr	dry out (s.t.)	pat	istg, pat
2	BE DRY	jav	garing	bare	verb	1	-i	na	na	na	na	pat	
2	BE DRY	mnb	neu	ao-	verb	1	-i/-ghoo	na	na	na	na	pat	
2	BE DRY	nsy	langu	bare	verb	1	-i	2	CAUS+PLUR	mnr	dry (s.t.) out	pat	istg, pat
2	BE DRY	nsy	langu	bare	verb	1	-kun	2	CAUS	mnr	dry (s.t.)	pat	istg, pat
2	BE DRY	sas	gero	bare?	verb	1	-an	2	CAUS	mnr	make (s.t.) dry	pat	istg, pat
2	BE DRY	sun	garing	bare	verb	1	-keun	2	CAUS	mnr	make (s.t.) dry	pat	istg, pat, inst=PP
2	BE DRY	sun	garing	bare	verb	1	-an/pang- -keun	na	na	na	na	pat	
2	BE DRY	sun	garing	bare	verb	1	pang- -keun	3	CAUS+BEN	ditr	dry (s.t.) for (s.o.)	pat	istg, ben, pat
2	BE DRY	yka	toho'	bare	verb	1	-an	na	na	na	na	pat	
3	BURN	blz	tunu	AV	verb	2	-i	2	OPT	no-chg	burn (s.t.)	agt, pat	agt, pat
3	BURN	blz	tunu	AV	verb	2	-kon	2	BEN	no-chg	help (s.o.) burn/roast (s.t.)	agt, pat	agt, pat, rec=poss
3	BURN	blz	tunu	AV	verb	2	-ii	na	na	na	na	agt, pat	
3	BURN	ind	bakar	AV	verb	2	-i	2	PLUR	no-chg	burn (s.t.) rep.	agt, pat	agt, pat
3	BURN	ind	bakar	AV	verb	2	-kan	3	BEN	ditr	roast for (s.o.) (s.t.)	agt, pat	agt, rec, pat
3	BURN	mnb	tunu	ae-	verb	2	-i	2	PLUR	no-chg	burn (many)	agt, pat	agt, pat
3	BURN	mnb	tunu	ae-	verb	2	-ghoo	na	na	na	na	agt, pat	
3	BURN	nsy	suah	AV	verb	2	-i	2	PLUR	no-chg	burn (many)	agt, pat	agt, pat
3	BURN	nsy	suah	AV	verb	2	-kun	3	BEN	no-chg	burn (s.t.) for (s.o.)	agt, pat	agt, ben, pat
3	BURN	sun	duruk	AV	verb	2	-an	2	PLUR+LEX	no-chg	make a bonfire (dur.)	agt, pat	agt, pat
3	BURN	sun	duruk	AV	verb	2	pang- -keun	3	BEN	ditr	burn (s.o.) (s.t.)	agt, pat	agt, ben, pat
3	BURN	sun	duruk	AV	verb	2	-keun	na	na	na	na	agt, pat	
3	BURN	yka	eggas	bare, N-	verb	2	-an	na	na	na	na	agt, pat	
4	SINK	bes	beghak	te-?	verb	1	-ka	2	CAUS	mnr	immerse (s.o. or s.t.)	thm	istg, thm
4	SINK	bes	beghak	te-?	verb	1	-i	2	GOAL+LEX	mnr	immerse o.s. in loc	thm	agt, goal
4	SINK	blz	lolop	mo-	verb	1	-i/-ii/-kon	na	na	na	na	thm	
4	SINK	ind	tenggelam	fos.NVOL	verb	1	-kan	2	CAUS	mnr	cause (s.t.) to sink	thm	istg, thm, goal=PP
4	SINK	ind	tenggelam	fos.NVOL	verb	1	-i	na	na	na	na	thm	
4	SINK	jav	kelem	AV	verb	1	-aké	2	CAUS	mnr	submerge (s.t.)	thm, loc=PP	istg, thm, loc=PP
4	SINK	jav	kelem	AV	verb	1	-i	na	na	na	na	thm, loc=PP	
4	SINK	mnb	tondu	a-	verb	1	-i	2	CAUS	mnr	bury at sea, sink (s.t.)	thm	istg, thm, goal=PP
4	SINK	mnb	tondu	a-	verb	1	-ghoo	na	na	na	na	thm	
4	SINK	nsy	khendum	NVOL	verb	1	-i	2	CAUS+PLUR	mnr	submerge (many)	thm	istg, thm
4	SINK	nsy	khendum	NVOL	verb	1	-kun	2	CAUS	mnr	submerge (s.t.)	thm	istg, thm, goal=PP
4	SINK	sas	selem	bare	verb	1	-an	2	CAUS	mnr	cause (s.t.) to sink	thm	istg, thm, goal=PP
4	SINK	sun	kerelep	NVOL	verb	1	-keun	2	CAUS	mnr	cause (s.t.) to sink	thm	istg, thm, goal=PP
4	SINK	sun	kerelep	NVOL	verb	1	-an/pang- -keun	na	na	na	na	thm	
4	SINK	yka	lenneb	pa- SF	verb	1	-an	2	CAUS	mnr	submerge (s.t.)	thm	istg, thm
5	ROLL	blz	tinda'	ming-	verb	2	-i/-ii/-kon	na	na	na	na	agt, thm	
5	ROLL	ind	gelinding	MID	verb	1	-kan	2	CAUS	mnr	roll (s.t.)	thm	istg, thm
5	ROLL	ind	gelinding	MID	verb	1	-i	na	na	na	na	thm	
5	ROLL	jav	glundhung	AV	verb	1	-aké	na	CAUS	mnr	roll (s.t.)	thm, path=cls	
5	ROLL	jav	glundhung	AV	verb	1	-i	na	na	na	na	thm, path=cls	
5	ROLL	mnb	lole	ae-	verb	1	-i	2	CAUS+PLUR	mnr	roll (many)	thm	istg, thm
5	ROLL	sun	gulutuk	AV	verb	1	-keun	2	CAUS	mnr	roll (s.t.)	thm	istg, thm
5	ROLL	sun	gulutuk	AV	verb	1	-an/pang- -keun	na	na	na	na	thm	

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr
5	ROLL	yka	liring	pa- SF	verb	1	-an	1	PLUR	no-chg	roll, turn rep.	thm	thm
5	ROLL	yka	liring	pa- SF	verb	1	-an	2	CAUS	mnr	To roll, turn (s.t.)	thm	istg, thm
7	BE HUNGRY	blz	kayo'	ma-	verb	1	-i/-ii/-kon	na	na	na	na	expr	
7	BE HUNGRY	ind	lapar	bare	verb	1	-kan	2	CAUS	mnr	cause hunger	expr	istg, expr
7	BE HUNGRY	ind	lapar	bare	verb	1	-i	na	na	na	na	expr	
7	BE HUNGRY	jav	ngelih	bare	verb	1	-aké/-i	na	na	na	na	expr	
7	BE HUNGRY	mnb	gharo	ao-	verb	1	-i/-ghoo	na	na	na	na	expr	
7	BE HUNGRY	nsy	lutuh	bare	verb	1	-i	2	CAUS	mnr	hunger (one's stomach)	expr	istg, expr
7	BE HUNGRY	nsy	lutuh	bare	verb	1	-kun	2	CAUS	mnr	make (s.o.) hungry	expr	istg, expr
7	BE HUNGRY	sun	lapar	bare	verb	1	-an	na	na	na	na	expr	
7	BE HUNGRY	yka	inusan	bare	verb	1	-an	na	na	na	na	expr	
8	BE SAD	bes	sedih	bare?	verb	1	-i	na	na	na	na	expr	
8	BE SAD	bes	sedih	bare?	verb	1	-ka	1	CAUS	caus-detr	be saddening	expr	istg
8	BE SAD	blz	siongo'	ma-	verb	1	-i/-ii/-kon	na	na	na	na	emtr	
8	BE SAD	ind	sedih	bare	verb	1	-kan	2	CAUS	mnr	make (s.o.) sad	emtr	istg, emtr
8	BE SAD	ind	sedih	bare	verb	1	-kan	1	CAUS	caus-detr	arouse feelings of sadness	emtr	istg
8	BE SAD	ind	sedih	bare	verb	1	-kan	2	TARG	mnr	be sad about (s.t.)	emtr	emtr, targ
8	BE SAD	ind	sedih	bare	verb	1	-i	na	na	na	na	emtr	
8	BE SAD	jav	sedih	bare	verb	1	-aké	2	CAUS	mnr	sadden, grieve (s.o.)	emtr	istg, emtr
8	BE SAD	jav	sedih	bare	verb	1	-aké	1	CAUS	caus-detr	sorrowful, causing sorrow	emtr	istg
8	BE SAD	jav	sedih	bare	verb	1	-i	2	CAUS	caus	sadden, grieve (s.o.)	emtr	istg, emtr
8	BE SAD	mnb	bhela	ao-	verb	1	-ghoo	1	CAUS	caus-detr	saddening, pitiful	emtr	istg
8	BE SAD	mnb	bhela	ao-	verb	1	-i	na	na	na	na	emtr	
8	BE SAD	nsy	sedih	bare	verb	1	-i	1	REAS	caus-detr	be saddened by (s.t.)	emtr	emtr, targ=cls
8	BE SAD	nsy	sedih	bare	verb	1	-kun	1	CAUS	caus-detr	saddening	emtr	istg
8	BE SAD	sas	aséq	bare?	verb	1	-an	1	PLUR	no-chg	be sad (hab.)	emtr	emtr
8	BE SAD	sun	sedih	bare	verb	1	-an	na	na	na	na	emtr	
8	BE SAD	yka	sugul	bare	verb	1	-an	na	na	na	na	emtr	
9	DIE	bes	mati	bare	verb	1	-i	na	na	na	na	pat	
9	DIE	bes	mati	bare	verb	1	-ka	na	CAUS	mnr	turn off (s.t.)	pat	
9	DIE	blz	pate	bare	verb	1	-i+pa-	2	CAUS	mnr	kill (s.o./s.t.)	pat	istg, pat
9	DIE	blz	pate	bare	verb	1	-ii/-kon	na	na	na	kill (s.o./s.t.)	pat	
9	DIE	ind	mati	bare	verb	1	-kan	2	CAUS	mnr	kill (s.o./s.t.)	pat	istg, pat
9	DIE	ind	mati	bare	verb	1	-i	na	na	na	na	pat	
9	DIE	jav	mati	bare	verb	1	-aké	2	CAUS	mnr	let/cause (s.t.) to die	pat	istg, pat
9	DIE	jav	mati	bare	verb	1	-i	2	CAUS	mnr	kill/put out(s.o./s.t.)	pat	istg, pat
9	DIE	mnb	mate	a-	verb	1	-ghoo	2	REAS	mnr	die of (reas)	pat	pat, reas=IO
9	DIE	mnb	mate	a-	verb	1	-i	na	na	na	na	pat	
9	DIE	nsy	patai	bare	verb	1	-i	2	CAUS+PLUR	mnr	kill/turn off (many)	pat	istg, pat
9	DIE	nsy	patai	bare	verb	1	-kun	2	CAUS	mnr	kill/turn off (s.t.)	pat	istg, pat
9	DIE	sas	maté	bare	verb	1	-an+q	3	BEN	ditr	kill (s.o./s.t.) for (s.o.)	pat	istg, ben, pat
9	DIE	sun	paéh	bare	verb	1	-an	2	CAUS	mnr	kill (s.o./s.t.)	pat	istg, pat
9	DIE	sun	paéh	bare	verb	1	pang- -keun+-an	3	CAUS+BEN	ditr	kill (s.o./s.t.) for (s.o.)	pat	istg, ben, pat
9	DIE	sun	paéh	bare	verb	1	-keun	na	na	na	na	pat	
9	DIE	yka	matey	bare	verb	1	-an	na	na	na	na	pat	
10	BE COLD	bes	dingin	bare?	verb	1	-i	na	na	na	na	pat	
10	BE COLD	bes	dingin	bare?	verb	1	-ka	2	CAUS	mnr	cool (s.t.) down	pat	istg, pat
10	BE COLD	blz	memel	bare	verb	1	-i	2	CAUS	mnr	make (s.t.) cold, cooler	expr	istg, pat
10	BE COLD	blz	memel	bare	verb	1	-ii/-kon	na	na	na	na	expr	
10	BE COLD	ind	dingin	bare	verb	1	-kan	2	CAUS	mnr	make (s.t.) cold	pat	istg, pat
10	BE COLD	ind	dingin	bare	verb	1	-i	na	na	na	na	pat	
10	BE COLD	jav	adhem	bare	verb	1	-aké	2	CAUS	mnr	let (s.t.) cool	pat	istg, pat
10	BE COLD	jav	adhem	bare	verb	1	-i	2	CAUS	mnr	make (s.t.) cooler	pat	istg, pat
10	BE COLD	mnb	rindima	a-	verb	1	-i/-ghoo	na	na	na	na	pat	
10	BE COLD	nsy	ngisun	unclear	verb	1	-i	2	CAUS+PLUR	mnr	cool down (many)	pat	istg, pat
10	BE COLD	nsy	ngisun	unclear	verb	1	-kun	2	CAUS	mnr	cool down (s.t.)	expr	istg, pat
10	BE COLD	sas	enyet	bare?	verb	1	-an	2	CAUS	mnr	make (s.t.) cold	expr	istg, pat
10	BE COLD	sun	tiis	bare	verb	1	-an	2	CAUS+INST	mnr	cool (s.t.) using s.t.	pat	istg, pat, inst=PP
10	BE COLD	sun	tiis	bare	verb	1	-keun	2	CAUS	mnr	let (s.t.) cool	pat	istg, pat
10	BE COLD	yka	hanggut	bare	unclear	1	-an+pa-	2	CAUS	mnr	let (s.t.) cool off	pat	istg, pat
11	FEEL PAIN	bes	sakit	bare	verb	1	-i	2	CAUS	mnr	hurt (s.o.) (generally)	expr	istg, expr
11	FEEL PAIN	bes	sakit	bare	verb	1	-ka	2	CAUS+LEX	mnr	hurt (s.o.) emotionally	expr	istg, expr
11	FEEL PAIN	blz	polos	ma-	verb	1	-i+pa-	2	CAUS	mnr	cause pain to (s.o./s.t.)	expr	istg, expr
11	FEEL PAIN	blz	polos	ma-	verb	1	-ii/-kon	na	na	na	na	expr	

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr
11	FEEL PAIN	ind	sakit	bare	verb	1	-i	2	CAUS	mnr	cause pain to (s.o.)	expr	istg, expr
11	FEEL PAIN	ind	sakit	bare	verb	1	-kan	2	CAUS	mnr	cause pain to (s.t.)	expr	istg, (expr)
11	FEEL PAIN	jav	lara	bare	verb	1	-aké	1	CAUS	caus-detr	cause pain to (s.o.)	expr	istg
11	FEEL PAIN	jav	lara	bare	verb	1	-i	2	CAUS	mnr	cause pain to s.o. or s.t.	expr	istg, expr
11	FEEL PAIN	jav	lara	bare	verb	1	-i	1	CAUS	caus-detr	painful, capable of causing pain	expr	istg
11	FEEL PAIN	mnb	lea	ao-	verb	1	-i/-ghoo	na	na	na	na	expr	
11	FEEL PAIN	nsy	sakit	bare	verb	1	-i	2	CAUS	mnr	cause pain to (s.o.)	expr	istg, expr
11	FEEL PAIN	nsy	sakit	bare	verb	1	-kun	2	CAUS	mnr	hurt (s.o.)	expr	istg, expr
11	FEEL PAIN	sas	sakit	bare	verb	1	-an	1	CAUS	caus-detr	be painful	expr	istg
11	FEEL PAIN	sun	nyeri	bare	verb	1	-keun	2	CAUS	mnr	hurt (s.o./s.t.)	expr	istg, expr
11	FEEL PAIN	yka	peddi'	bare	verb	1	-an	na	na	na	na	expr	
12	SCREAM	bes	mekik	AV?	verb	1	-i/-ka	na	na	na	na	comr	
12	SCREAM	blz	karo'	unclear	verb	1	-i/-ii/-kon	na	na	na	na	comr	
12	SCREAM	ind	teriak	MID	noun	1	-kan	2	CONT	mnr	say (s.t.) w/ a loud voice	comr	comr, cont
12	SCREAM	ind	teriak	MID	unclear	1	-i	2	ADDR	mnr	call to (s.o.) by screaming	comr	comr, addr
12	SCREAM	jav	bengok	AV	verb	1	-aké	2	CONT	mnr	shout (s.t.)	comr	comr, cont
12	SCREAM	jav	bengok	AV	verb	1	-i	2	ADDR	mnr	shout at, call to (s.o.)	comr	comr, addr
12	SCREAM	mnb	podea	a-	verb	1	-i	2	ADDR	mnr	shout at (s.o.)	comr	comr, addr
12	SCREAM	mnb	podea	a-	verb	1	-ghoo	na	na	na	na	comr	
12	SCREAM	nsy	pekik	AV	verb	1	-i	2	ADDR	mnr	yell, call out to (s.o.)	comr	comr, addr
12	SCREAM	nsy	pekik	AV	verb	1	-kun	2	ADDR	mnr	yell to (s.o.)	comr	comr, addr
12	SCREAM	sas	surak	bare?	verb	1	-an	2	ADDR	mnr	yell at (s.o.)	comr	comr, addr
12	SCREAM	sun	jerit	AV	verb	1	-keun	2	CONT	mnr	scream out (s.t.)	comr	comr, cont
12	SCREAM	sun	jerit	AV	verb	1	-an+RDP	1	PLUR	no-chg	scream rep.	comr	comr
12	SCREAM	sun	jerit	AV	verb	1	pang- -keun	na	na	na	na	comr	
12	SCREAM	yka	tilahak	N-, mag-	verb	1	-an	2	CAUS	mnr	make loud (one's voice)	comr	comr, inst
13	LAUGH	bes	tawe	te-	verb	1	-i	na	na	na	na	emtr	
13	LAUGH	bes	tawe	te-	verb	1	-ka	2	TARG+LEX	mnr	laugh at (s.t.)	emtr	emtr, targ
13	LAUGH	blz	lengke	bare	verb	1	-i	2	TARG+LEX	mnr	laugh at, make fun of (s.t.)	emtr	emtr, targ
13	LAUGH	blz	lengke	bare	verb	1	-ii/-kon	na	na	na	na	emtr	
13	LAUGH	ind	tertawa	fos.NVOL	verb	1	-i	2	TARG+LEX	mnr	laugh at, mock (s.o.)	emtr	emtr, targ
13	LAUGH	ind	tertawa	fos.NVOL	verb	1	-kan	2	TARG	mnr	laugh about (s.t./s.o.)	emtr	emtr, targ
13	LAUGH	ind	tertawa	fos.NVOL	verb	1	-kan	2	CAUS	mnr	cause (s.o.) to laugh	emtr	istg, emtr
13	LAUGH	jav	guyu	AV	unclear	1	-aké	2	CAUS	mnr	make (s.o.) laugh	emtr	istg, emtr
13	LAUGH	jav	guyu	AV	unclear	1	-i	2	CAUS+LEX	mnr	joke (w/ s.o.), make laugh	emtr	istg, emtr
13	LAUGH	mnb	futaa	a-	verb	1	-i	2	TARG	mnr	laugh at (s.t./s.o.)	emtr	emtr, targ
13	LAUGH	mnb	futaa	a-	verb	1	-ghoo	na	na	na	na	emtr	
13	LAUGH	nsy	lalang	bare	verb	1	-i	2	TARG	mnr	laugh at (s.t.)	emtr	emtr, targ
13	LAUGH	nsy	lalang	bare	verb	1	-kun	2	TARG	mnr	laugh at (s.t.)	emtr	emtr, targ
13	LAUGH	sas	keréréq	bare	verb	1	-an	2	TARG	mnr	laugh at (s.t./s.o.)	emtr	emtr, targ
13	LAUGH	sun	seuri	bare	verb	1	-an	2	TARG	mnr	laugh at, about (s.t./s.o.)	emtr	emtr, targ
13	LAUGH	sun	seuri	bare	verb	1	-an+RDP	1	PLUR	no-chg	laugh rep./dur.	emtr	emtr
13	LAUGH	sun	seuri	bare	verb	1	pang- -keun	na	na	na	na	emtr	
13	LAUGH	yka	saye	mag-, -um-	verb	1	-an	2	TARG	mnr	laugh at (s.t.)	emtr, targ=PP	emtr, targ
14	PLAY	bes	pusik	no-base	no-base	na	-i	na	na	na	na		
14	PLAY	bes	pusik	no-base	no-base	na	-ka	2	OBLIG+LEX	na	make fun of (s.o.)		agt, targ
14	PLAY	bes	pusik	no-base	no-base	na	-ka	2	OBLIG	na	play w/ (s.t.)		agt, inst
14	PLAY	blz	guas	mo-	verb	1	-i/-ii/-kon	na	na	na	na	agt	
14	PLAY	ind	main	bare	verb	1	-kan	2	INST	mnr	play w/ (s.t.), play (inst)	agt, actv/inst=NP	agt, inst
14	PLAY	ind	main	bare	verb	1	-kan	2	CAUS	mnr	make play (a player)	agt, inst=NP	istg, csnd
14	PLAY	ind	main	bare	verb	1	-i	na	na	na	na	agt, inst=NP	
14	PLAY	nsy	main	unclear	verb	1	-i	2	CAUS	mnr	make play (s.o.)	agt, actv/inst=NP, loc=PP	istg, csnd
14	PLAY	nsy	main	unclear	verb	1	-kun	2	INST	mnr	play w/ or use (s.t.)	agt, actv/inst=NP, loc=PP	agt, inst
14	PLAY	nsy	main	unclear	verb	1	-kun	2	CAUS	mnr	make play (s.o.)	agt, actv/inst=NP, loc=PP	istg, csnd
14	PLAY	sas	kedek	bare?	verb	2	-an	2	OPT	no-chg	play w/ (s.t.), play (inst)	agt, inst	agt, inst
14	PLAY	sun	coo	AV	verb	1	-an	2	PLUR	mnr	play w/ (many)	agt, actv/inst=NP	agt, inst
14	PLAY	sun	coo	AV	verb	1	-keun/pang- -keun	na	na	na	na	agt, actv/inst=NP	
14	PLAY	yka	dagey	mag-	noun	1	-an	na	na	na	na	agt, actv=cls	
15	LIVE/STAY	blz	dodongo	no-+RDP+ -um-	verb	1	-i	2	CAUS	mnr	look after, watch (s.o.)	thm, loc=PP	istg, pat
15	LIVE/STAY	blz	dodongo	no-+RDP+ -um-	verb	1	-kon	2	CAUS	mnr	give birth to (s.o.)	thm, loc=PP	istg, thm
15	LIVE/STAY	blz	dodongo	no-+RDP+ -um-	verb	1	-ii	na	na	na	na	thm, loc=PP	
15	LIVE/STAY	ind	tinggal	bare	verb	1	-i	2	LOC	mnr	inhabit (a house)	thm, loc=PP	thm, loc
15	LIVE/STAY	jav	dunung	RDP	verb	1	-i	2	LOC	mnr	live (s.w.)	thm, loc=PP	thm, loc
15	LIVE/STAY	jav	dunung	RDP	verb	1	-aké	na	na	na	na	thm, loc=PP	

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr
15	LIVE/STAY	mnb	late	ae-	verb	1	-i	2	LOC	mnr	inhabit/occupy (s.t.)	thm, loc=PP	thm, loc
15	LIVE/STAY	mnb	late	ae-	verb	1	-ghoo	na	na	na	na	thm, loc=PP	
15	LIVE/STAY	sun	cicing	bare	verb	1	-an	2	LOC	mnr	live in (a house)	thm, loc=PP	thm, loc
15	LIVE/STAY	sun	cicing	bare	verb	1	-keun/pang-	na	na	na	na	thm, loc=PP	
15	LIVE/STAY	yka	luma'	mag-	verb	1	-an	na	na	na	na	thm, loc=PP	
16	LEAVE	bes	tinggal	bare	verb	1	-ka	2	CAUS	mnr	leave (s.t./s.o.) behind	thm, loc=PP	istg, thm
16	LEAVE	bes	tinggal	bare	verb	1	-i	3	REC	ditr	leave behind (for s.o.) (s.t.)	thm, loc=PP	istg, rec, thm
16	LEAVE	ind	tinggal	bare	verb	1	-i	2	CAUS	mnr	leave (s.t.) behind	thm, loc=PP	istg, thm, rec=PP
16	LEAVE	ind	tinggal	bare	verb	1	-kan	2	CAUS	mnr	leave (s.t./s.o./s.w.) behind	thm, loc=PP	istg, thm
16	LEAVE	jav	tinggal	AV	verb	2	-aké	3	REC	ditr	leave/bequeath (s.t.) to s.o.	istg, thm	istg, rec, thm
16	LEAVE	jav	tinggal	AV	verb	2	-aké	2	OPT+THM	no-chg	leave behind (s.t./s.w.)	istg, thm	istg, thm
16	LEAVE	jav	tinggal	AV	verb	2	-i	3	CAUS	ditr	leave/bequeath (s.t.)	istg, thm	istg, rec, thm
16	LEAVE	jav	tinggal	AV	verb	2	-i	2	OPT+THM	no-chg	leave behind (s.t.)	istg, thm	istg, thm
16	LEAVE	nsy	tinggal	bare	verb	1	-i	2	CAUS+BEN/REAS	mnr	leave (s.t.) behind	thm, loc=PP	istg, thm, purp/ben=PP
16	LEAVE	nsy	tinggal	bare	verb	1	-kun	2	CAUS	mnr	leave (s.t.) behind	thm, loc=PP	istg, thm
16	LEAVE	sas	nyedi	bare?	vern	1	-an	na	na	na	leave	movr, src=PP	
16	LEAVE	sun	tinggal	bare	verb	1	-an	3	CAUS+REC	ditr	leave (s.t.) for (s.o.)	thm	istg, rec, thm
16	LEAVE	sun	tinggal	bare	verb	1	-keun	2	CAUS	mnr	leave behind (s.t./s.o./s.w.)	thm	istg, thm, thm=PP
16	LEAVE	sun	tinggal	bare	verb	1	pang-	na	na	na	na	thm	
16	LEAVE	yka	amban	pa- SF	verb	1	-an	2	CAUS	na	leave behind (s.t.) at s.w.	thm, loc=PP	istg, thm, loc=PP
16	LEAVE	yka	amban	pa- SF	verb	1	-an	1	PLUR	na	stay behind, lag rep.	thm, loc=PP	
17	COME	bes	datang	bare	verb	1	-i	2	GOAL+PLUR	mnr	visit (s.w.)	thm	movr, goal
17	COME	bes	datang	bare	verb	1	-ka	2	CAUS	mnr	welcome (s.o.) to s.w.	thm	istg, thm
17	COME	blz	taka	bare	verb	1	-kon	2	CAUS	mnr	bring (s.t.)	thm, goal=PP	istg, thm
17	COME	blz	taka	bare	verb	1	-i	2	GOAL	mnr	find, encounter (s.t.)	thm, goal=PP	thm, goal
17	COME	blz	taka	bare	verb	1	-ii	na	na	na	na	thm, goal=PP	
17	COME	ind	datang	bare	verb	1	-kan	2	CAUS	mnr	bring (s.t.) from s.w.	thm, goal=PP	istg, thm, src=PP
17	COME	ind	datang	bare	verb	1	-i	2	GOAL	mnr	arrive at (s.w.)	thm, goal=PP	thm, goal
17	COME	jav	teka	bare	verb	1	-aké	2	CAUS	mnr	make (s.t.) come/happen	thm, goal=PP	istg, thm
17	COME	jav	teka	bare	verb	1	-i	2	GOAL+PLUR	mnr	visit (many)	thm, goal=PP	thm, movr, goal
17	COME	mnb	mai	a-	verb	1	-ghoo	1	SOURCE	no-chg	come from s.w.	thm, goal=PP	thm, src=PP
17	COME	mnb	mai	a-	verb	1	-i	2	GOAL	mnr	visit, come to (s.o.)	thm, goal=PP	movr, goal
17	COME	nsy	khatung	bare	verb	1	-kun	2	CAUS	mnr	bring (s.o.) in	thm, goal=PP	istg, thm
17	COME	nsy	khatung	bare	verb	1	-i	2	GOAL	mnr	visit (s.o.)	thm, goal=PP	movr, goal
17	COME	sas	dateng	bare	verb	1	-an	2	CAUS	mnr	bring (s.t.) from s.w.	thm, goal=PP	istg, thm
17	COME	sun	datang	bare	verb	1	-keun	2	CAUS	mnr	bring, invite (s.o.)	thm, goal=PP	istg, thm, loc=PP
17	COME	sun	datang	bare	verb	1	pang-	3	CAUS+BEN	ditr	invite (s.o.) for (s.o.)	thm, goal=PP	istg, ben, thm
17	COME	sun	datang	bare	verb	1	-an	2	GOAL	mnr	visit (s.w.)	thm, goal=PP	movr, goal
17	COME	yka	pitu	bare	verb	1	-an	1	PLUR	mnr	come rep.	movr	movr
18	SING	bes	nyanyi	AV?	verb	1	-i	na	na	na	na	prfr, prfc=NP	
18	SING	bes	nyanyi	AV?	verb	1	-ka	2	PERF	mnr	sing (song)	prfr, prfc=NP	prfr, prfc
18	SING	ind	nyanyi	AV	verb	1	-kan	2	PERF	mnr	sing (song)	prfr, prfc=NP	prfr, prfc
18	SING	ind	nyanyi	AV	verb	1	-i	na	na	na	na	prfr, prfc=NP	
18	SING	jav	tembang	AV	noun	2	-aké	2	OPT+PERF	no-chg	sing (song/words)	prfr, prfc	prfr, prfc/cont
18	SING	jav	tembang	AV	noun	2	-aké	3	BEN	ditr	sing for (s.o.)	prfr, prfc	prfr, ben, prfc
18	SING	jav	tembang	AV	noun	2	-i	na	na	na	na	prfr, prfc	
18	SING	mnb	lagu	ae-	noun	1	-ghoo	2	BEN	mnr	sing to/for (s.o.)	prfr	prfr, ben=IO
18	SING	mnb	lagu	ae-	noun	1	-i	na	na	na	na	prfr	
18	SING	nsy	nyanyi	unclear	verb	1	-kun	2	CAUS	mnr	sing (song)	prfr, prfc=NP	prfr, prfc
18	SING	nsy	nyanyi	unclear	verb	1	-i	na	na	na	na	prfr, prfc=NP	
18	SING	sas	nyanyi	bare?	verb	1	-an	2	PERF	mnr	sing (song/words)	prfr	prfr, prfc
18	SING	sun	nyanyi	bare	verb	1	-keun	2	PERF	mnr	sing (song)	prfr, prfc=NP	prfr, prfc
18	SING	sun	nyanyi	bare	verb	1	pang-	2	BEN	ditr	sing (song) for (s.o.)	prfr, prfc=NP	prfr, ben, prfc
18	SING	sun	nyanyi	bare	verb	1	-an	na	na	na	na	prfr, prfc=NP	
18	SING	yka	kanta	N-, mag-	noun	2	-an	na	na	na	na	prfr, prfc=NP	
19	JUMP	bes	lumpat	bare?	verb	1	-i	2	PATH	mnr	jump over (s.t.)	movr	movr, path
19	JUMP	bes	lumpat	bare?	verb	1	-ka	2	CAUS	mnr	help (s.o.) jump over s.t.	movr	istg, path, thm
19	JUMP	blz	tosik	-um-	verb	1	-i	2	GOAL	mnr	leap onto (s.t.)	movr	movr, goal
19	JUMP	blz	tosik	-um-	verb	1	-ii/-kon	na	na	na	na	movr	
19	JUMP	ind	lompat	AV	verb	1	-kan	2	CAUS	mnr	make (s.t.) jump	movr, goal=PP	agt, thm
19	JUMP	ind	lompat	AV	verb	1	-i	2	LOC	mnr	jump over (s.t.)	movr, goal=PP	movr, loc
19	JUMP	jav	lumpat	AV	verb	1	-aké	2	CAUS	mnr	cause/help (s.t.) to jump	movr, goal=PP	agt, thm
19	JUMP	jav	lumpat	AV	verb	1	-i	na	PATH	mnr	jump over (s.t.)	movr, goal=PP	
19	JUMP	mnb	punda	a-	verb	1	-i	2	PATH	mnr	jump over/toward (s.t.)	movr, goal=PP, src=PP	movr, path/goal

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr
19	JUMP	mnb	punda	a-	verb	1	-ghoo	na	na	na	na	movr, goal=PP, src=PP	
19	JUMP	nsy	luncat	AV	verb	1	-i	2	PATH	mnr	jump over (s.t.)	movr	movr, loc
19	JUMP	nsy	luncat	AV	verb	1	-kun	2	CAUS	mnr	assist (s.o.) to jump	movr	istg, thm
19	JUMP	sas	selöntak	bare?	verb	1	-an	2	CAUS	mnr	make (s.t./s.o.) jump	movr	istg, thm
19	JUMP	sun	luncat	bare	verb	1	-keun	2	CAUS	mnr	make (s.t.) jump up	movr, src=PP	agt, thm
19	JUMP	sun	luncat	bare	verb	1	pang- -keun	na	na	na	na	movr, src=PP	
19	JUMP	sun	luncat	bare	verb	1	-an	2	PATH	mnr	jump over (s.t.)	movr	movr, path
19	JUMP	yka	tugpa'	pa- SF	verb	1	-an	1	PLUR	no-chg	jump toward s.t., to dive rep.	movr, goal=PP	movr, goal=PP
19	JUMP	yka	tugpa'	pa- SF	verb	1	-an	2	GOAL	mnr	jump at (s.t./s.o.)	movr, goal=PP	movr, goal
21	SIT	bes	duduk	bare	verb	1	-i	2	CAUS	mnr	put down (s.t.)	movr, loc=PP	istg, thm
21	SIT	bes	duduk	bare	verb	1	-ka	2	CAUS	mnr	put down (s.t.), seat (s.o.)	movr, loc=PP	istg, thm
21	SIT	blz	oruang	-um-	verb	1	-kon	2	PURP	mnr	sit for (purp)	movr, loc=PP	movr, purp
21	SIT	blz	oruang	-um-	verb	1	-i/-ii	na	na	na	na	movr, loc=PP	
21	SIT	ind	duduk	bare	verb	1	-i	2	LOC	mnr	sit on (s.t.)	movr, loc=PP	movr, loc
21	SIT	ind	duduk	bare	verb	1	-kan	2	CAUS	mnr	seat (s.o.)	movr, loc=PP	istg, thm, loc=PP
21	SIT	jav	lungguh	bare	verb	1	-aké	2	CAUS	mnr	seat (s.o.)	movr, goal=PP	istg, thm
21	SIT	jav	lungguh	bare	verb	1	-i	2	LOC	mnr	sit on (s.t.), occupy (s.w.)	movr, goal=PP	movr, loc
21	SIT	mnb	ngkora	ae-	verb	1	-i	2	LOC	mnr	sit on (s.t.)	movr, loc=PP	movr, loc
21	SIT	mnb	ngkora	ae-	verb	1	-ghoo	na	na	na	na	movr, loc=PP	
21	SIT	nsy	hedung	AV	verb	1	-i	2	LOC	mnr	sit on (s.t.)	movr, loc=PP	movr, loc
21	SIT	nsy	hedung	AV	verb	1	-kun	2	CAUS	mnr	sit (s.o.) down	movr, loc=PP	istg, thm
21	SIT	sas	tököl	bare	verb	1	-an	2	CAUS	mnr	make (s.o.) sit	movr, loc=PP	istg, thm
21	SIT	sas	tököl	bare	verb	1	-an	2	LOC	mnr	sit on (s.t.)	movr, loc=PP	movr, loc
21	SIT	sun	diuk	bare	verb	1	-keun	2	CAUS	mnr	seat (s.o.)	movr, loc=PP	istg, thm, loc=PP
21	SIT	sun	diuk	bare	verb	1	pang- -keun	2	BEN	dir	help (s.o.) sit	movr, loc=PP	istg, thm+ben
21	SIT	sun	diuk	bare	verb	1	-an	2	GOAL	mnr	sit on (s.t.)	movr, loc=PP	movr, loc
21	SIT	yka	tingkó'	N-, mag-	verb	1	-an	na	na	na	na	movr, loc=PP	
22	RUN	bes	laghi	MID	verb	1	-i	na	na	na	na	movr	
22	RUN	bes	laghi	MID	verb	1	-ka	2	THM	mnr	run away w/ (s.o.), elope	movr	movr, com
22	RUN	blz	tende'	RDP + -um-	verb	1	-kon	2	THM	mnr	run off w/ (s.t.)	movr	movr, com
22	RUN	blz	tende'	RDP + -um-	verb	1	-i/-ii	na	na	na	na	movr	
22	RUN	ind	lari	bare	verb	1	-kan	2	THM	mnr	run away w/ (s.t.)	movr, src=PP	movr, com
22	RUN	ind	lari	bare	verb	1	-kan	2	CAUS	mnr	make (s.t.) run quickly	movr, src=PP	istg, thm
22	RUN	ind	lari	bare	verb	1	-i	na	na	na	na	movr, src=PP	
22	RUN	jav	blandhang	AV	verb	1	-aké	2	CAUS	mnr	make (s.t.) go fast	movr	istg, thm
22	RUN	jav	playu	AV, bare	verb	1	-aké	2	THM	mnr	run off w/ (s.t.) of s.o. else	movr	movr, com
22	RUN	jav	playu	AV, bare	verb	1	-i	2	GOAL	mnr	run to (s.o.)	movr	movr, goal
22	RUN	mnb	tende	a-	verb	1	-i	2	GOAL	mnr	run to (s.o.)	movr	movr, goal
22	RUN	mnb	tende	a-	verb	1	-ghoo	na	na	na	na	movr	
22	RUN	sas	pelai	bare	verb	1	-an	2	THM	mnr	run away, elope w/ (s.t.)	movr	movr, com
22	RUN	sas	pelai	bare	verb	1	-an	2	CAUS	mnr	make (s.t.) run quickly	movr	istg, thm
22	RUN	sun	lumpat	bare	verb	1	-keun	2	CAUS	mnr	make (s.t.) move quickly	movr	istg, thm
22	RUN	sun	lumpat	bare	verb	1	-an	2	LOC	mnr	overtake (s.t./s.o.)	movr	movr, path
22	RUN	sun	lumpat	bare	verb	1	pang- -keun	na	na	na	na	movr	
22	RUN	yka	ubas	mag-	verb	1	-an	2	THM	mnr	run taking (s.t./s.o.)	movr, goal=PP	co-movr, thm
23	CLIMB	blz	lopon	-um-	verb	1	-i	2	PATH	mnr	climb (path)	movr	movr, path
23	CLIMB	blz	lopon	-um-	verb	1	-kon	2	THM	mnr	climb w/ (s.t./s.o.)	movr	movr, co-thm
23	CLIMB	blz	lopon	-um-	verb	1	-ii+-i	3	BEN	mnr	climb (s.t.) for (s.o.)	movr	movr, rec, path
23	CLIMB	ind	naik	bare	verb	1	-i	2	PATH	mnr	climb on (s.t.)	movr, path=PP	movr, path
23	CLIMB	ind	naik	bare	verb	1	-kan	2	CAUS	mnr	make (s.t.) climb, raise	movr, path=PP	istg, thm
23	CLIMB	jav	penék	AV	verb	2	-aké	2	CAUS	mnr	help (s.o.) climb	movr, path	istg, thm
23	CLIMB	jav	penék	AV	verb	2	-i	2	OPT+PATH	no-chg	climb (s.t.)	movr, path	movr, path
23	CLIMB	jav	penék	AV	verb	2	-aké	3	BEN	mnr	climb (a tree) for (s.o.)	movr, path	movr, ben, path
23	CLIMB	mnb	tada	a-	verb	1	-i	2	PATH	mnr	climb/step on (s.t.)	movr, path=PP	movr, path
23	CLIMB	mnb	tada	a-	verb	1	-ghoo	na	na	na	na	movr, path=PP	
23	CLIMB	nsy	cakak	AV	verb	2	-i	2	PLUR	no-chg	climb (s.t.)	movr, path	movr, path
23	CLIMB	nsy	cakak	AV	verb	2	-kun	2	CAUS	no-chg	help (s.o.) climb (s.t.)	movr, path	istg, thm, path=PP
23	CLIMB	sun	naék	bare	verb	1	-an	2	PATH	mnr	climb (path)	movr, path=NP/PP	movr, path
23	CLIMB	sun	naék	bare	verb	1	-keun	2	CAUS	mnr	cause (s.t.) to go up	movr, path=NP/PP	istg, thm
23	CLIMB	sun	naék	bare	verb	1	-an	2	LOC	mnr	climb (path)	movr, path=NP/PP	movr, path
23	CLIMB	sun	naék	bare	verb	1	pang- -keun	na	na	na	na	movr, path=NP/PP	
23	CLIMB	yka	pana'ik	N-, mag-	verb	2	-an	2	PLUR	no-chg	climb (path) rep.	movr, path	movr, path
24	COUGH	bes	iyak	AV?	verb	2	-i/-ka	na	na	na	na	expr, path	
24	COUGH	blz	kokoyon	bare	verb	1	-i/-ii/-kon	na	na	na	na	expr	

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr
24	COUGH	ind	batuk	bare	unclear	1	-i	2	GOAL	mnr	cough on (s.o.)	expr	agt, goal
24	COUGH	ind	batuk	bare	unclear	1	-kan	2	THM	mnr	cough up (s.t.)	expr	agt, thm
24	COUGH	jav	watuk	RDP	unclear	1	-aké	2	THM	mnr	cough up (s.t.)	expr	agt, thm
24	COUGH	jav	watuk	RDP	unclear	1	-i	2	CAUS	mnr	make (s.o.) cough	expr	agt, csnd
24	COUGH	mnb	hoda	a-	verb	1	-ghoo	2	THM	mnr	cough up (s.t.)	expr	agt, thm
24	COUGH	mnb	hoda	a-	verb	1	-i	na	na	na	na	expr	agt, thm
24	COUGH	nsy	hiyuk	bare	verb	1	-kun	2	THM	mnr	cough up (s.t.)	expr	agt, thm
24	COUGH	nsy	hiyuk	bare	verb	1	-i	na	na	na	na	expr	agt, thm
24	COUGH	sas	batok	bare?	verb	1	-an	2	THM	mnr	cough (s.t.) up	expr	agt, thm
24	COUGH	sun	batuk	bare	verb	1	-keun	2	THM	mnr	cough up (s.t.)	expr	agt, thm
24	COUGH	sun	batuk	bare	verb	1	-an/pang-	na	na	na	na	expr	agt, thm
24	COUGH	yka	hiket	mag-	noun	1	-an	2	THM	mnr	cough up (s.t.)	expr	agt, thm
25	BLINK	bes	kerjap	unclear	unclear	na	-i	na	na	na	na	movr	agt, thm
25	BLINK	bes	kerjap	unclear	unclear	na	-ka	2	THM	mnr	blink (eyes)	movr	agt, thm
25	BLINK	blz	kudap	-um-	unclear	1	-kon	2	OPT+THM	no-chg	blink (eyes)	movr	agt, thm
25	BLINK	blz	kudap	-um-	unclear	1	-i/-ii	na	na	na	na	movr	agt, thm
25	BLINK	ind	kedip	MID	unclear	1	-kan	2	THM	mnr	wink (eyes)	movr/thm	agt, thm
25	BLINK	ind	kedip	MID	unclear	1	-i	na	na	na	na	movr/thm	agt, thm
25	BLINK	jav	kedhèp	AV	unclear	1	-i	2	ADDR	mnr	wink/blink at (s.o.) as a sign	movr/thm	agt, addr
25	BLINK	jav	kedhèp	AV	unclear	1	-aké	na	na	na	na	movr/thm	agt, addr
25	BLINK	mnb	mpidu-mpidu	ae-	verb	1	-ghoo	2	THM	mnr	blink (eyes)	movr	agt, thm
25	BLINK	mnb	mpidu-mpidu	ae-	verb	1	-i	na	na	na	na	movr	agt, thm
25	BLINK	nsy	khijap	AV	verb	1	-kun	2	THM	mnr	blink (eyes)	thm	agt, thm
25	BLINK	nsy	khijap	AV	verb	1	-i	na	na	na	na	thm	agt, thm
25	BLINK	sas	kejit	bare?	verb	1	-an	2	ADDR	mnr	wink at (s.o.)	movr	agt, addr
25	BLINK	sun	kiceup	AV	verb	1	-an	2	GOAL	mnr	wink at (s.o.)	movr	agt, addr
25	BLINK	sun	kiceup	AV	verb	1	-keun/pang-	na	na	na	na	movr	agt, addr
25	BLINK	yka	keddem	mag-	verb	1	-an	2	ADDR	mnr	blink at (s.o.) as a sign	movr	agt, addr
26	SHAVE	bes	cukur	AV?	verb	2	-i	2	PLUR	no-chg	shave (many)	agt, pat	agt, pat
26	SHAVE	bes	cukur	AV?	verb	2	-ka	2	LEX	no-chg	shave (hair)	agt, pat	agt, pat
26	SHAVE	blz	kuur	no-base	no-base	na	-i	2	OBLIG+PAT	na	shave (hair)	agt, pat	agt, pat
26	SHAVE	blz	kuur	no-base	no-base	na	-ii/-kon	na	na	na	na	agt, pat	agt, inst
26	SHAVE	ind	cukur	AV	verb	2	-kan	2	INST	no-chg	shave w/ (s.t.)	agt, pat	agt, inst
26	SHAVE	ind	cukur	AV	verb	2	-i	na	na	na	na	agt, pat	agt, inst
26	SHAVE	jav	cukur	AV	verb	2	-aké	2	CAUS	caus-remp	have (s.o.)'s hair cut	agt, pat	istg, pat
26	SHAVE	jav	cukur	AV	verb	2	-i	2	OPT/PLUR	no-chg	shave/cut (hair) (hab.)	agt, pat	agt, pat
26	SHAVE	mnb	kuru	ae-	verb	2	-i	2	PLUR	no-chg	shave, cut (many)	agt, pat	agt, pat
26	SHAVE	mnb	kuru	ae-	verb	2	-ghoo	na	na	na	na	agt, pat	agt, pat
26	SHAVE	mnb	kuru	ae-	verb	2	-i	na	na	na	na	agt, pat	agt, pat
26	SHAVE	nsy	cukukh	MID	verb	1	-i	2	PAT	mnr	shave (s.t.)	agt	agt, pat
26	SHAVE	nsy	cukukh	MID	verb	1	-kun	2	INST	mnr	shave w/ (s.t.)	agt	agt, inst
26	SHAVE	sas	kuris	MID	verb	1	-an	3	BEN	ditr	shave one's head for (s.o.)	agt	agt, ben, pat
26	SHAVE	sun	cukur	AV	verb	2	-an	2	PLUR	no-chg	cut/shave (many)	agt, pat	agt, pat
26	SHAVE	sun	cukur	AV	verb	2	-keun	na	na	na	na	agt, pat	agt, pat
26	SHAVE	sun	cukur	AV	verb	2	pang-	2	BEN	ditr	cut (s.o.'s) hair for (s.o.)	agt, pat	agt, ben, pat
26	SHAVE	yka	urut	N-, mag-	verb	2	-an	2	PAT	mnr	shave (hair of head)	agt, pat	agt, pat
28	WASH	bes	basuh	AV	verb	2	-ka	3	BEN	ditr	wash (s.t.) for (s.o.)	agt, pat	agt, ben, pat
28	WASH	bes	basuh	AV	verb	2	-i	2	PLUR	no-chg	wash (s.t.)	agt, pat	agt, pat
28	WASH	blz	oso'	ming-	verb	1	-i	2	PAT	mnr	wash (s.t.)	agt	agt, pat
28	WASH	blz	oso'	ming-	verb	1	-ii/-kon	na	na	na	na	agt	agt, pat
28	WASH	ind	cuci	AV	verb	2	-kan	2	CAUS	remp	have (s.t.) washed by s.o.	agt, pat	istg, pat, csnd=PP
28	WASH	ind	cuci	AV	verb	2	-kan	2	BEN	no-chg	wash (s.t.) (of s.o.)	agt, pat	agt, pat, ben=poss
28	WASH	ind	cuci	AV	verb	2	-i	na	na	na	na	agt, pat	agt, pat
28	WASH	jav	kumbah	AV	verb	2	-aké	2	CAUS	remp	have (s.t.) washed by s.o.	agt, pat	istg, pat, csnd=PP
28	WASH	jav	kumbah	AV	verb	2	-i	na	na	na	na	agt, pat	agt, pat
28	WASH	mnb	ghome	ae-	verb	2	-i	2	PLUR	no-chg	wash (many)	agt, pat	agt, pat
28	WASH	mnb	ghome	ae-	verb	2	-ghoo	na	na	na	na	agt, pat	agt, pat
28	WASH	nsy	basuh	AV	verb	2	-i	2	OPT/PLUR	no-chg	wash (many)	agt, pat	agt, pat
28	WASH	nsy	basuh	AV	verb	2	-kun	3	BEN	ditr	wash (s.t.) for (s.o.)	agt, pat	agt, ben, pat
28	WASH	sas	bisoq	bare?	verb	2	-an	3	BEN	ditr	wash (s.t.) for (s.o.)	agt, pat	agt, ben, pat
28	WASH	sun	kumbah	AV	verb	2	-an	2	PLUR	no-chg	wash (many)	agt, pat	agt, pat
28	WASH	sun	kumbah	AV	verb	2	-keun	na	na	na	na	agt, pat	agt, pat
28	WASH	sun	kumbah	AV	verb	2	pang-	2	BEN	no-chg	wash (s.t.) of s.o. for them	agt, pat	agt, pat, ben=poss
28	WASH	yka	dekdak	N-	verb	2	-an	2	OPT+PAT	no-chg	wash (clothes), do laundry	agt, pat	agt, pat

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr	
29	EAT	bes	makan	AV	verb	2	-i	2	PLUR	no-chg	eat (s.t.) rep.	cnstr, cnstd	cnstr, cnstd	
29	EAT	bes	makan	AV	verb	2	-ka	2	CAUS+PAT	caus-rempp	feed (s.t.) to s.o.	cnstr, cnstd	istg, cnstd, cnstr=PP	
29	EAT	blz	kaan	AV	verb	2	-i+pa-	2	CAUS	ditr	feed (s.o.) w/ (s.t.)	cnstr, cnstd	istg, cnstr, cnstd	
29	EAT	blz	kaan	AV	verb	2	-ii/-kon	na	na	na	na	cnstr, cnstd	cnstr, cnstd	
29	EAT	ind	makan	bare	verb	2	-i	2	PLUR	no-chg	eat rep., eat (much)	cnstr, cnstd	cnstr, cnstd	
29	EAT	ind	makan	bare	verb	2	-i	2	CAUS	mnr	give feed to (animals)	cnstr, cnstd	istg, cnstr	
29	EAT	ind	makan	bare	verb	2	-kan	2	CAUS	caus-rempp	give/allow (s.t.) to be eaten	cnstr, cnstd	istg, cnstd, cnstr=PP	
29	EAT	jav	pangan	AV	unclear	2	-i	2	PLUR	no-chg	eat constantly	agt, pat	cnstr, cnstd	
29	EAT	jav	pangan	AV	unclear	2	-aké	na	na	na	na	agt, pat	agt, pat	
29	EAT	mnb	fumaa	a-	verb	2	-ghoo	2	COMIT+LEX	rempp	eat (rice) w/ (s.t.)	cnstr, cnstd	cnstr, com	
29	EAT	mnb	fumaa	a-	verb	2	-i	na	na	na	na	cnstr, cnstd	cnstr, cnstd	
29	EAT	nsy	kanik	AV	verb	2	-i	2	PLUR	no-chg	eat (many) / to eat (rep.)	cnstr, cnstd	cnstr, cnstd	
29	EAT	nsy	kanik	AV	verb	2	-kun	2	other-imp	no-chg	eat	cnstr, cnstd	cnstr, cnstd	
29	EAT	sas	kaken	bare	verb	2	-an	na	na	na	na	cnstr, cnstd	cnstr, cnstd	
29	EAT	sun	dahar	bare	verb	2	-an	2	PLUR	no-chg	eat (s.t.) rep., hab.	cnstr, cnstd	cnstr, cnstd	
29	EAT	sun	dahar	bare	verb	2	-keun/pang-	-keun	na	na	na	cnstr, cnstd	cnstr, cnstd	
29	EAT	yka	mangan	bare	verb	2	-an	2	PLUR	no-chg	eat (many, of many, rep.)	cnstr, cnstr	cnstr, cnstd	
30	HELP	bes	tulung	AV	verb	2	-ka	2	other-actv	ditr	help (s.o.) do (s.t.)	agt, ben, actv=cls	agt, ben, actv	
30	HELP	bes	tulung	AV	verb	2	-i	2	PLUR	no-chg	help (s.o.) (rep.)	agt, ben, actv=cls	agt, ben	
30	HELP	blz	tulung	no-base	no-base	na	-i	2	OBLIG	na	help (s.o.) do s.t.	agt, ben, actv=cls	agt, ben, actv=cls	
30	HELP	blz	tulung	no-base	no-base	na	-ii/-kon	na	na	na	na	agt, ben	istg, cnstd	
30	HELP	ind	bantu	AV	verb	2	-kan	2	CAUS	no-chg	bring (s.o.) in to help.	agt, ben	istg, cnstd	
30	HELP	ind	bantu	AV	verb	2	-i	na	na	na	na	agt, ben	istg, cnstd	
30	HELP	jav	bantu	AV	verb	2	-aké	2	CAUS	no-chg	have (s.o.) help w/ s.o.	agt, ben, purp=cls	istg, cnstd, ben=PP	
30	HELP	jav	bantu	AV	verb	2	-i	2	OPT	no-chg	help (s.o.) to help w/ actv	agt, ben, purp=cls	agt, ben, actv=cls	
30	HELP	mnb	tulumi	ae-	noun	2	-i/-ghoo	na	na	na	na	agt, ben	istg, cnstd	
30	HELP	nsy	tulung	AV	verb	2	-i	2	OPT	no-chg	help (s.o.)	agt, ben, actv=cls	agt, ben	
30	HELP	nsy	tulung	AV	verb	2	-kun	3	BEN	ditr	help (s.o.) for (s.o.)	agt, ben, actv=cls	agt, ben, ben	
30	HELP	sas	tulong	bare	verb	2	-an	3	BEN	ditr	help (s.o.) for (s.o.)'s favor	agt, ben	agt, ben, ben	
30	HELP	sun	bantos	AV	verb	2	-an	2	PLUR	no-chg	help (s.o.) rep.	agt, ben	agt, ben	
30	HELP	sun	bantos	AV	verb	2	-keun/pang-	-keun	na	na	na	agt, ben	istg, cnstd	
30	HELP	yka	tabang	N-, mag-	noun	2	-an	2	LEX	no-chg	help (s.o.) in actv	agt, ben, reas/purp=cls	agt, ben+com, actv=cls	
31	FOLLOW	blz	lolo'	mongo-	verb	2	-kon	2	LEX	no-chg	comply w/, grant (a wish)	co-movr, actv/co-movr	agt, targ, emtr=poss	
31	FOLLOW	blz	lolo'	mongo-	verb	2	-i/-ii	na	na	na	na	co-movr, actv/co-movr	co-movr, actv, co-movr=PP	
31	FOLLOW	ind	ikut	bare	verb	2	-i	2	LEX	mnr	follow (s.t.), to obey (teaching)	co-movr, actv, co-movr=PP	agt, targ	
31	FOLLOW	ind	ikut	bare	verb	2	-i	2	OPT	no-chg	participate in (an event)	co-movr, actv, co-movr=PP	co-agt, actv	
31	FOLLOW	ind	ikut	bare	verb	2	-kan	2	CAUS	rempp	include (s.o.) in an event	co-movr, actv, co-movr=PP	istg, cnstd, actv=PP	
31	FOLLOW	jav	èlu	AV	verb	2	-i	2	LEX	no-chg	accompany (s.o.)	co-movr, actv, goal=PP	co-movr, co-movr	
31	FOLLOW	jav	èlu	AV	verb	2	-aké	2	CAUS+LEX	ditr	have (s.o.) accompany/live w/ (s.o.)	co-movr, actv, goal=PP	istg, co-movr, co-movr	
31	FOLLOW	mnb	sonso	unclear	verb	2	-i	2	LEX	no-chg	follow and bring back (s.t.)	movr, path	movr, thm	
31	FOLLOW	mnb	sonso	unclear	verb	2	-ghoo	na	na	na	na	movr, path	co-movr, co-movr	
31	FOLLOW	nsy	ikhing	AV	verb	2	-i	2	OPT+LEX	no-chg	follow/accompany (s.o.)	co-movr, co-movr	co-movr, co-movr, actv=cls	
31	FOLLOW	nsy	ikhing	AV	verb	2	-kun	2	OPT+LEX	no-chg	accompany (s.o.) to do s.t.	co-movr, co-movr	co-movr, co-movr, actv=cls	
31	FOLLOW	sas	turut	bare	verb	2	-an	2	LEX	no-chg	follow (targ)	co-movr, targ=PP	movr, targ	
31	FOLLOW	sun	iring	AV	verb	1	-an	2	other-actv+LEX	mnr	join, sign up for (actv)	co-movr, goal/actv=PP	agt, actv	
31	FOLLOW	sun	iring	AV	verb	1	-an	2	other+LEX	mnr	accompany (s.t.)	co-movr, goal/actv=PP	co-movr, co-movr	
31	FOLLOW	sun	iring	AV	verb	1	-keun	2	other	mnr	accompany (s.o.)	co-movr, goal/actv=PP	co-thm, co-thm	
31	FOLLOW	sun	iring	AV	verb	1	pang-	-keun	na	na	na	co-movr, goal/actv=PP	co-movr, goal/actv=PP	
31	FOLLOW	yka	turul	N-, mag-	verb	2	-an	2	CAUS+THM	rempp	follow (s.o.) (to deliver s.t.)	co-movr, co-movr	movr, co-thm	
32	MEET	bes	temu	MID	verb	1	-i	2	THM+LEX	mnr	meet (s.o.)	co-thm, co-thm=PP	co-thm, co-thm	
32	MEET	bes	temu	MID	verb	1	-ka	2	THM+LEX	mnr	discover (s.t.)	co-thm, co-thm=PP	co-thm, co-thm	
32	MEET	blz	tuung	poo-	verb	1	-i	2	THM+LEX	mnr	meet w/ (s.o.), discover (s.t.)	co-thm	co-thm, co-thm	
32	MEET	blz	tuung	poo-	verb	1	-ii/-kon	na	na	na	na	co-thm	co-thm	
32	MEET	ind	temu	MID	verb	1	-i	2	THM	mnr	meet w/ (s.o.), encounter (s.t.)	co-thm, co-thm=PP	co-thm, co-thm	
32	MEET	ind	temu	MID	verb	1	-kan	2	THM	mnr	find, discover (s.t.)	co-thm, co-thm=PP	obsr, stim	
32	MEET	jav	temu	AV	verb	2	-i	2	GOAL	mnr	look for, go see (s.o.)	co-thm, co-thm	agt, thm	
32	MEET	jav	temu	AV	verb	2	-aké	2	CAUS+LEX	mnr	bring together (couple)	co-thm, co-thm	agt, co-thm	
32	MEET	jav	temu	AV	verb	2	-aké	2	THM	mnr	find, discover (s.t.)	co-thm, co-thm	agt, thm	
32	MEET	nsy	temu	MID	verb	1	-i	2	GOAL	mnr	meet (s.o.)	co-thm, co-thm=PP	co-thm, co-thm	
32	MEET	nsy	temu	MID	verb	1	-kun	2	CAUS	no-chg	introduce (s.o.)	co-thm, co-thm=PP	agt, co-thm	
32	MEET	sas	dait	MID	verb	1	-an+pe-	2	CAUS	mnr	make (people) meet	co-thm	istg, co-thm	
32	MEET	sun	pendak	AV	verb	2	-an	2	PLUR	no-chg	meet, run into (s.o.) often	co-thm, co-thm	co-thm, co-thm	
32	MEET	sun	pendak	AV	verb	2	-keun	na	na	na	na	co-thm, co-thm	co-thm, co-thm	
32	MEET	sun	pendak	AV	verb	2	pang-	-keun	3	BEN	ditr	find (s.t.) for (s.o.)	co-thm, co-thm	obsr, ben, stim

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr
32	MEET	yka	säbu	N-, mag-	verb	2	-an	na	na	na	na	co-thm, co-thm	
33	HUG	bes	kepit	AV?	verb	2	-ka	3	BEN	ditr	embrace (s.o.) for (s.o.)	movr, goal	movr, ben, goal
33	HUG	bes	kepit	AV?	verb	2	-i	2	PLUR	no-chg	embrace (many)	movr, goal	movr, goal
33	HUG	blz	rangkot	AV	verb	2	-i/-ii/-kon	na	na	na	na	movr, goal	
33	HUG	ind	peluk	AV	verb	2	-i	2	PLUR	no-chg	hug (s.o.) (rep.)	movr, goal	movr, goal
33	HUG	ind	peluk	AV	verb	2	-kan	2	CAUS	remp	hug (arm) around s.o.	movr, goal	movr/agt, thm, goal=PP
33	HUG	jav	kekep	AV	verb	2	-i	2	OPT+GOAL	no change	embrace (s.o.)	movr, goal	movr, goal
33	HUG	jav	kekep	AV	verb	2	-aké	na	na	na	na	movr, goal	
33	HUG	mnb	kakopu	a-	verb	1	-ghoo	2	GOAL	mnttr	embrace (s.t.)	co-movr	agt, goal
33	HUG	mnb	kakopu	a-	verb	1	-i	na	na	na	na	co-movr	
33	HUG	nsy	khakuk	AV	verb	2	-i	2	OPT+LEX	no-chg	hug (s.o.)	movr, goal	movr, goal
33	HUG	nsy	khakuk	AV	verb	2	-kun	2	OPT+LEX	no-chg	hug (s.t.)	movr, goal	movr, goal
33	HUG	sas	kapông	bare	verb	2	-an	3	BEN	ditr	hug (s.t.) for (s.o.)	movr, goal	movr, ben, goal
33	HUG	sun	tangkeup	AV	verb	2	-an	2	PLUR	no-chg	keep hugging, often hug (s.o.)	movr, goal	movr, goal
33	HUG	sun	tangkeup	AV	verb	2	-keun	2	CAUS	remp	hug (arm) around s.o.	movr, goal	agt, thm, goal=PP
33	HUG	sun	tangkeup	AV	verb	2	pang- -keun	na	na	na	na	movr, goal	
33	HUG	yka	kekkep	AV, mag-	verb	2	-an	2	OBLIG-PV	no-chg	To hug, embrace, clasp (s.t.)	agt, goal	agt, goal
34	SEARCH FOR	bes	cakagh	AV	verb	2	-ka	3	BEN	ditr	search for (s.t.) for (s.o.)	agt, thm	agt, thm, thm
34	SEARCH FOR	bes	cakagh	AV	verb	2	-i	2	PLUR	no-chg	search for (s.t.)	agt, thm	agt, thm
34	SEARCH FOR	blz	sarak	AV	verb	2	-ii	3	BEN	ditr	search for (s.t.) for s.o.	agt, thm	agt, rec, thm
34	SEARCH FOR	blz	sarak	AV	verb	2	-kon	2	BEN	no-chg	look for (s.t.) for s.o.	agt, thm	agt, thm, ben=PP
34	SEARCH FOR	blz	sarak	AV	verb	2	-i	na	na	na	na	agt, thm	
34	SEARCH FOR	blz	sarak	AV	verb	2	-i	na	na	na	na	agt, thm	
34	SEARCH FOR	ind	cari	AV	verb	2	-kan	2	OPT+THM	no-chg	look for, get (s.t.) for purp	agt, thm	agt, thm, purp=PP
34	SEARCH FOR	ind	cari	AV	verb	2	-kan	3	BEN	ditr	look for (s.o.) (s.t.)	agt, thm	agt, rec, thm
34	SEARCH FOR	ind	cari	AV	verb	2	-i	na	na	na	na	agt, thm	
34	SEARCH FOR	jav	golèk	bare	verb	2	-aké	3	BEN	ditr	get, seek (s.t.) for (s.o.)	agt, thm	agt, ben, thm
34	SEARCH FOR	jav	golèk	bare	verb	2	-i	2	OPT+THM+LEX	no-chg	look for, find (s.t.)	agt, thm	agt, thm
34	SEARCH FOR	mnb	ghondo	a-	verb	2	-ghoo	2	LOC+LEX	remp	delouse (s.o.)	agt, thm	agt, loc
34	SEARCH FOR	mnb	ghondo	a-	verb	2	-i	2	OPT+THM	no-chg	look for, search for (s.t.)	agt, thm	agt, thm
34	SEARCH FOR	nsy	kilum	AV	verb	2	-i	2	OPT+THM	no-chg	search for (s.t.)	agt, thm	agt, thm
34	SEARCH FOR	nsy	kilum	AV	verb	2	-kun	3	BEN	ditr	search for (s.t.) for (s.o.)	agt, thm	agt, ben, thm
34	SEARCH FOR	sas	pète	bare	verb	2	-an	3	BEN	ditr	look for (s.o.) (s.t.)	agt, thm	agt, ben, thm
34	SEARCH FOR	sun	pilari	AV	verb	2	-an	2	PLUR	no-chg	search for (s.t.) rep., dur.	agt, thm	agt, thm
34	SEARCH FOR	sun	pilari	AV	verb	2	-keun	na	na	na	na	agt, thm	
34	SEARCH FOR	sun	pilari	AV	verb	2	pang- -keun	3	BEN	ditr	search for (s.t.) for (s.o.)	agt, thm	agt, rec, thm
34	SEARCH FOR	yka	piha	N-, mag-	verb	2	-an	na	na	na	na	agt, thm	
35	THINK	bes	rupuk	MID	verb	1	-i	2	CONT+LEX	mnttr	think about (s.o.)	cgnr, cont=PP	agt, cont
35	THINK	bes	rupuk	MID	verb	1	-ka	na	na	na	na	cgnr, cont=PP	
35	THINK	blz	inau'	AV	verb	1	-i/-ii/-kon	na	na	na	na	cgnr, cont=cls	
35	THINK	ind	pikir	MID	verb	1	-i	2	CONT	mnttr	think about (s.t.)	cgnr, cont=PP	cgnr, cont
35	THINK	ind	pikir	MID	verb	1	-kan	2	CONT	mnttr	think/be concerned about (s.t.)	cgnr, cont=PP	cgnr, cont
35	THINK	jav	pikir	AV	verb	2	-i	2	LEX	mnttr	think/be concerned about (s.t.)	cgnr, cont	cgnr, cont
35	THINK	jav	pikir	AV	verb	2	-aké	2	LEX	mnttr	give thought to (s.t.)	cgnr, cont	cgnr, cont
35	THINK	sas	pikir	bare	verb	1	-an	2	CONT	mnttr	think/be concerned about (s.t.)	cgnr, cont=cls	cgnr, cont
35	THINK	sun	pikir	AV	verb	1	-an	2	CONT+PLUR	mnttr	think about (s.t.), esp. rep.	cgnr	cgnr, cont
35	THINK	sun	pikir	AV	verb	1	-keun	2	CONT+BEN	mnttr	think about (s.t.), esp. for s.o.	cgnr	cgnr, cont, ben=PP
35	THINK	sun	pikir	AV	verb	1	pang- -keun	3	BEN	ditr	join in thinking about (s.t.) for (s.o.)	cgnr	cgnr, ben, cont
35	THINK	yka	pikil	N-, mag-	verb	2	-an	na	na	na	na	cgnr, cont	
36	KNOW	bes	kenal	bare?	verb	1	-i	2	CONT+LEX	mnttr	recognize, identify (s.o.)	cgnr, cont=PP	cgnr, cont
36	KNOW	bes	kenal	bare?	verb	1	-ka	2	CAUS	mnttr	introduce (s.o.) to s.o.	cgnr, cont=PP	istg, cgnr, cont=PP
36	KNOW	blz	kanaal	AV	verb	2	-i/-ii/-kon	na	na	na	na	cgnr, cont	
36	KNOW	ind	kenal	bare	verb	1	-i	2	LEX	mnttr	recognize, identify (s.o.)	cgnr, targ=PP	cgnr, cont
36	KNOW	ind	kenal	bare	verb	1	-kan+per-	2	CAUS	mnttr	introduce (s.o.) to s.o.	cgnr, targ=PP	cgnr, targ=PP
36	KNOW	jav	kenal	bare/di-	verb	2	-aké	2	CAUS	caus-remp	introduce (s.o.) to s.o.	cgnr, cont	agt, cont, cgnr=PP
36	KNOW	jav	kenal	bare/di-	verb	2	-i	2	LEX	no-chg	get acquainted w/ (s.o.)	cgnr, cont	cgnr, cont
36	KNOW	mnb	pandehao	a-	verb	2	-i/-ghoo	na	na	na	na	cgnr, cont	
36	KNOW	nsy	cinung	bare	verb	1	-i	2	LEX	mnttr	recognize (s.o.)	cgnr, cont=PP	cgnr, cont
36	KNOW	nsy	cinung	bare	verb	1	-kun	2	CAUS	mnttr	introduce (s.o.) to s.o.	cgnr, cont=PP	istg, cgnr, cont=PP
36	KNOW	sas	kenal	na	verb	1	-an	2	CAUS	mnttr	introduce (s.o.) to s.o.	cgnr, cont=PP	istg, cgnr, cont=PP
36	KNOW	sun	wawuh	bare	verb	1	-an	2	PLUR	no-chg	get to know (s.o.)	cgnr, targ=PP	cgnr, cont
36	KNOW	sun	wawuh	bare	verb	1	-keun	2	CAUS	mnttr	introduce (s.o./s.t.) (to s.o.)	cgnr, targ=PP	agt, cont, cgnr=PP
36	KNOW	sun	wawuh	bare	verb	1	pang- -keun	na	na	na	na	cgnr, targ=PP	
37	LIKE/LOVE	bes	lemak	bare	verb	1	-i	na	na	na	na	emtr, targ=PP	

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr
37	LIKE/LOVE	bes	lemak	bare	verb	1	-ka	2	CAUS	mntnr	make (s.o.) happy	emtr, targ=PP	istg, emtr
37	LIKE/LOVE	blz	kira'	ming-	verb	1	-i+RDP	2	INTENS	na	like, desire, covet (s.t.)	emtr, targ=cls	emtr, targ
37	LIKE/LOVE	ind	suka	bare	verb	2	-i	2	OPT+TARG	no-chg	like (s.t.)	emtr, targ	emtr, targ
37	LIKE/LOVE	ind	suka	bare	verb	2	-kan	2	CAUS	mntnr	make (s.o.) happy	emtr, targ	istg, emtr
37	LIKE/LOVE	jav	seneng	bare	unclear	1	-aké	2	CAUS	mntnr	please, amuse (s.o.)	emtr, targ=PP	agt, emtr
37	LIKE/LOVE	jav	seneng	bare	unclear	1	-i	2	TARG	mntnr	like (s.t.)	emtr, targ=PP	emtr, targ
37	LIKE/LOVE	mnb	asi	a-, ae-	verb	1	-ghoo	2	TARG	mntnr	like, to be fond of (s.t.)	emtr, targ=IO	emtr, targ
37	LIKE/LOVE	mnb	asi	a-, ae-	verb	1	-i	na	na	na	na	emtr, targ=IO	
37	LIKE/LOVE	nsy	hawak	bare	verb	1	-i	2	LEX	no-chg	toy w/, trick (s.o.)	emtr, actv/targ=VP	agt, pat
37	LIKE/LOVE	nsy	hawak	bare	verb	1	-kun	2	CAUS	caus-rempp	increase desire for (s.t.)	emtr, actv/targ=VP	istg, actv/targ=cls
37	LIKE/LOVE	sun	resep	bare	verb	1	-an/-keun/pang- -keun	na	na	na	na	emtr, targ=PP	
37	LIKE/LOVE	yka	lasa	ma-	verb	1	-an	2	TARG	mntnr	love (s.o./s.t.)	emtr, targ=IO	emtr, targ
38	FEAR	bes	takut	bare	verb	1	-i	2	CAUS	mntnr	frighten (s.o.)	emtr, targ=PP	istg, emtr
38	FEAR	bes	takut	bare	verb	1	-ka	2	CAUS	caus-detr	be frightening	emtr, targ=PP	istg
38	FEAR	ind	takut	bare	verb	1	-i	2	CAUS	mntnr	scare (s.o.)	emtr, targ=PP	agt, emtr
38	FEAR	ind	takut	bare	verb	1	-kan	2	TARG	mntnr	be afraid of (s.t.)	emtr, targ=PP	emtr, targ
38	FEAR	ind	takut	bare	verb	1	-kan	2	CAUS	caus-detr	be frightening	emtr, targ=PP	istg
38	FEAR	mnb	pipi	ao-	verb	1	-i/-ghoo	na	na	na	na	emtr, targ=PP	
38	FEAR	nsy	khabai	bare	verb	1	-kun	2	CAUS	mntnr	scare (s.o.)	emtr, targ=VP	istg, emtr
38	FEAR	nsy	khabai	bare	verb	1	-i	na	na	na	na	emtr, targ=VP	
38	FEAR	sas	takut	bare	verb	1	-an	2	TARG	mntnr	be afraid of (s.t.)	emtr	emtr, targ
38	FEAR	sun	sieun	bare	verb	1	-an+RDP	2	CAUS	mntnr	scare (s.o.)	emtr, targ=PP	agt, emtr
38	FEAR	sun	sieun	bare	verb	1	-keun/pang- -keun	na	na	na	na	emtr, targ=PP	
38	FEAR	yka	talew	bare	verb	1	-an	na	na	na	na	emtr, targ=PP	
40	SMELL	blz	ook	AV	verb	2	-i/-ii/-kon	na	na	na	na	prcr, stim	
40	SMELL	ind	cium	AV	verb	2	-i	2	PLUR	no-chg	smell (s.t.) rep.	prcr, stim	prcr, stim
40	SMELL	ind	cium	AV	verb	2	-kan	2	CAUS	caus-rempp	make (s.t.) smelled by s.o.	prcr, stim	agt, stim, prcr=PP
40	SMELL	jav	ambu	N-	noun	2	-aké	3	CAUS	ditr	cause (s.o.) to smell (s.t.)	prcr, stim	istg, prcr, stim
40	SMELL	jav	ambu	N-	noun	2	-i	3	CAUS+REC	ditr	offer (s.t.) to (s.o.) to smell	prcr, stim	istg, prcr, stim
40	SMELL	mnb	wono	ae-	noun	2	-i/-ghoo	na	na	na	na	prcr, stim	
40	SMELL	nsy	sium	AV	verb	2	-i	2	PLUR+LEX	no-chg	follow (s.t.), obey	prcr, stim	obsr, stim
40	SMELL	nsy	sium	AV	verb	2	-kun	3	BEN	ditr	smell (s.t.) for (s.o.)	prcr, stim	obsr, ben, stim
40	SMELL	sas	ambuq	bare?	verb	2	-an	2	BEN	ditr	smell (s.t.) for (s.o.)	prcr, stim	obsr, ben, stim
40	SMELL	sun	ambeu	AV	verb	2	-an	2	PLUR	no-chg	rep. smell (s.t.)	prcr, stim	obsr, stim
40	SMELL	sun	ambeu	AV	verb	2	-keun	2	CAUS	caus-rempp	make (s.t.) smelled by s.o.	prcr, stim	istg, stim, prcr=PP
40	SMELL	sun	ambeu	AV	verb	2	pang- -keun	2	BEN	ditr	smell for (s.o.) (s.t.)	prcr, stim	prcr, ben, stim
40	SMELL	yka	ük	bare	verb	2	-an	na	na	na	na	prcr, stim	
42	SEE	bes	kinak	AV	verb	2	-i	2	INTENS	no-chg	go and see (s.o.)	prcr, stim	obsr, stim
42	SEE	bes	kinak	AV	verb	2	-ka	2	CAUS	no-chg	show (s.t.) to s.o.	prcr, stim	obsr, stim
42	SEE	blz	piile	mi-	verb	2	-i/-ii/-kon	na	na	na	na	prcr, stim	
42	SEE	ind	lihat	AV	verb	2	-i	2	INTENS	no-chg	scrutinise (s.t.)	prcr, stim	obsr, stim
42	SEE	ind	lihat	AV	verb	2	-kan	2	INTENS	no-chg	look at, view, stare at (s.t.)	prcr, stim	obsr, stim
42	SEE	jav	deleng	N-	verb	2	-aké	2	INTENS	no-chg	see, observe (s.t.)	prcr, stim	obsr, stim
42	SEE	jav	pirsa	N-	verb	2	-aké	3	CAUS	ditr	show (s.o.) (s.t.)	prcr, stim	agt, prcr, stim
42	SEE	jav	pirsa	N-	verb	2	-i	2	CAUS	no-chg	look at, watch (s.t.)	prcr, stim	obsr, stim
42	SEE	mnb	wora	a-, ae-	verb	2	-ghoo	2	INST+LEX	rempp	see/witness (s.t.)	prcr, stim	prcr, inst, stim=cls
42	SEE	mnb	wora	a-, ae-	verb	2	-i	na	na	na	na	prcr, stim	
42	SEE	nsy	liyak	AV	verb	2	-i	2	INTENS	no-chg	look intently, on purp at (s.o.)	prcr, stim	obsr, stim
42	SEE	nsy	liyak	AV	verb	2	-kun	na	na	na	na	prcr, stim	
42	SEE	sas	gitaq	bare	verb	2	-an	3	BEN	ditr	look at (s.t.) for (s.o.)	prcr, stim	obsr, ben, stim
42	SEE	sas	gitaq	bare	verb	2	-an	3	CAUS	ditr	show (s.o.) (s.t.)	prcr, stim	istg, prcr, stim
42	SEE	sun	tinggali	AV	verb	2	-an	2	PLUR	no-chg	look at (s.t.) carefully	prcr, stim	prcr, stim
42	SEE	sun	tinggali	AV	verb	2	-keun	2	CAUS	caus-rempp	show (s.t.) to s.o.	prcr, stim	agt, stim, prcr=PP
42	SEE	sun	tinggali	AV	verb	2	-keun	2	INTENS	no-chg	look at (s.t.)	prcr, stim	obsr, stim
42	SEE	sun	tinggali	AV	verb	2	pang- -keun	3	BEN	ditr	look at (s.t.) for (s.o.)	prcr, stim	obsr, ben, stim
42	SEE	yka	kite	bare, AV	verb	2	-an	na	na	na	na	prcr, stim	
43	TALK	bes	kicik	AV	verb	1	-i	2	ADDR	mntnr	talk to (s.o.)	comr	comr, addr
43	TALK	bes	kicik	AV	verb	1	-ka	2	CONT	mntnr	talk about (s.t.)	comr	comr, cont
43	TALK	blz	bisara	ba-	verb	1	-kon	2	CONT	mntnr	speak about (s.t.)	comr, addr=PP	comr, cont
43	TALK	blz	bisara	ba-	verb	1	-i/-ii	na	na	na	na	comr, addr=PP	
43	TALK	ind	bicara	MID	verb	1	-kan	2	CONT	mntnr	talk about (s.t.)	comr, cont=PP	comr, cont
43	TALK	ind	bicara	MID	verb	1	-i	na	na	na	na	comr, cont=PP	
43	TALK	jav	omong	bare	verb	1	-aké	2	CONT	mntnr	speak (about) (s.t.)	comr, cont=cls	comr, cont
43	TALK	jav	omong	bare	verb	1	-i	2	ADDR	ditr	tell (s.o.) (s.t.)	comr, cont=cls	comr, addr, cont

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr
43	TALK	mnb	bisara	a-	verb	1	-ghoo	2	CONT	mnr	talk about (s.t.)	comr, addr=PP, cont=cls	comr, cont
43	TALK	mnb	bisara	a-	verb	1	-i	na	na	na	na	comr, addr=PP, cont=cls	comr, cont
43	TALK	nsy	kicik	AV	verb	1	-i	2	ADDR	mnr	talk to (s.o.)	comr	comr, addr
43	TALK	nsy	kicik	AV	verb	1	-kun	2	CONT	mnr	say (s.t.)	comr	comr, cont=cls
43	TALK	sas	ra'os	AV	verb	1	-an	2	CONT	mnr	talk about (s.t.)	comr, addr=PP	comr, cont
43	TALK	sun	omong	AV	verb	1	-an	2	ADDR	mnr	advise, scold (s.o.)	comr, addr=PP, cont=cls	comr, addr, cont=cls
43	TALK	sun	omong	AV	verb	1	-keun	2	CONT	mnr	speak (about) (s.t.)	comr, addr=PP, cont=cls	comr, cont, targ=PP
43	TALK	sun	omong	AV	verb	1	pang- -keun	3	BEN	ditr	speak about (s.t.) for (s.o.)	comr, addr=PP, cont=cls	comr, ben, cont
43	TALK	yka	bissā	N-, mag-	verb	1	-an	2	CONT+LEX	mnr	talk, speak ill about (s.o.)	comr, addr=PP, cont=cls	comr, cont
44	ASK FOR	bes	pintak	AV?	verb	2	-ka	3	BEN	ditr	ask for (s.t.) for (s.o.)	comr, cont	comr, ben, cont
44	ASK FOR	bes	pintak	AV?	verb	2	-i	2	PLUR	no-chg	ask for (s.t.) rep.	comr, cont	comr, cont
44	ASK FOR	blz	ase'	AV	verb	2	-i/-ii/-kon	na	na	na	na	comr, cont	comr, cont
44	ASK FOR	ind	menta	bare	verb	2	-kan	2	BEN	no-chg	ask for (s.t.) for s.o.	comr, cont	comr, cont, ben=PP
44	ASK FOR	ind	menta	bare	verb	2	-i	3	ADDR	ditr	ask (s.o.) for (s.t.)	comr, cont	comr, addr, cont
44	ASK FOR	jav	juluk	AV	verb	2	-aké	2	BEN	no-chg	ask for (s.t.) for s.o.	comr, cont	comr, cont, ben=PP
44	ASK FOR	jav	juluk	AV	verb	2	-i	na	na	na	na	comr, cont	comr, cont
44	ASK FOR	mnb	salu	ae-	verb	2	-i	2	PLUR	no-chg	ask for (many)	comr, cont	comr, cont
44	ASK FOR	nsy	tumpang	AV	verb	2	-i	3	ADDR	ditr	request from (s.o.) (s.t.)	comr, cont	comr, addr, cont
44	ASK FOR	nsy	tumpang	AV	verb	2	-kun	2	BEN	ditr	request from (s.o.) (s.t.)	comr, cont	comr, ben, cont
44	ASK FOR	sas	èndeng	AV	verb	2	-an	3	BEN	ditr	ask for (s.t.) for (s.o.)	comr, cont	comr, ben, cont
44	ASK FOR	sun	menta	bare	verb	2	-an	2	PLUR	no-chg	ask for (s.t.) rep.	comr, cont, targ=PP	comr, cont, addr=PP
44	ASK FOR	sun	menta	bare	verb	2	-keun	na	na	na	na	comr, cont, targ=PP	comr, cont
44	ASK FOR	sun	menta	bare	verb	2	pang- -keun	3	BEN	ditr	ask for (s.t.) for (s.o.)	comr, cont, targ=PP	comr, ben, cont
44	ASK FOR	yka	pāku	N-, mag-	verb	2	-an	na	na	na	na	comr, cont	comr, cont
46	TELL	bes	cerite	MID	verb	1	-i	2	ADDR	mnr	tell (s.o.)	comr, addr=PP	comr, addr
46	TELL	bes	cerite	MID	verb	1	-ka	2	CONT	mnr	tell (s.t.)	comr, addr=PP	comr, cont
46	TELL	blz	tundun	AV	noun	2	-kon	2	OPT+CONT	no-chg	tell, narrate (s.t.)	comr, cont	comr, cont
46	TELL	blz	tundun	AV	noun	2	-ii+ -kon	2	LEX	no-chg	discuss (s.t.)	comr, cont	comr, cont
46	TELL	blz	tundun	AV	noun	2	-i	na	na	na	na	comr, cont	comr, cont
46	TELL	ind	cerita	MID	verb	1	-kan	2	CONT	mnr	tell (s.t.)	comr, cont=PP	comr, cont
46	TELL	ind	cerita	MID	verb	1	-i	2	ADDR	mnr	tell a story to (s.o.)	comr, cont=PP	comr, addr
46	TELL	jav	carita	no-base	noun	na	-aké	2	CAT+CONT	mnr	tell about (s.t.)	comr, cont	comr, cont
46	TELL	jav	carita	no-base	noun	na	-i	2	CAT+ADDR	mnr	tell s.o. (story)	comr, cont, addr=PP	comr, addr, cont=PP
46	TELL	mnb	tula	ae-	noun	1	-ghoo	2	CONT	mnr	tell about (s.t.)	comr, cont=cls	comr, cont, cont=cls
46	TELL	mnb	tula	ae-	noun	1	-ghoo	2	ADDR	mnr	tell (s.o.)	comr, cont=cls	comr, addr
46	TELL	mnb	tula	ae-	noun	1	-i	na	na	N	na	comr, cont=cls	comr, cont
46	TELL	sas	cerite	MID	noun	2	-an+ -q	2	OPT+CONT	no-chg	tell (story)	comr, cont	comr, cont
46	TELL	sun	carita	AV	noun	1	-an	2	CONT	mnr	inform, remind of s.t. to (s.o.)	comr, cont=cls	comr, addr, cont=PP
46	TELL	sun	carita	AV	noun	1	-keun	2	ADDR	mnr	tell (s.t.)	comr, cont=cls	comr, cont
46	TELL	sun	carita	AV	noun	1	pang- -keun	3	BEN	mnr	tell (story) for (s.o.)	comr, cont=cls	comr, ben, cont
46	TELL	yka	suwi-suwi	N-, mag-	noun	2	-an	2	OPT+CONT	no-chg	tell (story)	comr, cont	comr, cont
47	SAY	bes	kate	bare	verb	1	-i	na	na	na	na	comr, cont=cls	comr, cont
47	SAY	bes	kate	bare	verb	1	-ka	2	CONT	mnr	say (s.t.)	comr, cont=cls	comr, cont
47	SAY	blz	tae	bare	noun	1	-i/-ii/-kon	na	na	na	na	comr, cont=cls	comr, cont
47	SAY	ind	kata	MID	unclear	1	-i	2	ADDR+LEX	mnr	bad mouth/insult (s.o.)	comr, cont=PP, addr=PP	comr, addr, cont=PP
47	SAY	ind	kata	MID	unclear	1	-kan	2	CONT	mnr	say (s.t.)	comr, cont=PP, addr=PP	comr, cont
47	SAY	jav	kandha	bare	unclear	1	-i	2	ADDR	ditr	tell (s.o.) (s.t.)	comr, cont=cls	comr, addr, cont
47	SAY	jav	kandha	bare	unclear	1	-i	2	ADDR	mnr	give (s.o.) a talking to	comr, cont=cls	comr, addr
47	SAY	jav	kandha	bare	unclear	1	-aké	2	CONT	mnr	tell about, report (s.t.)	comr, cont=cls	comr, cont, addr=PP
47	SAY	mnb	koamba	a-	verb	1	-i/-ghoo	na	na	na	na	comr, addr=IO, cont=cls	comr, cont
47	SAY	sas	òngkat	bare	unclear	1	-an	na	na	na	na	comr, cont=cls	comr, cont
47	SAY	sun	ceuk	bare	unclear	1	-i/-keun/pang- -keun	na	na	na	na	comr, cont=cls	comr, cont
47	SAY	yka	pā'in	bare	verb	2	-an	2	ADDR+LEX	remp	teach, instruct (s.o.)	comr, cont, addr=PP	comr, addr, cont=cls
48	NAME	bes	dame	no-base	noun	na	-i	2	CAT+LEX	na	name (animal)	comr, cont	istg, refr
48	NAME	bes	dame	no-base	noun	na	-ka	2	CAT+LEX	na	name (person)	comr, cont	istg, refr
48	NAME	blz	ngaan	AV	noun	3	-i	3	CAUS	caus-remp	name (s.o.); call (s.t.) (s.t.)	comr, refr, cont	istg, refr, cont
48	NAME	blz	ngaan	AV	noun	3	-i	2	LEX	no-chg	mock, ridicule, taunt (s.o.)	comr, refr, cont	comr, refr
48	NAME	blz	ngaan	AV	noun	3	-ii/-kon	na	na	na	na	comr, refr, cont	comr, refr
48	NAME	ind	nama	MID	noun	1	-i	2	CAUS	mnr	name (s.o.) w/ a name	comr, cont=NP	agt, refr, cont=PP
48	NAME	ind	nama	MID	noun	1	-kan	2	CAUS+CONT	ditr	call (s.o.) (name)	comr, refr, cont=NP	agt, refr, cont
48	NAME	jav	aran	no-base	noun	na	-i	2	CAT+CONT	na	call, to name (s.t.) (s.t.)	comr, refr, cont	comr, refr, cont
48	NAME	mnb	nea	no-base	noun	na	-ghoo	2	CAT+CONT	na	have the name (X)	comr, cont	refr, refr
48	NAME	mnb	nea	no-base	noun	na	-i	na	na	na	na	comr, cont	refr, refr
48	NAME	sas	aran	no-base	noun	na	-an	3	CAT+CONT	na	name (s.o.) (s.t.)	comr, cont	agt, refr, cont

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr
48	NAME	sun	ngaran	no-base	noun	na	-an	3	CAT+CONT	na	name (s.o.) (s.t.)		agt, refr, cont
48	NAME	sun	ngaran	no-base	noun	na	-keun	na	na	na	na		
48	NAME	sun	ngaran	no-base	noun	na	pang-	3	CAT+BEN	na	name (s.o.) for (s.o.)		agt, ben, refr
48	NAME	yka	alen	no-base	noun	na	-an	2	CAT+CONT	na	name (s.o.), call by name		agt, refr, refr
49	MAKE	bes	buat	AV	verb	2	-ka	3	BEN	ditr	make (s.o.) (s.t.)	crtr, crtn	crtr, ben, crtn
49	MAKE	bes	buat	AV	verb	2	-i	na	na	na	na	crtr, crtn	
49	MAKE	blz	wawau	AV	verb	2	-ii	3	BEN	ditr	make (s.o.) (s.t.)	crtr, crtn	crtr, rec, crtn
49	MAKE	blz	wawau	AV	verb	2	-kon	3	PURP	no-chg	make (for purp)	crtr, crtn	crtr, purp, crtn
49	MAKE	blz	wawau	AV	verb	2	-i	na	na	na	na	crtr, crtn	
49	MAKE	ind	buat	AV	verb	2	-kan	3	BEN	ditr	make (s.o.) (s.t.)	crtr, crtn	crtr, rec, crtn
49	MAKE	ind	buat	AV	verb	2	-i	na	na	na	na	crtr, crtn	
49	MAKE	jav	gawé	AV	verb	2	-aké	2	CAUS	caus-remp	have (s.t.) made by s.o.	crtr, crtn	agt, crtn, crtr=PP
49	MAKE	jav	gawé	bare	verb	2	-aké	2	BEN	ditr	make (s.o.) (s.t.)	crtr, crtn	crtr, ben, crtn
49	MAKE	jav	gawé	bare	verb	2	-i	na	na	na	na	crtr, crtn	
49	MAKE	mnb	rabu	ae-	verb	2	na	na	na	na	na	crtr, crtn	
49	MAKE	mnb	rabu	ae-	verb	2	-i/-ghoo	3	na	na	make (s.o.) (s.t.)	crtr, crtn	crtr, crtn, ben=IO
49	MAKE	nsy	sanik	AV	verb	2	-i	2	LEX	no-chg	fix (s.t.)	crtr, crtn	agt, pat
49	MAKE	nsy	sanik	AV	verb	2	-kun	3	BEN	ditr	make (s.o.) (s.t.)	crtr, crtn	agt, rec, pat
49	MAKE	sas	piyaq	bare	verb	2	-an	3	BEN	ditr	make (s.o.) (s.t.)	crtr, crtn	agt, ben, crtn
49	MAKE	sun	jieun	AV	verb	2	-keun	3	INST	ditr	make (s.t.) into (s.t.)	crtr, crtn	crtr, matl, crtn
49	MAKE	sun	jieun	AV	verb	2	pang-	3	BEN	ditr	make (s.o.) (s.t.)	crtr, crtn	crtr, rec, crtn
49	MAKE	sun	jieun	AV	verb	2	-an	na	na	na	na	crtr, crtn	
49	MAKE	yka	hinang	N-	verb	3	-an	na	na	na	na	crtr, crtn, matl	
50	BREAK	bes	pecah	bare	verb	1	-i	na	CAUS+PLUR	mnttr	break (s.t.) (rep.)	pat	
50	BREAK	bes	pecah	bare	verb	1	-ka	2	CAUS	mnttr	break (s.t.)	pat	istg, pat
50	BREAK	blz	era'	bare	verb	1	-i/-ii/-kon	na	na	na	na	pat	
50	BREAK	ind	pecah	bare	verb	1	-kan	2	CAUS	mnttr	break, shatter (s.t.)	pat	istg, pat
50	BREAK	ind	pecah	bare	verb	1	-i	2	CAUS+PLUR	mnttr	break, shatter (s.t.) (rep.)	pat	istg, pat
50	BREAK	jav	pecah	bare	verb	1	-aké	2	CAUS+LEX	mnttr	break (s.t.) uninten.	pat	agt, pat
50	BREAK	jav	pecah	bare	verb	1	-i	2	PLUR	no-chg	break (many)	pat	agt, pat
50	BREAK	mnb	bhera	ae-	verb	2	-i	2	PLUR	mnttr	cut, break (many)	agt, pat	agt, pat
50	BREAK	mnb	bhera	ae-	verb	2	-ghoo	na	na	na	na	agt, pat	
50	BREAK	nsy	pechuh	bare	verb	1	-i	2	CAUS+PLUR	mnttr	break (s.t.)	pat	istg, pat
50	BREAK	nsy	pechuh	bare	verb	1	-kun	2	CAUS	mnttr	break (s.t.)	pat	istg, pat
50	BREAK	sas	belaq	bare?	verb	1	-an	2	CAUS	mnttr	break, shatter (s.t.)	pat	agt, pat
50	BREAK	sun	peupeus	bare	verb	1	-keun	2	CAUS	mnttr	break (s.t.)	pat	istg, pat
50	BREAK	sun	peupeus	bare	verb	1	pang-	3	CAUS+BEN	na	break (s.t.) for (s.o.)	pat	agt, rec, pat
50	BREAK	sun	peupeus	bare	verb	1	-an	na	na	na	na	pat	
50	BREAK	yka	pessa'	N-, mag-	verb	2	-an	na	na	na	na	agt, pat	
53	HIT	bes	pantuk	AV	verb	2	-ka	2	INST	remp	hit (inst) on s.t.	agt, pat	agt, inst, pat=PP
53	HIT	bes	pantuk	AV	verb	2	-i	2	PLUR	no-chg	hit (s.t.) rep.	agt, pat	agt, pat
53	HIT	blz	bobok	AV	verb	2	-kon	3	BEN	ditr	pound (s.o.) (s.t.)	agt, pat	agt, rec, pat
53	HIT	blz	bobok	AV	verb	2	-i	2	PLUR	no-chg	beat up (s.o.), hit (s.t.) rep.	agt, pat	agt, pat
53	HIT	blz	bobok	AV	verb	2	-ii	na	na	na	na	agt, pat	
53	HIT	blz	bobok	AV	verb	2	-kon	2	INST	remp	hit (inst) against s.t.	agt, pat	agt, inst, pat=PP
53	HIT	ind	pukul	AV	verb	2	-i	2	PLUR	no-chg	hit (s.t.) rep.	agt, pat	agt, pat
53	HIT	ind	pukul	AV	verb	2	-kan	2	INST	remp	hit (inst) on s.t.	agt, pat	agt, inst, pat=PP
53	HIT	jav	thuthuk	AV	verb	2	-aké	2	INST	remp	knock (inst) against s.t.	agt, pat	agt, inst
53	HIT	jav	thuthuk	AV	verb	2	-i	2	PLUR	no-chg	knock, rap on (s.t.) rep.	agt, pat	agt, pat
53	HIT	mnb	wogha	ae-	verb	2	-i	2	PLUR	no-chg	hit (many)	agt, pat	agt, pat
53	HIT	mnb	wogha	ae-	verb	2	-ghoo	na	na	na	na	agt, pat	
53	HIT	nsy	panggul	AV	verb	2	-i	2	PLUR	no-chg	hit (s.o.) rep.	agt, pat	agt, pat
53	HIT	nsy	panggul	AV	verb	2	-kun	2	INST	remp	hit (s.t.) w/ s.t.	agt, pat	agt, inst, pat=PP
53	HIT	sas	empok	bare	verb	2	-an	3	BEN	ditr	hit (s.o.) for (s.o.)	agt, pat	agt, ben, thm
53	HIT	sun	teungeul	AV	verb	2	-an	2	PLUR	no-chg	hit (s.t.) rep.	agt, pat	agt, pat
53	HIT	sun	teungeul	AV	verb	2	-keun	2	INST	remp	hit (s.t.) on s.t.	agt, pat	agt, inst, pat=PP
53	HIT	sun	teungeul	AV	verb	2	pang-	3	BEN	ditr	hit (s.t.) for (s.o.)	agt, pat	agt, ben, pat
53	HIT	yka	lubak	bare	verb	2	-an	2	PLUR	no-chg	hit (s.t.) rep.	agt, pat	agt, pat
54	TOUCH	blz	koyong	AV	verb	2	-i/-ii/-kon	na	na	na	na	movr, goal	
54	TOUCH	ind	sentuh	AV	verb	2	-kan	2	CAUS	remp	touch (s.t.) to s.w.	thm, goal	agt, thm, goal=PP
54	TOUCH	ind	sentuh	AV	verb	2	-i	na	na	na	na	thm, goal	
54	TOUCH	jav	senggol	AV	verb	2	-aké	2	CAUS	remp	touch (s.t.) to s.t.	thm, goal	agt, thm, goal=PP
54	TOUCH	jav	senggol	AV	verb	2	-i	na	na	na	na	thm, goal	
54	TOUCH	mnb	tabu	unclear	verb	2	-i/-ghoo	na	na	na	na	agt, goal	

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr
54	TOUCH	sun	antel	bare	verb	1	-keun	2	CAUS	remp	touch (s.t.) to (s.w.)	thm, goal=PP	agt, thm, goal=PP
54	TOUCH	sun	antel	bare	verb	1	-an/pang-	2	na	na	na	thm, goal=PP	
54	TOUCH	yka	teppil	-um-	verb	1	-an	na	na	na	na	thm, goal=PP	
55	CUT	bes	tetak	AV?	verb	2	-ka	3	BEN	ditr	cut (s.t.) for (s.o.)		agt, ben, pat
55	CUT	bes	tetak	AV?	verb	2	-i	2	PLUR	no-chg	cut (s.t.) (rep.)		agt, pat
55	CUT	blz	kolong	AV	verb	2	-ii/-kon	na	na	na	na		agt, pat
55	CUT	blz	kolong	AV	verb	2	-i	2	LEX	no-chg	cut off (s.t.)		agt, pat
55	CUT	ind	potong	AV	verb	2	-kan	2	BEN	ditr	cut (s.t.) for (s.o.)		agt, rec, pat
55	CUT	ind	potong	AV	verb	2	-i	2	PLUR	no-chg	cut (many)		agt, pat
55	CUT	ind	potong	AV	verb	2	-kan	2	INST	remp	cut w/ (inst)		agt, inst, pat=PP
55	CUT	jav	iris	AV+RDP	verb	2	-aké	3	BEN	ditr	cut (s.t.) for (s.o.)		agt, ben, pat
55	CUT	jav	iris	AV+RDP	verb	2	-i	2	PLUR+LEX	no-chg	cut up (s.t.)		agt, thm
55	CUT	nsy	tetuk	AV	verb	2	-i	2	PLUR	no-chg	cut (s.t.) (many)		agt, pat
55	CUT	nsy	tetuk	AV	verb	2	-kun	3	BEN	ditr	cut (s.t.) for (s.o.)		agt, rec, pat
55	CUT	sas	peleng	bare?	verb	2	-an	3	BEN	ditr	cut (s.t.) for (s.o.)		agt, ben, thm
55	CUT	sun	teukteuk	AV	verb	2	-an	2	PLUR	no-chg	cut (many) w/ scissors		agt, pat
55	CUT	sun	teukteuk	AV	verb	2	pang-	3	BEN	ditr	cut (s.t.) for (s.o.)		agt, ben, pat
55	CUT	sun	teukteuk	AV	verb	2	-keun	na	na	na	na		agt, pat
55	CUT	yka	peddang	bare	verb	2	-an	2	LEX	no-chg	cut off (s.t.)	agt, pat, inst=PP	agt, pat
56	TAKE	bes	ambik	AV	verb	2	-ka	3	BEN	ditr	take (s.t.) for (s.o.)	agt, thm	agt, ben, thm
56	TAKE	bes	ambik	AV	verb	2	-i	2	PLUR	no-chg	take (many)		agt, thm
56	TAKE	blz	ala	AV	verb	2	-ii	3	BEN	ditr	get (s.o.) (s.t.)		agt, rec, thm
56	TAKE	blz	ala	AV	verb	2	-kon	2	BEN	no-chg	get (s.t.) for (s.o.)		agt, thm, ben=poss
56	TAKE	blz	ala	AV	verb	2	-kon	2	PURP	ditr	take (s.t.) for (purp)		agt, thm
56	TAKE	blz	ala	AV	verb	2	-i	na	na	N	na		agt, thm
56	TAKE	ind	ambil	AV	verb	2	-kan	3	BEN	ditr	take (s.t.) for (s.o.)	agt, thm, src=PP	agt, rec, thm
56	TAKE	ind	ambil	AV	verb	2	-i	2	PLUR	no-chg	take (many, rep.)	agt, thm, src=PP	agt, thm, src=PP
56	TAKE	ind	ambil	AV	verb	2	-kan	3	SOURCE	ditr	take (s.t.) from src	agt, thm, src=PP	agt, src, thm
56	TAKE	jav	njupuk	AV	verb	2	-aké	3	BEN	ditr	take (s.o.) (s.t.)	agt, thm, rec=PP	agt, rec, thm
56	TAKE	jav	njupuk	AV	verb	2	-i	2	PLUR	ditr	take (many, rep.)	agt, thm, rec=PP	agt, thm
56	TAKE	mnb	ala	ae-	verb	2	-ghoo	2	BEN	ditr	take (s.t.) for (s.o.)	agt, thm	agt, rec=IO, thm
56	TAKE	mnb	ala	ae-	verb	2	-i	2	PLUR	no-chg	take (many)	agt, thm	agt, thm
56	TAKE	nsy	akuk	AV	verb	2	-i	2	PLUR	no-chg	take (s.t.) rep.	agt, thm	agt, thm
56	TAKE	nsy	akuk	AV	verb	2	-kun	3	BEN	ditr	take (s.o.) (s.t.)	agt, pat	agt, rec, thm
56	TAKE	sas	bait	bare	verb	2	-an	3	BEN	ditr	take (s.t.) for (s.o.)	agt, pat	agt, ben, thm
56	TAKE	sun	cokot	AV	verb	2	-an	2	PLUR	no-chg	take (s.t.) rep.	agt, thm	agt, thm
56	TAKE	sun	cokot	AV	verb	2	-keun	na	na	na	na	agt, thm	agt, thm
56	TAKE	sun	cokot	AV	verb	2	pang-	3	BEN	ditr	take (s.t.) for (s.o.)	agt, thm	agt, ben, thm
56	TAKE	yka	eddo'	N-, mag-	verb	2	-an	na	na	na	na	agt, thm	
57	TEAR	bes	cebagh	bare?	verb	1	-i	2	CAUS+PLUR	mnr	tear (s.t.) rep.	pat	agt, pat
57	TEAR	bes	cebagh	bare?	verb	1	-ka	2	CAUS	mnr	tear (s.t.)	pat	agt, pat
57	TEAR	ind	sobek	AV	verb	2	-i	2	PLUR	no-chg	tear (s.t.) into pieces	agt, pat	agt, pat
57	TEAR	ind	sobek	AV	verb	2	-kan	na	na	na	na	agt, pat	
57	TEAR	jav	suwék	AV	verb	2	-aké	2	CAUS+LEX	mnr	tear (s.t.) unintent.	agt, thm	istg, thm
57	TEAR	jav	suwék	AV	verb	2	-i	2	PLUR	mnr	tear (rep./dur.)	agt, thm	agt, pat
57	TEAR	mnb	wia	ao-	verb	1	-i/-ghoo	na	na	na	na	pat	
57	TEAR	nsy	cabik	AV	verb	1	-i	2	CAUS+PLUR	mnr	tear (s.t.) rep.	pat	istg, pat
57	TEAR	nsy	cabik	AV	verb	1	-kun	2	CAUS	mnr	tear (s.t.)	pat	istg, pat
57	TEAR	sun	soék	bare	verb	1	-keun	2	CAUS	mnr	tear (s.t.)	pat	istg, pat
57	TEAR	sun	soék	bare	verb	1	pang-	2	CAUS+BEN	ditr	tear (s.t.) for (s.o.)	pat	agt, ben, pat
57	TEAR	sun	soék	bare	verb	1	-an	na	na	na	na	pat	
57	TEAR	yka	garet	N-, mag-	verb	2	-an	na	na	na	na	agt, pat	
58	PEEL	bes	kubak	AV	verb	2	-ka	2	BEN	ditr	peel (s.t.) for (s.o.)	agt, pat	agt, ben, pat
58	PEEL	bes	kubak	AV	verb	2	-i	2	PLUR	no-chg	peel (many)	agt, pat	agt, pat
58	PEEL	blz	dudusi	AV	verb	2	-i/-ii/-kon	na	na	na	na	agt, pat	
58	PEEL	ind	kupas	AV	verb	2	-kan	2	BEN	no-chg	peel (s.t.) for s.o.	agt, pat	agt, pat, rec=PP
58	PEEL	ind	kupas	AV	verb	2	-i	2	PLUR	no-chg	peel (many)	agt, pat	agt, pat
58	PEEL	jav	oncék	unclear	verb	2	-i	2	OPT+PAT	no-chg	peel, pare (s.t.)	agt, pat	agt, pat
58	PEEL	mnb	bhensi	ae-	verb	2	-i	2	PLUR	no-chg	peel, tear (many)	agt, pat	agt, pat
58	PEEL	mnb	bhensi	ae-	verb	2	-ghoo	na	na	na	na	agt, pat	
58	PEEL	nsy	kubak	AV	verb	2	-i	2	PLUR	no-chg	peel (s.t.) rep.	agt, pat	agt, pat
58	PEEL	nsy	kubak	AV	verb	2	-kun	3	BEN	ditr	peel (s.o.) (s.t.)	agt, pat	agt, rec, thm
58	PEEL	sas	lukék	bare?	verb	2	-an	3	BEN	ditr	peel (s.t.) for (s.o.)	agt, pat	agt, ben, pat
58	PEEL	sun	pesék	AV	verb	2	-an	2	PLUR	no-chg	peel (many)	agt, pat	agt, pat

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr
58	PEEL	sun	pesèk	AV	verb	2	pang- -keun	2	BEN	ditr/rempp	peel (s.t.) for (s.o.)	agt, pat	agt, ben, pat
58	PEEL	sun	pesèk	AV	verb	2	-keun	na	na	na	na	agt, pat	
58	PEEL	yka	lagi'	N-, mag-	verb	2	-an	na	na	na	na	agt, pat, src=PP	
59	HIDE	bes	bajik	no-base	unclear	na	-i	2	OBLIG+PLUR	na	hide (many)		agt, thm
59	HIDE	bes	bajik	no-base	unclear	na	-ka	2	OBLIG+THM	na	hide (s.t.)		agt, thm
59	HIDE	blz	sapit	AV	verb	1	-i	2	other-prcr	na	hide o.s. from (s.o.)	agt, loc=PP	agt, prcr
59	HIDE	blz	sapit	AV	verb	1	-kon	2	CAUS	mnr	hide (s.t.)	agt, loc=PP	agt, thm
59	HIDE	blz	sapit	AV	verb	1	-ii	na	na	na	na	agt, loc=PP	
59	HIDE	ind	sembunyi	MID	verb	1	-kan	2	CAUS	mnr	hide (s.t.)	agt, loc=PP	agt, thm
59	HIDE	ind	sembunyi	MID	verb	1	-i	na	na	na	na	agt, loc=PP	
59	HIDE	jav	dhelik	AV	verb	1	-aké	2	CAUS	mnr	hide (s.t.)	agt, loc=PP	agt, pat
59	HIDE	mnb	febuni	ae-	verb	2	-i/-ghoo	na	na	na	na	agt, thm	
59	HIDE	nsy	seguk	bare	verb	1	-i	2	CAUS	mnr	hide (s.t.)	agt	agt, thm
59	HIDE	nsy	seguk	bare	verb	1	-kun	2	CAUS	mnr	hide (s.t.)	agt	agt, thm
59	HIDE	sas	seboq	bare?	verb	1	-an	3	CAUS+BEN	mnr	hide (s.t.) from/for (s.o.)	agt/thm	agt, ben, thm
59	HIDE	sun	salindung	AV	verb	1	-keun	2	CAUS	mnr	make (s.t.) hidden	agt, loc=PP	agt, thm, loc=PP
59	HIDE	sun	salindung	AV	verb	1	pang- -keun	3	CAUS+BEN	ditr	hide (s.t.) for (s.o.)	agt, loc=PP	agt, ben, thm
59	HIDE	sun	salindung	AV	verb	1	-an	2	PLUR	no-chg	hide (s.t.) rep. from s.o.	agt, loc=PP	agt, prcr
59	HIDE	yka	tapuk	pa- SF, -um-	verb	1	-an	1	PLUR	no-chg	hide o.s. rep.	agt, loc=PP	agt, loc=PP
59	HIDE	yka	tapuk	pa- SF, -um-	verb	1	-an	2	CAUS	mnr	hide (s.t.)	agt, loc=PP	agt, pat, loc=PP
61	GIVE	bes	enjuk	AV	verb	3	-i	3	PLUR	no-chg	give (s.o.) (s.t.)	agt, rec, thm	agt, rec, thm
61	GIVE	bes	enjuk	AV	verb	3	-ka	2	THM	detr	give (s.t.) to s.o.	agt, rec, thm	agt, thm, rec=PP
61	GIVE	blz	tarai	AV	verb	3	-i/-ii/-kon	na	na	na	na	agt, rec, thm	
61	GIVE	ind	beri	AV	verb	3	-kan	3	OBLIG-PV	no-chg	give (s.t.) to (s.o.)	agt, rec, thm	agt, rec, thm
61	GIVE	jav	wèh	unclear	verb	1	-aké	2	THM	mnr	give (s.t.) to s.o.	agt, rec=PP	agt, thm, rec=PP
61	GIVE	jav	wèh	unclear	verb	1	-i	3	REC	ditr	give (s.o.) (s.t.)	agt, rec=PP	agt, rec, thm
61	GIVE	mnb	waa	a-	verb	2	-ghoo	3	REC	ditr	give (s.o.) (s.t.)	agt, thm, rec=IO	agt, rec=IO, thm
61	GIVE	mnb	waa	a-	verb	2	-i	na	na	na	na	agt, thm, rec=IO	
61	GIVE	nsy	njuk	AV	verb	3	-i	3	OPT+REC	no-chg	give (s.o.) (s.t.)	agt, rec, thm	agt, rec, thm
61	GIVE	nsy	njuk	AV	verb	3	-kun	3	BEN	no-chg	give (s.t.) to (s.o.)	agt, rec, thm	agt, ben, thm
61	GIVE	sas	béng	bare	verb	3	-an	2	OPT+THM	no-chg	give (s.t.) to s.o.	agt, rec, thm	agt, thm, rec=PP
61	GIVE	sun	béré	AV	verb	3	-an	3	PLUR	no-chg	give (many, rep.)	agt, rec, thm	agt, rec, thm
61	GIVE	sun	béré	AV	verb	3	-keun	3	LEX	no-chg	hand over (s.t.) to (s.o.)	agt, rec, thm	agt, rec, thm
61	GIVE	sun	béré	AV	verb	3	pang- -keun	3	BEN	no-chg	give (s.t.) for (s.o.)	agt, rec, thm	agt, rec, thm
62	SEND	bes	kighim	AV?	verb	2	-i	3	REC	ditr	send (s.o.) (s.t.)	agt, thm, rec/goal=PP	agt, rec, thm
62	SEND	bes	kighim	AV?	verb	2	-ka	2	OPT+THM	no-chg	send (s.t.) to s.o./s.w.	agt, thm, rec/goal=PP	agt, thm, rec/goal=PP
62	SEND	blz	pakatu	AV	noun	2	-i/-ii/-kon	na	na	na	na	agt, thm, rec=PP	
62	SEND	ind	kirim	AV	verb	2	-i	3	REC	ditr	send (s.o.) (s.t.)	agt, thm, rec=PP	agt, rec, thm
62	SEND	ind	kirim	AV	verb	2	-kan	2	OBLIG-PV	no-chg	send (s.t.) to s.o./s.w.	agt, thm, rec=PP	agt, thm, rec=PP
62	SEND	jav	kirim	AV	verb	2	-aké	2	OPT+THM	no-chg	send (s.t.) to s.o./s.w.	agt, thm, goal=PP	agt, thm, goal=PP
62	SEND	jav	kirim	AV	verb	2	-i	3	REC	ditr	send (s.o.) (s.t.)	agt, thm, goal=PP	agt, rec, thm
62	SEND	mnb	pakatu	ae-	verb	2	-ghoo	2	REC	ditr	send (s.o.) (s.t.)	agt, thm, rec=PP	agt, rec=IO, thm
62	SEND	mnb	pakatu	ae-	verb	2	-i	na	na	na	na	agt, thm, rec=PP	
62	SEND	nsy	kikhim	AV	verb	2	-kun	2	OPT+THM	no-chg	send (s.t.) to s.o.	agt, thm, rec/goal=PP	agt, thm, rec=PP
62	SEND	nsy	kikhim	AV	verb	2	-i	3	REC	ditr	send (s.o.) (s.t.)	agt, thm	agt, rec, thm
62	SEND	sas	kirim	AV	verb	2	-an	3	REC	ditr	send to (s.o.) (s.t.)	agt, thm	agt, rec, thm
62	SEND	sun	kirim	AV	verb	2	-an	2	PLUR	no-chg	send (many, rep.)	agt, thm, rec=PP	agt, thm, rec=PP
62	SEND	sun	kirim	AV	verb	2	-keun	3	CAUS+THM	caus-rempp	send (s.t.) to s.o. (by s.o.)	agt, thm, rec=PP	istg, thm, rec=PP
62	SEND	sun	kirim	AV	verb	2	pang- -keun	3	BEN	ditr	send (s.t.) for (s.o.)	agt, thm, rec=PP	agt, ben, thm
64	THROW	bes	telemplak	fos.NVOL	verb	2	-i	2	PLUR	no-chg	throw at (s.t.) w/ s.t.	agt, goal, thm=PP	agt, goal, thm=PP
64	THROW	bes	telemplak	fos.NVOL	verb	2	-ka	2	THM	rempp	throw (s.t.) to s.t.	agt, goal, thm=PP	agt, thm, goal=PP
64	THROW	blz	balo'	AV	verb	2	-kon	2	THM	rempp	throw (s.t.)	agt, goal	agt, thm
64	THROW	blz	balo'	AV	verb	2	-kon	2	LEX	na	throw (s.t.) away	agt, goal	agt, thm
64	THROW	blz	balo'	AV	verb	2	-i/-ii	na	na	na	na	agt, goal	
64	THROW	ind	lempar	AV	verb	2	-i	2	GOAL	rempp	throw (many) at s.t.	agt, goal, thm=PP	agt, goal, thm=PP
64	THROW	ind	lempar	AV	verb	2	-kan	2	OPT+THM	no-chg	throw (s.t.) at s.t.	agt, goal, thm=PP	agt, thm, goal=PP
64	THROW	jav	balang	AV	verb	2	-i	2	GOAL	ditr	throw (s.t.) at (s.t.)	agt, thm, goal=PP	agt, goal, thm
64	THROW	jav	balang	AV	unclear	2	-aké	2	OPT+THM	no-chg	throw (s.t.) at s.t.	agt, thm, goal=PP	agt, thm, goal=PP
64	THROW	mnb	ghoro	ae-	verb	2	-i	2	PLUR	no-chg	throw/discard (many)	agt, thm, rec=IO	agt, rec=IO, thm
64	THROW	mnb	ghoro	ae-	verb	2	-ghoo	2	BEN	ditr	throw (s.t.) for/to (s.o.)	agt, thm, rec=IO	agt, rec=IO, thm
64	THROW	nsy	khambui	AV	verb	2	-i	2	OPT+GOAL/PLUR	no-chg	throw at (s.o./s.t.)	agt, goal	agt, goal
64	THROW	nsy	khambui	AV	verb	2	-kun	2	INST	rempp	throw (s.t.)	agt, goal	agt, thm, targ=PP
64	THROW	sas	saut	bare?	verb	2	-an	3	BEN	ditr	throw (s.t.) for (s.o.)	agt, thm	agt, ben, thm
64	THROW	sun	alung	AV	verb	2	-an	2	PLUR	no-chg	throw (s.t.) at s.t. rep.	agt, thm	agt, thm, goal=PP

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr
64	THROW	sun	alung	AV	verb	2	-keun	2	THM+GOAL	no-chg	throw (s.t.) at s.t.	agt, thm	agt, thm, goal=PP
64	THROW	sun	alung	AV	verb	2	pang- -keun	3	BEN	ditr	throw (s.t.) for (s.o.)	agt, thm	agt, thm, thm
64	THROW	yka	teppad	N-, mag-	verb	2	-an	2	OBLIG-PV+THM	no-chg	throw (s.t.)	agt, thm, goal=PP	agt, thm, rec=PP
64	THROW	yka	teppad	N-, mag-	verb	2	-an	3	REC	ditr	throw (s.o.) (s.t.)	agt, thm, goal=PP	agt, thm, thm
65	TIE	bes	kebat	AV?	unclear	2	-i	2	PLUR	no-chg	tie (s.t.)	agt, thm	agt, thm
65	TIE	bes	kebat	AV?	unclear	2	-ka	2	INST	remp	tie (rope) to s.t.	agt, thm	agt, inst, goal=PP
65	TIE	blz	koot	AV	noun	2	-kon	3	GOAL	ditr	tie (s.t.) to (s.t.)	agt, pat	agt, goal, thm
65	TIE	blz	koot	AV	noun	2	-kon	3	BEN	ditr	tie (s.t.) for (s.o.)	agt, pat	agt, thm, pat
65	TIE	blz	koot	AV	noun	2	-i/-ii	na	na	na	na	agt, pat	agt, pat
65	TIE	ind	ikat	AV	noun	2	-kan	2	INST	remp	tie (inst) to s.t.	agt, pat	agt, inst, goal=PP
65	TIE	ind	ikat	AV	noun	2	-kan	2	OPT+THM	no-chg	tie (s.t.) to s.t., tie up	agt, pat	agt, pat, goal=PP
65	TIE	ind	ikat	AV	noun	2	-i	na	na	na	na	agt, pat	agt, pat
65	TIE	jav	tali	no-base	noun	na	-aké	2	CAT+INST	na	fasten (rope) to s.t.	agt, inst, goal=PP	agt, inst, goal=PP
65	TIE	jav	tali	no-base	noun	na	-aké	2	CAT+THM+BEN	na	tie (s.t.) of s.o.'s to s.t.	agt, thm, goal=PP/poss	agt, thm, goal=PP/poss
65	TIE	jav	tali	no-base	noun	na	-i	2	CAT+THM	na	tie up, bind (s.t.)	agt, thm, goal=PP	agt, thm, goal=PP
65	TIE	nsy	ikuk	AV	verb	2	-i	2	PLUR	no-chg	tie (s.t.) (rep.)	agt, pat	agt, pat
65	TIE	nsy	ikuk	AV	verb	2	-kun	2	INST	remp	tie w/ (s.t.) on s.t.	agt, pat	agt, inst, pat=PP
65	TIE	sas	taliq	bare	verb	2	-an	3	GOAL	ditr	tie up (s.t.) w/ (s.t.)	agt, inst, goal=PP	agt, inst, goal=PP
65	TIE	sas	taliq	bare	verb	2	-an	3	BEN	ditr	tie (s.t.) for (s.o.)	agt, inst, goal=PP	agt, thm, pat
65	TIE	sun	tali	no-base	noun	na	-an	2	CAT+THM	na	tie up, tether (s.t.)	agt, pat	agt, pat
65	TIE	sun	tali	no-base	noun	na	-keun	2	CAT+THM	na	tie (s.t.) to s.t.	agt, pat, goal=PP	agt, pat, goal=PP
65	TIE	sun	tali	no-base	noun	na	pang- -keun+-an	2	CAT+BEN	na	tie (s.t.) for (s.o.)	agt, thm, pat	agt, thm, pat
66	PUT	bes	tepike	no-base	no-base	na	-i	2	BEN+LEX	ditr	put out for (s.o.) (s.t.)	agt, thm, thm	agt, thm, thm
66	PUT	bes	tepike	no-base	no-base	na	-ka	2	OBLIG+THM	na	put (s.t.) at s.w.	agt, thm, goal=PP	agt, thm, goal=PP
66	PUT	blz	kekl	AV	verb	2	-i/-ii/-kon	na	na	na	na	agt, thm, goal=PP	agt, thm, goal=PP
66	PUT	ind	taruh	AV	verb	2	-i	2	LOC	remp	fill (s.w.) w/ s.t.	agt, thm, goal=PP	agt, goal, thm=PP
66	PUT	ind	taruh	AV	verb	2	-kan	2	BEN	no-chg	put (s.t.) of s.o. at s.w. for them	agt, thm, goal=PP	agt, thm, ben=poss
66	PUT	jav	dékék	AV	verb	2	-i	2	GOAL	ditr	place/put/lay (s.t.) on (s.t.)	agt, thm, goal=PP	agt, goal, thm
66	PUT	jav	dékék	AV	verb	2	-aké	2	OPT+GOAL	no-chg	place/put/lay (s.t.) on s.t.	agt, thm, goal=PP	agt, thm, goal=PP
66	TIE	mnb	tapu	unclear	verb	2	-i	2	PLUR	no-chg	tie, tether (many)	agt, thm, goal=PP	agt, thm
66	TIE	mnb	tapu	unclear	verb	2	-ghoo	na	na	na	na	agt, thm, goal=PP	agt, thm, goal=PP
66	PUT	nsy	pik	AV	verb	2	-i	2	PLUR	no-chg	put (s.t.) on s.t.	agt, thm	agt, thm, goal=PP
66	PUT	nsy	pik	AV	verb	2	-kun	2	OPT+THM	no-chg	put (s.t.) on s.t.	agt, thm	agt, thm, goal=PP
66	PUT	sas	toloq	bare	verb	2	-an	3	BEN	ditr	put (s.t.) at (place) for s.o.	agt, thm, goal=PP	agt, goal, thm
66	PUT	sas	toloq	bare	verb	2	-an	3	GOAL	ditr	put (s.t.) (s.w.)	agt, thm, goal=PP	agt, thm, thm
66	PUT	sun	teundeun	AV	verb	2	pang- -keun	3	BEN	no-chg	put (s.t.) for (s.o.) at s.w.	agt, thm, goal=PP	agt, thm, thm
66	PUT	sun	teundeun	AV	verb	2	-an	3	LOC+PLUR	no-chg	put (s.t.) at (place)	agt, thm, goal=PP	agt, goal, thm=PP
66	PUT	yka	bettad	pa- SF	verb	2	-an	2	BEN	ditr	set (s.t.) for (s.o.)	agt, thm, goal=PP	agt, thm, thm
66	PUT	yka	bettad	pa- SF	verb	2	-an	2	GOAL	ditr	set (s.w.) (s.t.)	agt, thm, goal=PP	agt, loc, thm
66	TIE	yka	ingket	N-	noun	2	-an	2	OBLIG-PV+THM	no-chg	tie (s.t.), to tether (s.t.)	agt, thm	agt, thm
67	POUR	bes	tuang	AV?	verb	2	-i	2	THM+PLUR	no-chg	pour (s.t.) into s.t. (rep.)	agt, thm, goal=PP	agt, thm, goal=PP
67	POUR	bes	tuang	AV?	verb	2	-ka	2	OPT+THM	no-chg	pour (s.t.) into s.t.	agt, thm, goal=PP	agt, thm, goal=PP
67	POUR	blz	tawa'	AV	verb	2	-i	1	LEX	no-chg	pay compensation, indemnity	agt, thm	agt
67	POUR	ind	tuang	AV	verb	2	-i	2	GOAL	remp	pour into (s.t.) w/ liquid	agt, thm	agt, goal, thm=PP
67	POUR	ind	tuang	AV	verb	2	-kan	2	OPT+THM	no-chg	pour (s.t.) into s.t.	agt, thm	agt, thm, goal=PP
67	POUR	jav	cur	RDP	verb	1	-aké	2	CAUS+THM	mnttr	pour (s.t.) into s.t.	thm	agt, thm, goal=PP
67	POUR	jav	cur	RDP	verb	1	-i	2	CAUS+GOAL	mnttr	pour water on (s.t.), to water	thm	agt, goal, thm=PP
67	POUR	mnb	hobha	ae-	verb	2	-i	2	GOAL+PLUR	ditr	pour on, spill, spray (many)	agt, thm, goal=PP	agt, goal, thm
67	POUR	mnb	hobha	ae-	verb	2	-ghoo	na	na	na	na	agt, thm, goal=PP	agt, thm, goal=PP
67	POUR	nsy	tuyung	AV	verb	2	-i	2	THM+PLUR	no-chg	pour into (many, rep.)	agt, thm	agt, thm, goal=PP
67	POUR	nsy	tuyung	AV	verb	2	-kun	2	OPT+THM	no-chg	pour (s.t.) into s.t.	agt, thm	agt, thm, goal=PP
67	POUR	sas	tumpah	bare?	verb	1	-an	3	CAUS+THM+GOAL	ditr	pour (s.t.) into s.t.	thm	agt, goal, thm
67	POUR	sas	tumpah	bare?	verb	1	-an	2	CAUS+THM	mnttr	spill (s.t.)	thm	agt, thm, goal=PP
67	POUR	sun	kucur	AV	verb	1	-an	2	CAUS+GOAL+PLUR	mnttr	pour into (s.t.) w/ s.t. rep.	thm	agt, thm, goal=PP
67	POUR	sun	kucur	AV	verb	1	-keun	2	CAUS+THM	mnttr	pour (s.t.)	thm	agt, thm, goal=PP
67	POUR	sun	kucur	AV	verb	1	pang- -keun	2	CAUS+THM+BEN	ditr	pour (s.t.) for (s.o.)	thm	agt, thm, thm
67	POUR	yka	bu'us	N-, mag-	verb	2	-an	2	OBLIG-PV+THM	no-chg	spill/pour our (s.t.)	agt, thm	agt, thm
68	COVER	bes	tukup	AV?	unclear	2	-ka	2	INST	remp	use (s.t.) as a cover on s.t.	agt, pat	agt, inst, goal=PP
68	COVER	bes	tukup	AV?	unclear	2	-i	2	OPT+GOAL	no-chg	cover (s.t.) w/ s.t.	agt, pat	agt, goal, thm=PP
68	COVER	blz	talop	no-base	noun	na	-i	2	GOAL	na	cover (s.t.)	agt, pat	agt, goal, thm=PP
68	COVER	blz	talop	no-base	noun	na	-ii/-kon	na	na	na	na	agt, pat	agt, pat
68	COVER	ind	tutup	AV	noun	2	-kan	2	INTENS	no-chg	close (s.t.) up tightly	agt, pat	agt, pat
68	COVER	ind	tutup	AV	noun	2	-kan	2	INST	remp	use (s.t.) as a cover	agt, pat	agt, inst, goal=PP
68	COVER	ind	tutup	AV	noun	2	-i	2	GOAL	no-chg	put a lid on, cover (s.t.)	agt, pat	agt, goal, thm=PP

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	FUNCTION	SynChg	AfMeaning	BsStr	AfStr
68	COVER	ind	tutup	AV	noun	2	-i	2	PLUR	no-chg	close (many, rep.)	agt, pat	agt, pat
68	COVER	jav	tutup	AV	unclear	2	-aké	3	BEN	ditr	close (s.t.) for (s.o.)	agt, pat	agt, ben, pat
68	COVER	jav	tutup	AV	unclear	2	-aké	2	INST	ditr	use (s.t.) to cover (s.t.)	agt, pat	agt, goal/pat, inst
68	COVER	jav	tutup	AV	unclear	2	-i	2	OPT+GOAL	no-chg	close, cover (s.t.)	agt/thm, pat/loc	agt/thm, pat/goal
68	COVER	mnb	songko	ae-	verb	2	-ghoo	na	na	na	na	agt, pat	
68	COVER	mnb	songko	ae-	verb	2	-i	2	GOAL+PLUR	no-chg	cover (many) (or o.s.)	agt, pat	agt, pat
68	COVER	nsy	tukup	AV	unclear	2	-kun	2	INST	remp	cover w/ (s.t.) onto s.t.	agt, pat	agt, inst, goal=PP
68	COVER	nsy	tukup	AV	unclear	2	-i	2	OPT+GOAL	no-chg	cover (s.t.)	agt, pat	agt, goal
68	COVER	sas	tutup	bare?	unclear	2	-an	3	BEN	ditr	close (s.t.) for (s.o.)	agt/inst, pat/goal	agt, ben, pat
68	COVER	sun	tutup	AV	noun	2	-an	2	OPT+GOAL	no-chg	cover (s.t.)	agt/inst, pat/goal	agt, goal, thm=PP
68	COVER	sun	tutup	AV	noun	2	pang- -keun	2	BEN	ditr	cover (s.t.) for (s.o.)	agt, pat	agt, pat, ben=PP
68	COVER	sun	tutup	AV	noun	2	-keun	na	na	na	na	agt, pat	
68	COVER	yka	lekkebb	N-	noun	2	-an	2	OBLIG-PV+GOAL	no-chg	cover (s.t.) (as w/ a lid)	agt, pat	agt, pat/goal
69	FILL	bes	isi	AV?	unclear	2	-ka	2	OPT+THM	no-chg	pour (s.t.) into s.t.	agt, thm, goal=PP	agt, thm, goal=PP
69	FILL	blz	isii	AV	noun	2	-kon	2	THM	remp	fill (s.t.) into s.t.	agt, goal, thm=PP	agt, thm, goal=PP
69	FILL	blz	isii	AV	noun	2	-i/-ii	na	na	na	na	agt, goal, thm=PP	
69	FILL	ind	isi	AV	noun	2	-kan	2	THM	remp	put (s.t.) into s.t.	agt, goal, thm=PP	agt, thm, goal=PP
69	FILL	ind	isi	AV	noun	2	-i	na	na	na	na	agt, goal, thm=PP	
69	FILL	jav	isi	AV	noun	2	-aké	2	OPT+THM	remp	put (s.t.) into s.t.	agt, thm, goal=PP	agt, thm, goal=PP
69	FILL	jav	isi	AV	noun	2	-i	2	GOAL	ditr	fill, put (s.t.) into (s.t.)	agt, thm, goal=PP	agt, goal, thm
69	FILL	mnb	fokosinihi	ae-+fo-	verb	2	-i/-ghoo	na	na	na	na	agt, thm, goal=PP	
69	FILL	nsy	isi	AV	verb	2	-kun	2	BEN	mtr	fill (s.o.) (s.t.) in s.t.	agt/loc, thm	agt, ben, thm
69	FILL	nsy	isi	AV	verb	2	-i	na	na	na	na	agt/loc, thm	
69	FILL	sas	isiiq	bare?	unclear	2	-an	3	BEN+THM	ditr	put (s.t.) in s.t. for (s.o.)	agt, goal, thm=PP	agt, ben, thm
69	FILL	sas	isiiq	bare?	unclear	2	-an	2	THM	remp	put (s.t.) in s.t.	agt, goal, thm=PP	agt, thm, goal=PP
69	FILL	sun	eusi	AV	noun	2	-keun	2	CAUS	caus-remp	fill (s.t.) into s.t.	loc, thm	agt, thm, goal=PP
69	FILL	sun	eusi	AV	noun	2	pang- -keun	3	CAUS+BEN	ditr	fill (s.t.) for (s.o.) into s.t.	loc, thm	agt, ben, thm
69	FILL	sun	eusi	AV	noun	2	pang- -keun+-an	3	CAUS+BEN	ditr	fill (s.t.) for (s.o.) w/ s.t.	loc, thm	agt, ben, goal
69	FILL	sun	eusi	AV	noun	2	-an	2	CAUS+GOAL	remp	fill (s.t.) w/ s.t.	loc, thm	agt, goal, thm=PP
69	FILL	yka	penno'	N-	verb	2	-an	2	OBLIG-PV+THM	no-chg	To fill (s.t.) w/ s.t.	agt, goal, thm=PP	agt, goal, thm=PP
70	LOAD	bes	muat	AV?	verb	1	-i	2	LOC	remp	load (s.w.) w/ s.t.	goal, thm=PP	agt, goal, thm=PP
70	LOAD	bes	muat	AV?	verb	1	-ka	2	THM	mtr	load (s.t.) into s.w.	goal, thm=PP	agt, thm, goal=PP
70	LOAD	blz	woot	AV	verb	2	-i/-ii/-kon	na	na	na	na	agt, thm, goal=PP	
70	LOAD	ind	muat	AV	verb	2	-i	2	CAUS	remp	load (s.w.) w/ s.t.	loc, thm	agt, goal, thm=PP
70	LOAD	ind	muat	AV	verb	2	-kan	2	CAUS+THM	remp	load (s.t.) into s.w.	loc, thm, goal=PP	agt, thm, goal=PP
70	LOAD	jav	amot	AV	verb	2	-aké	2	CAUS+THM	caus-remp	load (s.t.) into s.w.	loc, thm	agt, thm, goal=PP
70	LOAD	jav	amot	AV	verb	2	-aké	3	CAUS+BEN	ditr	load (s.t.) for (s.o.)	loc, thm	agt, rec, thm
70	LOAD	jav	amot	m-	verb	2	-i	2	CAUS+GOAL	ditr	load (s.t.) in/on (s.t.)	loc, thm	agt, goal, thm
70	LOAD	mnb	ulea	ae-	verb	2	-i/-ghoo	na	na	na	na	loc, thm	
70	LOAD	nsy	muat	AV	verb	2	-i	2	GOAL	remp	load (s.t.) w/ s.t.	agt, thm, goal=PP	agt, goal, thm=PP
70	LOAD	nsy	muat	AV	verb	2	-kun	2	OPT+THM	no-chg	load (s.t.) on s.t.	agt, thm, goal=PP	agt, thm, goal=PP
70	LOAD	sas	buat	bare?	verb	2	-an	3	BEN	ditr	load (s.t.) for (s.o.)	agt, thm	agt, ben, thm
70	LOAD	sun	muat	AV	verb	2	-an	2	CAUS	remp	load (s.w.) w/ s.t.	loc, thm	agt, goal, thm=PP
70	LOAD	sun	muat	AV	verb	2	-keun	2	CAUS+THM	remp	load (s.t.) into s.w..	loc, thm	agt, thm, goal=PP
70	LOAD	sun	muat	AV	verb	2	pang- -keun	2	CAUS+BEN	no-chg	load (s.t.) for s.o.	loc, thm	agt, thm, ben=PP
70	LOAD	yka	pa-sakey	mag-	verb	2	-an	na	na	na	na	agt, thm	
71	BRING	bes	batak	AV	verb	2	-ka	3	BEN	ditr	bring (s.t.) for (s.o. else)	agt, thm	agt, ben, thm
71	BRING	bes	batak	AV	verb	2	-i	2	PLUR	no-chg	bring (many)	agt, thm	agt, thm
71	BRING	blz	wawa	AV	verb	2	-kon	3	BEN	ditr	bring for (s.o.) (s.t.)	agt, thm	agt, ben, thm
71	BRING	blz	wawa	AV	verb	2	-i	3	REC	ditr	bring or carry (s.o.) (s.t.)	agt, thm	agt, rec, thm
71	BRING	blz	wawa	AV	verb	2	-ii	na	na	na	na	agt, thm	
71	BRING	ind	bawa	AV	verb	2	-kan	2	BEN	ditr	bring (s.t.) for (s.o. else)	agt, thm, rec/goal=PP	agt, rec, thm
71	BRING	jav	gawa	AV	verb	2	-aké	2	BEN	ditr	bring (s.t.) for (s.o.)	agt, thm, rec/goal=PP	agt, rec, thm
71	BRING	jav	gawa	AV	verb	2	-i	2	OPT+THM	no-chg	bring (s.t.)	agt, thm, rec/goal=PP	agt, rec, thm
71	BRING	mnb	owa	ae-	verb	2	-ghoo	3	REC	ditr	bring for (s.o.) (s.t.)	agt, thm, goal=PP	agt, rec=IO, thm
71	BRING	mnb	owa	ae-	verb	2	-ghoo	3	PURP	ditr	bring (s.t.) for (purp)	agt, thm, goal=PP	agt, purp=IO, thm
71	BRING	mnb	owa	ae-	verb	2	-i	na	na	na	na	agt, thm, goal=PP	
71	BRING	nsy	batuk	AV	verb	2	-i	2	PLUR	no-chg	bring (s.t.) (rep.)	agt, thm	agt, thm
71	BRING	nsy	batuk	AV	verb	2	-kun	3	BEN	ditr	bring (s.t.) for (s.o.)	agt, thm	agt, rec, thm
71	BRING	sas	jauq	bare?	verb	2	-an	3	BEN	ditr	bring (s.t.) for (s.o.)	agt, thm	agt, ben, thm
71	BRING	sun	bawa	AV	verb	2	-an	2	PLUR	no-chg	bring (many, rep.)	agt, thm, rec/goal=PP	agt, thm
71	BRING	sun	bawa	AV	verb	2	-keun	3	CAUS+THM	caus-remp	have (s.t.) sent	agt, thm, rec/goal=PP	istg, thm
71	BRING	sun	bawa	AV	verb	2	-keun	2	LEX	no-chg	present (perf)	agt, thm, rec/goal=PP	prfr, prfc
71	BRING	sun	bawa	AV	verb	2	pang- -keun	3	BEN	ditr	bring (s.t.) for (s.o.)	agt, thm, rec/goal=PP	agt, rec, thm

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr
71	BRING	yka	bo'ò	N-, mag-	verb	2	-an	3	REC	ditr	bring (s.o.) (s.t.)	agt, thm	agt, rec, thm
71	BRING	yka	bo'ò	N-, mag-	verb	2	-an+pa-	3	CAUS+REC	ditr	send (s.t.) (by s.o.)	agt, thm	istg, rec, thm
72	PUSH	bes	undu	AV?	verb	2	-i	2	PLUR	mnr	push (many)	agt, thm	agt, thm
72	PUSH	bes	undu	AV?	verb	2	-ka	2	OPT+THM	no-chg	push (s.t.)	agt, thm	agt, thm
72	PUSH	blz	dudul	no-base	verb	na	-kon	2	OBLIG+THM	na	push (s.t.)		agt, thm
72	PUSH	blz	dudul	no-base	verb	na	-i/-ii	na	na	na	na		agt, thm
72	PUSH	ind	dorong	AV	verb	2	-kan	2	OPT+THM	no-chg	push, move (s.t.) forward	agt, thm, goal=PP	agt, thm
72	PUSH	ind	dorong	AV	verb	2	-i	na	na	na	na	agt, thm, goal=PP	agt, thm
72	PUSH	jav	dorong	AV	verb	2	-aké	2	OPT+THM	no-chg	push (s.t.) in a direction		agt, thm, goal=PP
72	PUSH	jav	dorong	AV	verb	2	-i	na	na	na	na	agt, thm	
72	PUSH	mnb	dhudhu	ae-	verb	2	-i	2	INTENS	no-chg	push/force (s.t.) into s.t.	agt, thm, goal=PP	agt, thm, goal=PP
72	PUSH	mnb	dhudhu	ae-	verb	2	-ghoo	na	na	na	na	agt, thm, goal=PP	
72	PUSH	nsy	tulak	AV	verb	2	-kun	2	OPT+THM	no-chg	shove (s.o.)	agt, thm	agt, thm
72	PUSH	nsy	tulak	AV	verb	2	-i	na	na	na	na	agt, thm	
72	PUSH	sas	jelek	bare?	verb	2	-an	3	BEN	ditr	push (s.t.) for (s.o.)	agt, thm	agt, ben, thm
72	PUSH	sun	surung	AV	verb	2	-keun	2	OPT+THM	no-chg	(s.t.) toward s.w.	agt, thm, goal=PP	agt, thm, goal=PP
72	PUSH	sun	surung	AV	verb	2	-an	2	LEX	no-chg	urge (s.o.) to do (s.t.)	agt, thm, goal=PP	agt, thm
72	PUSH	sun	surung	AV	verb	2	pang- -keun	3	BEN	ditr	push (s.t.) for (s.o.)	agt, thm, goal=PP	agt, ben, thm
72	PUSH	yka	de'en	N-	verb	2	-an	na	na	na	na	agt, thm	
72	PUSH	yka	luntud	N-	verb	2	-an	2	OBLIG-PV+THM	no-chg	push (s.t.)	agt, thm, goal=PP	agt, thm
73	DIG	bes	kajah	AV?	verb	2	-ka	3	BEN	ditr	dig for (s.o, s.t.)	agt, thm	agt, ben, thm
73	DIG	bes	kajah	AV?	verb	2	-i	2	PLUR	no-chg	dig (rep.)	agt, thm	agt, thm
73	DIG	blz	keke	AV	verb	2	-i/-ii/-kon	na	na	na	na	agt, loc	
73	DIG	ind	gali	AV	verb	2	-i/-kan	na	na	na	na	agt, thm	
73	DIG	jav	kudhuk	AV	verb	2	-i	2	OPT+THM+LEX	no-chg	dig (s.t.) up/out	agt, thm	agt, thm
73	DIG	jav	kudhuk	AV	verb	2	-aké	na	na	na	na	agt, thm	
73	DIG	mnb	dongka	ae-	verb	1	-i	na	PLUR	no-chg	dig up (many)	agt, loc=PP	
73	DIG	mnb	dongka	ae-	verb	1	-ghoo	na	na	na	na	agt, loc=PP	
75	STEAL	bes	maling	AV?	verb	2	-ka	3	BEN	ditr	steal (s.t.) for (s.o.)	agt, thm	agt, thm, thm
75	STEAL	bes	maling	AV?	verb	2	-i	2	PLUR	no-chg	steal (s.t.) (rep., hab.)	agt, thm	agt, thm
75	STEAL	blz	gamut	AV	verb	2	-i/-ii/-kon	na	na	na	na	agt, thm	
75	STEAL	ind	curi	AV	verb	2	-kan	3	BEN	ditr	steal (s.o.) (s.t.)	agt, thm	agt, ben, thm
75	STEAL	ind	curi	AV	verb	2	-kan	2	OPT+THM	no-chg	steal (s.t.) [uncommon]	agt, thm	agt, thm
75	STEAL	ind	curi	AV	verb	2	-i	na	na	na	na	agt, thm	
75	STEAL	jav	colong	AV	verb	2	-aké	3	BEN	ditr	steal (s.t.) for (s.o.)		agt, ben, thm
75	STEAL	jav	colong	AV	verb	2	-i	2	PLUR	no-chg	steal (s.t.) rep.	agt, thm	agt, thm
75	STEAL	mnb	mbolaku	ae-	verb	2	-i/-ghoo	na	na	na	na	agt, thm	
75	STEAL	nsy	paling	AV	verb	2	-i	2	PLUR	no-chg	steal (many)	agt, thm	agt, thm
75	STEAL	nsy	paling	AV	verb	2	-kun	3	BEN	ditr	steal (s.t.) for (s.o.)	agt, thm	agt, rec, thm
75	STEAL	sas	paling	bare?	verb	2	-an	3	BEN	ditr	steal (s.t.) for (s.o.)	agt, thm	agt, ben, thm
75	STEAL	sun	paling	AV	verb	2	-an	2	PLUR	no-chg	steal (rep., of many)	agt, thm	agt, thm
75	STEAL	sun	paling	AV	verb	2	pang- -keun	3	BEN	ditr	steal (s.t.) for (s.o.)	agt, thm	agt, rec, thm
75	STEAL	sun	paling	AV	verb	2	-keun	na	na	na	na	agt, thm	
75	STEAL	yka	tangkew	N-, mag	verb	2	-an	na	na	na	na	agt, thm	
76	GRIND	bes	isar	AV?	unclear	2	-ka	3	BEN	ditr	grind (s.t.) for (s.o.)	agt, pat	agt, rec, pat
76	GRIND	bes	isar	AV?	unclear	2	-i	2	PLUR	no-chg	grind (many)	agt, pat	agt, pat
76	GRIND	blz	gurinda	AV	noun	2	-i/-ii/-kon	na	na	na	na	agt, pat	
76	GRIND	ind	giling	AV	verb	2	-kan	2	CAUS+PAT	caus-remp	have (s.t.) milled	agt, pat	istg, pat
76	GRIND	ind	giling	AV	verb	2	-i	na	na	na	na	agt, pat	
76	GRIND	jav	giling	AV	verb	2	-aké/-i	na	na	na	na	agt, pat	
76	GRIND	mnb	gurinda	unclear	noun	2	-i/-ghoo	na	na	na	na	agt, pat	
76	GRIND	nsy	giling	AV	verb	2	-i	2	PLUR	no-chg	grind (many)	agt, pat	agt, pat
76	GRIND	nsy	giling	AV	verb	2	-kun	3	BEN	ditr	grind (s.t.) for (s.o.)	agt, pat	agt, rec, pat
76	GRIND	sas	giliq	bare?	verb	2	-an	3	BEN	ditr	grind (s.t.) for (s.o.)	agt, pat	agt, ben, pat
76	GRIND	sun	rendos	AV	verb	2	-an	2	PLUR	no-chg	grind, crush (many, rep.)	agt, pat, inst=PP	agt, pat, inst=PP
76	GRIND	sun	rendos	AV	verb	2	pang- -keun	2	BEN	ditr	grind (s.t.) for (s.o.)	agt, pat, inst=PP	agt, ben, pat
76	GRIND	sun	rendos	AV	verb	2	-keun	na	na	na	na	agt, pat, inst=PP	
76	GRIND	yka	giling	N-, mag-	verb	2	-an	na	na	na	na	agt, pat	
77	HEAR	bes	aning	AV	verb	2	-i	2	INTENS	no-chg	listen to (s.t.)	prcr, stim	obsr, stim
77	HEAR	bes	aning	AV	verb	2	-ka	na	na	na	na	prcr, stim	
77	HEAR	blz	rongor	AV	verb	2	-i/-ii/-kon	na	na	na	na	prcr, stim	
77	HEAR	ind	dengar	AV	verb	2	-kan	2	INTENS	no-chg	listen to (s.t.)	prcr, stim	obsr, stim
77	HEAR	ind	dengar	AV	verb	2	-kan	2	LEX	no-chg	pay heed to (s.t.)	prcr, stim	cgmr, cont
77	HEAR	ind	dengar	AV	verb	2	-i	na	na	na	na	prcr, stim	

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr
77	HEAR	jav	rungu	k-	verb	2	-aké	2	INTENS	no-chg	listen to (s.t.)	prcr, stim	obsr, stim
77	HEAR	jav	rungu	k-	verb	2	-i	na	na	na	na	prcr, stim	
77	HEAR	mnb	tingke	fe-	noun	2	-i/-ghoo	na	na	na	na	prcr, stim	
77	HEAR	nsy	dengul	AV	verb	2	-i	2	INTENS	no-chg	listen to (s.t.)	prcr, stim	obsr, stim
77	HEAR	nsy	dengul	AV	verb	2	-kun	2	INTENS	no-chg	listen to (s.t.)	prcr, stim	obsr, stim
77	HEAR	sas	dengah	bare?	verb	2	-an	3	BEN	ditr	listen to (s.t.) for (s.o.)	prcr, stim	obsr, ben, stim
77	HEAR	sun	dengé	AV	verb	2	-an+RDP	1	PLUR+LEX	detr	hear indistinctly (dur.)	prcr, stim	prcr
77	HEAR	sun	dengé	AV	verb	2	-keun	2	INTENS	no-chg	listen to (s.t.)	prcr, stim	obsr, stim
77	HEAR	sun	dengé	AV	verb	2	pang- -keun	2	BEN	no-chg	listen to (s.t.) for s.o.	prcr, stim	obsr, stim, ben=PP
77	HEAR	yka	kale	NVOL	verb	2	-an	na	na	na	na	prcr, stim	
78	TEACH	bes	ajagh	AV?	verb	2	-i	2	REC	ditr	teach (s.o.) (s.t.)	agt, cont	agt, cgnr, cont
78	TEACH	bes	ajagh	AV?	verb	2	-ka	na	na	na	na	agt, cont	
78	TEACH	blz	siso'	pi-	verb	2	-kon	2	CONT	ditr	na	agt, cgnr, cont=cls	agt, cgnr, cont
78	TEACH	blz	siso'	pi-	verb	2	-i/-ii	na	na	na	na	agt, cgnr, cont=cls	
78	TEACH	ind	ajar	AV	verb	3	-i	3	OPT+REC	no-chg	teach (s.o.) (s.t.)	agt, cgnr, cont	agt, cgnr, cont
78	TEACH	ind	ajar	AV	verb	3	-kan	2	OBLIG-PV+CONT	no-chg	teach (s.t.) to s.o. / (s.o.) (s.t.)	agt, cgnr, cont	agt, cgnr, cont
78	TEACH	jav	ajar	AV	verb	1	-aké	2	OPT+CONT	no-chg	teach (cont)	agt, cont, cgnr=PP	cgnr, cont, cgnr
78	TEACH	jav	ajar	AV	verb	1	-i	3	REC	ditr	teach (s.o.) (s.t.)	agt, cont, cgnr=PP	agt, cgnr, cont
78	TEACH	mnb	fo-guru	fo-	verb	2	-i/-ghoo	na	na	na	na	agt, crtn, ben=IO	
78	TEACH	nsy	ajakh	AV	verb	2	-i	2	GOAL	remp	teach (s.o.) (s.t.)	agt, cont	agt, cgnr, cont=PP
78	TEACH	nsy	ajakh	AV	verb	2	-kun	2	OPT+CONT	no-chg	teach (s.t.) to (s.o.)	agt, cont	agt, cgnr, cont=PP
78	TEACH	sas	ajah	bare?	verb	2	-an	3	BEN	ditr	teach (s.o.) for (s.o.)	agt, cgnr	agt, ben, cgnr
78	TEACH	sas	ajah	bare?	verb	2	-an	2	OPT+CONT	no-chg	teach (s.t.) to s.o.	agt, cont, cgnr=PP	agt, cont, cgnr=PP
78	TEACH	sun	ajar	AV	verb	3	-keun	3	OPT+CONT	no-chg	teach (s.o.) (s.t.) / (s.t.) to s.o.	agt, cgnr, cont	agt, cgnr, cont
78	TEACH	sun	ajar	AV	verb	3	pang- -keun	3	BEN	no-chg	teach (s.o.) for (s.o.)	agt, cgnr, cont	agt, ben, cgnr
78	TEACH	sun	ajar	AV	verb	3	-an	na	na	na	na	agt, cgnr, cont	
78	TEACH	yka	tolo'	paN-	verb	2	-an	2	REC	ditr	teach (s.o.) (s.t.)	agt, cont, cgnr=PP	agt, cont, cgnr=PP
79	COOK	bes	gulai	MID	verb	1	-ka	2	PAT	mntn	cook (s.t.)	crtr	crtr, crtn
79	COOK	blz	taring	AV	verb	2	-kon	2	BEN	no-chg	cook (s.t.) for (s.o.)	crtr, crtn	crtr, crtn, ben=PP
79	COOK	blz	taring	AV	verb	2	-ii	3	BEN	ditr	cook (s.o.) (s.t.)	crtr, crtn	crtr, ben, crtn
79	COOK	blz	taring	AV	verb	2	-i/-ii	na	na	na	na	crtr, crtn	
79	COOK	ind	masak	AV, bare	verb	2	-kan	3	BEN	ditr	cook (s.o.) (s.t.)	crtr, crtn	crtr, rec, crtn
79	COOK	ind	masak	AV, bare	verb	2	-i	na	na	na	na	crtr, crtn	
79	COOK	jav	masak	AV	verb	2	-aké	3	BEN	ditr	cook (s.o.) (s.t.)	crtr, crtn, ben=PP	crtr, rec, crtn
79	COOK	jav	masak	AV	verb	2	-i	na	na	na	na	crtr, crtn, ben=PP	
79	COOK	mnb	gau	ae-	verb	2	-i/-ghoo	na	na	na	na	crtr, crtn	
79	COOK	nsy	gulai	AV	verb	2	-kun	3	BEN	ditr	cook (s.o.) (s.t.)	crtr, crtn	agt, rec, pat
79	COOK	nsy	gulai	AV	verb	2	-i	na	na	na	na	crtr, crtn	
79	COOK	sas	periap	AV	verb	2	-an	3	BEN	ditr	cook (s.o.) (s.t.)	crtr, crtn	crtr, ben, crtn
79	COOK	sun	masak	AV	verb	2	pang- -keun	3	BEN	ditr	cook (s.o.) (s.t.)	crtr, crtn	crtr, rec, crtn
79	COOK	sun	masak	AV	verb	2	-an/-keun	na	na	na	na	crtr, crtn	
79	COOK	yka	bella	N-, mag-	verb	2	-an	na	na	na	na	crtr, crtn	
80	BOIL	blz	lua'	bare	verb	1	-i/-ii/-kon	na	na	na	na	pat	
80	BOIL	ind	mendidih	AV	verb	1	-kan	2	CAUS	mntn	make (s.t.) boil	pat	istg, pat
80	BOIL	ind	mendidih	AV	verb	1	-i	na	na	na	na	pat	
80	BOIL	jav	umob	bare	verb	1	-aké	2	CAUS	mntn	make (s.t.) boil	pat	istg, pat
80	BOIL	jav	umob	bare	verb	1	-i	na	na	na	na	pat	
80	BOIL	mnb	lua	a-	verb	1	-i	2	GOAL	mntn	boil over onto (s.t.)	pat	pat, goal
80	BOIL	mnb	lua	a-	verb	1	-ghoo	na	na	na	na	pat	
80	BOIL	sas	kelaq	bare	verb	1	-an	3	BEN	ditr	boil, cook (s.t.) for (s.o.)	pat	agt, ben, pat
80	BOIL	sun	golotrok	AV	verb	1	-keun	2	CAUS	mntn	make (s.t.) boil	pat	istg, pat
80	BOIL	sun	golotrok	AV	verb	1	-an/pang- -keun	na	na	na	na	pat	
80	BOIL	yka	bukal	N-, mag-	verb	1	-an	na	na	na	na	pat	
81	PLANT	bes	tanam	AV	verb	2	-i	2	LOC	remp	plant (s.w.) w/ s.t.	agt, thm, loc=PP	agt, loc, thm=PP
81	PLANT	bes	tanam	AV	verb	2	-ka	2	OPT+THM	no-chg	plant (s.t.)	agt, thm, loc=PP	agt, thm
81	PLANT	blz	asok	AV	verb	2	-i	3	LOC	ditr	plant (s.w.) w/ s.t.	agt, thm, loc=PP	agt, loc, thm
81	PLANT	blz	asok	AV	verb	2	-ii/-kon	na	na	na	na	agt, thm, loc=PP	
81	PLANT	ind	tanam	AV	verb	2	-i	2	LOC	remp	plant (s.w.) w/ s.t.	agt, thm, loc=PP	agt, loc, thm=PP
81	PLANT	ind	tanam	AV	verb	2	-kan	2	OPT+THM	no-chg	plant (s.t.) in s.w.	agt, thm, loc=PP	agt, thm, loc=PP
81	PLANT	jav	tandur	AV	verb	2	-i	3	LOC	ditr	plant (s.w.) w/ s.t.	agt, thm	agt, loc, thm
81	PLANT	mnb	tisa	a-	verb	2	-i	2	PLUR	no-chg	plant (many)	agt, thm, loc=IO	agt, thm
81	PLANT	mnb	tisa	a-	verb	2	-ghoo	na	na	na	na	agt, thm, loc=IO	
81	PLANT	nsy	takhuk	AV	verb	2	-i	2	PLUR	no-chg	plant (many)	agt, thm	agt, thm
81	PLANT	nsy	takhuk	AV	verb	2	-i	2	LOC	remp	plant (s.w.) w/ s.t.	agt, thm	agt, loc, thm=PP

Table E.1: Lexical data by base meaning (cont.)

ID	Meaning	ISO	Base	BsMorph	BsCls	BsV	Affix	AfV	Function	SynChg	AfMeaning	BsStr	AfStr
81	PLANT	nsy	takhuk	AV	verb	2	-kun	2	OPT+THM	no-chg	plant (s.t.)	agt, thm	agt, thm
81	PLANT	nsy	takhuk	AV	verb	2	-kun	3	BEN	ditr	plant (s.t.) for (s.o.)	agt, thm	agt, ben, thm
81	PLANT	sas	talet	bare	verb	2	-an	2	LOC	mntr	plant (s.t.) in (s.w.)	agt, thm, loc=PP	agt, loc, thm
81	PLANT	sun	pelak	AV	verb	2	-an	2	LOC+PLUR	remp	plant (s.w.)	agt, thm, loc=PP	agt, loc, thm=PP
81	PLANT	sun	pelak	AV	verb	2	-keun	2	OPT+THM	no-chg	plant (s.t.)	agt, thm, loc=PP	agt, thm
81	PLANT	yka	tanem	N-, mag-	verb	2	-an	2	LOC+PLUR	remp	plant (s.w.) w/ s.t.	agt, thm, loc=PP	agt, loc, thm=PP
82	BORROW	bes	pinjam	AV?	verb	2	-ka	2	CAUS+THM	caus-remp	lend (s.t.) to s.o. for (s.o.)	agt, thm	agt, thm, rec=PP
82	BORROW	bes	pinjam	AV?	verb	2	-i	2	CAUS+REC	ditr	lend (s.o.) (s.t.)	agt, thm	istg, rec, thm
82	BORROW	blz	bolos	AV	verb	2	-i	2	LEX	no-chg	replace (s.t.)	agt, thm	agt, thm
82	BORROW	blz	bolos	AV	verb	2	-ii	3	CAUS+REC	ditr	lend (s.o.) (s.t.)	agt, thm	istg, rec, thm
82	BORROW	blz	bolos	AV	verb	2	-kon	3	CAUS+THM	ditr	lend (s.t.) to (s.o.)	agt, thm	istg, rec, thm
82	BORROW	ind	pinjam	AV	verb	2	-i	3	CAUS+REC	ditr	lend (s.o.) (s.t.)	agt, thm	istg, rec, thm
82	BORROW	ind	pinjam	AV	verb	2	-kan	2	CAUS+THM	caus-remp	lend (s.t.) to s.o.	agt, thm	istg, thm, rec=PP
82	BORROW	ind	pinjam	AV	verb	2	-kan	3	BEN	ditr	borrow (s.t.) for (s.o.)	agt, thm	agt, ben, thm
82	BORROW	jav	silih	AV	verb	2	-aké	3	BEN	ditr	borrow (s.t.) for (s.o.)	agt, thm	agt, ben, thm
82	BORROW	jav	silih	AV	verb	2	-aké	2	CAUS+THM	caus-remp	lend (s.t.) (to s.o.)	agt, thm	istg, thm, rec=PP
82	BORROW	jav	silih	AV	verb	2	-i	3	CAUS+REC	ditr	lend (s.t.) to (s.o.)	agt, thm	agt, rec, thm
82	BORROW	mnb	ada	ao-	verb	2	-ghoo+pa-	3	REC	ditr	loan (s.o.) (s.t.)	agt, thm	agt, rec=IO, thm
82	BORROW	mnb	ada	ao-	verb	2	-i	na	na	na	na	agt, thm	
82	BORROW	nsy	pinjam	AV	verb	2	-i	3	CAUS+REC	ditr	lend (s.o.) (s.t.)	agt, thm	istg, rec, thm
82	BORROW	nsy	pinjam	AV	verb	2	-kun	3	BEN	ditr	borrow (s.t.) for (s.o.)	agt, thm	agt, ben, thm
82	BORROW	sas	singgaq	bare?	verb	2	-an	2	CAUS+THM	caus-remp	lend (s.t.) (to s.o.)	agt, thm	istg, thm, rec=PP
82	BORROW	sun	injeum	AV	verb	2	-an	2	PLUR	no-chg	borrow (s.t.) rep.	agt, thm, src=PP	agt, thm
82	BORROW	sun	injeum	AV	verb	2	-keun	2	CAUS+THM	caus-remp	lend (s.t.) to s.o. for (s.o.)	agt, thm	istg, thm, rec=PP
82	BORROW	sun	injeum	AV	verb	2	pang- -keun	3	CAUS+BEN	caus-remp	lend (s.t.) for (s.o.)	agt, thm	istg, ben, thm
82	BORROW	sun	injeum	AV	verb	2	pang- -keun	3	CAUS	ditr	borrow for (s.o.) (s.t.)	agt, thm	agt, ben, thm
82	BORROW	yka	indam	N-, mag-	verb	2	-an+pa-	3	REC	ditr	loan (s.o.) (s.t.)	agt, thm	istg, rec, thm

Appendix F. Glossary of terms

- A:** This grammatical macro-role represents the most agent-like argument in a basic transitive clause (see Comrie 1989).
- A-oriented:** A diathesis or voice construction is A-oriented if the most syntactically privileged argument in such a construction is the most agent-like argument, or A.
- Addressee:** The addressee semantic role refers to an entity that is the recipient (or intended recipient) of a message in an event of communication (Johnson & Fillmore 2000: 59).
- Agent:** The agent semantic role represents a participant that intentionally engages in an action by which some other entity is affected (see Van Valin & LaPolla 1997: 85).
- Applicative:** An applicative, or applicative construction (AC), is a clausal construction in which overt morphological marking on the verbal complex coincides with the selection of a non-agent, non-patient semantic role to map to a core argument in the clause.
- Applicative morpheme (AM):** A morpheme or morphological process which is observed to mark the verbal complex in an applicative construction. Such a morpheme may be polyfunctional; an AM that is part of the fixed form of one or more applicative constructions may also be associated with a non-applicative meaning, such as causative, pluractional, or intensive semantics, in other constructions.
- Base construction (BC):** A clausal construction in which no applicative morpheme marks the verbal complex. A base construction with the same lexical verb may be compared to an applicative construction exhibiting marking with an applicative morpheme.
- Beneficiary:** The beneficiary semantic role refers to an entity that is affected advantageously by an event without being the agent or primary affected participant of the event (Kittilä & Zúñiga 2010: 2).
- Aspect:** The term aspect, or aspectual refers to perspectives on the unfolding of a state or event in time (e.g. perfective, imperfective, progressive) (Comrie 1976).
- AV:** A-Voice (sometimes also called actor voice) describes a grammatical voice construction in which the A argument is the most syntactically privileged argument. In a symmetrical voice language, AV typically represents one of multiple transitive voice alternations available.
- Base:** In an applicative construction, base refers to the root or stem of the main verb in the clause. Applicative morphology is usually affixed to the base rather than some other part of the verbal complex. Some languages allow use of intransitive verbal bases only in applicatives, while others allow monotransitive or even ditransitive verbal bases. In some languages, bases may also belong to categories other than verb, especially nominal bases.

- Causative construction:** A causative construction is a type of diathetical operation that increases the semantic valency of a verb (Zúñiga & Kittilä 2019: 14). In a causative construction, an instigating causer is introduced to the argument structure of a verb as an A argument.
- Circumstantial role:** A circumstantial semantic role describes the circumstances under which a state or event occurs. I use this label to refer to the reason that a state or event occurs or the intended purpose for which a state or event is initiated. My use of the term circumstantial in reference to semantic role is narrower than the use of the term in the label *circumstantial voice*, see below.
- Circumstantial voice:** The term circumstantial voice (CV) refers to a category of transitive voice alternations in a symmetrical voice system, namely, alternations in which the pivot has the semantic role of beneficiary, recipient, theme, instrument, reason, or purpose. Circumstantial voice is defined in this manner because the mapping of such roles to the pivot relation are all associated with one reconstructed category of voice morphology in the verbal paradigm for Proto Austronesian (see Wolff 1973; Chen 2017: 103). In specific Philippine-type languages, voice constructions which fall under the circumstantial voice category may be referred to using different labels; this is particularly useful in languages which have multiple CV constructions, with each selecting a different subset of possible roles as pivot. Such labels as used in the literature on Austronesian languages are diverse and include: benefactive voice (BV), conveyance voice (ConV), instrumental voice (IV), theme voice (TV), dative voice (DV), and referent voice (RV), among others.
- Coding of arguments:** This term refers to properties of the syntactic forms by which clausal arguments are realized, including the case, agreement, and phrase structure by which constituents expressing clausal arguments are marked, and which may co-vary with grammatical relations (Bickel 2010: 28).
- Comitative:** The comitative semantic role describes a participant that accompanies another participant in an event.
- Construction:** A construction is a conventionalized pairing of a fixed form and consistent meaning (see e.g. Goldberg 1995).
- Content:** The content semantic role refers to the content or topic of act of communication or cognition.
- Diathesis:** The term diathesis refers to specific assignments or mappings of semantic roles to grammatical relations in a clausal construction (Zúñiga & Kittilä 2019: 4). A diathetical alternation exists between two clausal constructions that differ in mapping of semantic roles to grammatical relations. Diathetical alternations need not be distinguished by morphological marking on the predicate; when they are, such alternations may be referred to as grammatical voice alternations.
- Durative aspect:** An event, or state extends over some period of time.
- Experiencer:** The experiencer semantic role describes an entity, typically a person or other sentient being, that experiences a perceptual experience, emotion, sensation, or another internal state.
- Goal:** The goal semantic role refers to the ending location of an entity that changes location.

Grammatical relations: As used in this study, the term grammatical relations describes the syntactic relations between a clause and its arguments, and these relations are understood to be construction-specific (Bickel 2010). Grammatical relations are identified by the syntactic properties of arguments, including their coding properties, and their behavioral properties (syntactic behavior).

Habitual aspect: An aspectual category indicating that an action, state, or event is characteristic for some entity, or in some situational context.

Instigating causer: In a causative construction, an instigating causer is a semantic participant whose role is to incite an event or bring about a state, without being responsible for the performance of the event, or directly participating in the stative condition (see Zúñiga & Kittilä 2019: 15). This role is also sometimes called an external causer or external agent.

Instrument: The instrument semantic role refers to an inanimate entity manipulated to some effect (see Van Valin 2005: 58–59; Fillmore 1968).

Intensive: A semantic category indicating a greater than normal degree of a quality, especially force, potency, or effort.

Iterative aspect: An aspectual category indicating that an action or event is repeated during some period of time.

Location: The location semantic role describes the general or static location of an entity or event.

Locative expression: A locative expression serves as the basis for describing the spatial position of an entity or event. Certain semantic roles are locative expressions, such as location, source, path, direction, or goal.

Locative voice: The term locative voice (LV) refers to a category of transitive voice alternations in a symmetrical voice system, namely, alternations in which the pivot has the semantic role of a generalized or static location. In some languages, a phrase expressing the location of an event in time, i.e. a temporal phrase, may be also be the pivot.

Mood: Mood or modality refers to semantic meanings expressed by a speaker regarding the proposition of an utterance (e.g. actuality, potentiality, desiderative). Such functions may be expressed by grammatical categories in a language, which is grammatical mood or mode (e.g. realis/irrealis, indicative/optative/imperative) (Bybee & Fleischman 1995)

Morphological alignment: Morphological alignment refers to patterning of the morphological marking of nominal constituents expressing the core argument relations S, A, and P. Evidence of morphological alignment may include the use of case markers, articles, or other grammatical particles, as well as paradigmatic alternation in the forms of nouns and pronominal indexes when these co-vary with the core argument relations.

Path: The path semantic role describes the trajectory of an entity that changes location with respect to some ground over which the entity travels or landmark by which the entity travels.

P: This grammatical macro-role represents the most patient-like argument in a basic transitive clause (see Comrie 1989).

Passive: A type of diathesis or voice construction in which the P argument is syntactically privileged, while the A argument maps to a non-core relation or is suppressed. A passive is a type of P-oriented construction and typically may take a single core argument, which corresponds to P in basic transitive clauses.

P-oriented: A diathesis or voice construction is P-oriented if the most syntactically privileged argument in such a construction is the most patient-like argument, or P.

Patient: The patient semantic role refers to an entity that exists in an indicated state or condition, or undergoes a change of state or condition (Van Valin & LaPolla 1997: 85).

Peripheral semantic role: Peripheral semantic role refers to a non-agentive, non-patientive semantic role. This category excludes the roles that describe the two core arguments of a prototypical transitive verb, like ‘kill’ or ‘hit’, that is an agentive actor that intentionally engages in an action, and an affected patient that undergoes a change of state.

Performance: The performance semantic role refers to the experience generated by performers and perceived by the audience in a performing arts event.

Person-indexing: The use of bound forms on a verb or verbal complex, expressing grammatical categories for person (e.g. first, second, third person) and number (e.g. singular, plural) in reference to arguments of the verb (Haspelmath 2013). Agreement marking, argument co-reference on the verb, and the expression of clausal arguments as pronominal clitics or affixes are all types of person-indexing as defined here.

Philippine-type voice system: A Philippine-type voice system is a kind of symmetrical voice system characterized by alternations between more than one nonactor oriented transitive voice construction. Such alternations are distinguished by morphological marking on the verb and mapping of semantic roles to a syntactically-privileged grammatical relation.

Pivot: As used in this study, pivot refers a type of grammatical relation relevant in symmetrical voice systems. In such systems, the pivot represents the grammatical relation in a clause that shows unique syntactic privilege, by means of access to syntactic operations. This relation is sometimes referred to as the subject, and has also been called focus, trigger, or topic by various authors in the literature on Austronesian languages.

Pivot-neutral applicative construction: I use this term to refer to applicative constructions in symmetrical voice languages in which the applied phrase is not necessarily selected as the pivot argument, i.e. the most syntactically privileged argument in the clause. Pivot-neutral applicative constructions typically may combine with other major voice constructions (e.g. AV, PV, passive), and this conditions the mapping of the applied phrase to argument structure. In many but not all pivot-neutral applicative constructions investigated in this study, the applied phrase may be the pivot argument in P-oriented constructions, e.g. PV, passive.

Pivot-selecting applicative construction: I used this term to refer to applicative constructions in symmetrical voice languages in which the applied phrase always is selected as the pivot argument, i.e. the most syntactically privileged argument in the clause. Pivot-selecting applicative constructions generally include all voice constructions in Philippine-type voice systems that select a peripheral semantic role as the pivot, e.g. benefactive voice (BV), circumstantial voice (CV), instrumental voice (IV), and/or locative voice (LV) constructions.

Pluractional aspect: This category describes events comprised of multiple instances, situations, or participants, including repeated, iterative, and habitual events, and events with multiple actors and/or undergoers (see Wood 2007; Mattiola 2019).

Purpose: This semantic role refers to a future circumstance which motivates the occurrence of an event, i.e. the purposeful intentions that prompt an event’s occurrence.

PV: P-Voice (also sometimes called patient voice, or undergoer voice) describes a grammatical voice construction in which the P argument is the most syntactically privileged argument. In a symmetrical voice language, PV typically represents one of multiple transitive voice alternations available.

R: This grammatical macro-role refers to the argument in a ditransitive clause which is recipient-like, i.e. receives possession of a transferred entity (see Haspelmath 2015).

Reason: This semantic role describes the prior circumstances due to which an event occurs. This role is sometimes also called cause, or causal role, but in order to distinguish this role from that of instigating causers in causative constructions, I have avoided use of such labels.

Recipient: The recipient semantic role refers to the entity that receives possession of an entity in a transfer event.

S: This grammatical macro-role represents the sole core argument in a basic intransitive clause (see Comrie 1989).

Source: The source semantic role refers to the starting location of an entity that changes location.

Stimulus: The stimulus semantic role refers to the entity, event, or phenomenon that is perceived and/or brings about a sensation, or other internal state (see Van Valin & LaPolla 1997: 55)..

Syntactic alignment: Syntactic alignment refers to patterning of syntactic properties that co-vary with the core argument relations S, A, and P. Evidence of syntactic alignment includes word order and access to syntactic operations, which may be termed syntactic behavior.

Syntactic behavior: The term syntactic behavior or behavioral properties, is used to describe observable properties of clausal arguments that implicate syntactic structure. These include word order, and access to syntactic operations like relativization, control, raising, quantifier float, and reflexive binding (see Keenan 1976).

Symmetrical voice: A symmetrical voice system is characterized by alternations between more than one transitive voice construction, which are distinguished by morphological marking on the verb, and mapping of semantic roles to a syntactically-privileged grammatical relation.

T: This grammatical macro-role refers to the argument in a ditransitive clause that is theme-like, i.e. an entity which is transferred (see Haspelmath 2015).

Target: This semantic role refers to a participant towards which a emoter's feelings are directed (Van Valin 2005: 55).

TAM: This abbreviation refers to tense, aspect, and mood (or mode). This designation applies to grammatical categories in a language that reflect semantic dimensions of a state, event, or proposition.

Tense: The term tense refers to grammatical categories expressing the position of a state or event in time (e.g. past, present, future) (Comrie 1985b).

Theme: The theme semantic role refers to an entity that is located in space or undergoes a change of location, position, or possession. (see Van Valin & LaPolla 1997: 85).

References

- Adelaar, K. Alexander. 1984. Some Proto-Malayic affixes. *Bijdragen tot de Taal-, Land- en Volkenkunde* 140(4). 402–421. <http://www.jstor.org/stable/27863614>.
- Adelaar, K. Alexander. 1992. *Proto Malayic: The reconstruction of its phonology and parts of its lexicon and morphology* (Pacific Linguistics C 119). Canberra: Pacific Linguistics and Australia National University.
- Adelaar, K. Alexander. 1994. The classification of the Tamanic languages (West Kalimantan). In Thomas Dutton & Darrell T. Tryon (eds.), *Language contact and change in the Austronesian world* (Trends in Linguistics Studies and Monographs 77), 1–41. Berlin: Mouton de Gruyter.
- Adelaar, K. Alexander. 1995. Problems of definiteness and ergativity in Embaloh. *Oceanic Linguistics* 34(2). 375. <https://doi.org/10.2307/3623049>.
- Adelaar, K. Alexander. 2002. Salako morphology and the interrelation between voice, mood and aspect. In *Between worlds: Linguistic papers in memory of David John Prentice*, 1–27. Canberra: Pacific Linguistics and Australia National University.
- Adelaar, K. Alexander. 2005a. Malayo-Sumbawan. *Oceanic Linguistics* 44(2). 357–388. <https://doi.org/10.1353/ol.2005.0027>.
- Adelaar, K. Alexander. 2005b. *Salako or Badameá: Sketch grammar, text and lexicon of a Kanayatn dialect in West Borneo*. Vol. 2 (Frankfurter Forschungen zu Südostasien). Wiesbaden: Harrassowitz.
- Adelaar, K. Alexander. 2006. Where does Belangin belong? In Fritz Schulze & Holger Warnk (eds.), *Insular Southeast Asia: Linguistic and cultural studies in honour of Bernd Nothofer*, 65–84. Wiesbaden: Harrassowitz.
- Adelaar, K. Alexander. 2009. Towards an integrated theory about the Indonesian migrations to Madagascar. In Peter Peregrine, Ilia Peiros & Marcus Feldman (eds.), *Ancient human migrations: A multidisciplinary approach*, 149–172. Salt Lake City: University of Utah Press.
- Adelaar, K. Alexander. 2011. Javanese *-aké* and *-akən*: A short history. *Oceanic Linguistics* 50(2). 338–350.
- Adelaar, K. Alexander. 2012. Malagasy phonological history and Bantu influence. *Oceanic Linguistics* 51(1). 123–159.
- Adriani, N. 1928. Spraakkunstige schets van de taal der Mentawai-eilanden. *Bijdragen tot de taal-, land- en volkenkunde / Journal of the Humanities and Social Sciences of Southeast Asia* 84(1). 1–17. <https://doi.org/10.1163/22134379-90001493>.
- Adriani, Nicolaus & Samuel J. Esser. 1939. *Koelawische taalstudien* (Bibliotheca Celebica I, I, III). Bandoeng: Nix & Co.

- Aikhenvald, Alexandra. 2006. Serial verbs constructions in a typological perspective. In A. Y. Aikhenvald & R. M. W. Dixon (eds.), 1–68. Oxford, U.K.: Oxford University Press. https://doi.org/10.8/9310_Aikhenvald_2006_compressed.pdf.
- Aikhenvald, Alexandra Y. & R. M. W. Dixon. 2011. A typology of argument-determined constructions. In *Language at large: Essays on syntax and semantics* (Empirical Approaches to Linguistic Theory vol. 2), 44–85. Leiden; Boston: Brill. <https://doi.org/10.1163/ej.9789004206076.i-606.85>.
- Akamine, Jun. 2003. *A basic grammar of Southern Sinama*. Vol. A3-012 (ELPR A3-012). Osaka Gakuin University, Japan: Endangered Languages of the Pacific Rim, Faculty of Informatics.
- Alsina, Alex & Sam A. Mchombo. 1990. The syntax of applicatives in Chicheŵa: Problems for a theta theoretic asymmetry. *Natural Language & Linguistic Theory* 8(4). 493–506.
- Alsina, Alex & Sam A. Mchombo. 1993. Object asymmetries and the Chichewa applicative construction. In Sam A. Mchombo (ed.), *Theoretical aspects of Bantu grammar 1* (CSLI Lecture Notes 38), 17–45. Stanford, CA: CSLI Publications.
- Anceaux, J. C. 1952. *The Wolio language: Outline of grammatical description and texts* (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde ; deel 11). 's-Gravenhage: Martinus Nijhoff.
- Anclom. 2019. In *Wikikamus Basa Sunda*. <https://su.wiktionary.org/wiki/anclom> (5 February, 2024).
- Anderbeck, Karl & Jonathan McDowell. 2020. *The Malay lects of southern Sumatra* (JSEALS Special Publication 7). Honolulu, HI: University of Hawai'i Press.
- Andersen, Suree & T. David Anderson. 2005. Semantic analysis of the Moronene verbal prefix *moN-*. In I Wayan Arka & Malcolm Ross (eds.), *The many faces of Austronesian voice systems: Some new empirical studies* (Pacific Linguistics 571), 243–278. Canberra: Pacific Linguistics and Australia National University.
- Andersen, T. David. 2013. *Topicality and functional voice in Hebrew and Moronene, with application to translation* (SIL E-Books 56). SIL International. <https://www.sil.org/resources/publications/entry/53965>.
- Arka, I Wayan. 1993. *Morpholexical aspects of the -kan causative in Indonesian*. Sydney: University of Sydney MA thesis.
- Arka, I Wayan. 2003. *Balinese morphosyntax: A Lexical-Functional approach* (Pacific Linguistics 547). Canberra: Pacific Linguistics.
- Arka, I Wayan, Mary Dalrymple, Meladel Mistica, Suriel Mofu, Avery Andrews & Jane Simpson. 2009. A linguistic and computational morphosyntactic analysis for the applicative-*i* in Indonesian. In Miriam Butt & Tracey Holloway King (eds.), *Proceedings of the LFG09 Conference*, 85–105. Stanford, CA: CSLI Publications.
- Artawa, Ketut. 1998. Ergativity and Balinese Syntax: Part I. NUSA: Linguistic Studies of Indonesian and Other Languages in Indonesia (42).
- Austin, Peter. 2001. Verbs, valence and voice in Balinese, Sasak and Sumbawan. *La Trobe Papers in Linguistics* 11(3). 47–71.
- Aye, Daw Khin Khin. 2005. *Bazaar Malay: History, grammar and contact*. National University of Singapore MA thesis. <https://scholarbank.nus.edu.sg/handle/10635/15028>.
- Aznar, Jocelyn, Christian Döhler & Jozina Vander Klok (eds.). 2023. *Applicatives in Austronesian Languages* (NUSA, Linguistic Studies of Indonesian and Other Languages in Indonesia 74). <https://doi.org/10.15026/0002000084>.

- Baker, Mark C. 1988a. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.
- Baker, Mark C. 1988b. Theta theory and the syntax of applicatives in Chichewa. *Natural Language and Linguistic Theory* 6(3). 353–389. <https://doi.org/10.1007/BF00133903>.
- Barr, Donald F. 1988a. Da'a verbal affixes and clitics. In Hein Steinhauer (ed.), *Papers in western Austronesian linguistics, no. 4* (Pacific Linguistics A 79), 11–49. Canberra: Pacific Linguistics and Australia National University.
- Barr, Donald F. 1988b. Focus and mood in Da'a discourse. In *Papers in western Austronesian linguistics, no. 4* (Pacific Linguistics A 79), 77–129. Canberra: Pacific Linguistics and Australia National University. <https://doi.org/10.15144/PL-A79.77>.
- Behrens, Dietlinde (ed.). 2002. *Yakan-English Dictionary* (Philippine Journal of Linguistics : Special Monograph Issue 40, vol. 2). Manila: Linguistic Society of the Philippines. https://philippines.sil.org/resources/online_resources/yka.
- Belding, Joanna, Wyn D. Laidig & Sahabu Dg. Maingak. 2001. An initial description of Barang-Barang morphology. *NUSA: Linguistic Studies of Indonesian and Other Languages* 49. 1–60.
- Beratha, Ni Luh Sutjiati. 1992. *Evolution of verbal morphology in Balinese*. Canberra: Australian National University dissertation. <https://openresearch-repository.anu.edu.au/handle/1885/109364>.
- Bickel, Balthasar. 2010. Grammatical relations typology. In Jae Jung Song (ed.), *The Oxford Handbook of Language Typology*, 399–444. Oxford: Oxford University Press.
- Billings, Blaine & Bradley McDonnell. Forthcoming. Sumatran. *Oceanic Linguistics*.
- Bintoro. 1980. *Javanese transitive verbs: A Tagmemic grammar* (NUSA: Linguistic Studies of Indonesian and Other Languages in Indonesia 8). Jakarta, Indonesia: Badan Penyelenggara Seri Nusa, Universitas Atma Jaya, Jakarta.
- Blust, Robert. 2003. Three notes on early Austronesian morphology. *Oceanic Linguistics* 42(2). 438–478. <https://doi.org/10.2307/3623246>.
- Blust, Robert A. 1988. Sketches of the morphology and phonology of Bornean languages 2: Mukah (Melanau). In Hein Steinhauer (ed.), *Papers in western Austronesian Linguistics, no. 3* (Pacific Linguistics A 78), 151–216. Canberra: Pacific Linguistics and Australia National University.
- Blust, Robert A. 1991. The Greater Central Philippine Hypothesis. *Oceanic Linguistics* 30(2). 73–129.
- Blust, Robert A. 1999. Language, dialect, and riotous sound change: The case of Sa'ban. In *Papers from the 9th Annual Meeting of the Southeast Asian Linguistics Society*, 249–359. <http://sealang.net/sala/archives/pdf4/blust2001language.pdf>.
- Blust, Robert A. 2005. The linguistic macrohistory of the Philippines: Some speculations. In Hsiuchuan Liao & Carl Rubino (eds.), *Current Issues in Philippine Linguistics and Anthropology: Parangal kay Lawrence A. Reid*, 31–68. Manila: Linguistic Society of the Philippines and SIL Philippines.
- Blust, Robert A. 2010. The Greater North Borneo hypothesis. *Oceanic Linguistics* 49(1). 44–118.
- Bowden, John. 2001. *Taba: Description of a South Halmahera language* (Pacific Linguistics 521). Canberra: Pacific Linguistics and Australia National University. <https://doi.org/10.15144/PL-521>.
- Brainard, Sherri & Dietlinde Behrens. 2002. *A grammar of Yakan* (Special Monograph Issue 40). Manila: Linguistic Society of the Philippines.

- Breiman, Leo. 2001. Random forests. *Machine Learning* 45. 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Bresnan, Joan, Anna Cueni, Tatiana Nikitina & R. Harald Baayen. 2007. Predicting the dative alternation. In Gerlof Bouma & Zwarts (eds.), *Cognitive foundations of interpretation*, 69–94. Amsterdam: Royal Netherland Academy of Arts and Sciences.
- Bresnan, Joan & Lioba Moshi. 1990. Object asymmetries in comparative Bantu syntax. *Linguistic Inquiry* 21(2). 147–185.
- Bresnan, Joan & Lioba Moshi. 1993. Object asymmetries in comparative Bantu syntax. In Sam A. Mchombo (ed.), *Theoretical aspects of Bantu grammar 1* (CSLI Lecture Notes 38), 47–92. Stanford, CA: CSLI Publications. <http://hdl.handle.net/2123/12669>.
- Brown, Lea. 2001. *A grammar of Nias Selatan*. Sydney: University of Sydney dissertation. <http://hdl.handle.net/2123/12669>.
- Brown, Lea. 2005. Nias. In K. Alexander Adelaar & Nicholas P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar* (Routledge Family Language Series), 562–589. London; New York: Routledge.
- Brunelle, Marc. 2020. The loss of affixation in Cham. In David Gil & Antoinette Schapper (eds.), *Austronesian Undressed: How and why languages become isolating* (Typological Studies in Language), 97–118. Amsterdam; Philadelphia: John Benjamins Publishing Company. <https://doi.org/10.1075/tsl.129.02bru>.
- Busenitz, Robert L. & Daniel Bradbury (eds.). 2016. *Balantak dictionary*. SIL International. <https://www.webonary.org/balantak/> (8 February, 2024).
- Bybee, Joan L. & Suzanne Fleischman. 1995. *Modality in Grammar and Discourse* (Typological Studies in Language 32). Amsterdam; Philadelphia: John Benjamins Publishing. <https://doi.org/10.1075/tsl.32>.
- Campbell, Phillip J. 1989. *Some aspects of Pitu Ulunna Salu grammar: A typological approach*. Arlington, TX: University of Texas at Arlington MA thesis.
- Carochi, Horacio. 1981. *Arte de la lengua mexicana con la declaracion de los adverbios della*. Mexico City: Editorial Innovación.
- Chen, Victoria. 2017. *A reexamination of the Philippine-type voice system and its implications for Austronesian primary-level subgrouping*. Honolulu: University of Hawai'i at Mānoa dissertation.
- Chen, Victoria & Bradley McDonnell. 2019. Western Austronesian Voice. *Annual Review of Linguistics* 5(1). 173–195. <https://doi.org/10.1146/annurev-linguistics-011718-011731>.
- Chou, Shu Hsiu. 1999. *Melanau and Indonesian: A comparative grammatical study*. Northern Territory University MA thesis.
- Chung, Sandra. 1976. An object-creating rule in Bahasa Indonesia. *Linguistic Inquiry* 7(1). 41–87.
- Clayre, Beatrice. 1996. The changing face of focus in the languages of Borneo. *Pacific Linguistics. Series A. Occasional Papers* 84. 51–88.
- Clayre, Beatrice. 1997. Verb affixes in Berawan. In *Proceedings of the seventh International Conference on Austronesian Linguistics: Leiden 22-27 August 1994*, 232–255. Amsterdam: Rodopi.
- Clayre, Beatrice. 2005. Kelabitic languages and the fate of ‘focus’: Evidence from the Kerayan. In I Wayan Arka & Malcolm Ross (eds.), *The many faces of Austronesian voice systems: Some new empirical studies* (Pacific Linguistics 571), 17–57. Canberra: Pacific Linguistics and Australia National University.

- Clayre, Iain F. C. S. 1972. *A grammatical description of Melanau*. University of Edinburgh dissertation.
- Clynes, Adrian. 1995a. Sundanese. In Darrell T. Tryon (ed.), *Comparative Austronesian Dictionary: An introduction to Austronesian studies*, vol. 1 (Trends in Linguistics. Documentation. 10), 459–468. Berlin; New York: Mouton de Gruyter.
- Clynes, Adrian. 1995b. Topics in the phonology and morphosyntax of Balinese based on the dialect of Singaraja, North Bali. <https://doi.org/10.25911/5d77865d38e15>.
- Clynes, Adrian. 2001. Brunei Malay: An overview. *Occasional Papers in Language Studies* 7. 11–43.
- Clynes, Adrian. 2005. Belait. In *The Austronesian Languages of Asia and Madagascar* (Routledge Language Family Series), 429–455. London & New York: Routledge.
- Cohen, Nancy H. 1999. Nabay verb morphology. *SIL Electronic Working Papers* 1999–003. 1–50. <http://www.sil.org/silewp/1999/003/>.
- Cole, Peter & Min-Jeong Son. 2004. The argument structure of verbs with the suffix *-kan* in Indonesian. *Oceanic Linguistics* 43(2). 339–364.
- Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems* (Cambridge Textbooks in Linguistics). Cambridge ; Cambridge University Press.
- Comrie, Bernard. 1985a. Causative verb formation and other verb deriving morphology. In Timothy Shopen (ed.), *Language typology and syntactic description*, 1st edn., vol. 3. Cambridge ; Cambridge University Press.
- Comrie, Bernard. 1985b. *Tense* (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.
- Comrie, Bernard. 1989. *Language universals and linguistic typology: Syntax and morphology*. 2nd edn. Oxford: Blackwell.
- Connell, Timothy M. 2013. *A sketch grammar of Matéq: A Land Dayak language of West Kalimantan*. University of Canterbury MA thesis.
- Connors, Thomas J. 2008. *Tengger Javanese*. New Haven, CT: Yale University dissertation.
- Coolsma, Sierk. 1904. *Soendaneesche spraakkunst*. Leiden: A.W. Sijthoff.
- Coolsma, Sierk. 1913. *Soendaneesch-Hollandsch woordenboek*. Leiden: A.W. Sijthoff.
- Creissels, Denis. 2004. Non-canonical applicatives and focalization in Tswana. Paper presented at the First Syntax of the World's Languages Conference (SWL 1), 5–8 August, Leipzig.
- Creissels, Denis. 2010. Benefactive applicative periphrases: A typological approach. In Fernando Zuñiga & Seppo Kittilä (eds.), *Benefactives and Malefactives: Typological perspectives and case studies* (Typological Studies in Language 92), 29–70. Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Creissels, Denis. 2014. Cross-linguistic variation in the treatment of beneficiaries and the argument vs. adjunct distinction. *Linguistic Discovery* 12(2). <https://doi.org/10.1349/PS1.1537-0852.A.445>.
- Crowley, Terry. 2002. *Serial verbs in Oceanic: A descriptive typology*. Oxford: Oxford University Press.
- Crowley, Terry. N.d. Grammatical sketch of Enggano. ms.
- Dardjowidjojo, Soenjono. 1971. The *men-*, *men-kan* and *men-i* verbs in Indonesian. *Philippine Journal of Linguistics* 2(2). 71–84.
- Dardjowidjojo, Soenjono. 1974. The role of overt markers in some Indonesian and Javanese passive sentences. *Oceanic Linguistics* 13(1). 371–389.

- Davies, William D. 2005. The richness of Madurese voice. In I Wayan Arka & Malcolm Ross (eds.), *The many faces of Austronesian voice systems: Some new empirical studies*. Canberra: Pacific Linguistics.
- Davies, William D. 2010. *A grammar of Madurese* (Mouton Grammar Library 50). Berlin; New York: De Gruyter Mouton.
- del Rincón, Antonio. 1885. *Arte mexicana compuesta por el padre Antonio Del Rincón de la compañía de Jesus : Dirigido al ilustrissimo y reverendissimo S. Don Diego Romano Obispo de Tlaxcallan, y del consejo de Su Magestad, &c. en Mexico en casa de Pedro, Balli. 1595. Se reimprime en 1885 bajo el cuidado del Dr. Antonio Peñafiel*. Mexico City: Oficina tip. de la Secretaría de fomento.
- Diedrich, Daniela. 2018. *A grammar of Paku: A language of Central Kalimantan*. Melbourne: University of Melbourne dissertation.
- Dillon, John A. 1994. *A grammatical description of Tatana'*. Ann Arbor: University of Michigan MA thesis. ix 155.
- Dixon, R. M. W. 1972. *The Dyirbal language of North Queensland* (Studies in Linguistics 9). Cambridge: Cambridge University Press.
- Dixon, R. M. W. 2012. *Basic Linguistic Theory Volume 3: Further grammatical topics*. Vol. 3. Oxford: Oxford University Press.
- Dixon, R. M. W. & Alexandra Y. Aikhenvald (eds.). 2000. *Changing valency: Case studies in transitivity*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511627750>.
- Donohue, Mark. 1994. *Tukang Besi*. *Journal of the International Phonetic Association* 24(1). 39–41. <https://www.jstor.org/stable/44526156> (1 March, 2024).
- Donohue, Mark. 1996. *Bajau: A symmetrical Austronesian language*. *Language* 72(4). 782–793. <https://doi.org/10.2307/416102>.
- Donohue, Mark. 1999. *A grammar of Tukang Besi* (Mouton Grammar Library 20). Hawthorne, NY: Mouton de Gruyter.
- Donohue, Mark. 2001. Coding choices in argument structure: Austronesian applicatives in texts. *Studies in Language* 25(2). 217–254. <https://doi.org/10.1075/sl.25.2.03don>.
- Donohue, Mark. 2002. Voice in *Tukang Besi* and the Austronesian focus system. In Fay Wouk & Malcolm Ross (eds.), *The history and typology of western Austronesian voice systems*, 81–100. Canberra: Pacific Linguistics and Australia National University. <http://sealang.net/archives/pl/pdf/PL-518.81.pdf> (22 February, 2022).
- Donohue, Mark & Charles E. Grimes. 2008. Yet more on the position of the languages of Eastern Indonesia and East Timor. *Oceanic Linguistics* 47(1). 114–158. <https://doi.org/10.1353/ol.0.0008>.
- Durie, Mark. 1985. *A grammar of Acehnese on the basis of a dialect of North Aceh* (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 112). Dordrecht, Holland ; Foris Publications.
- Durie, Mark. 1988. Verb serialization and “verbal-prepositions” in Oceanic languages. *Oceanic Linguistics* 27(1/2). 1–23. <https://doi.org/10.2307/3623147>.
- Eades, Domenyk. 2005. *A grammar of Gayo: A language of Aceh, Sumatra* (Pacific Linguistics 567). Canberra: Pacific Linguistics.
- Eberhard, David & Charles D. Fennig (eds.). 2023. *Ethnologue: Languages of the world*. Twenty-sixth edn. Dallas, TX: SIL International. <http://www.ethnologue.com>.
- Eberhard, David M., Gary F. Simons & Charles D. Fennig (eds.). 2021. *Ethnologue: Languages of the world*. Twenty-fourth edn. Dallas, TX: SIL International.

- Edwards, Owen. 2012. *Grammatical functions in Tolaki*. Canberra: Australian National University MA thesis.
- Edwards, Owen. 2015. The position of Enggano within Austronesian. *Oceanic Linguistics* 54(1). 54–109. <https://doi.org/10.1353/ol.2015.0001>.
- Eringa, F. S. 1984. *Soendaas-Nederlands woordenboek*. Dordrecht, Netherlands: Foris Publications.
- Ernanda, N. 2017. *Phrasal alternation in Kerinci*. Utrecht: Leiden University dissertation. <https://hdl.handle.net/1887/49206>.
- Esser, Samuel J. 2011. *Phonology and morphology of Mori*. Trans. by David Mead (SIL E-Books 27). Dallas: SIL International.
- Evans, Bethwyn. 2003. *A study of valency-changing devices in Proto Oceanic* (Pacific Linguistics 539). Canberra: Pacific Linguistics and Australia National University.
- Evans, Donna. 1996. Causation in Kaili. In Hein Steinhauer (ed.), *Papers in Austronesian Linguistics, no. 3*, vol. 84 (Pacific Linguistics A), 173–189. Canberra: Research School of Pacific and Asian Studies, Australian National University.
- Evans, Donna (ed.). 2003. *Kamus Kaili-Ledo - Indonesia - Inggris*. Palu: Pemerintah Daerah Propinsi Sulawesi Tengah, Dina Kebudayaan dan Pariwisata. <https://www.sil.org/resources/archives/37514>.
- Ewing, Michael C. 1999. *The clause in Cirebon Javanese conversation*. Santa Barbara, CA: University of California, Santa Barbara dissertation.
- Ewing, Michael C. 2019. The predicate as a locus of grammar and interaction in colloquial Indonesian. *Studies in Language* 43(2). 402–443. <https://doi.org/10.1075/sl.17062.ewi>.
- Fillmore, Charles J. 1968. The case for case. In Emmon Bach & Robert Thomas Harms (eds.), *Universals in Linguistic Theory*, 1–88. New York: Hold, Rinehart and Winston.
- Fillmore, Charles J. & Collin F. Baker. 2012. A frames approach to semantic analysis. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, 2nd edn. (Oxford Handbooks), 313–339. Oxford; New York: Oxford University Press.
- Foley, William A. 2007. A typology of information packaging in the clause. In Timothy Shopen (ed.), *Language typology and syntactic description: Volume 1: Clause structure*, 2nd edn., vol. 1, 362–446. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511619427.007>.
- Foley, William A. 2008. The place of Philippine languages in a typology of voice systems. In Peter K. Austin & Simon Musgrave (eds.), *Voice and grammatical relations in Austronesian languages* (Studies in Constraint-Based Lexicalism), 22–44. Stanford, CA: Center for the Study of Language.
- Friberg, Barbara. 1991. Ergativity, focus and verb morphology in several South Sulawesi languages. In Ray Harlow (ed.), *VICAL 2: Western Austronesian and contact languages, parts I and II: Papers from the Fifth International Conference on Austronesian Linguistics*, 103–130. Auckland: Linguistic Society of New Zealand.
- Friberg, Timothy. 1995. Konjo. In Darrell T. Tryon (ed.), *Comparative Austronesian dictionary: An introduction to Austronesian studies*, vol. 1 (Trends in Linguistics. Documentation. 10), 563–571. Berlin; New York: Mouton de Gruyter.
- Gault, JoAnn Marie. 1999. *An ergative description of Sama Bangingi'* (Philippine Journal of Linguistics : Special Monograph Issue 46). Manila: Linguistic Society of the Philippines.
- Gerds, Donna B. 1988. *Object and absolutive in Halkomelem Salish* (Outstanding Dissertations in Linguistics). New York: Garland.

- Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure* (Cognitive Theory of Language and Culture). Chicago: University of Chicago Press.
- Goldberg, Adele E. 2006. *Constructions at work the nature of generalization in language* (Oxford Linguistics). Oxford: Oxford University Press.
- Goldberg, Adele E. 2013. Constructionist approaches. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of Construction Grammar*, 14–31. Oxford: Oxford University Press.
- Goldhahn, Dirk, Thomas Eckart & Uwe Quasthoff. 2012. Building large monolingual dictionaries at the Leipzig Corpora Collection: From 100 to 200 languages. In *Proceedings of the 8th International Language Resources and Evaluation (LREC'12)*. European Language Resources Association.
- Goudswaard, Nelleke. 2005. *The Begak (Ida'an) language of Sabah*. Utrecht: Leiden University dissertation.
- Gudai, Darmansyah H. 1985. *A grammar of Maanyan: A language of Central Kalimantan*. Canberra: Australian National University dissertation.
- Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank (eds.). 2022. *Glottolog 4.7*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://doi.org/10.5281/zenodo.7398962>.
- Hanafi, Nurachman. 1997. *A typological study of Sundanese*. Bundoora, VIC: La Trobe University dissertation.
- Hanson, Christopher John. 2003. *A grammar of Bugis based on the Soppeng dialect*. Bundoora, VIC: LaTrobe University dissertation.
- Hardeland, August. 1858. *Versuch einer Grammatik der Dajackschen Sprache*. Amsterdam: Frederik Muller.
- Hardjadibrata, R. R. 1985. *Sundanese: A syntactical analysis* (Pacific Linguistics D-65). Canberra: Pacific Linguistics and Australia National University.
- Hardjadibrata, R. R. 2003. *Sundanese-English dictionary*. Jakarta, Indonesia: PT Dunia Pustaka Jaya.
- Hartmann, Iren, Martin Haspelmath & Bradley Taylor (eds.). 2013. *Valency patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://valpal.info/> (12 November, 2021).
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687.
- Haspelmath, Martin. 2013. Argument indexing: A conceptual framework for the syntactic status of bound person forms. In Dik Bakker & Martin Haspelmath (eds.), *Language across boundaries: Studies in memory of Anna Siewierska*, 197–226. De Gruyter Mouton. <https://doi.org/10.1515/9783110331127.197>.
- Haspelmath, Martin. 2015. Ditransitive constructions. *Annual Review of Linguistics* 1(1). 19–41. <https://doi.org/10.1146/annurev-linguist-030514-125204>.
- Haspelmath, Martin & Thomas Müller-Bardey. 2004. Valency change. In Geert Booij, Christian Lehmann & Joachim Mugdan (eds.), *Morphology: A handbook on inflection and word formation*, vol. 2, 1130–1145. Berlin; New York: Mouton de Gruyter.
- Hemmings, Charlotte. 2013. Causatives and applicatives: The case for polysemy in Javanese. *SOAS Working Papers in Linguistics* 16(2013). 167–194.
- Hemmings, Charlotte. 2016. *The Kelabit language: Austronesian voice and syntactic typology*. University of London dissertation.

- Hendon, Rufus S. 1966. *The phonology and morphology of Ulu Muar Malay (Kuala Pilah District, Negri Sembilan, Malaya)* (Yale University Publications in Anthropology 70). New Haven, CT: Department of Anthropology, Yale University.
- Hernández-Green, Néstor. 2016. Registration vs. applicative constructions in Acazulco Otomi. *International Journal of American Linguistics* 82(3). 353–383.
- Himmelman, Nikolaus & John U. Wolff. 1999. *Toratán (Ratahan)* (Languages of the World. Materials 130). München: LINCOM Europa.
- Himmelman, Nikolaus P. 2002. Voice in western Austronesian: An update. In Fay Wouk & Malcolm Ross (eds.), *The history and typology of western Austronesian voice systems*, 7–16. Canberra: Pacific Linguistics.
- Himmelman, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In K. Alexander Adelaar & Nikolaus P. Himmelman (eds.), *The Austronesian languages of Asia and Madagascar* (Routledge Family Language Series), 110–181. New York: Routledge.
- Himmelman, Nikolaus P. & Sonja Riesberg. 2013. Symmetrical voice and applicative alternations: Evidence from Totoli. *Oceanic Linguistics* 52(2). 396–422. <https://doi.org/10.1353/ol.2013.0021>.
- Hogan, David W. 1988. *Urak Lawoi': Basic structures and a dictionary* (Pacific Linguistics C 109). Canberra: Pacific Linguistics and Australia National University.
- Hogan, David W. 1999. *Urak Lawoi'*. Vol. 268 (Languages of the World/Materials). München: LINCOM Europa.
- Hohaus, Vera & M. Ryan Bochnak. 2020. The grammar of degree: Gradability across languages. *Annual Review of Linguistics* 6(1). 235–259. <https://doi.org/10.1146/annurev-linguistics-011718-012009>.
- Holton, Gary. 2010. Person marking, verb classes, and the notion of alignment in Western Pantar (Lamma). In Michael C. Ewing & Marian Klamer (eds.), *East Nusantara: Typological and areal analyses* (Pacific Linguistics 618), 97–117. Canberra: Pacific Linguistics and Australia National University.
- Horne, Elinor Clarke. 1974. *Javanese-English dictionary*. New Have, CT; London: Yale University Press.
- Ikranagara, Kay. 1975. *Malayu Betawi grammar*. Honolulu, HI: University of Hawai'i at Mānoa dissertation.
- International Computer Science Institute. 2001. *Communication*. Framenet. <https://framenet2.icsi.berkeley.edu/fnReports/data/frame/Communication.xml> (15 December, 2023).
- James, Jeremiah Joy. 2017. *Central Sinama voice: A symmetrical analysis*. Graduate Institute of Linguistics MA thesis.
- Jerro, Kyle. 2016. *The syntax and semantics of applicative morphology in Bantu*. Austin, TX: University of Texas at Austin dissertation.
- Jerro, Kyle. 2021. Applied objects and the syntax–semantics interface. *Journal of Linguistics* 57(2). 365–403. <https://doi.org/10.1017/S0022226720000225>.
- Johnson, Christopher & Charles J. Fillmore. 2000. The FrameNet tagset for frame-semantic and syntactic coding of predicate-argument structure. In *NAACL 2000: Proceedings of the 1st North American Chapter of the Association for Computational Linguistics Conference*, 56–62. Association for Computational Linguistics. <https://dl.acm.org/doi/proceedings/10.5555/974305>.

- Jukes, Anthony. 2013. Voice, valence, and focus in Makassarese. In Alexander Adelaar (ed.), *Voice variation in Austronesian languages of Indonesia*. (NUSA 54), 67–84.
- Jukes, Anthony. 2020. *A grammar of Makasar* (Grammars and Sketches of the World's Languages: Mainland and Insular South East Asia). Leiden; Boston: Brill.
- Kähler, Hans. 1940. Grammatischer Abriss des Enganno. *Zeitschrift für Eingeborenen-Sprachen* 30. 81–117, 182–210, 296–320.
- Kähler, Hans. 1955. *Die Sichule-Sprache auf der Insel Simular an der Westküste von Sumatra*. Berlin: Reimer.
- Kaswanti Purwo, Bambang. 1995. The two prototypes of ditransitive verbs: The Indonesian evidence. In Werner Abraham, Talmy Givón & Sandra A. Thompson (eds.), *Discourse Grammar and Typology: Papers in Honor of John W. M. Verhaar* (Studies in Language Companion Series 27). Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Kaswanti Purwo, Bambang. 1997. The direct object in bi-transitive clauses in Indonesian. In Talmy Givón (ed.), *Grammatical relations: A functionalist perspective*, 233–252. Amsterdam; Philadelphia: John Benjamins Publishing Company.
- Kaufman, Daniel. 2009. Austronesian typology and the nominalist hypothesis. In K. Alexander Adelaar & Andrew Pawley (eds.), *Austronesian historical linguistics and culture history: A festschrift for Robert Blust*, 179–226. Canberra: Australian National University.
- Keenan, Edward L. 1976. Towards a universal definition of “subject”. In Charles N. LI (ed.), *Subject and topic*, 303–333. New York: Academic Press.
- Khairunnisa. 2022. *Diathesis, grammatical relations, and clitics in Ampenan Sasak*. Honolulu, HI: University of Hawai'i at Mānoa dissertation.
- Khairunnisa & Bradley McDonnell. in prep. Polyfunctionality and productivity of the verbal suffix *-an* in Ampenan Sasak.
- Kikusawa, Ritsuko. 2012. On the development of applicative constructions in Austronesian languages. *Bulletin of the National Museum of Ethnology* 36(4).
- King, Julie K. 1984–2000. Tombonuo grammar notes. Typescript.
- King, Julie K. 1991. Thematic continuity and development in Tombonuo narrative discourse. In *Thematic continuity and development in languages of Sabah* (Pacific Linguistics C 118), 45–74. Canberra: Australian National University.
- Kittilä, Seppo. 2005. Recipient-prominence vs. beneficiary-prominence. *Linguistic Typology* 9(2). 269–297. <https://doi.org/10.1515/lity.2005.9.2.269>.
- Kittilä, Seppo & Fernando Zúñiga. 2010. Introduction: Benefaction and malefaction from a cross-linguistic perspective. In *Benefactives and malefactives: Typological perspectives and case studies* (Typological Studies in Language 92), xi–xxxviii. Amsterdam: John Benjamins Publishing Company.
- Kiyosawa, Kaoru. 2006. *Applicatives in Salish languages*. Burnaby, BC: Simon Fraser University dissertation.
- Klamer, Marian. 1998. *A grammar of Kambera*. Berlin: De Gruyter Mouton.
- Klamer, Marian & Michael C. Ewing. 2010. The languages of East Nusantara: An introduction. In Michael C. Ewing & Marian Klamer (eds.), *East Nusantara: Typological and areal analyses* (Pacific Linguistics 618), 1–24. Canberra: Pacific Linguistics and Australia National University.
- Kroeger, Paul. 1988. Verbal focus in Kimaragang. In *Papers in western Austronesian Linguistics, no. 3* (Pacific Linguistics A-78), 217–240. Canberra: Pacific Linguistics.

- Kroeger, Paul. 2005. Kimaragang. In K. Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian Languages of Asia and Madagascar* (Routledge Language Family Series), 419–450. London ; New York: Routledge. <https://doi.org/10.4324/9780203821121-21>.
- Kroeger, Paul. 2007. Morphosyntactic vs. morphosemantic functions of Indonesian *-kan*. In Annie Zaenen, Jane Simpson, Tracey Holloway King, Jane Grimshaw, Joan Maling & Christopher D. Manning (eds.), *Architectures, rules, and preferences: Variations on themes of Joan Bresnan*, 229–251. Stanford, CA: CSLI Publications.
- Kroeger, Paul R. & Sonja Riesberg. Forthcoming. Voice and transitivity. In K. Alexander Adelaar & Antoinette Schapper (eds.), *The Oxford guide to the Malayo-Polynesian Languages of Southeast Asia* (Oxford Guides to the World's Languages). Oxford: Oxford University Press.
- Kulikov, Leonid & Jae Jung Song. 2010. Voice typology. In *The Oxford Handbook of Linguistic Typology*, 368–398. New York: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199281251.001.0001>.
- Kurniawan, Eri. 2013. *Sundanese complementation*. Iowa City, IA: University of Iowa dissertation. <https://doi.org/10.17077/etd.09n28b9j>.
- Kustian. N.d. *Nyawa*. Basa Sunda: Referensi artikel Bahasa Sunda. <https://basasunda.com/nyawa/> (31 December, 2021).
- Larish, Michael. 2005. Moken and Moklen. In K. Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar* (Routledge Language Family Series), 513–533. London ; New York: Routledge.
- Laskowske, Douglas C. 2016. *Voice in Bugis: An RRG perspective*. University of North Dakota MA thesis. <https://commons.und.edu/theses/2039>.
- Laskowske, Kathryn. 1994. Negation in Seko Padang. *NUSA: Linguistic studies of Indonesian and other languages in Indonesia* 36. 49–64.
- Laskowske, Thomas. 2006. The Seko languages of South Sulawesi: A reconstruction.
- Laskowske, Thomas & Arif. 2000. *Pedoman membaca dan menulis Bahasa Seko Padang*. Sulawesi Tengah, Indonesia: Panitia Bahasa Seko Padang.
- Lee, Li-Young. 1986. *Rose: Poems*. Brockport, NY: BOA Editions.
- Levin, Beth. 1993. *English verb classes and alternations*. Chicago: University of Chicago Press.
- Levinsohn, Stephen H. & John Wayne King. 1991. Participant reference in Tombonuo. In Stephen H. Levinsohn (ed.), *Thematic continuity and development in languages of Sabah* (Pacific Linguistics C 118), 75–92. Canberra: Australian National University.
- Liao, Hsiu-chuan. 2004. *Transitivity and ergativity in Formosan and Philippine languages*. Honolulu, HI: University of Hawai'i at Mānoa dissertation.
- Liaw, Andy & Matthew Wiener. 2002. Classification and regression by randomForest. *R News* 2(3). 18–22. <https://CRAN.R-project.org/doc/Rnews/>.
- Lobel, Jason William. 2013. Southwest Sabah revisited. *Oceanic Linguistics* 52(1). 36–68. <https://www.jstor.org/stable/43286760> (24 October, 2023).
- Loeb, Edwin M. 1928. Mentawai social organization. *American Anthropologist* 30. 408–433.
- Lovestrand, Joseph. 2021. Serial verb constructions. <https://doi.org/10.1146/annurev-linguistics-031920-115317>.
- Lumpat*. 2021. In *Wikikamus Basa Sunda*. <https://su.wiktionary.org/wiki/anclom> (13 February, 2024).
- Mahmood, Abdul Hamid. 1994. *Sintaksis dialek Kelantan*. Kuala Lumpur: Dewan Bahasa dan Pustaka, Kementerian Pendidikan Malaysia.

- Martens, Michael P. 1988a. Focus and discourse in Uma. In Hein Steinhauer (ed.), *Papers in western Austronesian linguistics*, no. 4 (Pacific Linguistics A 79), 239–256. Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University.
- Martens, Michael P. 1988b. Notes on Uma verbs. In *Papers in western Austronesian linguistics*, no. 4 (Pacific Linguistics A 79), 167–237. Pacific Linguistics.
- Martens, Michael P. 1990. Proto Kaili-Pamona: Reconstruction of the protolanguage of a language subgroup in Sulawesi. Hasanuddin University & The Summer Institute of Linguistics.
- Mattiola, Simone. 2019. *Typology of pluractional constructions in the languages of the world* (Typological Studies in Language 125). Amsterdam; Philadelphia: John Benjamins. <https://doi.org/10.1075/tsl.125>.
- Mayani, Luh Arik. 2013. *A grammar of Tajio: A language spoken in central Sulawesi*. Cologne: Universität zu Köln dissertation.
- McDonnell, Bradley. 2016. *Symmetrical voice constructions in Besemah: A usage-based approach*. Santa Barbara, CA: University of California, Santa Barbara dissertation.
- McDonnell, Bradley & Victoria Chen. 2022. The evolution of syntax in Western Austronesian. In *The Routledge Handbook of Asian Linguistics*, 11–32. London: Routledge. <https://doi.org/10.4324/9781003090205>.
- McDonnell, Bradley & Christina L. Truong. 2024. Applicative constructions in languages of western Indonesia. In Fernando Zuñiga & Denis Creissels (eds.), *Applicative constructions in the world's languages* (Comparative Handbooks of Linguistics 7), 971–1004. Berlin; Boston: De Gruyter Mouton.
- McDonnell, Bradley & Christina L. Truong. Forthcoming. The non-Malayic languages of Sumatra and the Barrier Islands. In K. Alexander Adelaar & Antoinette Schapper (eds.), *The Oxford guide to the Malayo-Polynesian Languages of Southeast Asia* (Oxford Guides to the World's Languages), 407–427. Oxford: Oxford University Press.
- McGinn, Richard. 1982. *Outline of Rejang syntax*. Vol. 14 (NUSA). Jakarta: Badan Penyelenggara Seri Nusa Universitas Katolik Atma Jaya Jakarta.
- McGinn, Richard. 2009. Out-of-Borneo subgrouping hypothesis for Rejang: Re-weighing the evidence. In K. Alexander Adelaar & Andrew Pawley (eds.), *Austronesian historical linguistics and culture history: a festschrift for Robert Blust* (Pacific Linguistics 601), 411–424. Canberra: Research School of Pacific and Asian Studies, Australian National University.
- McGinnis, Martha. 2008. Applicatives. *Language and Linguistics Compass* 2(6). 1225–1245. <https://doi.org/10.1111/j.1749-818X.2008.00078.x>.
- McKinnon, Timothy, Peter Cole & Gabriella Hermon. 2011. Object agreement and ‘pro-drop’ in Kerinci Malay. *Language* 87(4). 715–750.
- Mead, David. 1998. *Proto-Bungku-Tolaki: Reconstruction of its phonology and aspects of its morphosyntax*. Houston, TX: Rice University dissertation.
- Mead, David. 1999. Agreement marking in Bungku-Tolaki languages: Toward an historical account. Paper presented in the Rice University Department of Linguistics Colloquium Series, 28 Jan, Houston, TX. <https://www.sil.org/resources/archives/88049>.
- Mead, David. 2001. A preliminary sketch of the Bobongko language. *NUSA: Linguistic Studies of Indonesian and Other Languages* 49. 61–94.
- Mead, David. 2003. Evidence for a Celebic supergroup. In John Lynch (ed.), *Issues in Austronesian Historical Phonology* (Pacific Linguistics 550), 115–141. Canberra: Pacific Linguistics.

- Mead, David. 2005. Mori Bawah. In K. Alexander Adelaar & Nikolaus Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar* (Routledge Language Family Series), 683–708. London ; New York: Routledge.
- Mead, David. 2013. *Wotu grammar notes* (Sulawesi Language Data and Working Papers: Grammar Descriptions 1). Sulawesi Language Alliance. <http://sulang.org/resources/grammar-descriptions/001>.
- Miller, Mark. 2007. *A grammar of West Coast Bajau*. Arlington, TX: University of Texas at Arlington dissertation.
- Mills, Roger F. 1975. *Proto South Sulawesi and Proto Austronesian phonology*. Ann Arbor, MI: University of Michigan dissertation.
- Mithun, Marianne. 2001. Understanding and explaining applicatives. In Mary Andronis, Christopher Ball, Heidi Elston & Sylvain Neuvel (eds.), *CLS 37, The panels 2001: Proceedings from the parasessions of the thirty-seventh meeting of the Chicago Linguistic Society*, vol. 2, 73–97. Chicago: Chicago Linguistic Society.
- Mithun, Marianne & Wallace L. Chafe. 1999. What are S, A, and O ? *Studies in Language* 23(3). 569–596. <https://doi.org/10.1075/sl.23.3.05mit>.
- Moro, Francesca R. 2010. *A sketch grammar of Dampelas, a language of Central Sulawesi*. Leiden: Leiden University MA thesis.
- Morris, Max. 1900. *Die Mentawai-sprache*. Berlin: C. Skopnik.
- Mortensen, Christian J. 2021. *The Kemaloh Lun Bawang language of Borneo*. University of Hawai'i at Mānoa dissertation.
- Nababan, P. W. J. 1981. *A grammar of Toba-Batak* (Pacific Linguistics 37). Canberra, A.C.T: Research School of Pacific Studies, Australian National University.
- Nguyen, Tam Thi Minh. 2013. *A grammar of Bih*. Eugene, OR: University of Oregon dissertation.
- Nothofer, Bernd. 1975. *The reconstruction of Proto-Malayic-Javanic* (Verhandelingen van Het Koninklijk Instituut Voor Taal-, Land- En Volkenkunde 73). 's-Gravenhage: Martinus Nijhoff.
- Nothofer, Bernd. 1980. *Dialektgeographische Untersuchungen in West-Java und im westlichen Zentral-Java*. Wiesbaden: Harrassowitz.
- Nothofer, Bernd. 1986. The Barrier Island languages in the Austronesian language family. In Paul Geraghty, Lois Carrington & Stephen A. Wurm (eds.), *FOCAL II: Papers from the Fourth International Conference on Austronesian Languages*, 89–107. Canberra: Research School of Pacific and Asian Studies, Australia National University. <http://dx.doi.org/10.15144/PL-C94.87>.
- O'Herin, Brian. 2001. Abaza applicatives. *Language* 77(3). 477–493. <https://doi.org/10.1353/lan.2001.0181>.
- Oglobin, Alexander K. 2005. Javanese. In K. Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*. New York: Routledge.
- Omar, Asmah Haji. 1981. *The Iban language of Sarawak: A grammatical description*. Kuala Lumpur: Dewan Bahasa dan Pustaka, Kementerian Pendidikan Malaysia.
- Omar, Asmah Haji. 1983. *The Malay peoples of Malaysia and their languages*. Kuala Lumpur: Dewan Bahasa dan Pustaka, Kementerian Pendidikan Malaysia.
- Pacchiarotti, Sara. 2020. *Bantu applicative constructions* (Standard Monographs in African Languages). Stanford, CA: CSLI Publications.
- Pacchiarotti, Sara & Fernando Zuñiga. 2022. Introduction. In *Applicative morphology: Neglected syntactic and non-syntactic functions* (Trends in Linguistics. Studies and Monographs [TiLSM] 373), 1–17. Berlin; Boston: De Gruyter Mouton.

- Pallesen, A. Kemp. 1985. *Culture contact and language convergence* (Linguistic Society of the Philippines Special Monograph Issue 24). Manila: Linguistic Society of the Philippines Special Monograph Issue.
- Pappas, Leah. 2022. *A multi-modal study of spatial language in Hawu in typological perspective*. Honolulu, HI: University of Hawai'i at Mānoa dissertation. <https://hdl.handle.net/10125/102246>.
- Pawley, Andrew & Lawrence A. Reid. 1979. The evolution of transitive construction in Austronesian. In Paz Buenaventura Naylor (ed.), *Austronesian studies: Papers from the Second Eastern Conference on Austronesian languages* (Michigan Papers of South and Southeast Asia 15), 103–130. Ann Arbor: University of Michigan. <http://hdl.handle.net/10125/33014>.
- Payne, Thomas E. & Thomas Laskowske. 1997. Voice in Seko Padang. In *Essays on language function and language type: Dedicated to T. Givón*, 423–736. Amsterdam; Philadelphia: John Benjamins.
- Payne, Thomas Edward. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge ; New York: Cambridge University Press.
- Pearson, Joel Matthew. 2001. *The clause structure of Malagasy: A minimalist approach*. Los Angeles: University of California, Los Angeles dissertation.
- Percival, W.K. 1981. *A grammar of the urbanised Toba-Batak of Medan* (Pacific Linguistics B 76). Canberra: Pacific Linguistics and Australia National University.
- Perek, Florent. 2015. *Argument structure in usage-based construction grammar: Experimental and corpus-based perspectives*. Vol. 17 (Constructional Approaches to Language). Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/cal.17>.
- Peterson, David A. 1999. *Discourse-functional, historical, and typological aspects of applicative constructions*. Berkeley, CA: University of California, Berkeley dissertation.
- Peterson, David A. 2007. *Applicative constructions* (Oxford Studies in Typology and Linguistic Theory). Oxford: Oxford University Press.
- Polinsky, Maria. 2013. Applicative constructions. In Matthew Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <https://wals.info/chapter/109> (24 February, 2021).
- Post, Mark W. & Yankee Modi. 2022. Applicatives in Macro-Tani languages (Trans-Himalayan, Eastern Himalaya): Forms, functions and historical origins. In *Applicative morphology: Neglected syntactic and non-syntactic functions* (Trends in Linguistics 373), 299–328. Berlin; Boston: De Gruyter Mouton.
- Prentice, David J. 1969. Verbal inflection in Sabah Murut. *Pacific Linguistics. Series A. Occasional Papers* (20). 9–21.
- Prentice, David J. 1995. Timugon Murut. In Darrell T. Tryon (ed.), *Comparative Austronesian dictionary: An introduction to Austronesian studies*, vol. 1 (Trends in Linguistics. Documentation 10), 381–391. Berlin; New York: Mouton de Gruyter.
- Probst, Philipp, Marvin N. Wright & Anne-Laure Boulesteix. 2019. Hyperparameters and tuning strategies for random forest. *WIREs Data Mining and Knowledge Discovery* 9(3). e1301. <https://doi.org/10.1002/widm.1301>.
- Pusat Bahasa (Indonesia). 2007. *Kamus besar bahasa Indonesia*. 3rd edn. Jakarta: Balai Pustaka Pub.
- Pylkkänen, Liina. 2002. *Introducing arguments*. Cambridge, MA: Massachusetts Institute of Technology dissertation.

- Quick, Phil. 2007. *A grammar of the Pendau language of central Sulawesi, Indonesia* (Pacific Linguistics 590). Canberra: Pacific Linguistics and Australia National University.
- R Core Team. 2023. *R: A language and environment for statistical computing*. Manual. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>.
- Rappaport Hovav, Malka & Beth Levin. 1998. Building verb meanings. In Miriam Butt & Wilhelm Geuder (eds.), *The projection of arguments: lexical and compositional factors* (CSLI Lecture Notes 83). Stanford, CA: CSLI Publications.
- Rasoloson, Janie & Carl Rubino. 2005. Malagasy. In K. Alexander Adelaar & Nikolaus Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar* (Routledge Language Family Series), 456–488. London ; New York: Routledge.
- Ravindranath, Maya & Abigail C. Cohn. 2014. Can a language with millions of speakers be endangered? *Journal of the Southeast Asian Linguistics Society (JSEALS)* 7 (2014): 64–75. <https://openresearch.repository.anu.edu.au/handle/1885/11968> (11 March, 2024).
- Reid, Lawrence A & Hsiu-chuan Liao. 2004. A brief syntactic typology of Philippine languages. *Language and Linguistics* 5(2). 433–490.
- Rensch, Calvin R., Carolyn M. Rensch, Jonas Noeb & Robert Sulis Ridu. 2012. *The Bidayuh language: Yesterday, today and tomorrow*. Vol. 33 (SIL E-Books). Dallas: SIL International. 517.
- Riesberg, Sonja. 2014a. Passive actors are not adjuncts—Consequences for the distinction between symmetrical and asymmetrical voice alternations. In I Wayan Arka & N. L. K. Mas Indrawati (eds.), *Argument realisations and related constructions in Austronesian languages: Papers from 12-ICAL, Vol. 2* (Asia-Pacific Linguistics 013; Studies on Austronesian Languages 002), 281–302. Canberra: Asia-Pacific Linguistics. <http://hdl.handle.net/1885/12059>.
- Riesberg, Sonja. 2014b. *Symmetrical voice and linking in western Austronesian languages* (Pacific Linguistics 646). Berlin: De Gruyter Mouton.
- Rigg, Jonathan. 1862. *A dictionary of the Sunda language of Java*. Batavia: Lange.
- Robins, R. H. 1959. Nominal and verbal derivation in Sundanese. *Lingua* 8. 337–369.
- Robson, Stuart & Singgih Wibisono (eds.). 2002. *Javanese-English dictionary*. Hong Kong: Periplus Editions.
- Ross, Malcolm. 2002. The history and transitivity of western Austronesian voice and voice-marking. In Fay Wouk & Malcolm Ross (eds.), *The history and typology of Western Austronesian voice systems* (Pacific Linguistics 518), 17–62. Canberra: Pacific Linguistics and Australia National University.
- Ross, Malcolm. 2009. Proto Austronesian verbal morphology: A reappraisal. In K. Alexander Adelaar & Andrew Pawley (eds.), *Austronesian historical linguistics and culture history: A festschrift for Robert Blust*, 295–326. Canberra: Pacific Linguistics.
- Ross, Malcolm. 2012. In defense of Nuclear Austronesian (and against Tsouic). *Language and Linguistics* 13(6). 1253–1330.
- Ruppenhofer, Josef, Michael Ellsworth, Miriam R. L. Petruck, Christopher R. Johnson, Collin F. Baker & Jan Scheffczyk. 2016. *FrameNet II: Extended theory and practice*. Berkeley, CA: International Computer Science Institute.
- Satjadibrata, R. 1948. *Kamoes basa Soenda*. Jakarta: Bale Poestaka.
- Schachter, Paul. 1984. Semantic-role-based syntax in Toba Batak. In Paul Schachter (ed.), *Studies in the Structure of Toba Batak* (UCLA Occasional Papers in Linguistics 5), 122–149.

- Sercombe, Peter. 2006. The Eastern Penan language of Borneo. In *Essays in honour of Professor James T. Collins: Reflections in Southeast Asian seas*, vol. 1, 1–34. Pontianak: STAIN Pontianak Press.
- Shibatani, Masayoshi & Prashant Pardeshi. 2002. The causative continuum. In *The grammar of causation and interpersonal manipulation* (Typological Studies in Language 48), 85–126. Amsterdam: John Benjamins.
- Shiohara, Asako & I Wayan Arka. Forthcoming. Balinese, Sasak, and Sumbawa. In K. Alexander Adelaar & Antoinette Schapper (eds.), *The Oxford Guide to the Malayo-Polynesian Languages of Southeast Asia* (Oxford Guides to the World's Languages). Oxford: Oxford University Press.
- Shore, Susan. 2016. *Kahirámpuá laluita Básá Behóá - Básá Indonesia* (Behoa - Indonesian Dictionary). Trial Edition. Palu.
- Siewierska, Anna. 1999. Languages with and without objects: The Functional Grammar approach. *Languages in Contrast* 1(2). 173–190. <https://doi.org/10.1075/lic.1.2.05sie>.
- Sirk, Ülo. 1983. *The Buginese language* (Languages of Asia and Africa). Moscow: Nauka. 111.
- Sirk, Ülo. 1996. On the history of transitive verb suffixes in the languages of western Indonesia. In Hein Steinhauer (ed.), *Papers in Austronesian linguistics, no. 3* (Pacific Linguistics A 84), 191–205. Canberra: Pacific Linguistics.
- Smith, Alexander D. 2017. The Western Malayo-Polynesian problem. *Oceanic Linguistics* 56(2). 435–490. <https://doi.org/10.1353/ol.2017.0021>.
- Smith, Alexander D. 2018. The Barito linkage hypothesis, with a note on the position of Basap. *Journal of the Southeast Asian Linguistics Society* 11. 13–34.
- Sneddon, J. N., K. Alexander Adelaar, Dwi Noverini Djénar & Michael C. Ewing. 2010. *Indonesian reference grammar*. 2nd edn. Crows Nest, N.S.W: Allen & Unwin.
- Sneddon, James N. 1993. The drift towards open final syllables in Sulawesi languages. *Oceanic Linguistics* 32(1). 1–44.
- Sneddon, James N. 2006. *Colloquial Jakartan Indonesian*. Canberra: Pacific Linguistics and Australia National University.
- Sommerlot, Carly Joann. 2020. *On the syntax of West Kalimantan: Asymmetries and A'-movement in Malayic and Land Dayak languages*. Arlington, TX: University of Texas at Arlington dissertation. <https://rc.library.uta.edu/uta-ir/handle/10106/29430>.
- Soravia, Giulio. 2007. *The Alas language (Northern Sumatra)*. Bologna, IT: Università di Bologna, Alma Mater Studiorum. <http://doi.org/10.6092/unibo/amsacta/2367>.
- Soriente, Antonia. 2013. Undergoer voice in Borneo: Penan, Punan, Kenyah and Kayan languages. In K. Alexander Adelaar (ed.), *Voice variation in Austronesian languages of Indonesia* (NUSA 54(31)), 175–203. <http://hdl.handle.net/10108/71811>.
- Starosta, Stanley, Andrew Pawley & Lawrence A Reid. 1982. The evolution of focus in Austronesian. In Amran Halim, Lois Carrington & Stephen A. Wurm (eds.), *Papers from the Third International Conference on Austronesian Linguistics, vol. 2: Tracking the travellers* (Pacific Linguistics C 75), 145–170. Canberra: Australia National University. <http://hdl.handle.net/10125/33001>.
- Stefanowitsch, Anatol & Stefan Th Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8(2). 209–243. <https://doi.org/10.1075/ijcl.8.2.03ste>.
- Stoel, Ruben. 2005. *Focus in Manado Malay: Grammar, particles, and intonation*. Leiden: CNWS Publications.

- Sударsono, P. D. 2002. *Description of the Bakatik Dayak language*. Bundoora, VIC: LaTrobe University dissertation.
- Tadmor, Uri. 1995. *Language contact and systemic restructuring: The Malay dialect of Nonthaburi, central Thailand*. Honolulu, HI: University of Hawai'i at Mānoa dissertation.
- Tamsyah, Budi Rahayu. 1997. *Kamus Indonesia-Sunda*. Bandung: Pustaka Setia.
- Thompson, Sandra A. 1990. Information flow and dative shift in English discourse. In Jerald A. Edmondson, Crawford Feagin & Peter Mühlhäusler (eds.), *Development and diversity: Language variation across time and space: A festschrift for Charles-James N. Bailey* (SIL International Publications in Linguistics 93), 239–253. Dallas, TX: Summer Institute of Linguistics and University of Texas at Arlington. <https://www.sil.org/resources/archives/8425> (23 March, 2021).
- Thurgood, Graham. 1999. *From ancient Cham to modern dialects: Two thousand years of language contact and change* (Oceanic Linguistics Special Publication 28). Honolulu: University of Hawai'i Press.
- Thurgood, Graham. 2005. Phan Rang Cham. In K. Alexander Adelaar & Nikolaus Himmelmann (eds.), *The Austronesian Languages of Asia and Madagascar* (Routledge Language Family Series), 489–512. London ; New York: Routledge.
- Thurgood, Graham. 2010. Hainan Cham, Anong, and Eastern Cham: Three languages, three social contexts, three patterns of change. *Journal of Language Contact* 3(2). 39–65. <https://doi.org/10.1163/19552629-90000019>.
- Thurgood, Graham, Ela Thurgood & Li Fengxiang. 2014. *A grammatical sketch of Hainan Cham: History, contact, and phonology*. Berlin; Boston: De Gruyter Mouton. <https://doi.org/10.1515/9781614516040>.
- Tjia, Johnny. 2007. *A grammar of Mualang: An Ibanic language of Western Kalimantan, Indonesia*. Utrecht: Leiden University dissertation. <http://hdl.handle.net/1887/11862>.
- Townsend, Joel. 2017. *Reduplication in Serudung Murut*. Leiden: Leiden University MA thesis.
- Truong, Christina L. in prep. How does vowel harmony develop? A case study from Behoa, a language of Central Sulawesi.
- Truong, Christina L. & Bradley McDonnell. 2021. Exploring western Indonesian applicative constructions through the lexicon. Paper presented at the 15th International Conference on Austronesian Languages (ICAL 15), 28 June–2 July, Palacký University, Olomouc, Czech Republic. <https://sinofon.cz/wp-content/uploads/2021/06/15ICAL-2021-Book-of-Abstracts.pdf>.
- Truong, Christina L. & Bradley McDonnell. 2022. Neglected functions of western Indonesian applicative morphology. In Sara Pacchiarotti & Fernando Zuñiga (eds.), *Applicative morphology: Neglected syntactic and non-syntactic functions* (Trends in Linguistics. Studies and Monographs [TiLSM] 373), 405–436. Berlin; Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110777949>.
- Tsuchida, Shigeru. 1976. *Reconstruction of Proto-Tsouic phonology* (Tokyo: Study of Languages and Cultures of Asia & Africa Monograph Series 5). Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Tsukida, Naomi. 2009. *Se dekku go (Taiwan) no bunpou [Grammar of Seediq language (Taiwan)]*. University of Tokyo dissertation.
- Valkama, Kari. 1993. *Studies in Duri grammar: Person marking, basic constituent order and grammatical relations and voice in Duri*. University of Helsinki MA thesis.
- Vamarasi, M. Kana. 1999. *Grammatical relations in Bahasa Indonesia* (Pacific Linguistics D-93). Canberra: Research School of Pacific Studies, Australian National University.

- van den Berg, René. 1991. Muna historical phonology. *NUSA: Linguistic Studies of Indonesian and Other Languages* 33. 3–28.
- van den Berg, René. 1995. Verb classes, transitivity, and the definiteness shift in Muna: A counterexample to the Transitivity Hypothesis? *Oceanic Linguistics* 34(1). 161–189.
- van den Berg, René. 2013. *A grammar of the Muna language* (SIL E-Books 52). Dallas, TX: SIL International. <https://www.sil.org/resources/archives/52170> (22 February, 2022).
- van den Berg, René. 2020. *The Busoa language of Southeast Sulawesi: Grammar sketch, texts, vocabulary* (SIL Language and Culture Documentation and Description 48). Dallas: SIL International.
- van den Berg, René & Robert L. Busenitz. 2012. *A grammar of Balantak, a language of Eastern Sulawesi* (SIL E-Books 40). SIL International. <https://www.sil.org/resources/publications/entry/49492>.
- van den Berg, René & La Ode Sidu Marafad (eds.). 2016. *Muna dictionary*. Web edition. SIL International. <https://www.webonary.org/muna/>.
- van der Tuuk, H.N. 1971 [1864-1867]. *A grammar of Toba Batak [Tobasche spraakunst]*. Trans. by Jeune Scott-Kemball (Translation Series, Koninklijk Instituut Voor Taal-, Land- En Volkenkunde 13). The Hague: Nijhoff.
- Van Syoc, Wayland Bryce. 1959. *The phonology and morphology of the Sundanese language*. Ann Arbor, Michigan: University of Michigan.
- Van Valin, Robert D & Randy J. LaPolla. 1997. *Syntax, structure, meaning and function*. Cambridge: Cambridge University Press.
- Van Valin, Robert D., Jr. 2005. *Exploring the syntax-semantic interface*. Cambridge: Cambridge University Press.
- Vander Klok, Jozina & Bethwyn Evans. 2022. The evolution of non-syntactic functions of applicatives: *-i* suffixation in Javanese and neighbouring languages. In Sara Pacchiarotti & Fernando Zuñiga (eds.), *Applicative morphology: Neglected syntactic and non-syntactic functions* (Trends in Linguistics. Studies and Monographs [TiLSM] 373), 437–475. De Gruyter Mouton.
- Villierius, Sophie. 2019. *Development of Surinamese Javanese. Language contact and change in a multilingual context*. Utrecht: Radboud Universiteit Nijmegen dissertation.
- Walker, Dale F. 1976. *A grammar of the Lampung language: The Pesisir dialect of Way Lima* (NUSA 2). Jakarta: Universitas Katolik Indonesia Atma Jaya.
- Walton, Charles. 1986. *Sama verbal semantics: Classification, derivation and inflection* (Special Monograph Issue 25). Manila: Linguistic Society of the Philippines.
- Wechsler, Stephen & I Wayan Arka. 1998. Syntactic ergativity in Balinese: An argument structure based theory. *Natural Language & Linguistic Theory* 16(2). 387–442. <https://doi.org/10.1023/A:1005920831550>.
- Wickham, Hadley. 2017. *Tidyverse: Easily install and load the 'Tidyverse'*. Manual. <https://CRAN.R-project.org/package=tidyverse>.
- Wickham, Hadley. 2019. *Stringr: Simple, consistent wrappers for common string operations*. Manual. <https://CRAN.R-project.org/package=stringr>.
- Wolff, John U. 1973. Verbal inflection in Proto-Austronesian. In Andrew B. Gonzalez (ed.), *Parangal kay Cecilio Lopez: Essays in honor of Cecilio Lopez on his seventy-fifth birthday* (Philippine Journal of Linguistics: Special Monograph Issue 4), 71–91. Quezon City: Linguistic Society of the Philippines.

- Wolff, John U. 1996. The development of the passive verb with pronominal prefix in Western Austronesian languages. In Bernd Nothofer (ed.), *Reconstruction, classification, description: Festschrift in honor of Isidore Dyen*, 15–40. Hamburg: Abera.
- Wood, Esther. 2007. *The semantic typology of pluractionality*. Berkeley, CA: University of California, Berkeley dissertation.
- Woollams, Geoff. 1996. *A grammar of Karo Batak, Sumatra* (Pacific Linguistics 130). Canberra: Research School of Pacific Studies, Australian National University.
- Yanti. 2010. *A reference grammar of Jambi Malay*. Newark, DE: University of Delaware dissertation.
- Yanti, Timothy Mckinnon, Peter Cole & Gabriella Hermon. 2018. The phonological basis of syntactic change in Kerinci. *Oceanic Linguistics* 57(2). 433–483.
- Zeitoun, Elizabeth & Stacy Fang-ching Teng. 2016. Reassessing the position of Kanakanavu and Saaroa among the Formosan languages. *Oceanic Linguistics* 55(1). 163–198.
- Zobel, Erik. 2020. The Kaili–Wolio Branch of the Celebic Languages. *Oceanic Linguistics* 59. 297–346.
- Zuñiga, Fernando & Denis Creissels. 2024. Applicative constructions: An introductory overview. In Fernando Zuñiga & Denis Creissels (eds.), *Applicative constructions in the world's languages* (Comparative Handbooks of Linguistics 7), 1–56. Berlin; Boston: De Gruyter Mouton.
- Zuñiga, Fernando & Seppo Kittilä. 2019. *Grammatical voice* (Cambridge Textbooks in Linguistics). Cambridge ; New York: Cambridge University Press.