



Universiteit
Leiden
The Netherlands

Traces of language contact: The Flores-Lembata languages in eastern Indonesia

Fricke, H.L.A.

Citation

Fricke, H. L. A. (2019, November 13). *Traces of language contact: The Flores-Lembata languages in eastern Indonesia*. LOT dissertation series. LOT, Amsterdam. Retrieved from <https://hdl.handle.net/1887/80399>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/80399>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/80399> holds various files of this Leiden University dissertation.

Author: Fricke, H.L.A.

Title: Traces of language contact: The Flores-Lembata languages in eastern Indonesia

Issue Date: 2019-11-13

Traces of language contact
The Flores-Lembata languages in eastern
Indonesia

Published by

LOT
Kloveniersburgwal 48
1012 CX Amsterdam
The Netherlands

phone: +31 20 525 2461
e-mail: lot@uva.nl
<http://www.lotschool.nl>

Cover illustration: Language map at Museum Blikon Blewut at Ledalero on Flores and a woven sarong made by Ursula Blikon in Bakan on Lembata, both in eastern Indonesia. The map shows the line between languages distinguished by different types of possessive constructions.

ISBN: 978-94-6093-329-5
NUR: 616

Copyright © 2019 Hanna Lotte Anneliese Fricke. All rights reserved.

Traces of language contact
The Flores-Lembata languages in eastern
Indonesia

Sporen van taalcontact
De Flores-Lembatatalen in Oost-Indonesië
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus Prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 13 november 2019
klokke 10.00 uur

door

Hanna Lotte Anneliese Fricke

geboren 16 juni 1987
te Winsen (Luhe), Duitsland

Promotor: Prof. dr. Marian Klamer

Copromotores: Dr. Owen Edwards
Dr. Eugénie Stapert

Promotiecommissie: Prof. dr. Juliette Blevins (City University of New York)
Prof. dr. Maarten Kossmann
Dr. Francesca Moro
Prof. dr. Pieter Muysken (Radboud Universiteit)

This research was funded by the Dutch Research Council (NWO) as part of the VICI project *Reconstructing the past through languages of the present: The Lesser Sunda Islands* by Prof. dr. Marian Klamer (project number: 277-70-012).

Contents

Acknowledgements	xv
List of Abbreviations	xix
1 Introduction	1
1.1 Overview	1
1.2 Eastern Flores and the Solor Archipelago	4
1.2.1 Geographic location	4
1.2.2 Livelihood and beliefs	5
1.2.3 Divisions in the Lamaholot society	6
1.2.3.1 Overview	6
1.2.3.2 Weaving and non-weaving communities	6
1.2.3.3 Clan systems within a community . . .	7
1.2.3.4 Socio-political division: Paji and Demon	8
1.2.3.5 Conclusions	11
1.2.4 History	11
1.2.4.1 Overview	11
1.2.4.2 Alorese migration to Alor and Pantar . .	12
1.2.4.3 Migration from Lepad Batan	13
1.2.4.4 Colonialism and Christianisation	16
1.2.4.5 The Indonesian national language . . .	17
1.2.4.6 Summary	18
1.3 The wider areal context	19
1.3.1 Linguistic diversity	19

1.3.2	Genetic diversity	21
1.4	Research questions and aims	23

I Grammatical description 27

2	Introduction to Part I	29
2.1	Location and speakers	29
2.2	Data collection and processing	32
2.3	The Central Lembata Corpus	33
2.4	Citation codes	34
2.5	Transcription and glossing conventions	35
3	A grammar of Central Lembata	37
3.1	Structure of this grammar	37
3.2	Phonology	37
3.2.1	Phoneme inventory	37
3.2.2	Phonotactics	39
3.2.2.1	Distribution of single phonemes	39
3.2.2.2	Consonant clusters	40
3.2.3	Stems	45
3.2.4	Complex words	48
3.2.5	Stress	49
3.2.6	Summary	50
3.3	Nouns and noun phrases	50
3.3.1	Noun types	50
3.3.1.1	Inalienable nouns	52
3.3.1.2	Alienable nouns	56
3.3.2	Suffixes on alienable nouns	60
3.3.3	Coda alternation: origin and regional context	63
3.3.3.1	Central Lembata	64
3.3.3.2	Central Lamaholot	66
3.3.3.3	Western and Eastern Lamaholot	67
3.3.3.4	Languages of Timor	69
3.3.3.5	Summary	70
3.3.4	Noun phrases	71

3.3.4.1	Basic noun phrases	71
3.3.4.2	Inalienable possessive construction . .	73
3.3.4.3	Alienable possessive constructions . . .	75
3.3.4.4	The numeral slot	76
3.3.4.5	Relative clauses	77
3.3.4.6	Phrase-final slots	78
3.3.5	Demonstratives in noun phrases	78
3.3.6	Nominalising morphology	79
3.3.6.1	The infix <i>-n-</i> and its allomorph <i>b-</i>	79
3.3.6.2	The suffix <i>-k</i>	84
3.3.7	Summary	84
3.4	Pronominals	85
3.4.1	Overview	85
3.4.2	Core argument pronominals	87
3.4.2.1	Free pronouns	87
3.4.2.2	Pronominal clitics	88
3.4.2.3	Pronominal prefixes	90
3.4.2.4	Pronominal suffixes	91
3.4.2.5	Alignment of core arguments	97
3.4.3	Possessor pronouns and suffixes	98
3.4.4	Summary	101
3.5	Verbs	102
3.5.1	Overview	102
3.5.2	Verb classes	102
3.5.2.1	G-suffix verbs	103
3.5.2.2	K-suffix verbs	104
3.5.2.3	Suffixless verbs	107
3.5.3	Grammaticalised verbs	107
3.5.3.1	The verb <i>-ora</i> 'join; be with'	108
3.5.3.2	The verb <i>maa</i> 'say'	110
3.5.4	The verbalising prefix <i>g-</i>	111
3.5.5	Summary	111
3.6	Basic clausal syntax	112
3.6.1	Simple verbal clauses	112
3.6.2	Temporal adverbs	115
3.6.3	Aspectual adverbs	118
3.6.4	Negation and prohibition	121

3.6.5	Multi-verb sequences	122
3.6.6	Nominal predicates	125
3.6.7	Adjectival root predicates	127
3.6.8	Existential clauses	129
3.6.9	Questions	130
3.6.10	Clause and phrase boundary markers	134
3.6.11	Clause combining	136
3.6.12	Summary	138
3.7	Location	138
3.7.1	Locationals	138
3.7.2	Locational verb phrases	142
3.7.2.1	Locative nouns	143
3.7.2.2	Deictic motion verbs	144
3.7.3	Summary	149
3.8	Summary and conclusions	149

II Historical phonology and lexical innovations 151

4	Introduction to Part II	153
4.1	The Flores-Lembata languages	153
4.1.1	Overview	153
4.1.2	Sika	156
4.1.3	Western Lamaholot	156
4.1.4	Central Lamaholot	159
4.1.5	Eastern Lamaholot	160
4.1.6	Kedang	160
4.2	Language sample and data sources	160
4.3	Methodology	164
4.4	Data representation	168
4.4.1	Transcription conventions	168
4.4.2	Organisation of tables	169
5	Historical phonology and subgroupings	171
5.1	Introduction	171
5.2	Reflexes of PMP sounds	175

5.2.1	PMP voiceless stops *p, *t, *k and *q	175
5.2.1.1	Initial and intervocalic position	175
5.2.1.2	Initial stops with PMP *ma-	180
5.2.1.3	Final stops PMP *p, *t, *k and *q	182
5.2.2	PMP voiced obstruents *b, *d, *z, *j and *g	185
5.2.2.1	Overview	185
5.2.2.2	Split: PMP *b > PFL *b/*v	186
5.2.2.3	Merger: PMP *d/*j/*z > PFL *d	188
5.2.2.4	Options for the PFL reflex of PMP *d/*j/*z	193
5.2.2.5	Final PMP *-d and *-j	196
5.2.2.6	Evidence for PFL *dʒ with no PMP source	197
5.2.2.7	The retention of PMP *g	197
5.2.3	PMP fricatives *h and *s	198
5.2.3.1	Loss of PMP *h	198
5.2.3.2	Split: PMP *s > PFL *s/*h	201
5.2.4	PMP nasals *m, *n and *ŋ	204
5.2.5	PMP liquids *l and *R	207
5.2.6	PMP glides *w *y	209
5.2.7	PMP vowels	211
5.2.8	PMP vowel-glide sequences in final position	216
5.2.9	The Proto-Flores-Lembata phoneme inventory	219
5.2.10	Reconstructed initial clusters in PFL	221
5.3	Subgroups within Flores-Lembata	222
5.3.1	Evidence for low-level subgroups	223
5.3.1.1	Sika	223
5.3.1.2	Western Lamaholot	224
5.3.1.3	Central Lamaholot	224
5.3.1.4	Eastern Lamaholot	226
5.3.1.5	Kedang	226
5.3.2	No good evidence for mid-level subgroups	226
5.4	Evidence for a Flores-Lembata subgroup	228
5.5	Evidence for a Bima-Lembata subgroup	229
5.6	Conclusions	232
6	Inheritance and innovation in the lexicon	233
6.1	Introduction	233
6.2	Proto-Flores-Lembata reconstructions	235

6.2.1	Overview	235
6.2.2	PFL reconstructions with PMP sources	236
6.2.3	PFL reconstructions without PMP sources	242
6.2.4	Irregular reflexes in individual subgroups	244
6.2.5	Discussion	248
6.3	Unreconstructible regular lexeme sets	250
6.3.1	Overview	250
6.3.2	Lamaholot-Kedang lexeme sets	250
6.3.3	Lamaholot-Sika lexeme sets	253
6.3.4	Lamaholot lexeme sets	255
6.3.5	Discussion	257
6.4	Irregular lexeme sets	259
6.4.1	Possible loans	260
6.4.2	Lexeme sets with unexplained correspondences	262
6.5	Conclusions	265

III Morpho-syntactic innovations 269

7	Introduction to Part III	271
7.1	Language sample and data sources	271
7.2	Methodology	276
7.3	Transcription and glossing conventions	277
8	Innovations in the nominal domain	279
8.1	Introduction	279
8.2	Alienability	281
8.2.1	Overview	281
8.2.2	No alienability distinction	284
8.2.3	Alienability distinction	286
8.2.3.1	Lamaholot	286
8.2.3.2	The Austronesian languages of Timor	290
8.2.3.3	The Timor-Alor-Pantar languages	291
8.2.4	The emergence of the alienability distinction	292
8.3	Property nouns	295
8.3.1	Overview	295

8.3.2	No property nouns	297
8.3.3	Property nouns in the FL languages	299
8.3.3.1	Sika	299
8.3.3.2	Lamaholot	302
8.3.3.3	Alorese	303
8.3.3.4	Kedang	304
8.3.3.5	Summary	304
8.3.4	Property nouns in the AN languages of Timor . . .	305
8.3.5	Possessed property nouns	306
8.3.6	The rise of property nouns in Flores-Lembata . . .	308
8.3.6.1	Summary	308
8.3.6.2	Evidence for contact-induced change .	310
8.3.6.3	Evidence for internally-driven change .	313
8.3.6.4	Conclusions	315
8.4	Plural marking on nouns	316
8.4.1	Overview	316
8.4.2	Plural in the Flores-Lembata languages	318
8.4.2.1	Central Lembata	318
8.4.2.2	Western Lamaholot	319
8.4.2.3	Sika and Kedang	320
8.4.3	Plural in the AN languages of Timor	320
8.4.4	Plural in the Timor-Alor-Pantar languages	321
8.4.5	The emergence of plural marking	321
8.5	Conclusions	323
9	Innovations in the noun phrase	327
9.1	Introduction	327
9.2	Pre-nominal possessor	328
9.2.1	Overview	328
9.2.2	N-Poss	331
9.2.3	Poss-N	332
9.2.3.1	The Flores-Lembata languages	332
9.2.3.2	The AN languages of Timor	334
9.2.3.3	The Timor-Alor-Pantar languages	335
9.2.4	The emergence of Poss-N	336
9.3	Post-nominal locatives	337
9.3.1	Overview	337

9.3.2	LOC-N	340
9.3.3	N-LOC	342
9.3.3.1	The Flores-Lembata languages	342
9.3.3.2	The AN languages of Timor	347
9.3.3.3	The Timor-Alor-Pantar languages	348
9.3.4	Etymology of the Flores-Lembata locative nouns	350
9.3.5	Suffixes on Flores-Lembata locative nouns	354
9.3.6	The emergence of N-LOC	355
9.4	Post-nominal numerals	357
9.4.1	Overview	357
9.4.2	NUM-N	359
9.4.3	N-NUM	361
9.4.3.1	The Central Flores languages	361
9.4.3.2	The Flores-Lembata languages	361
9.4.3.3	The AN languages of Timor	363
9.4.3.4	Non-AN Timor-Alor-Pantar languages	364
9.4.3.5	Summary	365
9.4.4	Etymology of numerals in Flores-Lembata	365
9.4.5	The rise of post-nominal numerals	369
9.5	Conclusions	370
10	Innovations in the clause	373
10.1	Introduction	373
10.2	Clause-final deictic motion verbs	375
10.2.1	Overview	375
10.2.2	Medial DMV on Flores, Sumba and Timor	377
10.2.3	Final DMV in the Flores-Lembata languages	378
10.2.3.1	Overview	378
10.2.3.2	The Lamaholot varieties	381
10.2.3.3	Kedang	383
10.2.3.4	Sika	384
10.2.4	Final DMV in the Alor-Pantar languages	385
10.2.5	Etymology of the Flores-Lembata DMV	386
10.2.5.1	The Flores-Lembata 'go' verbs	386
10.2.5.2	The Flores-Lembata 'come' verbs	388
10.2.6	The rise of clause-final deictic motion verbs	389
10.3	Clause-final negation	391

10.3.1	Overview	391
10.3.2	Theoretical background: Jespersen Cycle	393
10.3.3	Pre-predicate negation on Flores and Sumba	394
10.3.4	Negation in the Flores-Lembata languages	395
10.3.4.1	Overview of negation patterns	395
10.3.4.2	Pre-predicate single negation	396
10.3.4.3	Double negation	397
10.3.4.4	Clause-final single negation	398
10.3.5	Negation in the AN languages of Timor	399
10.3.6	Etymology of the Flores-Lembata negators	401
10.3.6.1	Overview of negators	401
10.3.6.2	Pre-predicate negators	401
10.3.6.3	Clause-final negators	402
10.3.7	The rise of clause-final negation	403
10.4	Conclusions	406

11	Discussion and conclusions	409
11.1	Overview	409
11.2	Non-Austronesian substrate features	410
11.3	Contact scenarios	414
11.4	Summary of main conclusions	419
11.5	Directions for further research	420

Appendices	425

A	Central Lembata phoneme minimal pairs	425
A.1	Overview	425
A.2	The plosives /p, t, k/ and /b, d, g/	426
A.3	The nasals /m, n, ŋ/	426
A.4	The fricatives /v, s/ and the affricate /dʒ/	427
A.5	The liquids /l, r/	429
A.6	The high vowels /i, u/	429
A.7	The mid vowels /e, ə, o/	431
A.8	The low vowel /a/	432

B	Flores-Lembata lexeme sets	435
C	Orthographies and glosses from other sources	531
C.1	List of orthographic representations	531
C.2	List of glosses from other sources	534
	Bibliography	541
	English summary	563
	Nederlandse samenvatting	565
	Ringkasan dalam bahasa Indonesia	567
	Curriculum Vitae	571

Acknowledgements

Without my team of supervisors, the completion of this dissertation would not have been possible. Thanks to my promotor Marian Klammer, I discovered eastern Indonesia and its linguistic diversity. During the writing process, she always had great advice in structuring my thoughts and chapters. Eugénie Stapert contributed an excellent “outside” perspective from someone with expertise in contact linguistics but not familiar with Austronesian languages and the area of eastern Indonesia. Owen Edwards added valuable experience and knowledge on the languages of the region, in particular on etymologies. In addition, I am grateful to having been able to learn from him how to make and edit language maps. My deepest thanks to all three of you!

Throughout all the stages of my research, I have always profited from the feedback and discussion with my colleagues at LUCL and elsewhere. First of all, I am grateful for the mutual support in our project *Reconstructing the past through languages of the present*, collaborations and regular meetings. Thanks to the project members Owen Edwards, Gereon Kaiping, Marian Klammer, Francesca Moro and George Saad, and to the affiliated PhDs and the student assistants! Many thanks also to Felix Ameka, Nicolas Brucato, Maarten Kossmann, Maarten Mous, Sophie Villerius and all other attendees of the Anthropological and Descriptive Linguistics Discussion Group, the Indonesian Languages Meetings and the LUCL Writing Group. Special thanks to Matthias Pache and Alex Palmér for providing feedback on individual chapters of my dissertation. I am also thankful for the great audiences of the Austronesian and Papuan Languages and Linguistics Conferences APLL8 in London, APLL9 in Paris and APLL11 in Leiden, and the audience of the East

Nusantara Conference ENUS 7 in Kupang for their feedback on my talks. I also very much appreciated the feedback of the reading committee and I am thankful that they took the time to read and to comment on my dissertation.

Several people have been of great help in introducing me to new software programs and always being open to technical questions on my way of learning to use ELAN, FLEX, LaTeX and other programs: Amanda Delgado, Owen Edwards, Gereon Kaiping, Martin Kroon and Sara Petrollino, thank you all!

Without my fellow PhD colleagues, the five years at LUCL would not have been the same. Thank you so much: Astrid van Alem, Bushra Alkhammes, Jermy Balukh, Kate Bellamy, Chams Bernard, Hang Cheng, Amanda Delgado, Carmen Ebner, Ernanda, Aliza Glasbergen-Plas, Adrian Gomes, Han Hu, Elena Karvovskaya, Saskia Lensink, Min Liu, Arjan Mossel, Nazarudin, Ami Okabe, Matthias Pache, Arum Perwitasari, Sara Petrollino, Bobby Ruijgrok, George Saad, Menghui Shi, Gulnaz Sibgatullina, Yunus Sulistyono, Xander Vertegaal, Cesko Voeten, Jiang Wu, Nurenzia Yannuar. I retain great memories of my time as a PhD Council member and of the many different social events and PhD Lunches in the Common Room. My office mates at Room 004 at Wijkplaats 4, many thanks for your company in the first half of my time at LUCL. Special thanks to Nurenzia Yannuar for our time together at our lovely office at Eyckhof 2 during the last two years, for always motivating and supporting each other and enjoying time together. Many thanks also in particular to Eri Yoga Sidharta for the help with the cover picture, to Xander Vertegaal for the Dutch summary and to Yunus Sulistyono for the Indonesian summary of this dissertation.

I am deeply thankful to the speakers of the Central Lembata community in the villages of Kalikasa, Bakan and Lewaji on the island of Lembata in Indonesia, especially Wilhelm Wuwur, Angela “Cici” Lamak and Endang Ujan who put a lot of effort into helping me to learn their language. Without them and the whole speaker community, I could never have realised this dissertation, in particular the first part which is solely based on data that I collected on Lembata. *Terima kasih rai-rai mio swaolmi*. I am also thankful to my family in Hewa for introducing me to Guru Wilem and Ibu Lin in Kalikasa. My fieldwork trips to Indonesia would not have been as easy and enjoyable without the whole supportive network that I had in Kupang and on Java: June Jacob, all staff members of UNKRIS, Marcelinus Akoli and Elvis Bin Toni in Kupang, Obing Katubi and all staff members at LIPI in Jakarta. Svenja, Marc, Katja

and Yunan in Jogja, thank you so much for again and again welcoming me to your lovely home! Siti Makmuria for your and your families' kindness and help during my administrative visit to Jakarta. Thank you so much, all of you!

Finally, my deepest thanks to my parents and my sister who could not have been a better support. My friends from Germany, Hendrickje, Saskia, and Miriam, I am honoured that you are coming to join my defense and celebrate with me here in Leiden. I am thankful to have all of you! I am also grateful for the time together with my aunt and uncle in Rijswijk and my cousins close by who provided a great family support away from home. I am most thankful to Mikael Soge for joining me here in the Netherlands and always supporting me through all phases of my research and life.

List of Abbreviations

Glosses

1	First person
2	Second person
3	Third person
1DET	First person determiner
2DEM	Proximal demonstrative to second person
3DET	Third person determiner
ABS	Absolutive case
ANAPH	Anaphoric
ART	Article
ATTR	Attributive marker
CAUS	Causative
CLF	Classifier
COMPL	Completive
COM	Comitative
DET	Determiner
DISC	Discourse marker
DIST	Distal
EPN	Epenthetic
ERG	Ergative case
EXCL	Exclusive
FIN	End of phrase marker
HUM	Human referent

xx

INCEP	Inceptive, inchoative
INCL	Inclusive
IPFV	Imperfective
LINK	Linker
LOC	Locative
L	Long form of an alternating noun
MOD	Mood marker
NEG	Negation, negator
NMLZ	Nominalizer
NOM	Nominative case
ORD	Ordinal number
PFV	Perfective
PL	Plural
POSS	Possessor
PRES	Present
PROG	Progressive
PROH	Prohibitive
PROX	Proximate
Q	Interrogative
REAL	Realis
REL	Relativizer
SBJ	Subject
SG	Singular
SPEC	Specific
STAT	Stative
S	Short form of an alternating noun
TOP	Topic
VIS	Visible

Languages and subgroups

AL	Alorese
AN	Austronesian
AP	Alor-Pantar
CL	Central Lamaholot
EL	Eastern Lamaholot
FL	Flores-Lembata

KD	Kedang
LH	Lamaholot
PAL	Proto-Alorese
PAN	Proto-Austronesian
PAP	Proto-Alor-Pantar
PCEMP	Proto-Central-Eastern-Malayo-Polynesian
PCL	Proto-Central Lamaholot
PCMP	Proto-Central-Malayo-Polynesian
PFL	Proto-Flores-Lembata
PMP	Proto-Malayo-Polynesian
PTAP	Proto-Timor-Alor-Pantar
PWL	Proto-Western Lamaholot
SK	Sika
TAP	Timor-Alor-Pantar
WL	Western Lamaholot

Other abbreviations

A	Transitive subject argument
adj.	Adjective
C	Consonant
DEM	Demonstrative slot
DISC	Discourse particle slot
DMV	Deictic motion verb
Env.	Environment
L	Liquid = [l,r]
Loc	Locative nouns or topological noun (in the NP)
lit.	Literally, literal meaning
LOC.N	Locative noun slot
LOC.V	Locational verb slot
N	Noun slot
n.	Noun
n.a.	Not applicable
NUM	Numeral (in the NP)
NEG	Negation slot
NP	Noun phrase
NUM	Numeral slot

P	Object argument
Poss	Nominal possessor (in the NP)
pers. comm.	Personal communication
prep.	Preposition
S	Intransitive subject argument
V	Verb slot
V	Vowel
v.	Verb
V1	The first verb in a multi-verb sequence
V2	The second verb in a multi-verb sequence
VP	Verb phrase
Ø	Zero phoneme or morpheme

Citation codes

C1	Breakfast Conversation
C2	Children Conversation
C3	Random Conversation
C4	Grandmother conversation 3
C5	Grandmother conversation 1
C6	Grandmother conversation 2
F1	Frogstory recorded in Kalikasa
F2	Frogstory recorded in Lewaji
FH1	Event and Position Stimuli 1
FH2	Event and Position Stimuli 2
FH3	Event and Position Stimuli 3
FH4	Event and Position Stimuli 4
FH5	Event and Position Stimuli 5
FH6	Event and Position Stimuli 6
FH7	Event and Position Stimuli 7
HC_SR	Hewa Collection: Spatial Relations
I1	Weaving in Kalikasa
I2	Weaving in Lewaji
I3	Weaving in Bakan
L1	Legend of the Stone Boat (Watu Tena)
L2	Origin story of Lewaji
L3	Origin story of Bakan

L4	The Ata Uja Batu Koti clan
M1	Family description
M2	History of Kalikasa
M3	About life in former times
N1	Woodchopper
NB	Notebook 2016-2019
N2	Chore girl
N3	The origin of fire and crop plants
N4	The origin of crop plants
NT	Paper notes 2016
S1	Surrey Stimuli 1
S2	Surrey Stimuli 2
S3	Surrey Stimuli 3
S4	Surrey Stimuli 4
S5	Surrey Stimuli 5
S6	Surrey Stimuli 6
S7	Surrey Stimuli 7
S8	Surrey Stimuli 8
T1	Marriage customs
T2	Death customs
T3	Planting and harvesting customs
T4	Traditional commodities and death customs
UM	Unrecorded materials 2015-2016

CHAPTER 1

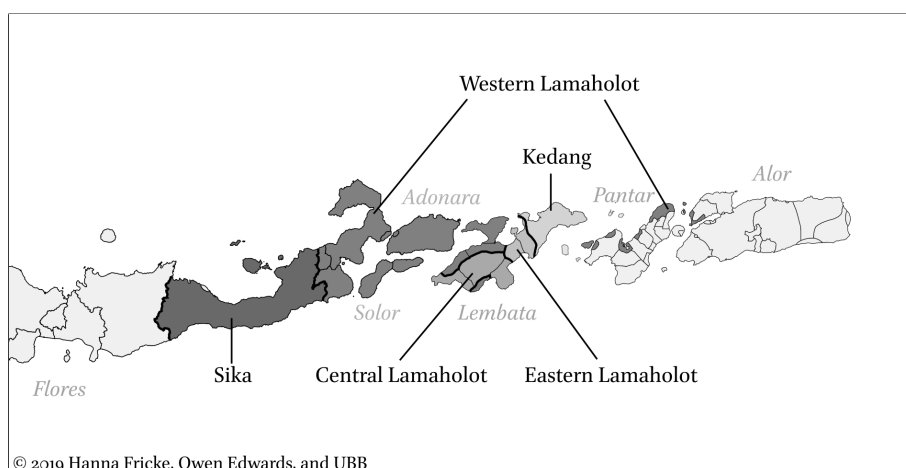
Introduction

“Languages have a “story” to tell: it is our job to find these stories and figure them out.” (Hyman 2001:22)

1.1 Overview

The overall aim of this dissertation is to reconstruct the history of the Flores-Lembata languages including traces of contact-induced change in these languages. The Flores-Lembata languages are spoken in eastern Indonesia and constitute a low-level subgroup within the Austronesian language family. The Flores-Lembata group can be divided into five linguistically defined subgroups. These are: Sika, Western Lamaholot, Central Lamaholot, Eastern Lamaholot and Kedang, as shown on the map in Figure 1.1. Each of these subgroups includes one or more languages. Proposed language boundaries are indicated as lines on the map. A more detailed introduction on the Flores-Lembata languages is §4.1.

Figure 1.1: The Flores-Lembata subgroups



This dissertation is arranged in three parts: Part I Grammatical description, Part II Historical phonology and lexical innovations, and Part III Morpho-syntactic innovations.

Part I fills a gap in the documentation of the Flores-Lembata languages by providing a descriptive grammar of the language Central Lembata which is part of the Central Lamaholot subgroup. In Chapter 2 of Part I, I provide an introduction to this grammar and Chapter 3 constitutes the grammatical description itself. This grammar is the first extensive description of a language belonging to the Central Lamaholot subgroup. In order to be able to carry out the comparative work on the Flores-Lembata family in Part II and Part III of this dissertation, it was essential to add to the description of the Flores-Lembata languages and describe a variety of Central Lamaholot. Also the Eastern Lamaholot subgroup is largely undescribed but the scope of this dissertation only allows for the description of one language. Central Lamaholot was chosen because, based on wordlists collected by Keraf (1978a), this subgroup appeared to be more diverse and innovative, both in lexicon and morphology than Eastern Lamaholot. The other subgroups, Sika, Western Lamaholot and Kedang, are relatively well documented and described (cf. §4.1). The language of Central Lembata within the Central Lamaholot subgroup was targeted because of previously established contacts with speakers of this language which provided an important starting point for field research. Phonologically conservative but innovative in mor-

phology and lexicon, Central Lembata adds considerably to the knowledge about the Flores-Lembata languages which are the object of investigation in the following parts.

Part II concerns the historical phonology and the lexicon of the Flores-Lembata languages. Chapter 4 introduces the dataset and the methodology used for the analysis of phonology and lexicon. Chapter 5 is a historic-comparative study of the Flores-Lembata phonology showing the regular reflexes of Austronesian sounds in the Flores-Lembata languages and revealing the internal structure of the family, as well as the family's higher-level affiliation in the Bima-Lembata subgroup. The results of this study contribute to a more fine-grained understanding of the history of Austronesian languages in eastern Indonesia. Chapter 6 examines the origin of lexical items in the Flores-Lembata languages and provides evidence for a lexical substrate.

Part III examines eight morpho-syntactic features and their origin in contact. These eight features of the Flores-Lembata languages are atypical for Austronesian languages. Chapter 7 is an introduction to the dataset and the methodology used for the detection of contact-induced structural features. In the Chapters 8, 9 and 10, the eight features are described and their potential of being the result of contact with non-Austronesian languages of the area is evaluated.

The ultimate aim of this dissertation is the synthesis of the results from Part II and III in Chapter 11 and the discussion of them against the background of language contact theory.

This dissertation contributes to three fields: (i) to the documentation of the linguistic diversity in Indonesia, (ii) to the reconstruction of the history of the Austronesian languages in eastern Indonesia, and (iii) to the systematic study of morpho-syntactic change due to contact across language families. As such, this dissertation contributes to the field of descriptive and contact linguistics as well as to a reconstruction of the linguistic and social history of eastern Indonesia.

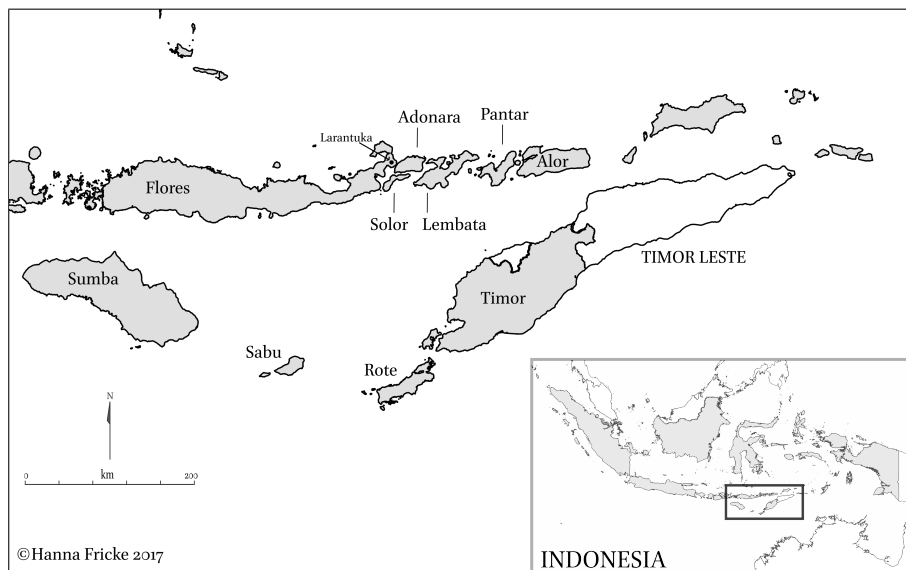
This chapter is a general introduction to this dissertation and is structured as follows. In §1.2, I provide background information on eastern Flores and the Solor Archipelago where the Flores-Lembata languages are spoken. In §1.3, the case study of Flores-Lembata is located within its wider areal and linguistic context. In §1.4, the research questions for this dissertation are laid out and connected to the following chapters.

1.2 Eastern Flores and the Solor Archipelago

1.2.1 Geographic location

The island of Flores and the adjacent Solor Archipelago are part of the Lesser Sunda islands located on the southern hemisphere in island Southeast Asia, north-west of Australia. Politically, Flores and the Solor Archipelago are part of the Indonesian province of Nusa Tenggara Timur which also includes the islands of Sumba, western Timor and several smaller surrounding islands, as shown on the map in Figure 1.2. The country Timor-Leste is located in the eastern part of the island of Timor.

Figure 1.2: The Indonesian province of Nusa Tenggara Timur



The island of Flores extends about 350 km from west to east and only about 50 km from north to south. The Solor archipelago includes three smaller islands, Solor, Adonara and Lembata, which cover a distance of about 100 km from the western tip of Solor to the eastern tip of Lembata. In the west, Flores has high mountains which reach altitudes over 2,000 m, as well as areas of highlands, while in the east of the island the mountains reach altitudes of over 1,500 m. Most of the mountains of Flores, as well as in the Solor

Archipelago, are active volcanoes. The climate in the Lesser Sunda islands is tropical, showing little variation in temperature, but characterised by monsoon winds which bring heavy rain from November to March. The remaining time of the year is mainly dry. Original vegetation consists of dry forest that is less dense than equatorial rain forest (cf. Bellwood 2017:16-17). Nowadays, crop fields have replaced huge parts of forest, but large-scale plantations are still very rare on these islands, probably also due to the mountainous landscape.

1.2.2 Livelihood and beliefs

The cultural features presented in this section and in §1.2.3.3 further below are to a large extent summarised from the answers to a cultural questionnaire I conducted in the Central Lembata speaking village of Lewaji on the island of Lembata. The interviewee Servasius Boli was the headmaster of the local school and native to Lewaji. The questionnaire is archived in the Central Lembata Corpus (Fricke 2019). Most of these cultural features described here are representative for the whole area of eastern Flores and the Solor Archipelago.

Traditionally, inhabitants of eastern Flores and adjacent islands hunt with bow and arrow or by using traps and slingshots. But as the number of wild animals is declining, people more and more rely on animal husbandry with pigs, goats, chicken and dogs. Horticulture, extensive agriculture based on rainy seasons and intensive irrigation agriculture is found throughout the area. Agriculture based on irrigation is more common in flat areas, while extensive agriculture is found in the mountains. The main crops are rice, corn, roots and tubers, such as cassava. Rice is a rather new crop in the area. About 50 years ago, it was not yet the main staple which it has become today. The main fruits and vegetables are papaya, banana, mango, avocado, cassava leaves, eggplant and different kinds of beans. People also plant coffee, cacao, cashew, coconuts and candlenuts and sell them to merchants. In coastal areas, people go fishing or collect sea food at low tides. At specific seasons, whale hunting is carried out by the people of Lamalera (Barnes 1996). Men and women are both involved in farming activities. Hunting is traditionally a male domain, while women are occupied with weaving cloth.

Traditional life is organised around ceremonies and rituals related to farming, birth, marriage and death. Most of these ceremonies and rituals

are kept up until today. Important elements in ceremonies and rituals are betel nuts and palm wine. The traditional creator in the Lamaholot areas of eastern Flores and the Solor Archipelago is *Lera Wulan*, the 'sun-moon' god and the ruler of the sky. His counterpart is *Tana Ekan*, the 'land-soil' goddess and the ruler of earth. Together they symbolise a conceptual dichotomy representing the male and the female side, the sky and the earth (Arndt 1951; Aritonang and Steenbrink 2008:73).¹ Traditional beliefs include ancestor worship by regular offerings of little amounts of food and other valuable things, which is still practised in villages nowadays. Since missionary activities started in the 16th century, Catholicism has spread over eastern Flores and the Solor Archipelago and is now the religion of the majority. However, there are also villages, especially in the coastal areas of Lembata, which have been Islamic for centuries. Protestant Christian minorities are only found in towns.

1.2.3 Divisions in the Lamaholot society

1.2.3.1 Overview

In the following sections, I discuss three types of potential socio-cultural lines of division attested especially in the Lamaholot groups. These are (i) a contrast between traditionally weaving and non-weaving communities (cf. §1.2.3.2), (ii) a contrast between native and immigrant clans (cf. §1.2.3.3) and (iii) a division of villages into two groups of socio-political alliances: Paji and Demon (cf. §1.2.3.4). In §1.2.3.5, I conclude that these three binary contrasts do not line up nor do they match with the three linguistic Lamaholot subgroups (cf. §1.1). Nevertheless, this information is interesting for the investigation of language contact in the past, as such divisions might have originated in a division of two linguistically different groups.

1.2.3.2 Weaving and non-weaving communities

The process of producing woven cloth goes from harvesting cotton, spinning threads, producing natural colors to dye threads, often using complex *ikat* techniques, up to weaving the cloth on a wooden loom (cf. Barnes 1989).

¹ The same dichotomy is also found in the mythology of the Tana 'Ai region, the south-eastern edge of the Sika speaking area (Lewis 1988:76).

However, not all areas in eastern Flores and the Solor Archipelago are traditionally weaving areas. Barnes (1987) states that Adonara and east Solor are traditionally non-weaving areas, while eastern Flores and west Solor have strong weaving traditions. On Lembata, it appears that some coastal areas developed strong weaving skills, probably because they were strategically located for trade, while this is not the case in mountain areas. In the Kedang area and on the Mingar peninsular, both on Lembata, weaving was traditionally prohibited. Reasons for this prohibition cannot be retrieved from the sources available (Barnes 1987:21). This division between weaving and non-weaving communities may point to a contrast between Austronesian and non-Austronesian culture, as weaving cloth has been considered as a typical Austronesian cultural feature (Blust 2013:24).

1.2.3.3 Clan systems within a community

Socially, the inhabitants of the region are organised in clans. A community with a common language is composed of several clans. Each clan has its own history of origin which is orally transmitted and distinguishes, on the one hand, native clans, which, according to local legends, emerged out of the earth or the mountains at place, and, on the other hand, immigrant clans coming from overseas (cf. §1.2.4).² Clan association is inherited paternally and clan heads are men. Thus, clans are usually patrilineal and patrilocal (woman moves to the village of the husband's family after marriage) but there are a few matrilineal clan structures found in the Tana 'Ai region of Sika (Lewis 1988:111). There are specific rules connected to clan membership and potential marriage partners. Generally, a person is not allowed to marry someone from his or her own clan. People can marry a person from another

² For the Western Lamaholot area, Arndt (1938:29) distinguishes native clans that originate from the mountains, called *ilé jadi* 'mountain born', and immigrant clans from other islands. In the Central Lamaholot villages Bakan and Lewaji, I have documented a similar division between native clans that emerged out of the earth and immigrant clans that came from other islands, such as Kroko Pukén or Awo Lolón, from where they had to flee due to natural disasters. In Bakan, there are four native clans, i. e. Blikon, Krésaj, Lamak and Ata Uja Wai Lolo, and at least two immigrant clans, i. e. Ata Wuwur and Ata Uja Bat Koti (L3:378-379). In Lewaji, there is just one native clan, i. e. Pewut. Several new clans came from the surrounding villages and from further away to settle with them (L2:5,37-38,45-47). This distinction between native and immigrant clans most likely holds for the whole Lamaholot area, possibly also for Kedang and Sika.

clan of the same village or from a different village.

In the Central Lembata area, a triangular marriage system is found. The clans in a village are grouped into three units with a more or less equal amount of members. This means that big clans may form their own unit, while small clans are grouped together. An example is taken from the village of Lewaji which has the following units: (i) members of the Wawin clan, (ii) members of the Tukan clan, and (iii) members of the Bakeor clan, the Pewut clan and the Lamadua clan. Following the Lewaji marriage rules, Wawin men marry Tukan women, Tukan men marry Bakeor/Pewut/Lamadua women and Bakeor/Pewut/Lamadua men marry Wawin women. This kind of triangular marriage system is also found in other communities with different clans being involved.

The male family members of potential wives (the father or an uncle of the potential wife) are referred to, as well as addressed, using the Central Lembata word *opo* or with the Malay equivalent *om* (< Dutch *oom* 'uncle'). The people addressed as *opo* or *om* are highly respected. Marriages are preceded by negotiations between the two clans of the potential partners. The clan of the groom has to pay a bride price to the bride's family. This is traditionally an ivory tusk. The bride's family gives jewellery and woven cloths to the groom's family in return. If all duties are paid, the wife moves to the husband's family but they will normally build their own house. However, modern times have changed settlement patterns and married couples might choose their place of living more freely, sometimes also moving to a town or another island for work. However, marriage rules about whom to marry are still taken rather seriously and breaking them is considered bad.

1.2.3.4 Socio-political division: Paji and Demon

Throughout the Lamaholot area, the population is divided into two different socio-political groups: Paji and Demon.³ Each Lamaholot village belongs either to the group of Paji or of Demon. There is neither such division in the Sika area, nor in Kedang (Arndt 1938:11; Barnes 1987:19)

Arndt (1938:29) suggests that the division between Paji and Demon goes very far back in time, but the names Paji and Demon were only given to

³ Alternative names are Demonara and Pajinara (or Painara in languages that underwent *dʒ > y). The word *nara* is Javanese, ultimately Sanskrit *nāra* 'human' (Arndt 1938:29,35).

these groups under the influence of the Hindu Javanese kingdom of Majapahit in the 14th century. According to origin myths from Adonara and Solor, the Demon-Paji division goes back to a fight between two brothers named Demon and Paji (Arndt 1938:3-5). The Demon-Paji division appears to be mythologically grounded in Adonara and Solor but more political and less mythical on Lembata (Barnes 1987:18). Kedang only became part of this political division when the king of Adonara gained power over Kedang and thus forced Kedang to become part of the Paji alliance (Barnes 1987:19).

Arndt (1938:10,22) suggests that Paji and Demon were not brothers but unrelated groups of people. The following features of Paji and Demon are collected from orally transmitted stories and described by Arndt (1938). The Paji, also called Beda or Pati by the Demons in some myths, do not show pride in themselves but they rather seem to be without courage and are depressed (Arndt 1938:3). They are a mix of different people (Arndt 1938:40). They lived on Kroko Puken but also in eastern Flores (Arndt 1938:5). In eastern Flores they lived in big numbers, initially outnumbering the incoming Demon (Arndt 1938:10). Nowadays in eastern Flores, the Paji have almost completely been replaced by Demon. They mainly live at the island's edges, such as Tanjung Flores, and also in the southwest and northeast of Adonara, east Solor, the Minggar peninsular on Lembata, the Lerek peninsular on Lembata, and on the Lewotolo peninsular (= Ile Ape) on Lembata (Arndt 1938:39; Barnes 1987:16). It is said that they do not give offerings to the Earth, neither have a village temple (*korke*) or that they build their temples poorly compared to the Demon (Arndt 1938:19,22,28).

The Demon, also called Pagong, speak proudly of themselves but degrade the Pajis (Arndt 1938:3). They lived on Kroko Puken together with Paji and got into a fight. Then they left to eastern Flores, where they took the land from the Paji and the Paji had to flee (Arndt 1938:9-10). They had a village temple (*korke*), music, songs and dances (Arndt 1938:22).

When the new religions Islam and Christianity began to compete in the Solor Archipelago about the time of the 16th century, the Paji, living in coastal areas, converted to Islam and had contact with Javanese traders. Later they became allies of the Dutch who were the competitors of the Portuguese. The Demon lived more in the interior and were open to Christianity which made them allies of the Portuguese. This choice of religion and political alliances reinforced the antagonism between the two groups (Aritonang and Steenbrink 2008:77).

At the end of the 16th century, the enmity became even more political when both sides joined a different local king. Despite conflicting evidence about the details of the story, it can be concluded that Igo, the king of Larantuka, a town at the eastern edge of Flores, got into a fight with his younger brother Enga. Eventually, Enga joined the Paji and became king of the Paji in Adonara and Igo stayed the leader of the Demon in Larantuka (Arndt 1938:26,44-47).

This strong enmity between the two groups led to many fights and victims on both sides. Some myths connect this to the god Lera Wulan (cf. §1.2.2) who asked for human skulls as offerings. If he did not receive them, the rain would not come (Arndt 1938:9). During the past decades, the enmity has lost ground and nowadays people live together peacefully although each village still recalls to which side they belonged.

Barnes (1987:21) describes an association of Demon affiliation with *ikat* weaving and use of red dye, and an association of Paji affiliation with little or no *ikat* weaving, the use of black dye and weaving restricted to coastal areas and done predominantly for trade. However, there are several exceptions to this binary association. Especially on Lembata but also in eastern Flores, as well as in small areas of Solor and Adonara, the association above does not hold.

The beginning of the division of Paji and Demon is difficult to connect to historic dates. From Portuguese documents, it is known that in 1624 the division was already in place on the island of Lembata (Barnes 1987:18). As the Paji-Demon division appears to be more mythologically grounded on Solor and Adonara, but less on Lembata, it is likely that it did not originate from Lembata. If the division was already in place in the early 17th century, its origins must lay further back in time. In fact, the politicisation of the division due to a fight between the royal brothers Iga and Enga in the kingdom of Larantuka has been dated to the end of the 16th century. This means that the division must have its roots in times earlier than that. As Arndt (1938) suggests that the current labels Paji and Demon came with the Hindu-Javanese influence but that the division into two groups is much older, it must go back to time before the 14th century. However, this hypothesis is impossible to prove with the current data available.

1.2.3.5 Conclusions

A socio-political division such as the Paji-Demon distinction could have originated in an ethnic-linguistic division between incoming speakers of Austronesian languages (=Demon and weaving) on the one hand and speakers of native non-Austronesian languages (=Paji and non-weaving) on the other hand. However, this remains difficult to prove for the Lamaholot case for the following reasons. The sources on the history of the Paji-Demon distinction are mythical origin stories and there is no evidence for a fundamental linguistic division between the groups. People belonging to both, the Paji and Demon group, are spread over all three subgroups in the Lamaholot area. This could only be explained when assuming that the Paji adopted the Austronesian language of the Demon, and only after that this language diversified into the three subgroups of Lamaholot. An additional complication for the Paji/non-Austronesian/native versus Demon/Austronesian/immigrant hypothesis is that none of the two groups declares themselves as being entirely native or immigrant, a distinction in origin prominently found in clan origin stories throughout the whole area (cf. §1.2.3.3 and §1.2.4). Among the native clans, some belong to the Paji and others to the Demon (Arndt 1938:33). Again, one could suggest that the distinction between native and immigrant clans is younger than the Paji-Demon distinction. Among the immigrant clans, there are some that say the Paji-Demon distinction already existed on the island of Kroko Puken from where they fled before arriving in the Solor Archipelago (Arndt 1938:9). Nevertheless, the Demon-Paji division, in addition to the weaving-non-weaving division (Barnes 1987) and the native-immigrant clan distinction mentioned earlier, can be regarded as another piece of evidence for a socio-cultural divide in the history of the Lamaholot people.

1.2.4 History

1.2.4.1 Overview

In this section, I provide an overview of the history of eastern Flores and the Solor Archipelago drawing from oral and written sources. The first two sections concern migration histories and the last two section concern colonial history and the role of the Malay language and later the Indonesian

language in the area.

The history of this area is characterised by migrations of local groups of people and shifting power relations of external forces from Java and Europe. All over the area, oral history reports that there are native clans that emerged out of the ground or out of the mountains and immigrant clans that have come from other islands (cf. §1.2.3.3). §1.2.4.2 introduces the Alorese, a Western Lamaholot group that migrated to the east and settled in the Alor Archipelago which is adjacent to the Solor Archipelago. §1.2.4.3 discusses the migration from the islands Lepad and Batan to the Solor Archipelago which reoccurs in origin stories of clans and in legends. §1.2.4.4 is about the establishment of European settlements, mainly Portuguese, with the aims of Christianisation, trade and gaining political power. §1.2.4.5 describes the more recent influence of the national language Indonesian since the country became independent in 1945. §1.2.4.6 is a summary of the topics discussed in this section on history.

The information on the history of Flores and the Solor Archipelago is collected from published sources and from my own recordings made on Lembata. My own records are indicated with citation codes pointing to recordings in the Central Lembata Corpus (cf. §2.4). A list of citation codes used in the corpus and in this dissertation is given in the list of abbreviations on page xxii.

1.2.4.2 Alorese migration to Alor and Pantar

Speakers of Alorese (a Western Lamaholot language) have been settled since the 14th century in coastal areas of Pantar and Alor which are located to the east of the Solor Archipelago (Stokhof 1975:8; Klamer 2011:8-15; Wellfelt 2016:248-249). Shared sound changes and innovations between Alorese and other Western Lamaholot varieties shown in Chapter 5 of this dissertation confirm that the Alorese language is clearly part of the Western Lamaholot subgroup (cf. §4.1.3).

Oral history of the Alorese on Pantar tells about the arrival of Javanese on Pantar and their establishment of kingdoms in the 14th century (Moro 2018:180). It remains unclear whether the founder of the Alorese kingdoms were actually Javanese or if these founders were the Western Lamaholot people coming to Pantar and Alor. There are two possible scenarios. First, the Western Lamaholot people came to Pantar in the 14th century and foun-

ded kingdoms based on legends relating to Javanese kings. This is not unlikely as at that time the whole area was influenced by the Hindu-Javanese kingdom Majapahit and there is no doubt of legends emerging from there. Second, it could also be that there were actually people related to Majapahit coming to Pantar and initiating the foundation of kingdoms among the Alorese in the 14th century. These Majapahit people were not necessarily ethnically Javanese but may have been allies of the Majapahit kingdom which had its centre on Java. This scenario would presuppose that the Alorese themselves had been already on Pantar in coastal areas before that time and recognised the newcomers as their leaders. This would also mean that the new leaders adapted to the Alorese language rather than introducing their own language. The number of Alorese speakers might have been higher than the number of incoming speakers of other languages, or Alorese could have been already a lingua franca in the Alor Archipelago used by speakers of Alor-Pantar languages as a second language to communicate with people on the coast. Thus, Alorese became the language of the kingdoms in the coastal areas as well as on smaller off-shore islands.

1.2.4.3 Migration from Lepad Batan

Many clans in the Solor Archipelago trace their origins back to the island Kroko Puken, also called Lepad Batan, from where they had to flee due to a natural disaster about 500 years ago (Vatter 1932:9-10; Barnes 1982). These immigrant clans are also referred to as *téna mau* lit. 'boat float' (Ataladjar 2015:10-11).⁴ Lepad and Batan are two small islands located towards the east of the Solor Archipelago, in the strait between Lembata and Pantar (Ataladjar 2015:17). The inhabitants of these two islands fled from a flood, maybe caused by a tsunami or a volcanic eruption. In most myths about how this disaster came about a *naga* snake is involved. People fled with boats to Alor, Pantar, Lembata, Adonara, Solor and Flores, some even to Timor (Ataladjar 2015:21-22). Some of these boats became stones according to the legends, and are still there, such as the *watu téna* 'stone boat' on a beach of southern Lembata (L1, Ataladjar 2015:32-34).

The core clans from Lamalera, a village on the south coast of Lembata, even relate their history further back to Seram Goram in the Moluccas and

⁴ Ataladjar (2015:8) does not provide the meaning of *mau* but according to Pampus (1999) *mao* means 'float' in WL-Lewolema.

Luwuk on Sulawesi. They came as allies of the Hindu-Javanese kingdom of Majapahit which extended its power over most parts of present-day Indonesia at its peak between 1350 and 1389, including Flores and the Solor Archipelago (Aritonang and Steenbrink 2008:73; Barnes 2013:39-40; Ataladjar 2015:10,12-13). According to oral history, the ancestors of the clans from Luwuk came as people of Gajah Mada, the prime minister of the Majapahit empire at that time, to Seram Goram and then further to Lepan Batan. In the legend, it is said that Gajah Mada himself came to Lepan Batan together with the people from Luwuk. When the clans from Luwuk came to Lembata, they found people there and they started to live with them together peacefully.

The Nagarakertagama chronicles of Majapahit written in 1365 suggest that the places of origin, such as Luwuk and Seram Goram, transmitted by oral history are plausible. The chronicles provide a list of Majapahit dependencies in eastern Indonesia, among others these are Solor, Pantar (there called Galiyao), Timor, Seram (also: Seran) and Gorom (also: Gurun, Goran) in the Moluccas and Luwuk on Sulawesi (Pigeaud 1960:34; Pigeaud 1962:34; Barnes 1982; Hägerdal 2012:22; Ataladjar 2015:9-10).

In the 16th century, the islands of Lepan and Batan are mentioned in the records of European sailors. The islands were located on the route of trade ships from Europe to Timor and Moluccas and it is known that Magellan's ship "Victoria" passed Lepan Batan on 9 January 1522 (Ataladjar 2015:19). Francisco Albo, the helmsman on Magellan's ship Victoria, writes in his diary about their travel through the Alor Archipelago in January 1522.⁵ He mentions two little islands in between Pantar and Lembata which have been identified as Lepan and Batan by Le Roux (1928:12). Albo writes in Spanish *y son habitadas* 'and (they) are inhabited'. Le Roux (1928:14) doubts that his statement is related to the islands of Lepan and Batan but claims that Albo meant the islands of Pantar and Lembata which Albo also mentions in the same sentence. This assumption by Le Roux is based on the fact that Lepan and Batan were no longer inhabited at the time when Le Roux was writing, thus in the early 20th century. However, we know from oral history

⁵ At that time Magellan himself had already died in a fight on an island of the Philippines. Only one of his five ships, the Victoria, finally completed the journey around the world. The Victoria was under the command of Juan Sebastian de Elcano and the diaries of Antonio Pigafetti in Italian and of Francisco Albo in Spanish have become the most important documents on the journey of this ship (Le Roux 1928:1).

that people had to flee from these islands due to a natural disaster, so it is likely that Lepad and Batan were still inhabited when the “Victoria” passed in 1522. Ataladjar (2015:36) uses Albo’s statement as a basis to assume that the disaster must have happened after 1522.⁶ In the history of Larantuka, people from Lepad Batan are mentioned. These people were already there when Sira Napang became king of Larantuka in 1525 (Ataladjar 2015:37). From this, Ataladjar (2015:37) concludes that the disaster of Lepad Batan must have happened between 1522 and 1525. Nowadays, Lepad is a coral island below sea level, whereas Batan is about 250 m above sea level and its land is fertile, although neither of the islands has any fresh water source. On Batan, there is still an old ruined village which points to past inhabitation.

In the stories, Lepad Batan is described as fertile and having a safe harbor called Leffo Hajjo. The people had a good life on Lepad Batan and they had trade connections to Munaseli and Pandai on Pantar (Ataladjar 2015:15-16). According to stories from Lembata, crucial elements of the present-day Lamaholot culture were found on Lepad Batan (Ataladjar 2015:18). The people of Lepad Batan already believed in the god Lera Wulan Tana Ekan (cf. §1.2.2), built a *korke* and had *nubanara* stones. The *korke* is a temple to worship Lera Wulan Tana Ekan (‘Sun-Moon Land-Soil’) and the ancestors (cf. §1.2.2). The *nubanara* stones are an important element in rituals.

The people on Lepad Batan could have been a group of Western Lamaholot speakers similar to the Alorrese that have settled in the coastal areas of Pantar and Alor, close to the island Lepad and Batan, since at least the 14th century (cf. §1.2.4.2). This does not exclude the possibility that there was an additional minority of people from the Moluccas as reported by Lamalera’s oral history. If indeed the people on Lepad Batan were descendants of Western Lamaholot speakers, this could explain why the immigrants from Lepad Batan have integrated so well with the Lamaholot in the Solor Archipelago. Clans that claim to originate from Lepad Batan are linguistically not distinct from other clans in the Lamaholot area. Linguistic boundaries rather follow village boundaries and do not mark clan distinctions.

⁶ Vatter (1932:9-10) had suggested mid or end of the 17th century as the date for the Lepad Batan disaster which was based on traditional counting of generations. On the other hand, Barnes (1982:411) suggested 1450 also based on generation counting.

1.2.4.4 Colonialism and Christianisation

In the middle of the 16th century, Dominican priests from Portugal, followed by Portuguese traders and soldiers, came to the Solor Archipelago via Malacca where they had settled earlier (Aritonang and Steenbrink 2008:75). The Dominicans established their first settlement on the island of Solor in 1561. Portuguese priests, but also laymen marrying local women, started to christianise the local population (Abdurachman 2008). When the Dutch conquered the Solor fortress in 1613, some Portuguese and christianised locals moved to Larantuka. An additional influx of Christian people to Larantuka came in the 17th century when Portuguese, Christian Malays and Christian Chinese fled, via Makassar, from Malacca which had been conquered by the Dutch in 1641 (Aritonang and Steenbrink 2008:87). These migrants from from the west are also known as *sina jawa malaka* or *lau wai* ‘from the water’ (Ataladjar 2015:10-11). After the fall of Malacca in 1641, Larantuka gained more and more importance and additional settlements of incoming Christians were established in Wureh on Adonara and Konga in eastern Flores.

At this time the Malay variety of the town Larantuka started to develop as the people from Malacca spoke a variety of Malay. Malay became the main language of the mixed population of Larantuka, including indigenous Lamaholot, migrants from Rote, Sabu, Sulawesi and Ternate, as well as Portuguese marchants, solidiers and priests. Over time, this language developed into its own variety of Larantuka Malay (Steinhauer 1991:181). The speakers themselves call their language Nagi (Paauw 2008:66). Larantuka Malay is also reported to be spoken in several villages around Konga bay in eastern Flores and in the village Wure on Adonara (Steinhauer 1991:180; Paauw 2008:69). The strong influence from Portuguese and the direct descentance from Peninsular Malay (Malacca) makes the history of Larantuka Malay fundamentally different from the history of other Malay varieties in eastern Indonesia, such as Alor Malay or Kupang Malay (Paauw 2008:11). In the wider Solor Archipelago, the local languages of the Flores-Lembata family were the main means of communication. Larantuka Malay was not used much, as Portuguese missionaries experienced difficulties in communicating with people from the villages (Aritonang and Steenbrink 2008:76). Other varieties of Malay were not used either until a variety of Malay, different from Larantuka Malay, was introduced in primary education in the first half of the 20th century (Aritonang and Steenbrink 2008:244).

The decline of the Portuguese and the rise of the Dutch as colonial powers in eastern Indonesia during the 18th and 19th century reduced the Dominican influence on Flores and in the Solor Archipelago. In the early 20th century, Catholic missionaries from the Societas Verbi Divini (SVD), mainly German, started to become active in the area and cooperated with the Dutch colonial government in developing Flores and the Solor Archipelago (Aritonang and Steenbrink 2008:94,244). The first primary schools in the area with Malay as language of instruction were established by SVD missionaries. Since then bilingualism in local languages and Malay started to spread to wider parts of the society.

Nowadays the languages of the colonial powers, Portuguese and Dutch, are no longer used in eastern Flores and the Solor Archipelago, but loanwords, such as Central Lembata *kédéra* 'chair' from Portuguese *cadeira* [ka'dejra] 'chair' or *bal* 'ball' from Dutch *bal* 'ball' remain. These words are different from the corresponding Indonesian words *kursi* 'chair' and *bola* 'ball', which are also loans but with different histories.

1.2.4.5 The Indonesian national language

In 1945 Indonesia declared independence. With the independence movement in the first half of the 20th century, Malay was chosen to become the national language and since then, under the name 'Indonesian', the language was standardised further and developed into a full national language for administration and education (Sneddon 2003). In this dissertation, I use the term 'Indonesian' to refer to the local variety of Indonesian that people speak in eastern Flores and the Solor Archipelago. This variety diverges in various respects from Indonesian that is spoken in the west of the country, such as on Java, as well as from written Indonesian. It also diverges from Lantaka Malay (cf. §1.2.4.4). The local variety of Indonesian used in eastern Flores and in the Solor Archipelago has not yet been studied in particular.

Initially, Indonesian was a second language that was only acquired when children started to go to school. Nowadays, Indonesian, or more precisely its local variety, is an additional first language next to the local languages, even the only first language for many children. Code-switching between Indonesian and local languages is common practice and many loan words from Indonesian have been added to the lexicon of the local languages. According to the results of a cultural questionnaire that I conducted in the village

Lewaji on Lembata (cf. §1.2.2), couples with different local language background either use the language of the husband or Indonesian as their family language. In case the couple lives in the village of the husband, the wife will learn the local language of the village. However, when they live somewhere else, they may rather use Indonesian as their family language. Nowadays, some couples with different language backgrounds, even though living in the village of the husband, may chose Indonesian as their family language. Indonesian is continuing to gain more and more importance, while the local languages become more marginalised. During the political period of the New Order, in the 1960ies to the 1990ies, local languages were forbidden at schools with the aim of strengthening the new national language Indonesian. These days, local languages are not suppressed anymore. Nevertheless, local languages are used less and less due to the fact that they are only useful in the context of traditional village life. If asked, people usually agree that it is a pity that the local languages are used less by younger people.

1.2.4.6 Summary

The eastward migration of the Alorese speakers to the islands of Alor and Pantar in the 14th century is the earliest migration movement known for the region of eastern Flores and its adjacent islands. Further, oral history tells that in the 16th century a devastating natural disaster forced a large number of people to flee from the small islands Lengan and Batan located between the island of Pantar and the island of Lembata. These people settled throughout the Solor Archipelago and until today, they trace back their clan origin to Lengan Batan. Also in the 16th century, the first Catholic priests from Portugal arrived in the Solor Archipelago and began their missionary work. In the 17th century, the town of Larantuka became a center of Christianity and settlement for migrant groups from different parts of Island Southeast Asia, such as from Malacca. The 17th and 18th century were characterised by a rise of Dutch colonial power. In the 20th century, a new Catholic order, the Societas Verbi Divini (SVD), began their missionary work in the area. Until today, SVD is the most wide-spread Catholic missionary order in eastern Flores and the Solor Archipelago. Since independence in 1945, the Indonesian national language, a variety of Malay, has gained more and more importance due to its use in administration and education. This has also caused a decline in use of local languages.

1.3 The wider areal context

1.3.1 Linguistic diversity

The Flores-Lembata languages of eastern Flores and the Solor Archipelago are located in eastern Indonesia, an area of linguistic diversity and contact. Eastern Indonesia is characterised by the presence of Austronesian languages and languages of non-Austronesian ('Papuan') families which have co-existed and influenced each other for about 3,500 years.⁷ This contact has led to linguistic features diffusing between languages regardless of their genetic affiliation (Klamer et al. 2008:10,136; Ewing and Klamer 2010a). This contact zone of Austronesian and non-Austronesian languages has been studied as a linguistic area labelled East Nusantara (Klamer et al. 2008; Ewing and Klamer 2010b; Holton and Klamer 2017). A slightly smaller linguistic area has been named Wallacea (Schapper 2015). Both proposed areas are characterised by a specific set of linguistic features that have diffused over language family boundaries and are attested in Austronesian as well as non-Austronesian languages.

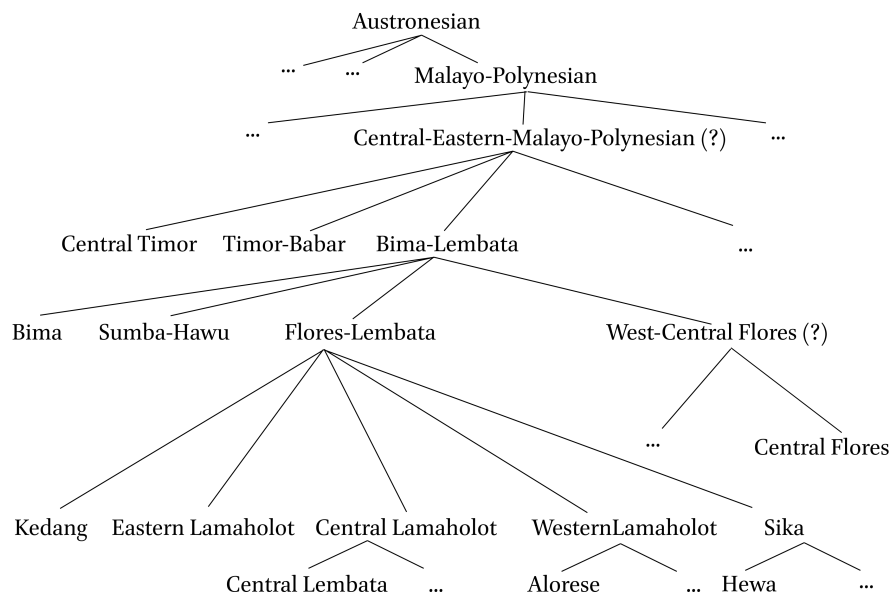
The area studied in particular for this dissertation is part of this contact zone and covers the province of Nusa Tenggara Timur (NTT) and the country of Timor-Leste (cf. Figure 1.2 on page 4). This area has been chosen as it covers the Flores-Lembata languages and its western as well as eastern neighbours of both Austronesian and non-Austronesian language families. In the area studied for this dissertation, there are languages of two language families: Austronesian languages and Timor-Alor-Pantar (TAP) languages which are Papuan.⁸

⁷ The non-Austronesian language families of this area are often referred to by the cover term 'Papuan' although they do not form a genealogical unit with a common ancestor. The term 'Papuan' has been used because most of the non-Austronesian language families are found on the island of New Guinea and in its vicinity. The western part of New Guinea belongs to Indonesia and is called (West) Papua, while the eastern part of the island constitutes the country Papua New Guinea. In this dissertation, the term 'Papuan' is used interchangeably with 'non-Austronesian' as it refers to all languages of the region that do not belong to the Austronesian languages family.

⁸ The Timor-Alor-Pantar languages are now the western-most non-Austronesian languages spoken in eastern Indonesia. Until 1815, the now extinct Tambora language was spoken on Sumbawa island which is located further west than TAP. This language was probably not related to either the Austronesian or Timor-Alor-Pantar (Donohue 2007).

The Austronesian (AN) language family is, in terms of number of languages and geographic spread, one of the biggest language families worldwide. The 1,200 Austronesian languages are spoken from Taiwan in the north to New Zealand in the south, and from Madagascar in the west to the Easter Islands and Hawai'i in the east. Virtually the whole area in between these points is populated by speakers of Austronesian languages, with the exception of the island of New Guinea and Australia (Blust 2013:1-3). The AN languages spoken in the area of study, displayed in Figure 1.3, are part of the subgroup of Malayo-Polynesian which includes all AN languages outside of Taiwan, the homeland of the Austronesian language family (Blust 2013:30).

Figure 1.3: The AN languages of the area: current stage of classification



Further subgrouping within Malayo-Polynesian is still debated (Blust 1993; Donohue and Grimes 2008; Blust 2009; Klamer 2019). This includes the putative AN subgroup Central-Eastern-Malayo-Polynesian (CEMP) which is supposed to include the area of study for this dissertation and all Austronesian languages further north and east, excluding only the island of Sulawesi.

This dissertation provides evidence for a Bima-Lembata subgroup encompassing the languages of Sumba, Sabu, Bima, Western Flores, Central

Flores and Flores-Lembata (cf. §5.5). Within the Bima-Lembata subgroup Sumba-Hawu (Blust 2008), Central Flores (Elias 2017b) and Flores-Lembata (§5.4) have been identified as innovation-defined subgroups. Further likely subgroups within Bima-Lembata are Bima and West-Central Flores (*Flores Barat* in Fernandez 1996). The details of the internal structure of Bima-Lembata remain to be investigated.

Two further low-level subgroups within (CE)MP can be established on the island of Timor: Central Timor and Timor-Babar (Edwards 2018b; Edwards 2019). Central Timor encompasses the languages Welaun, Tokodede, Mambae and Kemak on the island of Timor. Timor-Babar includes the remaining AN languages of the island of Timor as well as languages further to the east.

The languages of the Timor-Alor-Pantar (TAP) family are spoken on the islands of Alor and Pantar, as well as in East Timor (Klamer 2017; Schapper 2014b). Several hypotheses about the relation of the TAP languages to other non-Austronesian language families in the wider area have been proposed but none of these has yet been demonstrated on the basis of enough evidence to draw definite conclusions (Holton and Robinson 2017b). At the current stage of research, the TAP languages are most likely either an isolate family or are related to the languages of West Bomberai (Holton and Robinson 2017b:183-184). Within the TAP languages, the Alor-Pantar languages of the islands of Alor and Pantar form a clear subgroup excluding the TAP languages of Timor (Holton and Robinson 2017a). Subgroupings of the Timor-Alor-Pantar languages have been proposed in Holton et al. (2012), Schapper et al. (2012); Schapper et al. (2017) and Kaiping and Klamer (to appearb).

1.3.2 Genetic diversity

Archaeology and population genetics reveal two major waves of modern human (*homo sapiens*) migration into island Southeast Asia: an earlier arrival of non-Asian populations and a later influx of Asian populations (Hudjashov et al. 2017:2447).⁹ Descendants of the earliest wave of migrants are still living on New Guinea highlands and in Australia. But also in parts of the Philippines and eastern Indonesia a high degree of ancestry from this

⁹ Other human species, *homo erectus* and *homo floresiensis*, have been attested in island Southeast Asia preceding the first occurrence of *homo sapiens* (Bellwood 2017:34-85).

earlier population is attested (Bellwood 2017:86-87).

The first modern humans entered the region in the Late Pleistocene about 50,000 years ago — colloquially referred to as Late Ice Age, a time when the islands of Sumatra, Java and Borneo still formed an extension of the Asia continent called the Sunda shelf. Much later, about 4,000-5,000 years ago, the Malayo-Polynesian branch of the Austronesian languages and its speakers moved from Taiwan into island Southeast Asia and further west and east (cf. Karafet et al. 2010:1833; Bellwood 2017:181; Hudjashov et al. 2017:2439-2440).¹⁰ Some later migration phases are connected to the spread of Hinduism starting almost 2,000 years ago, the later spread of Islam several hundred years ago and more recent Chinese influence (Karafet et al. 2010:1841-1842).

Traces of all migration phases are still present in the modern day genome of the Indonesian people (Karafet et al. 2010:1842). There appears to be a clear boundary between Asian genomes in the west and non-Asian genomes in the east between the islands of Bali and Flores, more specifically between Sumbawa and Flores (Karafet et al. 2010:1837-1838; Hudjashov et al. 2017:2447). This genome boundary coincides with the Wallace line, a boundary of transition between Asian and Australian types of flora and fauna.

The genetic sampling results of Hudjashov et al. (2017) provide more specific data on the Indonesian province of Nusa Tenggara Timur and the country of Timor-Leste which are the focus of this dissertation. The sampling results from the eastern part of Flores (Bama), from Lembata and from Timor show an almost half-half division of the Papuan and the Austronesian genome component. This genetic profile is very similar to the sampling results on the islands of Alor and Pantar which show only a slightly higher Papuan component. Towards the west, the central and western part of Flores (Bena and Rampasasa) and Sumba show a smaller Papuan component than found further east. On Sumba, the Austronesian component is the biggest. On Flores, the admixture is more diverse, also including other Southeast

¹⁰ The Austronesian expansion may have contributed to the Neolithic transition in island Southeast Asia. But there is also evidence for neolithic influence from mainland Southeast Asia, such as domesticated pigs, and also New Guinea has been recognised as a center of neolithic innovations. For example, the cultivation of bananas in the region most likely has its origins in New Guinea and was taken towards the west from there (Hudjashov et al. 2017:2440).

Asian components (Hudjashov et al. 2017:2442). Overall, the eastern Indonesian islands sampled in Hudjashov et al. (2017), all show a very similar picture of genetic contact. Following different types of analyses, Hudjashov et al. (2017:2445) deduce a date for the latest admixture of around more or less 2,000 years ago between two populations: one “non-Asian/Papuan-like” type and one “Indonesian/Philippine-like” (Austronesian) type. There is a crucial difference between the Austronesian component of western Flores and Sumba on the one hand and the Austronesian component of Lembata, Alor, Pantar and Timor on the other hand. The source for the Austronesian admixture on western Flores and Sumba contains a larger proportion of Java and Bali components, while the Austronesian admixture on Lembata, Alor, Pantar and Timor is dominated by Sulawesi inheritance (Hudjashov et al. 2017:2447).

The dates of the most recent major genetic admixture between the “Austronesian” and “Papuan” components in eastern Indonesia contrast with the archaeological signals of the Austronesian expansion in the region that date back to about one thousand years earlier (Hudjashov et al. 2017:2447-2448). This suggests that there were either several migration waves out of Taiwan which blur the genetic signal, or that culture and technology spread earlier than genetic contact took place (Hudjashov et al. 2017:2448).

1.4 Research questions and aims

The Flores-Lembata languages in eastern Indonesia are located in an area that is linguistically mainly Austronesian, with the exception of the Timor-Alor-Pantar family (cf. §1.3.1). In contrast, the genetic profile of the population appears to be mixed with almost equal “Austronesian-like” and “Papuan-like” components (cf. 1.3.2). Assuming that usually languages match with genomes, this miss-match of genome and language family raises the hypothesis of language shift: speakers of non-Austronesian languages may have shifted to Austronesian languages. Language shift can leave traces of the languages that are disappearing in the structure of the languages people are shifting to, especially when preceded by a period of bilingualism (Thomason and Kaufman 1988; Muysken 2010; Ross 2013).

Based on the hypothesis that languages match with genomes and non-Austronesian speakers shifted to incoming Austronesian languages, non-

Austronesian features are expected to be found in the Austronesian languages of the area. Previous research has indeed reported non-Austronesian traits, specifically in some of the Austronesian languages of Nusa Tenggara Timur and Timor-Leste. In particular for Western Lamaholot, a subgroup of Flores-Lembata, a whole set of non-Austronesian structural features has been reported by Klamer (2012a). Further examples of non-Austronesian traces in Austronesian languages of this area include (i) sound correspondences with no Austronesian source (Edwards 2016b), (ii) plural words (Moro 2018), and (iii) non-decimal numeral systems (Schapper and Hammarström 2013; Schapper and Klamer 2017).

With this background, this dissertation aims to examine in more detail the contact history of the Flores-Lembata family by taking a regional comparative perspective. This family is particularly interesting because the Flores-Lembata languages are genealogically closely related to their western neighbours: the Austronesian languages of Western and Central Flores (cf. Chapter 5). However, morpho-syntactically, they are mixed and share features with their eastern neighbours of the Timor-Alor-Pantar family, as well as with the Austronesian languages of Timor (cf. Chapter 8, 9 and 10).

The grammatical description of the Central Lembata language in Part I provides the description of a language belonging to the previously undescribed Central Lamaholot subgroup of the Flores-Lembata family, as laid out in §1.1. After having filled this documentation gap, Part II investigates the phonological and lexical history of the Flores-Lembata family and answers the following research questions:

- (1) How are the Proto-Malayo-Polynesian (PMP) proto-sounds reflected in the Flores-Lembata languages?
- (2) What is the evidence for (i) subgroups within Flores-Lembata, (ii) Flores-Lembata as a subgroup, and (iii) Bima-Lembata as a higher-level subgroup including Flores-Lembata and other Austronesian languages?
- (3) Which Flores-Lembata lexical items are inherited from an Austronesian ancestor?
- (4) Which Flores-Lembata lexical items are not inherited from an Austronesian ancestor?
- (5) Which Flores-Lembata lexical items can be reconstructed to Proto-Flores-Lembata?

Questions (1) and (2) are addressed in Chapter 5 and questions (3) to (5) are addressed in Chapter 6. Chapter 5 is a top-down approach to the historical phonology of Flores-Lembata. In this chapter, I show that Proto-Malayo-Polynesian (PMP) sounds are regularly reflected in the languages of Flores-Lembata. Based on exclusively shared sound changes, I provide evidence for Flores-Lembata as a subgroup, as well as for five lower-level subgroups within this family. In addition, I propose the change of PMP *b > w in the same set of lexical items as evidence for a higher-level subgroup Bima-Lembata that includes, next to Flores-Lembata, the Austronesian languages of Central and Western Flores, of Sumba, Sabu and Bima.

Chapter 6 describes the Flores-Lembata lexicon with a focus on words inherited from PMP, on the one hand, and words that cannot be shown to have a PMP origin, on the other hand. I show that Proto-Flores-Lembata (PFL) has a comparably small component of non-Austronesian vocabulary, while the individual subgroups of Flores-Lembata have added more new vocabulary to different extents. There is an increase of non-Austronesian vocabulary towards the geographic centre of the Flores-Lembata subgroup which includes the three Lamaholot subgroups.

Part III of this dissertation aims at the comparative analysis of grammatical innovations in the languages of Flores-Lembata to detect possible contact-induced changes. The Flores-Lembata languages underwent a considerable number of structural innovations that are not attested in their closest genealogical relatives further west, the other languages of the Bima-Lembata subgroup. In contrast, most of these features are present in Austronesian languages of Timor, as well as in the non-Austronesian Timor-Alor-Pantar languages. The following research questions are addressed in Part III.

- (6) Which structural features in the Flores-Lembata languages are innovations?
- (7) Which structural innovations can be attributed to contact-induced change?

Chapter 8 is concerned with three morpho-syntactic innovations. (i) The Lamaholot subgroups have developed an alienability distinction in their possessive constructions. (ii) All Flores-Lembata have reanalysed property words as nouns. (iii) Central Lamaholot also innovated plural marking on nouns via grammaticalisation of the third person plural pronoun.

Chapter 9 describes three word order changes in the noun phrase that are found in all Flores-Lembata languages. (i) Possessor nouns are pre-nominal, while (ii) locative nouns and (iii) numerals are post-nominal. This is the opposite to the inherited Austronesian word order in the noun phrase. However, this is the word order found in the non-Austronesian languages of the area.

Chapter 10 discusses the innovation of two clause-final elements in the languages of Flores-Lembata. (i) All Flores-Lembata languages have clause-final deictic motion verbs, such as ‘come’ and ‘go’, although the general constituent order is verb-medial. In the Lamaholot subgroups and in Kedang, but not in Sika, these clause-final verb can also encode elevation. (ii) The Lamaholot subgroups, but not Kedang and Sika, have innovated clause-final negators. While Eastern Lamaholot and Central Lamaholot still retain also a pre-predicate negator at the same time, Western Lamaholot has only clause-final negation.

In Chapter 11, I synthesise the results from Part II and Part III and provide a hypothesis on the contact scenario that possibly led to the innovations in lexicon and grammar established in the preceding chapters by answering the last two research questions:

- (8) What kind of scenario led to the lexical and structural changes in the Flores-Lembata languages?
- (9) What kind of language(s) was/were the contact language(s)?

I argue that the Flores-Lembata languages acquired most of their innovative features through contact-induced change via language shift preceded by long-term language mixing in a bilingual community. The now extinct contact languages were typological very similar to the Timor-Alor-Pantar (TAP) languages but there is not enough lexical evidence to assume that the contact languages were a branch of TAP.

Part I

Grammatical description

CHAPTER 2

Introduction to Part I

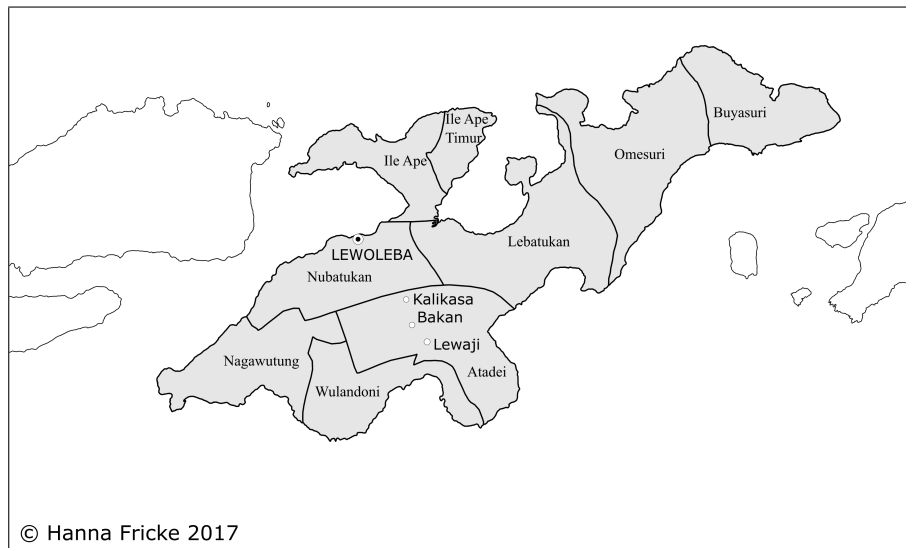
2.1 Location and speakers

Central Lembata is spoken by about 3,000 people in six villages in the central mountains on the island of Lembata: Kalikasa (Katakeja), Bakan (Ile Kerbau), Lewaji (Dori Pewut), Kolilerek (Tubuk Rajan), Waimuda (Nuba Boli) and Mudalerek (Nogo Doni).¹ All six Central Lembata villages are part of the Atadei district (Indonesian: *kecamatan*) in the Lembata regency (Indonesian: *kabupaten*).² The number of speakers is calculated according to recent census data (Badan Pusat Statistik Kabupaten Lembata 2018). The map in Figure 2.1 shows the three biggest villages of the Central Lembata speaking area. The map also contains district names and boundaries in the Lembata regency. Kalikasa is the administrative capital of the Atadei district.

¹ The names in brackets are the official village names only used in very formal or written contexts, while the other names are the original names that have been in use since long before these villages became officially administrative units of *desa* (Indonesian: ‘village’). The official *desa* names were created by village elders and are not used in daily conversation.

² Lembata has been a single regency since 1999. Previously, it was part of Flores Timur which is located towards the west of Lembata, encompassing the eastern tip of Flores and the islands Adonara and Solor.

Figure 2.1: Fieldwork locations on Lembata



In the Atadei district, infrastructure is still developing. During the time of my fieldwork trips, electricity was produced by generators but only for a few hours at night. Kalikasa had mobile coverage whereas the other two villages, Bakan and Lewaji, did not. The inhabitants of the Atadei district are mainly farmers. There are only a few places of employment like in schools or the medical center in Kalikasa. This is why many people leave the area to work in the regency's capital Lewoleba or even further away in Kalimantan, Java or Malaysia. Roads are available reaching all villages in the district with daily running trucks to transport passengers and goods. However the roads are generally in bad states which makes travelling slow and uncomfortable.

Kalikasa is officially named Katakeja 'I am the one in the middle' in the language of Central Lembata. This refers to its central position on the island of Lembata. The word Kalikasa denotes a cactus plant which is found in the area and which used to be planted around the village for protection. The official name of the village Bakan is Ile Kerbau which means 'buffalo hill'. This name was chosen because the hills around Bakan used to be full of buffaloes (L2:117). According to oral history, the word *bakan* has three meanings (i) 'plain', (ii) 'a kind of grass' and (iii) 'older siblings' (cf. L2, L3). The village Lewaji is officially named Dori Pewut meaning 'joining the bees'. This goes back to the origin story of the village Lewaji which tells about the Pewut

'bee' clan as the only native clan of the village that emerged out of the earth and other immigrant clans that fled from a tsunami that had destroyed their island Lapan Batan and joined the Pewut 'bee' clan (L2:36-48, §1.2.4.3). The name Lewaji is short for *lewu wajinu* 'the village of the younger siblings'. This goes back to an event in oral history that caused the inhabitants of Lewaji to be seen as younger siblings, thus subordinate to those of Bakan, which they had to address from then on as 'older sibling' (L2:110). It becomes clear that the official names are the newer names of the villages because their meaning is transparent in the local language, whereas the meanings of the older names are already more obscure and only known by certain people, mainly elders. Sometimes the older names have several possible meanings, as in the case of Bakan.

Based on oral history (L3:398-408), the oldest village among the Central Lembata villages is Bakan and people have come from there and founded the other surrounding villages of Kolilerek, Kalikasa, Waimuda and Mudalerek. However, for Lewaji, it has been said that there were people already living there with which the Bakan people started to live together (L2:37-39).

The Atadei district reaches further southeast and includes about 16 villages where the neighbouring language Atadei Painara (also: Eastern Atadei, or South Lembata) is spoken.³ Based on information from a sociolinguistic questionnaire archived in the Central Lembata Corpus (Fricke 2019) and on information given in Keraf (1978a), the variety of Central Lembata is also spoken in a few more villages towards the west. Among those are Boto (Laba Limut) and Udek in the Nagawutung district and Lewuka (Belobao) in the Wulandoni district. Nowadays, for administrative and infrastructural reasons, there is not much contact between the Central Lembata people from Atadei and those from the other districts.

In previous publications, the Central Lembata area was considered to be part of the Lamaholot dialect chain (Arndt 1937; Keraf 1978a; Grangé 2015b), although very little information was available on the languages of that area. During my fieldwork, it became evident that the speakers of Central Lembata do not consider their language to be Lamaholot. Lamaholot for them is the language spoken in the northern part of Lembata and on the other is-

³ According to district records the Atadei Painara speaking villages are: Karangora (Ile Kimok), Waipei, Atalojo, Bauraja, Waiwejak, Lewokurang, Paulolo (Leba Ata), Lewokoba, Watuwawer (Ata Kore), Atawolo (Lusi Lame), Brenai, Lerek, Lamanuna, Duli, Tobilolong and Baulolong. Not all of these are administrative units of *desa*.

lands to the west, Adonara, Solor and Eastern Flores, thus the language that has been linguistically classified as Western Lamaholot (cf. §4.1). I chose the language name Central Lembata as a geographical label to cover the six Atadei villages described above. The speakers themselves do not have a specific name for their language variety. They would refer to it as *bahasa daerah* (Indonesian: ‘local language’), or combine the Indonesian term *bahasa* ‘language’ with a village name, such as *Bahasa Bakan* ‘Bakan language’.

2.2 Data collection and processing

The fieldwork to collect data for the grammar of Central Lembata in Chapter 3 was conducted in three villages: Kalikasa (Katakeja), Bakan (Ile Kerbau) and Lewaji (Dori Pewut) in the Central Lembata speaking area (cf. §2.1). The central part of Lembata was chosen as field site as it shows the highest linguistic diversity of the Lamaholot area, and the varieties spoken there had not been studied before (cf. §1.1 and 4.1.4). The data collection in the field was carried out for a total of about 9 months distributed over the years of 2015, 2016 and 2018. Out of these 9 months, two months were spent working with native speakers of Central Lembata in Yogyakarta, a student town on Java.

Recordings of Central Lembata speech were stored in ‘wav’ format for audio data and in ‘mp4’ format for video data. All audio recordings, in most cases with corresponding video, were transcribed using the linguistic annotation software ELAN produced by the Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands, accessible at <https://tla.mpi.nl/tools/tla-tools/elan> (Wittenburg et al. 2006).⁴ Transcription was done by me together with native speakers of Central Lembata. While transcribing, I made notes on word meanings. ELAN produced ‘eaf’ files with time-aligned annotations and these were exported as ‘flextext’ files and then imported into a FLEX project that contains a lexical database. FLEX is free-ware provided by SIL at <https://software.sil.org/fieldworks>. The program allows linguists to build a text corpus with glossings and translations that are connected to a lexical database that can be built and expanded while glossing. In the FLEX corpus, the texts were glossed and translated into Eng-

⁴ One recording of a conversation could not be transcribed due to bad quality (cf. C4).

lish and in most cases also Indonesian. This was also done with help of native speakers when needed. Finally, the glossed and translated transcriptions were re-imported into ELAN to link them to the recordings. These eaf files are archived together with their corresponding recordings in the Central Lembata Corpus (cf. §2.3). For some of the transcriptions, the exported glossings are still work in progress.

2.3 The Central Lembata Corpus

The data collected is archived at <https://hdl.handle.net/1839/ed1aa6f0-93da-4a44-8ddb-66221cbc01ce> in the Central Lembata Corpus at the MPI Language Archive in Nijmegen, the Netherlands. Table 2.1 gives an overview of the records in the corpus. It contains video and audio recordings of speakers native to Kalikasa, Bakan, Lewaji and surrounding villages. The time indications refer to the length of the whole recordings. The time of recorded Central Lembata free speech amounts to approximately 7,5 hours in form of narratives (~ 4,5 hours) and conversations (~ 3 hours). The total recordings are much longer because parts of the recordings are in Indonesian and there are also recordings that do not contain free speech but other contents, such as wordlists or elicited sentences.

The main elicitation tasks used were the Surrey Stimuli (Fedden et al. 2010) and the Event and Position List (Moro and Fricke submitted). Both elicitation lists are collections of short video clips and some pictures that the speaker had to describe using Central Lembata. The wordlist and the cultural questionnaire contributed to the LexiRumah database (Kaiping et al. 2019) and the CultureRumah database (Kaiping and Klamer to appear).

Apart from recordings, a lexical database of about 1,800 Central Lembata lexical entries is part of the corpus. A community dictionary based on this lexical database will be published with a local publisher in Indonesia.

Table 2.1: Overview of records in the Central Lembata Corpus

Genre	Recorded time	Speakers
Conversation	4 hours	11 females; 6 males 6 <30 years; 3 <60 years; 2 >60 years; 6 unknown
Narratives	5 hours	4 females; 10 males 2 <30 years; 8 <60 years; 4 >60 years
Surrey Stimuli	2 hours	3 females; 5 males 3 <30 years; 5 <60 years
Event and Position List	2,5 hours	2 females; 4 males 2 <30 years; 4 <60 years
Other elicitations	4,5 hours	1 female; 2 males 2 <30 years; 1 <60 years
Wordlist (600 items)	2 hours	1 male 1 <60 years
Sociolinguistic questionnaire	1,5 hours	2 males 2 <60 years
Cultural questionnaire	40 minutes	1 male 2 <60 years
Total	22 h 10 min	

2.4 Citation codes

Each recording session in my data collection has a citation code which is deduced from the original file name of the associated recording. A list of citation codes is found in the list of abbreviations on page xxii. These codes are used when citing examples from the corpus in this dissertation and elsewhere. An example of a citation would be L1:50. This means that the example sentence or phrase is found in the recording with the code L1 in line number 50. The line corresponds to the segment number in the corresponding ‘eaf’ file that can be opened with the program ELAN, as well as to the line in the flextext file readable by FLEX (cf. §2.2). In the FLEX corpus, as well as in the archive, the titles of the texts in the FLEX corpus all start with

the citation code. In the electronic version of this dissertation, the citation codes are linked to the Language Archive where the recordings are publically accessible (cf. §2.3).

A few examples in this dissertation are taken from the Hewa Collection collected by me in 2013 in the village Hewa, in the Wulanggitang district of the Flores Timur regency, Indonesia. This collection is archived at the MPI Language Archive in Nijmegen, the Netherlands (Fricke et al. 2019). The coding system for examples from this corpus follows the same strategy, with the difference that HC (=Hewa Collection) is the beginning of the code and that the line number does not correspond to an ELAN annotation but to a line in the corresponding Toolbox text file. An example is the code HC_SR:31 which refers to line 31 in the text SR (=Spatial relations) in the Hewa Collection.

2.5 Transcription and glossing conventions

All linguistic examples in the grammar of Central Lembata are either transcribed using the International Phonetic Alphabet (IPA) or a practical orthography with each character representing one phoneme. This orthography is laid out in §7.3. IPA transcriptions, in phonemic (/ /), are used in the section on the Central Lembata phonology, while the practical orthography, in italic font, is used in all other sections of the grammar.

For abbreviations in glosses, I follow the Leipzig glossing rules with my own additions whenever necessary (Max Planck Institute for Evolutionary Anthropology 2015). A list of all glosses and their meanings is given in the list of abbreviations on page xix.

CHAPTER 3

A grammar of Central Lembata

3.1 Structure of this grammar

In this grammar, I describe the phonology of Central Lembata (§3.2), nouns and noun phrases (§3.3), pronouns (§3.4), verbs (§3.5) and basic clausal syntax (§3.6). §3.7 is a description on location as this is a prominent feature in the language that interacts with several parts of the grammar. §3.8 is a summary of the chapter.

3.2 Phonology

3.2.1 Phoneme inventory

Central Lembata has 14 phonemic consonants displayed in Table 3.1 and 6 phonemic vowels listed in Table 3.2. Minimal pairs that demonstrate the phonemic status of the Central Lembata consonants and vowels are given in Appendix A. In the phoneme tables of §3.2.1, orthographic representations diverging from IPA symbols are given in angle brackets.

Table 3.1: Consonant inventory of Central Lembata

	Labial	Coronal	Dorsal
Voiceless plosive	p	t [t̚]	k
Voiced plosive	b	d	g
Nasal	m	n	ŋ <ng>
Trill		r	
Fricative	v <w>	s	
Affricate		dʒ <j>	
Lateral		l	

Central Lembata has 6 plosives that appear in voiceless-voiced pairs. These are bilabial /p b/, coronal /t d/ and dorsal /k g/. The voiceless plosive /t/ of the coronal pair is articulated dentally as [t̚], whereas /d/ is alveolar. The 3 nasals in the consonant inventory of Central Lembata occur in the same places of articulation as the stops. The two fricatives, voiceless coronal /s/ and voiced labiodental /v/, and the voiced coronal affricate /dʒ/ form a natural class. They all appear in all positions of the word and lack a voicing contrast. In the phonology of Central Lembata, the affricate /dʒ/ could also be analysed as a fricative, the voiced counterpart of /s/. This becomes evident from morpho-phonological devoicing of /dʒ/ to /s/ when following a plosive. This is the case when the plural suffix /dʒa/ is added to a plosive-final root, such as /ra.mut/ ‘root’. The affricate /dʒ/ is transformed into /s/ and the plosive is deleted which leads to the word /ra.mu.sa/ ‘root-PL’ (cf. §3.3.2). The labiodental fricative /v/ can also be realised as a labiodental approximate [v̥] with lower lip and upper teeth slightly more apart. Finally, Central Lembata has two phonemic liquids, the voiced lateral /l/ and the voiced coronal trill /r/.

Table 3.2: Vowel inventory of Central Lembata

	Front	Central	Back
High	i		u
Mid	e <é>	ə <e>	o
Low		a	

Central Lembata has two high vowels, the unrounded front vowel /i/ and the rounded back vowel /u/, three mid vowels, the unrounded front vowel /e/, the central vowel /ə/ and the rounded back vowel /o/, and finally one low vowel, the central vowel /a/. The phonemes /e/ and /o/ can have more open variants [ɛ] and [ɔ], especially in closed syllables. This alternation is free variation rather than based on a strict allophonic rule.

3.2.2 Phonotactics

3.2.2.1 Distribution of single phonemes

In Table 3.3 and 3.4, I indicate all permitted positions for single consonants and vowels. A plus sign marks permitted positions for phonemes and a minus marks disallowed positions. Consonants and vowels in Central Lembata have a wide distribution, only a few distribution constraints apply, summarised in (1).

- (1) Distribution constraints of Central Lembata phonemes
 - a) The voiced plosives /b, d, g/ cannot appear in the coda.
 - b) The nasal /ŋ/ cannot appear in the root-initial onset.
 - c) A schwa /ə/ cannot appear root-finally.

Table 3.3: Distribution constraints of consonants

Position	p	b	t	d	k	g	m	n	ŋ	v	s	dʒ	r	l
syllable level														
onset	+	+	+	+	+	+	+	+	+	+	+	+	+	+
coda	+	-	+	-	+	-	+	+	+	+	+	+	+	+
root level														
initial	+	+	+	+	+	+	+	+	-	+	+	+	+	+
medial	+	+	+	+	+	+	+	+	+	+	+	+	+	+
final	+	-	+	-	+	-	+	+	+	+	+	+	+	+

Table 3.4: Distribution constraints of vowels

Position	i	u	e	ə	o	a
syllable level						
initial	+	+	+	+	+	+
medial	+	+	+	+	+	+
final	+	+	+	+	+	+
root level						
initial	+	+	+	+	+	+
medial	+	+	+	+	+	+
final	+	+	+	-	+	+

Although being allowed in all positions, the phoneme /dʒ/ is marginal in initial position. In my corpus, it only occurs in four native words, which are /dʒəma/ ‘night; time period’, /dʒua/ ‘two’, /dʒene/ ‘upwards’ and /dʒune/ ‘downwards’.

Schwa cannot appear root-finally but it can be found in non-final open syllables or in final syllables of derived words. The latter is rare but found in derivatives that result from consonant deletion, such as /iməŋ/ > /imə/ ‘place’ or /kudʒəŋ/ > /kudʒə/ ‘cooking pot’. More about consonant insertion and deletion in alienable nouns is found in §3.3.1.2 later on. In the phonetic realisation of consonant-final words, schwa can be epenthetically added at the end of the word. Those schwas are not phonemic.

3.2.2.2 Consonant clusters

Central Lembata has several clusters of two consonants. Clusters of more than two consonants do not occur. In careful speech, consonant clusters can be broken up by the insertion of a non-phonemic schwa. Table 3.5 provides an overview of the consonant clusters found in onset position in my corpus of Central Lembata. A dash represents a cluster that is not attested.

Table 3.5: Onset consonant clusters in Central Lembata

C1→ C2↓	/p/	/b/	/t/	/k/	/g/	/s/
/p/	-	-	-	/kp/	-	/sp/
/b/	-	-	-	/kb/	-	/sb/
/t/	/pt/	-	-	/kt/	-	-
/d/	-	/bd/	-	/kd/	-	-
/k/	/pk/	-	-	-	-	/sk/
/g/	-	-	-	-	-	/sg/
/v/	/pv/	/bv/	-	/kv/	/gv/	/sv/
/s/	/ps/	-	-	/ks/	-	-
/dʒ/	-	/bdʒ/	-	-	-	-
/m/	-	-	/tm/	/km/	/gm/	/sm/
/n/	/pn/	/bn/	/tn/	/kn/	/gn/	/sn/
/l/	/pl/	/bl/	/tl/	/kl/	/gl/	/sl/
/r/	/pr/	/br/	/tr/	/kr/	/gr/	/sr/

Several of these non-attested clusters can be explained by the regular constraints listed in (2) below, while some of the gaps cannot be explained by the same restrictive rules. These gaps may be due to a lack of data.

- (2) Combination constraints of consonants in word-initial clusters.
- The first consonant of a cluster has to be a plosive or the voiceless fricative /s/.
 - If the first consonant is voiced, the second consonant cannot be voiceless.
 - Two consonants only distinguished by voice or by nasality cannot be combined.
 - The plosive /d/ cannot be the first consonant of a cluster.
 - The affricate /dʒ/ is only combined with preceding /b/.

Universally, syllables are often built according to a sonority hierarchy which means that sonority increases rather than decreases towards the nucleus of the syllable (Blevins 1995; Parker 2002:8). However, the Constraints (2a) and (2b) can only be explained partially by this universal tendency. Plosives tend to precede other consonant in a cluster (Constraint 2a), and voiceless

elements precede voiced phonemes in the onset of a syllable (Constraint 2b). But not all clusters appear to be sonority based. Clusters such as /kp/, /sp/, /sb/ or /sg/ violate the sonority hierarchy.

Further, Constraint (2c) could arise due to the articulatory similarity of the the two consonants involved. At first glance, the cluster /tn/ appears to be an exception to this. However, this cluster is also distinguished in place of articulation, as the phoneme /t/ is articulated as dental [t̪] and /n/ is alveolar. For Constraint (2d) and (2e) no explanation could yet be found.

Table 3.6 lists an example of each consonant cluster and the number of lexical items that contain this cluster in my corpus.

Table 3.6: Consonant clusters with frequencies and examples

Cluster	Number of lexemes	Example	Gloss
/k/-initial			
/kn/	36	/kna.ver/	‘door’
/kl/	25	/klo.rer/	‘morinda tree’
/kv/	19	/kva.lek/	‘shoulder’
/kr/	18	/kro.kor/	‘tree sp.’
/kb/	15	/kbe.kek/	‘strong’
/km/	13	/kmi.ker/	‘flying fox’
/ks/	11	/ksi.lap/	‘lightening’
/kd/	7	/kdu.uk/	‘spin’
/kp/	7	/kpa.e/	‘be silent’
/kt/	4	/ktu.mav/	‘cloth louse’
TOTAL	155		
/s/-initial			
/sn/	29	/snə.gur/	‘smoke’
/sl/	7	/sla.e/	‘be clean’
/sb/	6	/sbu.ur/	‘scatter’
/sv/	5	/sve.gir/	‘uneven’
/sr/	4	/sru.dəm/	‘kneel’
/sp/	3	/spa.a.ti/	‘four of them’
/sm/	2	/sme.i/	‘blood’
TOTAL	56		

Cluster	Number of lexemes	Example	Gloss
<i>/g/-initial</i>			
<i>/gr/</i>	15	<i>/gri.di/</i>	'be angry'
<i>/gl/</i>	13	<i>/gla.sa/</i>	'play'
<i>/gm/</i>	6	<i>/gma.di/</i>	'call'
<i>/gn/</i>	6	<i>/gne.tem/</i>	'pass on'
<i>/gv/</i>	4	<i>/gve.sa/</i>	'pull'
TOTAL	44		
<i>/b/-initial</i>			
<i>/bl/</i>	28	<i>/blə.bət/</i>	'cage'
<i>/bn/</i>	16	<i>/bni.sak/</i>	'broken'
<i>/br/</i>	6	<i>/bra.koŋ/</i>	'lush'
<i>/bdʒ/</i>	3	<i>/bdʒa.e/</i>	'hillwards'
<i>/bd/</i>	3	<i>/bdo.rɨŋ/</i>	'follower'
<i>/bv/</i>	1	<i>/bva.rak/</i>	'hot'
TOTAL	37		
<i>/p/-initial</i>			
<i>/pl/</i>	11	<i>/pli.rar/</i>	'k. o. basket'
<i>/pr/</i>	9	<i>/prə.vak/</i>	'thick'
<i>/pn/</i>	7	<i>/pna.it/</i>	'bitter'
<i>/pt/</i>	2	<i>/pto.de/</i>	'pound'
<i>/ps/</i>	2	<i>/psa.at/</i>	'curse'
<i>/pv/</i>	1	<i>/pvə.li/</i>	'pay (ceremonial)'
TOTAL	32		
<i>/t/-initial</i>			
<i>/tn/</i>	11	<i>/tni.saŋ/</i>	'hole'
<i>/tm/</i>	7	<i>/tmə.lar/</i>	'flea'
<i>/tl/</i>	1	<i>/tlu.por/</i>	'aubergine'
<i>/tr/</i>	1	<i>/trə.kət/</i>	'stick'
TOTAL	20		

In contrast to the complex onsets listed above, a complex coda is not possible in Central Lembata. Word-medially, sequences of consonants can oc-

cur but they are different from onset clusters for two reasons. Word-medial sequences are always divided by a syllable boundary, or even a morpheme boundary, where the first consonant is the coda of the preceding syllable and the second consonant is the onset of the next syllable. Thus, word-medial sequences only occur at a synchronic morpheme boundary or in loan words.

In the examples in Table 3.6, it can be seen that there is a tendency for /n/, /l/ and /r/ to be the most frequent second consonant. Taking the sonority hierarchy into account, it is not surprising that /l/ and /r/ are frequently attested as the second member of a cluster. This is because liquids are higher in sonority than plosives for example and thus usually found closer to the nucleus of the syllable (Blevins 1995).

Many of the other clusters may be the results of historic derivational processes. The nasal /n/ as a second consonant is highly frequent. For most cases, there is evidence that the word is a derivative using the infix *-n-* (cf. §3.3.6.1). Another striking observation is the high frequency of /k/-initial clusters. This may have been caused by a fossilised prefix *k-*. This could be a reflex of PAN **ka-* ‘formative for abstract nouns (often deverbal)’ or PAN **ka-* ‘formative for abstract nouns of quality (Blust and Trussel 2010).¹ Due to the high frequency of /p/ and /t/ as initial consonants in clusters with another non-liquid consonant, I suggest that these plosives might also be historic derivational prefixes. Possible sources are PMP **pa-* ‘causative prefix’ and PMP **ta-* ‘prefix marking spontaneous or involuntary action’ (Blust and Trussel 2010). A more systematic analysis of the words with /p/- and /t/-initial consonant cluster is needed to confirm this putative connection.

Another possible origin of initial consonant clusters is the loss of an unstressed antepenultimate vowel, such as in /tməlar/ ‘flea’ (< PFL **təməla* < PMP **qatiməla* ‘flea’). This could explain why consonant clusters only appear word-initially (cf. (3a) in §3.2.3).

¹ Reflexes of PAN **ka-* are attested in Austronesian languages of the region. Examples of reflexes are the nominalising affix *k(a)- ... (-k)* in Tetun (van Klinken 1999:81), the nominalising prefix *ka-* in Kambera (Klamer 1998:259-260) and the nominalising circumfix *?...-?* in Amarasi (Edwards 2016a:121). Edwards (2016b:68) also reconstructs **ka-...-t* for Pre-Uab Meto.

3.2.3 Stems

Stems are canonically disyllabic with a template of (C)(C)V.(C)V(C). From this template, I derive the distribution constraints for syllable types in Central Lembata in (3).

- (3) Distribution constraints of syllables types in a stem
- a) Consonant clusters are only allowed stem-initially.
 - b) Only root-final syllables can have a coda.

Table 3.7 provides a list of possible stem types including examples. Examples of identical vowel sequences are given in addition to non-identical vowel sequences whenever attested. The stem types V.V and CCV.V are not attested with identical vowel sequences.

Table 3.7: Stem types in Central Lembata

Stem type	Example	Gloss
V.V	/o.i/	‘correct’
V.VC	/u.ak/	‘stream’
	/a.an/	‘what’
V.CV	/o.pi/	‘buy’
V.CVC	/e.vel/	‘tongue’
CV.V	/pa.i/	‘tooth’
	/li.i/	‘leg; foot’
CV.CV	/pi.ra/	‘how many’
CV.VC	/ba.en/	‘morning’
	/pu.ur/	‘blow’
CV. CVC	/pa.dʒuk/	‘stick’
CCV.V	/sme.i/	‘blood’
CCV.CV	/pli.ra/	‘k. o. basket’
CCV.VC	/klu.ok/	‘cooked rice’
	/kdu.uk/	‘spindle’
CCV.CVC	/bni.sak/	‘broken’

There are no phonemic diphthongs in Central Lembata. The nucleus of a syllable can only contain one vowel. However, sequences of two identical vow-

els can occur, such as in /pu.ur/ [pu:r] ‘blow’. Phonetically, these sequences are realised as long vowels, no glottal stop can be inserted in between them.

The general rule of disyllabic stems can have two exceptions. First, function words, such as shortened forms as *la* ‘LOC’ from *lané* ‘LOC’ and discourse marker, such as *bo* ‘DISC’ or *ro* ‘FIN’, can have an exceptional monosyllabic CV structure. But no word consisting of only a single vowel is attested.

Second, words with a final sequence of two identical vowels, such as /tu.'lu.u/ ‘repair’, realised as [tu.'lu:], are underlyingly trisyllabic but only have two syllables on the surface. The underlying trisyllabic structure leads to ultimate stress in the phonetic realisation of these words (cf. §3.2.5). However, there are also a few words in my corpus (listed in Table 3.8) that, synchronically, have to be regarded as trisyllabic stems but cannot be explained by a structure with two identical vowels.

Table 3.8: Trisyllabic stems in Central Lembata

Trisyllabic stem	Gloss	
Loanwords		Origin
/ker.ba.or/	‘buffalo’	< Malay <i>kerbau</i> ‘buffalo’
/ke.de.ra/	‘chair’	< Portuguese <i>cadeira</i> ‘chair’
/i.dʒo.ni/	‘green’	< Malay <i>hijau</i> ‘green’
/re.ka.do/	‘invite guest’	< Portuguese <i>recado</i> ‘message’
/sər.vu.e/	‘struggle’ (v.)	? < Portuguese <i>serviço</i> ‘service, work, labour’
/ta.ba.kor/	‘tabacco’	< Portuguese <i>tabaco</i> ‘tabacco’
Plant or animal names		
/kal.ve.or/	‘balam tree’	
/mak.dʒa.va/	‘bean sp.’	
/səl.va.lak/	‘earthworm’	
/tak.be.ser/	‘pumpkin sp.’	
/le.ŋa.rar/	‘sea weed’	
/lak.bo.rit/	‘spider’	
/səl.ga.er/	‘tomato’	
/kal.ka.sar/	‘tree sp.’	

Trisyllabic stem	Gloss
Others	
/nər.mo.i/	‘correct’
/lak.va.e/	‘couple’
/sbu.a.i/	‘eat’ (archaic)
/pə.vu.dʒa/	‘foam’
/tok.ba.dʒu/	‘rainbow’
/səl.da.e/	‘randomly’
/na.ke.te/	‘recently’
/kra.po.sa/	‘tinea’ (disease)
/nə.gə.ro/	‘yesterday’

Many of the trisyllabic stems in Table 3.8 are either loanwords (as indicated) or can be analysed as historic compounds. There are two pieces of evidence for this. Firstly, in many of these words, the first syllable does not conform to the phonotactical rule which disallows closed syllables in non-final position of the word (cf. rules in (3)). Thus, the first syllable is probably historically an independent prosodic word. Secondly, many of the trisyllabic words denote plant or animal names which often are compounds due to the representation of taxonomies, where the name specifies to which subcategory a certain animal or plant belongs. For the two tree names /kal.ka.sar/ and /kal.ve.or/, even the first syllable matches which could point to a word for tree or a certain tree category.

For those trisyllabic stems that currently can neither be explained by their foreign origin nor their status as animal or plant name, I suspect that they either are historic compounds with synchronically non-analysable components or that they have incorporated morphology that is not discoverable anymore. Although the history of most historic compounds or morphologically complex words remains opaque, a few can be explained as follows. The word /lak.va.e/ ‘couple’ has two components, the first part going back to /la.ki/ ‘male’ (< PMP *laki ‘man’) and the second one to /va.e/ ‘female’ (< PMP *bahi ‘woman’). In Central Lembata /la.ki/ ‘male’ is only used for animals but in other varieties of Lamaholot it can also be used for humans. The word /va.e/ ‘female’ is still found in Central Lembata but it is considered archaic. The word /kra.po.sa/ ‘tinea’ could be a stem with a plural suffix *-ja* realised as /sa/ (cf. §3.3.1.2). The word /pə.vu.dʒa/ ‘foam’ appears to have an

unknown prefix /p/ as it etymologically comes from Proto-Flores-Lembata (PFL) *vuda ‘foam’ (< PMP *bujəq ‘foam’).

3.2.4 Complex words

In Central Lembata, morphologically complex words can have more than two syllables. For words with more than two syllables, the last syllable is usually a suffix, as for instance in (4).

- (4) /ko.po.ŋu/
kopong-u
 child-SPEC
 ‘the child’

For the suffixes *-u* ‘SPEC’ and *-ja* ‘PL’ that attach to alienable nouns, morpho-phonological rules apply to avoid a vowel sequence or to resolve consonant clusters. When the suffix *-u* ‘SPEC’ is attached to a vowel-final noun stem, the liquid [r] is inserted between the stem-final vowel and the suffix, such as in *snaé-[r]-u* ‘shawl-SPEC’. When the suffix *-ja* ‘PL’ is attached to a consonant-final noun stem, different rules apply to solve the consonant cluster, such as the deletion of *j* in *kopong + -ja* → *kopong-a* ‘child-PL’ (cf. §3.3.2).

In contrast to these two suffixes that attach to alienable nouns, the use of pronominal suffixes on verbs can result in a word-medial cluster at the morpheme boundary, such as in *panaw-ga* ‘walk-1SG’. But vowel sequences at the morpheme boundary are not possible with pronominal suffixes. In those cases a nasal is inserted, such as in *grati-[ŋg]-a* ‘call-3SG’ (cf. §3.4.2.4). Also the use of possessor suffixes that attach to inalienable nouns can lead to word-medial clusters, such as in *najan-ga* ‘name-1SG.POSS’ (cf. §3.4.3).

There are also some words that have four and five syllables but cannot be analysed further into morphemes, listed in (5).

- (5) Words with more than three syllables
 /te.de.ra.kum/ ‘chitter’
 /se.ge.re.ge.te/ ‘walk drunk’
 /sə.pə.rə.pu.tə/ ‘wince because of an ant bite’

These words refer to repetitive movements of the body. This suggests that the high amount of syllables is iconic for these movements.

3.2.5 Stress

Stress in Central Lembata is marked by increased vowel length and higher pitch. In unstressed syllables, vowels can be reduced to schwa. In Central Lembata, the penultimate syllable of the word is stressed, unless the penultimate contains the vowel schwa. This becomes clear in suffixed nouns, such as *kopong* /'ko.pɔŋ/ 'child' → *kopong-u* /ko.'po.ŋu/ 'child-SPEC', which are both stressed on the penultimate syllable.

In most words, the underlying penultimate stress corresponds to the penultimate vowel on the surface level, as /ko.'po.ŋu/ 'child' is realised as [ko.'po.ŋu] on the surface. In both representations, the stress falls on the penultimate vowel. However, in trisyllabic stems with final sequences of two identical vowels, the surface realisation shows ultimate stress, as in /tu.'lu.u/ 'repair' which is realised as [tu.'lu:]. A list of minimal pairs that are, on the surface, distinguished by stress and vowel length are given in (6). Underlyingly, the pairs are distinguished by the number of syllables in the word.

(6) Surface minimal pairs by vowel length

	Phonemic	Phonetic	Gloss
a.	/tu.lu/	[tu.lu]	'frog (edible)'
	/tu.'lu.u/	[tu.'lu:]	'repair'
b.	/bo.ti/	[bo.ti]	'carry (a person)'
	/bo.'ti.i/	[bo.'ti:]	'make a small hill of sand or rice grains'
c.	/do.ri/	[do.ri]	'follow (tr.)'
	/do.'ri.i/	[do.'ri:]	'follow (intr.)'
d.	/ta.li/	[ta.li]	'again'
	/ta.'li.i/	[ta.'li:]	'add'
e.	/nu.bu/	[nu.bu]	'grow'
	/nu.'bu.u/	[nu.'bu:]	'shoot (n.)'
f.	/ba.tu/	[ba.tu]	'throw'
	/ba.'tu.u/	[ba.'tu:]	'pull something out of the soil'

Schwa is the only penultimate vowel that cannot be stressed. Words with penultimate schwa are stressed at the ultimate syllable, such as /bə.'ka/ 'to fly'. This even applies when all vowels in the word are schwa, such as in /('kva.ru) ə.'bət/ 'bunch (of maize plants)' or /rə.'kət/ 'sharp' where the stress falls on the ultimate schwa.

However, for consonant-final words apparent free variation between penultimate and ultimate stress is found. One example is /ke.dak/ 'big' which is sometimes realised as [ˈke.dak] and sometimes as [ke.ˈdak]. This variation can be explained by two competing rules. The general rule of penultimate stress leads to the realisation of [ˈke.dak]. But there appears to be a tendency to stress heavy final syllables with CVC structure which leads to the realisation of [ke.ˈdak]. To repair the non-penultimate stress in the second realisation, an epenthetic schwa can be added to a consonant-final word, such as in /ke.dak/ [ke.ˈda.kə] 'big'. This results again in regular penultimate stress.

3.2.6 Summary

Central Lembata has 14 consonant /p, b, t, d, k, g, m, n, ŋ, r, v, s, dʒ, l/ and 6 vowel /i, u, e, ə, o, a/ phonemes. Voiced plosives cannot appear in coda position, the nasal /ŋ/ cannot appear initially and phonemic schwa is not allowed root-finally. A wide range of consonant clusters is found in word-initial position. Many of them are a result of fossilised morphology. The template of a stem is (C)(C)V.(C)V(C). Stems are generally disyllabic, while the morpho-syntactic word can also be trisyllabic due to affixation. Central Lembata words have penultimate stress, unless the penultimate is a schwa.

3.3 Nouns and noun phrases

3.3.1 Noun types

Central Lembata distinguishes between two main types of nouns: inalienable nouns and alienable nouns. This semantic division is reflected in the grammatical properties of the nouns, summarised in Table 3.10.

Table 3.10: Noun types in Central Lembata

	Inalienable	Alienable
Semantics	Body parts Kinship terms Parts of wholes Property concepts	All others
Suffixes	Possessor	Plurality Specificity None
Subtypes	None	Simple (40%) Coda alternating (60%)

In semantic terms, inalienable nouns refer to entities that are in a possessive relationship with another entity that cannot be resolved (Payne 1997:105). These nouns have referents that do not exist without the entity by which they are possessed or of which they are part. The two entities cannot be separated under normal circumstances. Most alienable nouns can also be possessed but, if they are, the connection between possessor and posses- sum can be dissolved. Inalienable nouns are typically body part nouns, kin- ship terms and parts of wholes. From a cross-linguistic perspective, these are expected categories for inalienable possession (Payne 1997:105; Shopen 2007:185). In Central Lembata, in addition to these three categories, prop- erty concepts are also expressed by inalienable nouns.² There are also some Central Lembata nouns that might be considered inalienable from a seman- tic point of view but which are grammatically alienable. Examples include the body part nouns *latar* ‘head hair’ and *tusor* / *tusu* ‘breast’.

The semantic distinction between inalienable and alienable nouns is grammaticalised in the choice of suffixes available for the noun. Nouns in Central Lembata have only one slot for suffixes. Inalienable and alienable nouns each have their own set of suffixes that can be used in this slot. Inali- enable nouns take possessor suffixes (cf. §3.4.3), whereas alienable nouns

² §3.6.6 and §3.3.4 provide further evidence for the word class of property nouns. In §8.3.5, I argue that the expression of property concepts by possessed nouns is an innovation in Lamaholot.

can take a specific suffix, a plural suffix or no suffix (cf. §3.3.1.2). Furthermore, inalienable nouns are formally coherent, whereas alienable nouns have two formal subtypes. Alienable nouns are subdivided into simple nouns and coda alternating nouns. Around 40% of alienable nouns are simple nouns which have only one form per lexeme, whereas around 60% of alienable nouns are coda alternating nouns, and have a long and a short form of each lexeme.

The association of a noun with the alienable or inalienable type is not entirely lexicalised. There are a few cases where a noun can behave as inalienable or alienable depending on the context in which it occurs. An example is *knawer* ‘door’. This word can be either seen as an independent entity, in which case it will behave as alienable with all associated features, or as part of a whole, namely the door of a house or a hut, in which case it will behave inalienably.

3.3.1.1 Inalienable nouns

Examples of inalienable nouns from each of the four semantic fields are given in Table 3.11. For each field, one consonant-final and one vowel-final form is provided.

Table 3.11: Inalienable nouns

	Lexeme	Gloss
Body parts	<i>lotor</i>	‘knee’
	<i>mata</i>	‘eye’
Kinship terms	<i>anak</i>	‘child’
	<i>ina</i>	‘mother’
Parts of wholes	<i>wikil</i>	‘branch’
	<i>tawa</i>	‘stem’
Property concepts	<i>kédak</i>	‘big.one’
	<i>weru</i>	‘new.one’

Inalienable nouns take a possessor suffix to indicate the possessor (§3.4.3). For body parts and kinship terms, pronominal suffixes of all persons can be employed, such as in (7) and in (8). In these examples, the use of a third

person plural possessor suffix *-ja* is exemplified with the body part term *niju* ‘nose’, while the kinship term *awa* ‘spouse’ is marked for a first person singular possessor with the possessor suffix *-ga*.

- (7) *niju-ja*
 nose-3PL.POSS
 ‘their noses’ (CI:76)

- (8) *awa-ga*
 spouse-1SG.POSS
 ‘my wife’ (CI:189)

Unlike body parts and kinship terms, parts of wholes and property concepts can only be possessed by third person possessors and obligatorily appear in a phrasal construction with a dependent noun preceding them, such as *una* ‘house’ in (9). These construction are described further in §3.3.4.2 on noun phrases with inalienable head nouns.

- (9) *una weru-ja*
 house new.one-3PL.POSS
 ‘new houses’ (NB:34)

Property nouns are either inherently nominal, such as *letes* ‘cold.one’ or derived from an adjectival root, such as *kédak* ‘big.one’ < *kéda* ‘big’ (cf. §3.3.6). The first type of property nouns are analysed as nouns because they appear in the same syntactic slot as the nominalised adjectival roots, thus as modifiers of nouns or as predicates (cf. §3.3.4.2 and §3.6.6).

Table 3.12 contains property nouns that are inherently nominal and not derived from an adjectival root. Table 3.13 contains property nouns that are derived from an adjectival root with the same meaning as the derived property noun. Most of the derived property nouns in Table 3.13 are derived using the nominalising infix *-n-* or the suffix *-k*. The infix *-n-* is a frequent nominalising suffix which also appears in other contexts, e. g. deriving a tool or result from an action verb (cf. §3.3.6). The suffix *-k* only appears in connection with property concepts in Central Lembata.

The property nouns in the tables are sorted by their semantic properties following categories proposed by Dixon (1982:16).

Table 3.12: Underived property nouns

Property noun	Gloss	Property noun	Gloss
PHYSICAL PROPERTY		HUMAN PROPENSITIES	
<i>gom</i>	‘deep’	<i>kruweng</i>	‘ashamed’
<i>kmoson</i>	‘dull’	<i>kruit</i>	‘ashamed’
<i>klemur</i>	‘fat’	<i>délong</i>	‘bold’
<i>mebur</i>	‘fresh’	<i>klébit</i>	‘deaf; stupid’
<i>brakong</i>	‘lush’	<i>geli</i>	‘jealous’
<i>groju</i>	‘many leaves’	<i>kmalas</i>	‘lazy’
<i>gai</i>	‘painful; strong, hard’	<i>bnogel</i>	‘lazy’
<i>reket</i>	‘sharp’	<i>kwasan</i>	‘rich’
<i>krogong</i>	‘skinny’	<i>knipék</i>	‘blind’
<i>klésar</i>	‘smooth; even, flat’	<i>kmiran</i>	‘stingy’
<i>gmelos</i>	‘smooth’	<i>blenger</i>	‘stupid’
<i>molur</i>	‘straight’	<i>klebéng</i>	‘stupid’
<i>kbékék</i>	‘strong’		
<i>plekok</i>	‘weak’		
AGE		DIMENSION	
<i>kbélek</i>	‘elder’	<i>klépar</i>	‘narrow’
<i>weru</i>	‘new’	<i>duang</i>	‘big and old’
<i>okim</i>	‘old (things)’		
<i>nuja</i>	‘young, fresh (fruits)’		
COLOUR		OTHERS	
<i>mitem</i>	‘black; dark’	<i>kpolot</i>	‘round’
<i>kberis</i>	‘blue, green bird sp.’	<i>klopor</i>	‘round’
<i>remak</i>	‘dark’	<i>blurek</i>	‘round and tiny’
<i>ijoni</i>	‘green’	<i>gésak</i>	‘different’
<i>méran</i>	‘red’	<i>blakun</i>	‘naked’
<i>kumas</i>	‘yellow’		
SPEED			
<i>gnesik</i>	‘fast’		

Table 3.13: Derived property nouns

Adjectival root	Derived noun	Gloss
PHYSICAL PROPERTY		
<i>pait</i>	<i>pnait</i>	'bitter'
<i>berat</i>	<i>bnerat</i>	'heavy'
<i>kati</i>	<i>knating</i>	'hot'
<i>gewa</i>	<i>knewak</i>	'rotten'
<i>peju</i>	<i>pnejuk</i>	'salty'
<i>gilu</i>	<i>kniluk</i>	'sour'
<i>seru</i>	<i>sneruk</i>	'sweet'
<i>seba</i>	<i>snebar</i>	'wet'
<i>slae</i>	<i>slaek</i>	'clean'
<i>manga</i>	<i>mangak</i>	'hard; thick'
<i>klea</i>	<i>kleak</i>	'light (weight)'
<i>mila</i>	<i>milan</i>	'dirty'
<i>meli</i>	<i>meling</i>	'quiet; no people'
<i>mipi</i>	<i>mipiw</i>	'thin'
DIMENSION		
<i>keda</i>	<i>kedak</i>	'big'
<i>dae</i>	<i>daek</i>	'near'
<i>koda</i>	<i>kodak</i>	'long' (time)
<i>doa</i>	<i>doak</i>	'long; far'
<i>meگو</i>	<i>megon/megok</i>	'short' (thing)
<i>lere</i>	<i>lerak</i>	'short (vertical); flat'
<i>ana</i>	<i>anak</i>	'small'
<i>golo</i>	<i>golok</i>	'tall; high'
COLOR		
<i>buja</i>	<i>bujak</i>	'white'
HUMAN PROPENSITIES		
<i>susa</i>	<i>snusan</i>	'poor; sad'

Adjectival root	Derived noun	Gloss
OTHERS		
<i>sukar</i>	<i>snukar</i>	‘difficult’
<i>rai</i>	<i>raik</i>	‘many’

From the tables, it becomes visible that the underived property nouns show a high diversity in semantic domains expressed and at the same time the category of DIMENSION is highly underrepresented, whereas the property concepts that have an adjectival root have a clear tendency to express properties of PHYSICAL PROPERTY and DIMENSION.

Property concepts expressed by nouns are an innovation in Central Lembata and the whole Flores-Lembata family. This innovation and its history is discussed in §8.3.

3.3.1.2 Alienable nouns

Alienable nouns, subdivided into simple and coda alternating nouns, cover various semantic fields. Simple nouns have one form for each lexeme, whereas coda alternating nouns have a long form and a short form for each lexical item. The division into simple and coda alternating nouns is lexically determined. However, historically the phonological structure of the stem plays an important role to determine whether a noun is simple or coda alternating. In this section, I focus on the synchronic forms and distributions of alienable nouns. The historic development of alienable nouns and their subdivision into simple nouns and coda alternating nouns is discussed in §3.3.3.

Table 3.14 provides examples of simple alienable nouns, which cover around 40% of the alienable nouns in the Central Lembata lexicon. Simple alienable nouns are mainly consonant-final but there are also a few vowel-final nouns that fall into this category, such as *snaé* ‘shawl’.

Table 3.15 gives examples of coda alternating alienable nouns, which encompass around 60% of all alienable nouns in the lexicon. Coda alternating nouns have two synchronic realisations of the same lexeme. One form is consonant-final, such as for example *aor* ‘dog’ or *buser* ‘cotton bow’, and another form is shorter and vowel-final, such as *au* ‘dog’, or consonant-final

but with only one vowel which is phonetically long, such as *buus* [bu:s] ‘cotton bow’.³ The short forms with long vowels are marginal. In glosses of this dissertation, the long form is marked with \L (long) and the short form is marked with \s (short) but only if the distinction is relevant for the context of discussion.

Table 3.14: Simple alienable nouns

Lexeme	Gloss
<i>manuk</i>	‘chicken’
<i>kébol</i>	‘sugar palm’
<i>taum</i>	‘indigo plant’
<i>gerep</i>	‘young woman’
<i>apur</i>	‘lime’
<i>emut</i>	‘dust’
<i>snaé</i>	‘shawl’

Table 3.15: Coda alternating alienable nouns

Long form	Short form	Gloss
<i>aor</i>	<i>au</i>	‘dog’
<i>witér</i>	<i>witi</i>	‘goat’
<i>lisor</i>	<i>liso</i>	‘rice plant’
<i>buser</i>	<i>buus</i>	‘cotton bow’
<i>piring</i>	<i>piri</i>	‘plate’
<i>kopong</i>	<i>kopo</i>	‘young person; child’
<i>ékén</i>	<i>éke</i>	‘bamboo stairs’
<i>binén</i>	<i>biné</i>	‘woman, female’
<i>unan</i>	<i>una</i>	‘house’
<i>karaj</i>	<i>kara</i>	‘grain’

³ In the case of *buser* / *buus* ‘cotton bow’, the long form is the historically underlying form as this word goes back to PMP *busuR. The short form is thus generated by deletion of the final consonant /r/ which leads to a schwa-final word. As final schwa is dispreferred in the syllable structure, this is possibly the reason why the schwa is dropped and the /u/ is lengthened instead.

The two shapes of a coda alternating noun are used in different syntactic contexts. Different distribution rules (10) apply for the two shapes of coda alternating nouns in a noun phrase (NP) and nouns as part of a verb phrase (VP).

- (10) Distribution rules for the two forms of coda alternating nouns
- a) Noun Phrase: In non-final position, coda alternating nouns occur as short forms, see examples in (11). In final position, coda alternating nouns occur as long forms, see examples in (12).
 - b) Verb phrase: object nouns (which are always in final position of the VP) occur as short forms, see examples in (13). Long forms cannot occur within the VP.

In the following, I illustrate the two rules with examples. Within the noun phrase, the choice of long or short form is only based on the ‘core’ part of the NP, including nouns and numerals (cf. Table 3.22 in §3.3.4). Non-final elements, such as *kopo* ‘child\’s’ in (11a) and *biné* ‘woman\’s’ in (11b) appear in the short form as they are followed by another element in the NP, i. e. the inalienable noun *anak* ‘little.one’ in (11a) and the numeral *tuné* ‘one’ in (11b).

- (11) a. *Kopo anak wo ro kreka-nga.*
 child\’s little.one DIST FIN startle-3SG
 ‘That little.one child got frightened.’ (F1:26)
- b. *Biné tuné na=soga kaén.*
 female\’s one 3SG=hold cloth
 ‘A woman holds a piece of cloth.’ (FH1:28)

NP-final nouns always appear in the long form, such as *kopong* ‘child\’L’ in (12a) and in (12b) and also *lamén* ‘male\’L’ in (12c). Demonstratives, such as *wo* ‘DIST’, and end of phrase particles, such as *ro* ‘FIN’, are outside of the ‘core’ NP. Therefore, the noun *kopong* ‘child\’L’ in (12b) is the only element in the NP.

- (12) a. *Kopong gelé-na.*
 child\’L lie.down-3SG
 ‘The child sleeps.’ (F2:4)

- b. *Kopong wo ro kreka-nga.*
 child\L DIST FIN startle-3SG
 ‘That child is frightened.’ (F2:25)
- c. *Kopo lamén soro-nga woné kursi paap.*
 child\s male\L hide-3SG DIST chair side
 ‘A boy hides next to the chair.’ (FH1:37)

In verb phrases, such as in (13), the object noun occurs in the short form, regardless of the presence of absence other elements in the NP, such as *tu* ‘one’ in (13c) or *wo* ‘DIST’ in (13d).

- (13) a. *Kam=paraw kopo.*
 1PL.EXCL=feed child\s
 ‘We bring up (our) children.’ (T3:290)
- b. *Go durum ika.*
 1SG see fish\s
 ‘I sell fish.’ (NT:3)
- c. *Na=soga kaju tu.*
 3SG=hold wood\s one
 ‘She holds a wooden stick.’ (FH1:7)
- d. *Papi-nga witi wo, we ta=kaan-a kia wé.*
 burn-3SG goat\s DIST so 1PL.INCL=eat-3SG INCEP PROX
 ‘Let’s grill that goat, so that we can eat it now.’ (L3:191)

As soon as an object noun is fronted, such as *lisor* ‘rice\L’ in (14), it is not part of the VP any more and thus follows the rule for coda alternating nouns in the NP. The fronted object now appears in the long form because it is in NP-final position.

- (14) *Bo lisor bo ro bo da=durum.*
 DISC rice\L DISC FIN DISC 3PL=sell
 ‘And the rice, they sell [it].’ (M3:8)

The rules for the distribution of short and long forms of coda alternating nouns only apply to nouns that are not marked for specificity or plural number by means of suffixes (cf. §3.3.2). Coda alternating nouns with a suffix only have one form and it is the long form that takes the suffix (cf. §3.3.2).

3.3.2 Suffixes on alienable nouns

Simple nouns, such as *manuk* ‘chicken’, and the long forms of coda alternating nouns, such as *aor* ‘dog’, can be optionally marked with the suffixes *-ja* ‘PL’ for plurality or *-u* ‘SPEC’ for specificity. The specificity suffix *-u* is of unknown origin. It could be related to *tu* ‘one’ with an irregular loss of initial *t*. The plural suffix *-ja* has developed from the third person plural pronoun *da* (cf. §8.4.2.1). Nouns marked for specificity or plurality mainly occur as NP-final nouns and only rarely as object nouns. In the VP, plurality or specificity of the object noun is in most cases not overtly marked, compare example (13a).

The suffix *-u* ‘SPEC’ is the main means to mark an alienable noun as singular specific. A noun that is marked as specific indicates that the referent of the noun is a particular referent which is known to the speaker. Table 3.16 lists examples of nouns marked as specific using the suffix *-u* ‘SPEC’. As most simple nouns and all coda alternating free forms are consonant-final, the suffix *-u* ‘SPEC’ simply attaches to the stem, such as for the words *kopong-u* ‘child\L-SPEC’, *aor-u* ‘dog\L-SPEC’, *manuk-u* ‘chicken\L-SPEC’. When a simple noun is vowel-final, such as *snaé* ‘shawl’, the consonant [r] is inserted in between the final vowel of the stem and the suffix *-u* ‘SPEC’. The etymological origin of the inserted [r] remains unclear (cf. §3.3.3.1).

Table 3.16: Specificity suffix *-u* on alienable nouns

Specific noun	Morphemes	Meaning
<i>kopongu</i>	<i>kopong</i> -u	‘the child’
<i>aoru</i>	<i>aor</i> -u	‘the dog’
<i>manuku</i>	<i>manuk</i> -u	‘the chicken’
<i>snaéru</i>	<i>snaé</i> [r] -u	‘the shawl’

The plural suffix *-ja* ‘PL’ can mark the noun for plural number, as exemplified in Table 3.17. The plural suffix is used when there is no other indication of plurality in the noun phrase, e. g. by numerals or quantifiers. It can be used for all kinds of alienable nouns, including animates and inanimates, but not for concepts that inherently cannot have a plural meaning, such as *kelem* ‘sky’ or *luwak* ‘sun’.

Table 3.17: Plural suffix *-ja* on alienable nouns

Plural noun	Morphemes	Meaning
<i>koponga</i>	<i>kopong</i> <i>-ja</i>	‘children’
<i>aoja</i>	<i>aor</i> <i>-ja</i>	‘dogs’
<i>manusa</i>	<i>manuk</i> <i>-ja</i>	‘chickens’
<i>snaéja</i>	<i>snaé</i> <i>-ja</i>	‘shawls’

As most simple nouns and all coda alternating long forms are consonant-final, the use of the plural suffix *-ja* ‘PL’ results in an underlying consonant cluster which is resolved according to the morpho-phonological rules listed in (15) to (17), each with a number of examples.

According to the rule in (15), the final consonants *j, m, ng, l, w, s* of a noun taking a plural suffix are retained but they cause the initial *j* of the suffix to be dropped.

- (15) Deletion of *j* in *-ja* ‘PL’ following *j, m, ng, l, w, s*

$j \rightarrow \text{zero} / j, m, ng, l, w, s \_a$

<i>uaj -ja</i>	$\rightarrow$ <i>uaja</i>	‘rattan-PL’
<i>wetem -ja</i>	$\rightarrow$ <i>wetema</i>	‘millet-PL’
<i>kopong -ja</i>	$\rightarrow$ <i>koponga</i>	‘child-PL’
<i>kadal -ja</i>	$\rightarrow$ <i>kadala</i>	‘fence-PL’
<i>ktumaw -ja</i>	$\rightarrow$ <i>ktumawa</i>	‘louse.cloth-PL’
<i>temus -ja</i>	$\rightarrow$ <i>temusa</i>	‘whale-PL’

According to the rule in (16) the *j* of the plural suffix is devoiced to *s* after voiceless stops *p, t, k* and subsequently these voiceless stops are deleted.

- (16) Devoicing of *j* in *-ja* ‘PL’ after *p, t, k* and deletion of *p, t, k*

1. $j \rightarrow s / \_ \{p, t, k\}$

2. $p, t, k \rightarrow \text{zero}$

<i>bepap -ja</i>	$\rightarrow$ <i>bepasa</i>	‘lizard-PL’
<i>ramut -ja</i>	$\rightarrow$ <i>ramusa</i>	‘root-PL’
<i>manuk -ja</i>	$\rightarrow$ <i>manusa</i>	‘chicken-PL’

For nouns ending in *r* or *n*, the rule in (17) applies. This means that the consonant *j* of the suffix *-ja* is realised when the final consonant of the noun is either *n* or *r*. These coda consonants are dropped and *j* remains.

- (17) Deletion of nominal coda *n, r* before *-ja* ‘PL’

n, r → zero / *_ja*

<i>liar -ja</i>	→ <i>liaja</i>	‘ginger-PL’
<i>aor -ja</i>	→ <i>aoja</i>	‘dog-PL’
<i>aper -ja</i>	→ <i>apeja</i>	‘fire-PL’
<i>utan -ja</i>	→ <i>utaja</i>	‘bean-PL’
<i>angin -ja</i>	→ <i>angija</i>	‘wind-PL’

The fact that *r* is seemingly dropped here is probably related to the fact that final *r* is historically not part of the stem but inserted (cf. §3.3.3.1). In some cases, the final *n* could be a historic genitive suffix, and thus historically not be part of the stem. However, in both examples given above /n/ is historically part of the stem: *utan* < PMP *qutan ‘small, wild herbaceous plants; scrub-land, bush’ and *angin* < PMP *haŋin ‘wind’.

The liquids *l* and *r* do not always behave according to the general rules listed above. With 10 out of 12 occurrences in corpus, the final consonant *l* normally follows the rule in (15), thus causing deletion of *j* in the suffix. However, two *l*-final nouns in my corpus follow the rule in (17) which means that *l* is deleted and the initial *j* is kept.⁴ With 116 out of 128 occurrences in my corpus, the final consonant *r* usually follows the rule in (17). However, in 12 cases in my corpus the final consonant *r* follows the rule in (15) and is retained.⁵

The fact that the suffixes *-ja* ‘PL’ and *-u* ‘SPEC’ are mutually exclusive and the observation that nouns with the specificity suffix *-u* are by default singular leads to the question whether the two suffixes are a plural variant and

⁴ Examples of nouns that delete final *l* when pluralised and thus violate the rule in (15): *nukél* → *nukéja* ‘a kind of weaving tool’, *bnuél* → *bnuéja* ‘thread winder’.

⁵ Examples of nouns that retain final *r* when pluralised and thus violate the rule in (17): *naker* → *nakera* ‘roof’, *latar* → *latara* ‘hair’, *klobar* → *klobara* ‘part of corn cob’, *apur* → *apura* ‘lime’, *labur* → *labura* ‘shirt’, *kelir* → *kelira* ‘part of banana plant’, *lusir* → *lusira* ‘needle’, *kebur* → *kebura* ‘part of coconut’, *kalar* → *kalara* ‘traditional bracelet’, *kunur* → *kunura* ‘breadfruit’, *user* → *usera* ‘flesh’, *snokar* → *snokara* ‘digging stick’.

a singular variant of the same category. This could be e. g. number or specificity. Analyzing this category as number would suggest that a morphologically unmarked noun can be either read as singular or plural. However, it appears to be the case that unmarked nouns, at least as heads of a noun phrase, are always singular. Analyzing this category as specificity, with a singular and a plural morpheme, would require showing that marked head nouns are specific and that unmarked head nouns are non-specific, unless there is another element such as a demonstrative or a possessor present that yields a specific reading. However, this does not seem to be the case, as there are non-specific nouns marked with *-ja* ‘PL’, such as *laméja* in (18).

- (18) *Bo kalau lamé-ja bo grati maa ama ...*
 DISC if man-PL DISC call say father
 ‘Concerning men, we call them *ama*.’ (C1:130)

The plural noun *laméja* ‘man.PL’ has a clear non-specific meaning, as the meaning of the sentence concerns men in general and no specific group of men. So far, no evidence has been found to analyse both suffixes as a singular and a plural allomorph of the same category. A more systematic data collection is needed to ultimately answer this question.

Number marking is an innovation in Central Lembata and the whole Central Lamaholot subgroup. In particular, the marking of plurality on nouns appears to be an areal feature, possibly of non-Austronesian origin. This feature and its history is discussed in more detail in §8.4.

3.3.3 Coda alternation: origin and regional context

This section is a historic and comparative discursus on the emergence of coda alternation in Central Lembata. I first propose a diachronic explanation for coda alternation in Central Lembata in §3.3.3.1 which appears to be an innovation. It becomes evident that in Central Lembata coda alternation is connected to the innovation of the specificity suffix.

In the following, I examine how wide-spread the phenomenon of coda alternation is among the varieties of Lamaholot. In §3.3.3.2, I show that coda alternation, as well as number marking through suffixes, as described for Central Lembata in §3.3.2, is found in all varieties of Central Lamaholot. This suggests that these features developed in Proto-Central Lamaholot. In §3.3.3.3, I show that Western Lamaholot and Eastern Lamaholot do not

have these features but there is a slight possibility that these languages had coda alternation and/or number marking in the past but lost the features. In §3.3.3.4, I point out that coda alternation in Central Lembata shows striking similarities with morphological metathesis and cases of consonant insertion attested in Austronesian languages of Timor.

3.3.3.1 Central Lembata

In the following, I propose a hypothesis on the development of coda alternation which is attested in around 60% of the alienable nouns in Central Lembata. The remaining 40% of alienable nouns in this language do not show coda alternation, thus have only one form (cf. §3.3.1.2). There are three steps that led to the development of coda alternation in Central Lembata, as demonstrated in Table 3.18.

Table 3.18: The development of coda alternation in Central Lembata

<i>r</i> -insertion	<i>ili</i> ‘mountain’ + <i>-u</i> ‘SPEC’	→ <i>ile-r-u</i> ‘mountain-SPEC’
Reanalysis	<i>iler-u</i> ‘mountain-SPEC’	→ <i>iler</i> ‘mountain’
		→ 2 stems: <i>ili/iler</i> ‘mountain’
Analogy	<i>ikan</i> ‘fish’	→ <i>ika/ikan</i> ‘fish’

First, the liquid *r* was inserted before the specificity suffix *-u* ‘SPEC’ when the suffix attached to a vowel-final stem, such as *ili* ‘mountain’ (< PMP *qilih ‘mountain’), see also §3.3.2. The specific form of the word *ili* ‘mountain’ is then *iler-u* ‘mountain-SPEC’. Generally, high vowels *i* and *u* are lowered to *e* and *o* in this process. Therefore, the final vowel *i* of *ili* ‘mountain’ is lowered to *e*.

The inserted *r* is of unknown origin.⁶ Central Lembata does not allow a vowel-final stem to take a vowel-initial suffix. Sequences of two vowels are allowed stem-internally but not at a morpheme boundary (cf. 3.2.2). In the pronominal suffix paradigms (cf. §3.4.2.4), this is solved by the insertion of a nasal *ng* in between the verbal stem and the suffix *-a* ‘3SG’ or *-i* ‘3PL’, such

⁶ A possible origin could be a previously lost stem-final consonant (Blevins 2008:93-97). However, as most of the nouns with inserted *r* go back to a historically vowel-final stem, this appears unlikely.

as in *tula* ‘make’ which becomes *tula-ng-a* ‘make-3SG’. It is not clear why in the verbal case a nasal is inserted, while in the case of nominal stem and suffix *-u* ‘SPEC’ the liquid *r* is inserted.

As the second step, the inserted *r* is reanalysed as being part of the stem, which leads to a new additional longer form of the stem, here *iler* ‘mountain’, that exists in addition to the original short form *ili* ‘mountain’.

As the last step, many consonant-final nouns, mainly those ending in nasals *n* or *ng*, drop their final consonant in analogy to gain an additional form, such as *ikan* ‘fish’ (< PMP *hikan ‘fish’) which has the two forms *ikan* and *ika*. However, this process appears to be only semi-productive and is not applied to all C-final nouns in Central Lembata. Only around 60% of the alienable nouns show coda alternation. There is a clear pattern towards certain consonants being dropped more easily than others. The majority of alienable nouns ending in the nasals *n* and *ng* have dropped their final consonants to gain a vowel-final form, with only very few exceptions for *ng* that were not dropped.⁷ A few nouns that have final consonants other than *n* and *ng* also have dropped their final consonant to gain coda alternation. There are sporadic examples of dropped final *k*, *t*, *r*, *l* and *m* in my corpus.⁸ The consonants *j* and *w* are rare in final position. My corpus of Central Lembata contains seven alienable nouns with final *j*, out of which four are coda alternating, and one alienable noun ending in *w* which is also coda alternating.⁹

Synchronically, both forms of the nouns, the original form and the new form, are still used. Within the NP, the short form appears in non-final pos-

⁷ In my corpus, there are 4 out of 51 nouns ending in *ng* that do not show coda alternation, thus did not drop their final consonant to gain a second form. These are: *nileng* ‘k.o. jewelery’, *giwang* ‘earring’, *tuteng* ‘frog sp.’, *klewang* ‘long machete’ (loan from Malay).

⁸ Coda alternating nouns with final *k*, *t*, *l*, *m*: *weka/wekak* ‘bird sp.’, *kapa/kapak* ‘axe’, *gaja/gajak* ‘wound’, *suri/surit* ‘weaving sword’, *bnue/bnuel* ‘thread winder’, *peri/perim* ‘bamboo sp.’. To identify coda alternating nouns with final *r* that come from *r*-final nouns that dropped their final consonant is more difficult because, if no etymology is known, the same words could be vowel-final nouns with *r* insertion. Nevertheless, a few cases could be identified: *wuli/wulir* ‘bunch’ (In *wulir* the high vowel *i* before *r* is not lowered, this shows that the *r* is part of the stem and not inserted.), *ula/ular* ‘snake’ (PMP *hulaR ‘snake’).

⁹ Alienable nouns with final *j*: *blawaj* ‘stretching tool’, *blebaj* ‘k.o. basket’, *punaj* ‘snake sp.’, *ena/enaj* ‘soil’, *kara/karaj* ‘grain’, *kresa/kresaj* ‘tree sp.’, *ua/uaj* ‘rattan’. Alienable noun with final *w*: *ktuma/ktumaw* ‘cloth louse’.

ition and within the VP, the short form appears in object position following the verb (cf. §3.3.1.2). In these positions, specificity marking on the noun with the suffix *-u* ‘SPEC’ is not allowed (non-final NP position), or strongly dispreferred (post-verbal object position). Assuming that the use of the specificity suffix initiated the additional C-final stems of V-final alienable nouns, as laid out above, it is logical that these new forms with inserted *r* appear in positions where specificity marking is allowed or frequently used.

I propose that the alternation started with vowel-final nouns. When the pattern of long forms and short forms in their specific syntactic positions was regularised, C-final nouns started to drop their final consonant in analogy to the system that had developed in V-final nouns and started to follow the same distribution pattern. This suggests that the language was in the process of regularising the system of coda alternation to all alienable nouns. However, at some point, this process ceased. Alternatively, it can be proposed that the process is only semi-productive with a tendency to apply to vowel-final and *n/ng*-final nouns and less to others. There are still a few rather recent Malay loans, such as *piring* ‘plate’, *kopi* ‘coffee’, *kerbau* ‘buffalo’ or *kusing* ‘cat’, which have coda alternation. Therefore, the fossilisation of the process must have occurred when Malay had already been in use as a lingua franca in the region, which means not longer than 400 years ago (cf. §1.2.4.4).

3.3.3.2 Central Lamaholot

In this section, I show that coda alternation, as well as number marking (cf. §3.3.2) is not only found in Central Lembata, the language described in this grammar sketch, but in all varieties of the Central Lamaholot subgroup (cf. §4.1.4). Although the Central Lamaholot varieties are only poorly documented, the 33 wordlists collected by Keraf (1978a:262-297) throughout the Lamaholot area provide evidence that traces of plural marking, specificity marking and coda alternation are found in all varieties of Central Lamaholot, as shown in Table 3.19. This suggests that both phenomena can be reconstructed to Proto-Central Lamaholot. To what extent the systems are active and how they differ in detail among the varieties of Central Lamaholot remains to be explored. Table 3.19 shows a selection of nouns in Central Lamaholot (CL) varieties and their corresponding PMP forms. Central Lembata long and short forms are added for comparison.

Table 3.19: Comparative data for C insertion in CL varieties (Keraf 1978)

Variety	‘dog’	‘snake’	‘head louse’	‘stone’	‘mountain’
PMP	*asu	*hulaR	*kutu	*batu	*qilih
Central Lembata	<i>au/aor</i>	<i>ula/ular</i>	<i>kutu/kutor</i>	<i>watu/wator</i>	<i>ili/iler</i>
Lewotala	<i>aoru</i>	<i>ularu</i>	<i>kutoja</i>	<i>fatoja</i>	<i>ileru</i>
Imulolo	<i>aor</i>	<i>ular</i>	<i>kutor</i>	<i>fator</i>	<i>iler</i>
Lewopenutu	<i>aoru</i>	<i>ularu</i>	<i>kutoj</i>	<i>fatoru</i>	<i>ile</i>
Mingar	<i>aoru</i>	<i>ular</i>	<i>kutoru</i>	<i>wato</i>	<i>ileru</i>
Lewuka	<i>aor</i>	<i>ular</i>	<i>kutu</i>	<i>wator</i>	<i>iler</i>
Kalikasa	<i>au</i>	<i>ula</i>	<i>kutu</i>	<i>watu</i>	<i>ile</i>
Lewokukun	<i>aoru</i>	<i>ularu</i>	<i>kutoru</i>	<i>watoru</i>	<i>ileru</i>
Painara	<i>aho</i>	<i>ulay</i>	<i>kutoy</i>	<i>wator</i>	<i>iler</i>

C = consonant, CL = Central Lamaholot

The nouns in the table have diverse forms. There are vowel- and consonant-final forms, forms with the specificity suffix *-u*, as *iler-u* ‘hill-SPEC’ in Mingar, and plural forms with the suffix *-ja*, as *kuto-ja* ‘louse-PL’ in Lewotala.

Identifying the presence of morphology and coda alternation in the other Central Lamaholot varieties is only possible because the morphology of Central Lembata is known. There are no indications in the Keraf wordlists on these variety of forms. The sole occurrence of only one form per concept in the data provided by Keraf (1978a) suggests that the author was neither aware of the coda alternation nor the suffixes. The words must have been elicited without analysis of the morphology.

3.3.3.3 Western and Eastern Lamaholot

In this section, I address the question of whether the phenomenon of coda alternation was more wide-spread than Central Lamaholot. Based on inspection of the wordlists in Keraf (1978a) it appears that the phenomena of coda alternation as well as number marking (cf. §3.3.2) are exclusive to Central Lamaholot and not found in Western or Eastern Lamaholot varieties, as the Western and Eastern Lamaholot varieties, in Table 3.20 and Table 3.21, have very regular shapes of nouns. The nouns in the table, which are

cognates of the Central Lamaholot nouns shown in the previous section, all are vowel-final and do not show any suffixes or inserted final *r*.

Table 3.20: Comparative data for nouns in WL varieties (Keraf 1978)

Variety	‘dog’	‘snake’	‘head louse’	‘stone’	‘mountain’
PMP	*asu	*hulaR	*kutu	*batu	*qilih
Pukaunu	<i>aho</i>	<i>ula</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Lewolema	<i>aho</i>	<i>ula</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Waibalun	<i>aho</i>	<i>ula’</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Baipito	<i>aho</i>	<i>ula’</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Bama	<i>aho</i>	<i>ula’</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Lewolaga	<i>aho</i>	<i>ula</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Tanjung	<i>aho</i>	<i>ula’</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Ritaebang	<i>aho</i>	<i>ula</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Lewotobi	<i>aho</i>	<i>ula’</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Kiwangona	<i>aho</i>	<i>ula’</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Dulhi	<i>aho</i>	<i>ula’</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Horowura	<i>aho</i>	<i>ula’</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Waiwadan	<i>aho</i>	<i>ula</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Watan	<i>aho</i>	<i>ula’</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Lamakera	<i>aho</i>	<i>ula</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Botun	<i>aho</i>	<i>ula</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Ile Ape	<i>aho</i>	<i>ula</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Lamalera	<i>ao</i>	<i>ula</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Mulan	<i>aho</i>	<i>ula</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Belang	<i>ao</i>	<i>ula</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Wuakerong	<i>aho</i>	<i>ula</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Lamahora	<i>aho</i>	<i>ula</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>
Merdeka	<i>aho</i>	<i>ula</i>	<i>kuto</i>	<i>wato</i>	<i>ilé</i>

WL = Western Lamaholot

Table 3.21: Comparative data for nouns in EL varieties (Keraf 1978)

Variety	‘dog’	‘snake’	‘head louse’	‘stone’	‘mountain’
PMP	*asu	*hulaR	*kutu	*batu	*qilih
Lamatuka	<i>aho</i>	<i>ula</i>	<i>uto</i>	<i>wato</i>	<i>ilé</i>
Lewoeleng	<i>aho</i>	<i>ula</i>	<i>uto</i>	<i>wato</i>	<i>ilé</i>

EL = Eastern Lamaholot

3.3.3.4 Languages of Timor

On Timor, several Austronesian languages are attested that have lexemes with two forms. Similar to the coda alternating nouns in Central Lembata (cf. §3.3.3.1), these two forms show phonological alternations in their coda and they are used in different morpho-syntactic environments. Therefore, they can be classified as cases of coda alternation. In the Timor cases, the second form is historically derived from its base by metathesis or by consonant insertion. Consonant insertion — the insertion of *r* before the suffix *-u* ‘SPEC’ — is also attested in Central Lembata, but metathesis is not (cf. §3.3.3.1).

Cases of metathesis that create additional forms of lexemes for specific morpho-syntactic contexts are summarised in Edwards (2016a:31-67). Based on current knowledge and documentation, the following languages in the Timor region are reported to have metathesised and unmetathesised forms with morphological function: Leti, Roma, Mambae, Helong and languages of the Meto cluster.¹⁰ Although differing in details, in these languages there is a tendency for a final versus non-final distribution of the two forms on the phrase level. For example in Roma (Steven 1991), the unmetathesised form, such as *krahan* ‘house’, occurs when an attributive modifier follows the noun in the phrase, such as in (19a), or in object position, such as *hiwit* ‘machete’ in (19b). Unmodified nouns occur in the metathesised form, such as *hiwti* ‘machete’ in (19c).

¹⁰ Based on initial descriptions by Joseph Lovestrand (personal communication), it appears that the language Kodi spoken on Sumba shows very similar patterns of metathesis. Possibly, this pattern is more wide-spread in the region than currently documented.

(19) Roma

- a. *krahan* *popotna*
 house\UNMETATHESISED large
 ‘large house’ (Steven 1991:67)
- b. *N-la* *n-dahal* *hiwit-a*.
 3SG-go 3SG-search machete\UNMETATHESISED-EPN
 ‘He searched for a machete.’ (Steven 1991:67)
- c. *Hiwti* *ta-walli*.
 machete\METATHESISED NEG-exist
 ‘There wasn’t any machete.’ (Steven 1991:67)

To some extent, this pattern is similar to Central Lembata. In the Roma examples above, the unmetathesised form would correspond to the short form in Central Lembata, while the metathesised form would correspond to the long form (cf. §3.3.1.2).

Cases of final consonant insertion that create an additional form of a lexeme are attested in Amfo’an, a variety of Meto, and in varieties closely related to Amfo’an (Culhane 2019). An interesting parallel between the consonant insertion in Amfo’an and consonant insertion in Central Lembata is the following. In addition to the insertion of regular *g* and *j* in Amfo’an, Culhane (2019:75, 77) finds the insertion of *l* in certain words unexplained. Historically, the inserted *l* comes from **r*. Also Central Lembata inserts the liquid *r* which is of unknown origin. This putative connection remains to be examined further.

In the TAP family, coda alternating forms of a lexemes in different morpho-syntactic distributions are not attested. Reduced nouns in Makalero have some similarities in distribution with the coda alternation cases described in this section. However, only three lexemes with reduced forms are attested in Makalero: *ara-* ← *arapau* ‘buffalo’, *oma-* ← *omar* ‘stilt house’ and *nu-* ← *nunu* ‘lips’ (Huber 2011:120). These reduced forms dropped their final consonant or final syllable. They are found in nominal compounds and as verbal complements.

3.3.3.5 Summary

Coda alternation is attested throughout Central Lamaholot as well as in some Austronesian languages of Timor. The similarities in function, i. e. two forms

for different morpho-syntactic environments, and, for some languages also in strategy, i.e. consonant insertion in Central Lembata and Amfo'an, are striking. Nevertheless, there is no evidence to assume a shared innovation for this phenomenon in Central Lembata and the languages of Timor. Thus, coda alternation may be seen as a “universal” strategy attested in the region to mark an opposition in syntactic position for certain word classes, mainly nouns.

3.3.4 Noun phrases

3.3.4.1 Basic noun phrases

A noun phrase (NP) in Central Lembata has the following slots.

Table 3.22: Noun phrase template

NP slots	Core		Subordinate	Final	
Elements	N (N) (N)	NUM	relative clause	DEM	DISC

This NP template shows all possible noun phrase slots and the elements that can be inserted in these slots. In a noun phrase, all elements apart from one noun (N) are optional.

The core slot contains up to three nouns. If no numeral (NUM) is present, the final noun in the core slot can take a suffix, whereas all preceding nouns cannot be marked morphologically. The suffixes available for inalienable nouns (possessor) and for alienable nouns (plural and specificity) have been discussed in §3.3.1. With a numeral in the final position of the core slot, no morphological marking of the preceding noun is allowed (cf. §3.3.4.4). The core slot of the NP can be followed by a relative clause which is semantically subordinated but structurally simply juxtaposed (cf. §3.3.4.5). Following the subordinate slot, there is a final slot which can contain a demonstrative (DEM) and one or both of the discourse particles (DISC) *ro* ‘FIN’ and *bo* ‘DISC’ or less frequently *di* ‘DISC’ (cf. §3.3.4.6).

The simplest noun phrase contains a noun alone, either inalienable (20) or alienable (21).

- (20) *Ina-ga Lélú.*
 mother-1SG.POSS NAME
 'My mother is Lelu.' (M1:4)

- (21) *Kopong gelé-na.*
 child\1 sleep-3SG
 'The child sleeps.' (F2:4)

Inalienable head nouns are marked with a suffix, such as *-ga* '1SG.POSS' in (20). Alienable head nouns are optionally marked with *-u* 'SPEC', such as in (22) or *-ja* 'PL', such as in (23).

- (22) *Aor-u gelé dori-nga.*
 dog-SPEC sleep follow-3SG
 'The dog also sleeps.' (F1:5)

- (23) *Kopong-a da=maluw-i.*
 child-PL 3PL=be.hungry-3PL
 'The children are hungry.' (L3:189)

In general, most nouns preceding the main noun in the core NP are possessors of inalienable nouns and in that case the core NP is an inalienable possessive construction (cf. §3.3.4.2). Additional nouns preceding an alienable noun are more restricted. Only a hyponym-hypernym (kind-category) relation can be expressed between an alienable nouns and another noun preceding it. The alienable noun then expresses a kind and is preceded by another noun expressing a category. This is illustrated in (24) where the alienable head noun *granak* 'young male' expresses that the referent of *kopo granak* is 'a male kind' of the category 'child'.

- (24) *Kopo granak kéda kéi.*
 child\1 young.male big PFV
 'The boy is already grown up.' (C1:92)

As the head noun in this kind of construction is an alienable noun, it can be marked for plural, as in (25), or for specificity, as in (26).

- (25) *Bo dané kopo grana-sa d-awa.*
 DISC 3PL child\1 young.male-PL 3PL-still
 'They were still boys.' (M2:104)

- (26) ... *ta=dapat manuk telur-u.*
 1PL.INCL=get chicken egg-SPEC
 ‘... we will get eggs.’ (C3:115)

The sentence in (27) is an example of a noun phrase which has three nouns and a numeral in the core NP.

- (27) *Kopo lamé anak tuné na=diro duu anem tuné.*
 child male small.one one 3SG=PROG see thing one
 ‘A boy is looking at something.’ (S3:27)

This complex NP includes an inalienable possessive construction with *kopo lamé* ‘boy’ as the possessor and *anak* ‘small.one’ as the inalienably possessed noun.

3.3.4.2 Inalienable possessive construction

An inalienable head noun and its preceding possessor noun can express various semantic relations. These are possessum-possessor, meronym-holonym (part-whole), hyponym-hypernym (kind-category) and property relations.

In an inalienable possessive construction, the inalienable noun expresses the possessum, e.g. *ulu* ‘head’, and is marked with a possessor suffix (cf. §3.4.3), while the preceding noun denotes the possessor, e.g. *witi* ‘goat’, such as in (28).

- (28) *witi ulu-n*
 goat\ s head-3SG.POSS
 ‘the goat’s head’ (NT:2)

In this construction, it is possible to replace the possessor noun with proclitic, such as *mo*= ‘2SG’ in (29) or *kam*= ‘1PL.EXCL’ in (30).

- (29) *Mo=lima-m nao ga re?*
 2SG=hand-2SG.POSS kind where now
 ‘How is your hand now?’ (C6:177)
- (30) *kam=ina-mi*
 1PL.EXCL=mother-1PL.EXCL.POSS
 ‘our mothers’ (M3:116)

When the possessor is expressed by a preceding nominal, free pronoun or clitic, the possessor suffix on the inalienable noun can be left unexpressed, as in (31).

- (31) *rusa ulu*
 deer\’s head
 ‘the deer’s head’ (F1:42)

Part-whole relations are expressed by a construction that is parallel to an inalienable possessive construction. In addition to the pronominal possessor suffix, a nominal possessor obligatorily precedes the inalienable noun, as exemplified in (32). In (32), the inalienable head noun *lolo* ‘leaf’ refers to a meronym, a part of whole. Its holonym is expressed by *skajo* ‘tuber\’s’. The meronym *lolo* ‘leaf’ is marked by a possessor suffix, such as the possessed noun in a possessive construction, while the holonym *skajo* ‘tuber’ in the dependent slot remains unmarked, in the same way as the possessor in a possessive construction.

- (32) *skajo lolo-nu*
 tuber\’s leaf-3SG.POSS
 ‘a tuber leaf’ (C5:18)

Finally, an inalienable head noun can express a kind, such as *snipan* ‘side’, of a certain category, such as *éka* ‘garden\’s’ in (33) where the use of the possessor suffix *-i* ‘3PL’ on the head noun *snipan* ‘side’ yields a plural reading of the noun phrase.

- (33) *Da=muul-i waak-i éka snipan-i.*
 3PL=plant-3PL finish-3PL garden side-3PL.POSS
 ‘They finished planting the side gardens.’ (T3:57)

In Central Lembata, property nouns, such as *kédak* ‘big.one’ in (34), are inalienable nouns from a structural perspective. This means that they can be marked with possessor suffixes, such as *-i* ‘3PL’ in (34).

- (34) *naluk kédak-i*
 vegetable big.one-3PL.POSS
 ‘big vegetables’ (C5:24)

Here, the inalienable noun *kédak* ‘big.one’ is preceded by the noun *naluk* ‘vegetable’ referring to a category ‘vegetables’, namely ‘the big ones’. To represent this structure, this NP could be literally translated with ‘the big ones of vegetables’ meaning ‘big vegetables’.

The singular suffix *-n* ‘3SG.POSS’ can only attach to vowel-final stems. Therefore, consonant-final inalienable nouns, such as *anak* ‘small.one’ in (35), are not marked for third person singular.

- (35) *Kopo anak kerka-nga.*
 child small.one startle-3SG
 ‘The little child gets frightened.’ (F1:37)

Inalienable nouns ending in a vowel can be optionally marked with a possessor suffix, such as the 3SG.POSS suffix in (36).

- (36) *labur weru(-n)*
 shirt new.one-3SG.POSS
 ‘new shirt’ (NB:34)

For third person possessors in singular and in plural, the third person possessor suffix can be left out. This is observed especially in nouns expressing parts of wholes and property concepts where the possessor is also expressed in a preceding nominal. However, the pattern of this variation remains unclear and more research is needed.

3.3.4.3 Alienable possessive constructions

Nominal possessors of alienable nouns cannot be part of the core NP. Only free alienable possessor pronouns (cf. §3.4.3) can be found in the slot preceding an alienable noun. The NPs in (37) and (38) are examples of free possessor pronouns preceding the alienable noun within the core NP.

- (37) *go=kopong*
 1SG.POSS=child
 ‘my child’ (L3:195)
- (38) *Mo m-eti moé tapo-ja?*
 2SG 2SG-bring 2SG.POSS coconut-PL
 ‘Did you bring your coconuts?’ (C3:131)

In an alienably possessed noun phrase, a possessor pronoun is obligatory and cannot be replaced by a nominal possessor. However, a nominal possessor can optionally be added as an independent noun phrase which precedes the possessed noun phrase as in (39).

- (39) *Jon no Méri da=unan-u*
 NAME and NAME 3PL.POSS=house-SPEC
 'John and Mary's house' (N1:2)

In this example, the possessor NP *Jon no Méri* precedes the alienably possessed NP. In the alienably possessed NP, the possessor clitic *da=* '3PL.POSS' precedes the alienable noun *unan-u* 'house-SPEC'.

3.3.4.4 The numeral slot

After the noun in the core NP, a numeral (cf. §3.5.2.2) can occur as shown in (40).

- (40) *klié kora tuné*
 bird\s owl\s one
 'an owl' (F1:36)

In this core NP, there are two nouns followed by a numeral. The coda alternating noun *klié* 'bird\s' appears in the short form as it is modified by *kora* 'owl' and the *kora* 'owl' appears also in the short form as it is modified by the numeral *tuné* 'one' (cf. §3.3.1.2). The reason to analyse the numeral as part of the core NP is that it yields the short form of the preceding coda alternating noun. This is different from demonstratives which do not yield the short form of a preceding alternating noun and are thus considered to be outside of the core NP.

In noun phrases with inalienable nouns, the numeral slot can also be filled, such as in (41). However, the use of a numeral does not permit any possessor marking on the preceding inalienable noun, here *ulu* 'head'.

- (41) *rusa ulu tuné*
 deer\s head one
 'a deer's head' (F1:14)

This omission of the possessor suffix on an inalienable nouns appears parallel to the observation that coda alternating alienable nouns followed by a numeral, as *kora* ‘owl\’s’ in (40) followed by *tuné* ‘one’, have to appear in the short form which cannot take any suffixes.

Any numeral that is higher than one is obligatorily preceded by the general classifier *ua* ‘CLF’ as in (42).

- (42) *Skajo ua telu da=taru jé méja lolo.*
 tuber\’s CLF three 3PL=place upwards table top
 ‘Three cassava roots are placed on the table.’ (FH3:25)

The combination of classifier and numeral behaves in the same way as the numeral ‘one’ by not allowing any morphological marking on the noun and yielding the short form of coda alternating nouns, such as *skajo* ‘tuber\’s’.

3.3.4.5 Relative clauses

Central Lembata does not have a native marker for relative clauses. Usually, juxtaposition of clauses (cf. §3.6.11) is used rather than embedding. However, under the influence of the national language Indonesian, the Indonesian relative marker *yang* has entered Central Lembata speech and is very productive.

A relative clause is placed in the REL slot following dependent, head, and numerals, while preceding the demonstrative and slot for discourse markers. A relative clause can be used to relativise subject noun phrases (43) as well as object noun phrases (44) but it can also function as a nominaliser (45).

- (43) *Tité, yang sogu una biasan-i di, tité*
 1PL.INCL REL hold house normal-3PL.POSS DISC 1PL.INCL
kaa t-olu.
 1PL.INCL.eat 1PL.INCL-precede
 ‘Us, who have an ordinary house, we eat first.’ (T3:265)
- (44) *Buku, yang go opi wo ro, weli.*
 book REL 1SG buy DIST FIN expensive
 ‘The book that I bought was expensive.’ (NB:84)

- (45) *Kaju mekul bo yang klop.*
 wood part DISC REL round
 ‘The piece of wood [is] the round one.’ (S7:35)

The functions of the relativiser *yang* shown here in Central Lembata sentences are the same functions that the word *yang* has in Indonesian. Also the syntactic position of the relative clause before the demonstrative matches with the position of an Indonesian relative clause.

3.3.4.6 Phrase-final slots

The last two slots in an NP are the demonstrative slot and the slot for discourse particles. An example of a NP with demonstrative and the discourse particle *bo* in the final slots is given in (46).

- (46) *Magu tu ré bo na tobé-nga woné una lerit.*
 man one PROX DISC 3SG sit-3SG DIST house edge
 ‘This man is sitting at the edge of the house.’ (S5:42)

Here, the noun phrase is headed by the alienable noun *magu* ‘man’ followed by a numeral, and then the proximate demonstrative *ré* in the demonstrative slot and the discourse particle *bo* the slot for end particles. Demonstratives are discussed in more detail in §3.3.5 and §3.7.1. Discourse particles are introduced in §3.6.10.

3.3.5 Demonstratives in noun phrases

Central Lembata demonstratives (cf. §3.7.1) can replace a noun or a noun phrase, as seen in (47) with *wé* ‘PROX’, but also appear adnominally to locate a noun phrase in space, as seen in (48) with *wo* ‘DIST’. In example (47), the demonstrative *wé* ‘PROX’ is functioning as a subject NP, whereas in (48), the demonstrative *wo* ‘DIST’ locates the object NP *kaju tawa kédak* ‘big tree’ in space and follows the noun phrase.

- (47) *Wé da=lewor.*
 PROX 3PL.POSS=village
 ‘This is their village.’ (NT:9)

- (48) ... *ké dekul-a, dekul-a kaju tawa kédak wo* ...
 so crash-3SG crash-3SG tree stem big DIST
 ‘... so he crashes, he crashes into that big tree ...’ (S2:16)

The distal demonstratives, such as *wo* ‘DIST’ in (49), and the demonstrative *nakété*, such as in (50), can also be used anaphorically. In (49), the distal demonstrative *wo* ‘DIST’ specifies that the storage house had been mentioned before.

- (49) *Wétak wo tité bérsi-nga kéi.*
 storage.house DIST 1PL.INCL clean-3SG already
 ‘We already cleaned that storage house.’ (T3:170)

In (50), the use of the anaphoric demonstrative *nakété* indicates that the speaker is talking about the same child as he had already mentioned in the discourse.

- (50) *Kopong-u nakété diré léda la ama-n.*
 child-SPEC ANAPH stand lean LOC father-3SG.POSS
 ‘That child stands leaning against his father.’ (S2:2)

3.3.6 Nominalising morphology

Derivation by means of prefixes and infixes has been described for several Lamaholot varieties (Arndt 1937:6-16; Keraf 1978a:185-222; Pampus 1999:31-36). The most frequent nominalising affixes in Central Lembata are the general nominaliser *-n-* with its allomorph *b-* (§3.3.6.1) and the suffix *-k* (§3.3.6.2) which is only used to derive property nouns. These derivational processes appear to be fossilised nowadays. In some cases, only the derived form is still in use while its base has become obsolete. Most consonant clusters in Central Lembata probably go back to derived forms (cf. Table 3.6). In the following sections, I provide examples of each nominalisation affix.

3.3.6.1 The infix *-n-* and its allomorph *b-*

The infix *-n-* and the prefix *b-* are allomorphs of a nominalisation affix.¹¹ They typically nominalise a verb or an adjectival root. In a few cases, the

¹¹ As similar system of nominalisation morphology with even more allomorphs of the infix *-n-* has been described for Leti, an Austronesian language spoken to the east of Timor island (Blevins 1999).

base can also be a noun. The allomorphs are distributed according to the initial phoneme of the base, as laid out in Table 3.23. The prefix *b-* is preceding bases starting with *n*, *d*, *w* and *l*. In all other cases, the infix *-n-* is inserted directly after the word initial consonant. The choice of the allomorph appears to be related partially to the place of articulation of the initial consonant of the base. The prefix *b-* is used with stems starting with labial (*w*) or coronal (*n, d, l*) consonants, while the infix *-n-* is attested with consonants of labial, coronal and also dorsal places of articulation.

A nominalised vowel-final verb can take an additional final consonant, *k*, *r*, *ng* or *n*, during the derivation process, such as *k* in *wara* ‘burn’ → *bwarak* ‘in burning condition’. In some cases, it is unclear whether the final consonant belongs to the verbal stem (but is not realised when the verb appears in its bare form) or if it is indeed suffixed during the derivation process.

Table 3.23: The nominalisation infix *-n-* and its allomorph *b-*

Allomorph	Initial C	Base	Derivative
<i>b-</i>	<i>n</i>	<i>néka</i> ‘shine’	<i>bnéken</i> ‘light’
	<i>d</i>	<i>dori</i> ‘follow’	<i>bdorin</i> ‘the following (person)’
	<i>w</i>	<i>wara</i> ‘burn’	<i>bwarak</i> ‘in burning/ hot condition’
	<i>l</i>	<i>lawaj</i> ‘stretch’ (weaving process)	<i>blawaj</i> ‘stretching tool’
<i>-n-</i>	any other	<i>tiba</i> ‘ladle sth’	<i>tniba</i> ‘scoop’

Infixation with *-n-* leads to the morpho-phonological processes in Table 3.24. An initial *g* of the verb becomes voiceless when the infix is added. Initial bilabial *b* or *p* can lead to the merger of the nasal infix and the initial consonant into a single initial nasal *m*. The alveolar stop *t* followed by the infix can be deleted. However, this rule is only sporadically found. An example of an exception is *tungen* ‘cover’ → *tnungen* ‘covered’ where *t* is retained.

Table 3.24: Morpho-phonological processes following infixation

Process	Base	Derivative
$g > k / \_n-$	<i>gorot</i> ‘sweep’	<i>knorot</i> ‘broom’
$-n- > m- / \#\{p, b\}_-$	<i>balok</i> ‘mill sth.’	<i>malok</i> ‘cotton mill’
	<i>pujé</i> ‘follow (object)’	<i>mujéng</i> ‘new, following’
$-n- > n- / \#t_-$	<i>tukar</i> ‘climb’	<i>nukar</i> ‘steep condition’

There are several generalisations that can be drawn from the semantics of the bases and their nominal derivatives. When an action or action-process verb is nominalised, the derived noun will denote either the actor of the action, the tool to perform the action or the result of the process. In a few cases, the base is not an action verb but a noun. Examples of nominalised action verbs that express actions, tools or results are listed in Table 3.25.

Table 3.25: Nominalised action verbs

Base (v.)	Gloss	Derivative (n.)	Gloss
Action		Actor	
<i>dori</i>	‘follow’	<i>bdorin</i>	‘the following (person)’
<i>kiok</i>	‘bleep’	<i>kniok</i>	‘chicks’
<i>lalan</i>	‘way’ (n.)	<i>blalan</i>	‘walker’
<i>n-olu</i>	‘3SG-precede’	<i>bnolung</i>	‘ancestors’
<i>téré</i>	‘fill up’	<i>tnéréng</i>	‘person that fills sth. up’
Action-Process		Tool	
<i>balok</i>	‘mill’ (v.)	<i>malok</i>	‘cotton mill’
<i>bosil</i>	‘crash candlenuts’	<i>mosil</i>	‘candlenut crusher’
<i>géséng</i>	‘rub’	<i>knéséng</i>	‘a tool to make fire (by rubbing)’
<i>gorot</i>	‘sweep’	<i>knorot</i>	‘broom’
<i>lawaj</i>	to stretch (weaving process)	<i>blawaj</i>	‘stretching tool’
<i>sapé</i>	‘hang’	<i>snapéng</i>	‘hanger’

Base (v.)	Gloss	Derivative (n.)	Gloss
<i>sawit</i>	'break/cut'	<i>snawit</i>	'kind of tool'
<i>sigi</i>	'prick'	<i>snigi</i>	'toothpick'
<i>sipat</i>	'plant'	<i>snipat</i>	'dibble stick'
<i>subat</i>	'close' (v.)	<i>snubat</i>	'cap'
<i>tiba</i>	'ladle'	<i>tniba</i>	'scoop'
Action-Process		Result/Property	
<i>békén</i>	'let ripe'	<i>bnékén</i>	'stored to get ripe'
<i>bisak</i>	'break'	<i>bnisak</i>	'broken'
<i>bitol</i>	'open' (v.)	<i>bnitol</i>	'open' (adj.)
<i>bukat</i>	'open' (v.)	<i>bnukat</i>	'open' (adj.)
<i>diri</i>	'stand'	<i>bdiring</i>	'standing'
<i>dori</i>	'follow (human)'	<i>bdoring</i>	'the following'
<i>géngar</i>	'fry'	<i>knéngar</i>	'fried'
<i>kuluk</i>	'seed' (n.)	<i>knuluk</i>	'round'
<i>néka</i>	'shine'	<i>bnéken</i>	'light'
<i>n-olu</i>	'3SG-precede'	<i>nolung</i>	'old, preceding'
<i>paken</i>	'name' (v.)	<i>maken</i>	'name'
<i>pujé</i>	'follow (object)'	<i>mujéng</i>	'new, following'
<i>suda</i>	'order; command' (v.)	<i>snudang</i>	'slave'
<i>taki</i>	'make a ladder'	<i>tnaki</i>	'tree stairs'
<i>tejak</i>	'break'	<i>tnejak</i>	'broken'
<i>tetuk</i>	'crash'	<i>netuk</i>	'crashed'
<i>tisa</i>	'tier up'	<i>tnisang</i>	'torn; broken'
<i>tukar</i>	'climb'	<i>nukar</i>	'steep'
<i>tungen</i>	'cover' (v.)	<i>tnungen</i>	'covered'

When a verb denoting an involuntary process is nominalised, the derived noun will express the result of this process. Examples are given in Table 3.26.

Table 3.26: Nominalised verbs of involuntary processes

Base (v.)	Gloss	Derivative (n.)	Gloss
Involuntary process		Result	
<i>belok</i>	'appear'	<i>melok</i>	'hole'
<i>diri</i>	'stand'	<i>bdiring</i>	'in standing condition'
<i>mataj-</i>	'die'	<i>maték</i>	'dead'
<i>segu</i>	'smoke' (fire)	<i>snegur</i>	'smoke'
<i>seman</i>	'get rotten'	<i>sneman</i>	'something that is rotten'
<i>sobaj</i>	'swell'	<i>snobaj</i>	'in swollen condition'
<i>wara</i>	'burn'	<i>bwarak</i>	'in burning/hot condition'

Finally, adjectival roots can also be nominalised with the infix *-n-*. In this case, the semantic of the word does not change but it is rather the syntax of the phrase that determines whether the property concept has to be expressed as an adjective or a noun. More details on the grammar of property concepts can be found in §3.3.1.1 and §3.6.6.

Table 3.27: Nominalised adjectival roots

Base (adj.)	Derivative (n.)	Gloss
<i>berat</i>	<i>bnerat</i>	'heavy'
<i>gewa</i>	<i>knewak</i>	'rotten'
<i>gilu</i>	<i>kniluk</i>	'sour'
<i>kati</i>	<i>knating</i>	'hot'
<i>koda</i>	<i>knodan</i>	'long' (time)
<i>pait</i>	<i>pnait</i>	'bitter'
<i>peju</i>	<i>pnejuk</i>	'salty'
<i>seba</i>	<i>snebar</i>	'wet'
<i>seru</i>	<i>sneruk</i>	'sweet'
<i>susa</i>	<i>snusan</i>	'poor'
<i>sukar</i>	<i>snukar</i>	'difficult'

3.3.6.2 The suffix *-k*

The suffix *-k* attaches mainly to adjectival roots, and in a few cases also to nouns, as exemplified in Table 3.28. The suffix *-k* yields the meaning of ‘being like the object denoted’ (for nominal bases) or ‘being in the condition of the property denoted’ (for adjectival bases). The nominalised adjectival roots are property nouns which are semantically not different from their roots but functionally distinct (cf. §3.3.1.1 and §3.6.6).

Table 3.28: Derived property nouns with the suffix *-k*

Basis		Derived property noun	
Nominal base			
<i>rema</i>	‘night’ (obsolete)	<i>remak</i>	‘dark’
<i>watu</i>	‘stone’	<i>watuk</i>	‘frozen (stone condition)’
Adjectival root			
<i>ana</i>	‘small’	<i>anak</i>	‘small (condition)’
<i>buja</i>	‘white’	<i>bujak</i>	‘white (condition)’
<i>daé</i>	‘close by’	<i>daék</i>	‘close (condition)’
<i>doa</i>	‘far away’	<i>doak</i>	‘far (condition)’

3.3.7 Summary

Central Lembata has a grammatical distinction between inalienable and alienable nouns. Inalienable nouns take obligatory possessor suffixes, while alienable nouns optionally take a plural or a specificity suffix. The class of alienable nouns can be subdivided into two formal classes: simple nouns that have only one form, and coda alternating nouns that have one long and one short form. The use of the two forms of coda alternating nouns is syntactically conditioned. Long forms appear in final position of a pre-verbal core NP, while short forms are used in non-final position of a core NP and in any position of a post-verbal object NP. A Central Lembata noun phrase has a core part with up to three nouns and a numeral. Following the core NP, there is a final slot for demonstratives and discourse markers. A relative clause can be inserted in between the core NP and the NP-final slot.

Only the noun in the final position of the core NP can be morphologically marked. Demonstratives differentiate distal and proximate distance. They can replace a noun phrase or locate a noun phrase in space or discourse. In Central Lembata two main nominalising affixes are attested: the infix *-n-* nominalises verbs and the suffix *-k* nominalises adjectival roots and, to a lesser extent, also nouns.

3.4 Pronominals

3.4.1 Overview

Free and bound pronominals can replace a noun or a noun phrase. Only bound pronominals can occur with a co-referential noun phrase. In this section on pronominals, I discuss free pronouns, pronominal clitics and pronominal affixes. The Central Lembata pronominals are summarised in Table 3.29 and 3.30. Core argument pronouns, clitics and affixes denote a core argument of a clause, i. e. intransitive subject (S), transitive subject (A) or object (P). Possessor pronouns, clitics and suffixes denote the alienable or inalienable possessor of a possessed entity. In addition, pronominal proclitics and suffixes are relevant for mood and aspect marking as I show in the respective sections below (cf. §3.4.2.2 and §3.4.2.4). In the tables, free variants are separated by a slash, while conditioned allomorphs are separated by a semicolon. Several overlapping forms can be observed across the paradigms. In the sections hereafter, I discuss each pronoun set separately in more detail.

Table 3.29: Overview of Central Lembata core argument pronominals

	S/A/P Free	S/A Proclitics	S/A Prefixes	S/P Suffixes	
				G-set	K-set
1SG	<i>goné</i> / <i>go</i>	<i>ka=</i>	<i>k-</i>	<i>-ga</i>	<i>-ka</i>
2SG	<i>moné</i> / <i>mo</i>	<i>ma=</i>	<i>m-</i>	<i>-gu; -u</i>	<i>-ku</i>
3SG	<i>nané</i>	<i>na=</i>	<i>n-</i>	<i>-nga; -a</i>	<i>-na</i>
1PL.INCL	<i>tité</i>	<i>ta=</i>	<i>t-</i>	<i>-sa</i>	<i>-sa</i>
1PL.EXCL	<i>kamé</i>	<i>kam=</i>	<i>m-</i>	<i>-mi</i>	<i>-mi</i>
2PL	<i>mio</i>	<i>ma=</i>	<i>m-</i>	<i>-mi</i>	<i>-mi</i>
3PL	<i>dané</i>	<i>da=</i>	<i>d-</i>	<i>-ngi; -i</i>	<i>-Ø</i>

Table 3.30: Overview of Central Lembata possessor pronominals

	Alienable		Inalienable		
	Free	Proclitic	Free	Proclitic	Suffix
1SG	<i>goé</i>	<i>go=</i>	<i>goné</i>	<i>go=</i>	<i>-ga</i>
2SG	<i>moé</i>	<i>mo=</i>	<i>moné</i>	<i>mo=</i>	<i>-mu</i>
3SG	<i>naé</i>	<i>na=</i> / <i>né=</i>	<i>nané</i>	<i>na=</i>	<i>-nu; -V; -Ø</i>
1PL.INCL	<i>tité</i>		<i>tité</i>		<i>-sa</i>
1PL.EXCL	<i>kamé</i>	<i>kam=</i>	<i>kamé</i>	<i>kam=</i>	<i>-mi</i>
2PL	<i>mio</i>		<i>mio</i>		<i>-mi</i>
3PL	<i>daé</i>	<i>da=</i> / <i>dé=</i>	<i>dané</i>	<i>da=</i>	<i>-ja; -i</i>

When comparing the pronominal paradigms, several identical forms can be observed. The free pronouns for core arguments and for inalienable possessors are identical, with the exception that the core argument pronoun set includes two short forms for 1SG and 2SG. These cannot be analysed as proclitics because they also appear in post-verbal object position. The free pronouns denoting alienable possessors diverge in four out of seven forms from the other free pronoun sets.

The proclitics are in most cases abbreviations of the full free pronouns. Exceptions are the proclitics *ka=* ‘1SG’, *ma=* ‘2SG’, *ta=* ‘1PL.INCL’ and *ma=* ‘2PL’ which are only found in the proclitic set referring to core arguments and

not in the set referring to possessors. The initial consonants of these proclitics are the same as in the free forms, except for *ka*= ‘1SG’ which also diverges in the consonant, but the vowel is generalised to /a/ which is not found in the long form. Functionally, the subset of proclitics that are not directly derived from the free forms only appears in irrealis contexts.

In the suffix paradigms, the same forms occur for 1PL.INCL, 1PL.EXCL and 2PL across all paradigms, while the same form for 1SG is found in the inalienable possessor suffixes and in one of the S/P suffix sets.

The question must be asked whether the identical forms are homophones or syncretic, that is two different forms with the same phonological shape or one form with two meanings. In this grammar sketch, I treat them as homophones. In the glossing, I differentiate possessor pronouns from core argument pronouns, by glossing forms such as *go* either as ‘1SG’ or ‘1SG.POSS’ depending on their meaning in the context.

In the following sections, I first discuss core argument pronouns in §3.4.2, followed by a section on alignment of core arguments in 3.4.2.5. Possessor pronouns are introduced in 3.4.3.

3.4.2 Core argument pronominals

3.4.2.1 Free pronouns

In Table 3.31, I present the Central Lembata free pronouns. The set of full disyllabic pronouns is complete, whereas the set of monosyllabic free forms just contains pronouns for the first and second person singular.

Table 3.31: Free pronouns

	Full	Short
1SG	<i>goné</i>	<i>go</i>
2SG	<i>moné</i>	<i>mo</i>
3SG	<i>nané</i>	-
1PL.INCL	<i>tité</i>	-
1PL.EXCL	<i>kamé</i>	-
2PL	<i>mio</i>	-
3PL	<i>dané</i>	-

Historically, the ultimate syllable *né* found in all singular forms and in the third person plural is an addition, it cannot be reconstructed to a higher level than Proto-Lamaholot. This syllable appears to have been added to make the whole set of pronouns disyllabic. The same ultimate syllable is also found in locationals (cf. §3.7.1). The origin and meaning of *né* remains unknown.

Free pronouns have the same distribution as nouns, with the restriction that pronouns cannot be modified by other nouns or numerals. All full and short free pronouns can appear as S, A and P arguments, thus in both subject and object position. This makes them different from the proclitics discussed in 3.4.2.2 further below.

Free pronouns can also stand in isolation or be combined with other elements, such as a demonstrative, in a noun phrase as in (51) and (52). In these examples, the full pronouns *nané* ‘3SG’ and *tité* ‘1PL.INCL’ are combined with the proximate demonstratives *ré* and *wé* respectively in the same way as regular nouns are combined with demonstratives, as for example *aor wé* glossed as ‘dog PROX’ and translated as ‘this dog’.

- (51) *Nané ré surit.*
 3SG PROX sword
 ‘This is a weaving sword.’ (12:67)
- (52) *Tité wé t-o ga t-ai?*
 1PL.INCL PROX 1PL.INCL-to where 1PL.INCL-go
 ‘Where are we going to?’ (13:92)

3.4.2.2 Pronominal clitics

The Central Lembata proclitics given in Table 3.32 denote S and A subject arguments. The proclitics can be optionally preceded by a co-referential free pronoun or a full NP. Central Lembata shows a full paradigm of proclitic pronouns. However functionally, the set is not homogeneous. A subset of proclitic pronouns can only appear in irrealis contexts.

Table 3.32: Pronominal S/A clitics

	Proclitic	Context
1SG	<i>ka=</i>	Irrealis
2SG	<i>ma=</i>	Irrealis
3SG	<i>na=</i>	Realis & Irrealis
1PL.INCL	<i>ta=</i>	Irrealis
1PL.EXCL	<i>kam=</i>	Realis & Irrealis
2PL	<i>ma=</i>	Irrealis
3PL	<i>da=</i>	Realis & Irrealis

Three pronouns, *na=* ‘3SG’, *kam=* ‘1PL.EXCL’ and *da=* ‘3PL’, can appear in all contexts, i. e. in realis and irrealis sentences. An example is (53), where the proclitic *na=* ‘3SG’ attaches to the intransitive verb *géji* ‘enter’, indexing an S argument on the verb.

- (53) *Kopo biné tu na=géji wo una.*
 child\’s female\’s one 3SG=enter DIST house\’s
 ‘A girl enters the house.’ (FH1:15)

In (54), the proclitic *na=* ‘3SG’ is the only element denoting the A argument of the first clause with the verb *pakit* ‘peel’.

- (54) *na=pakit-a pa na=gaa*
 3SG=peel-3SG then 3SG=eat
 ‘She peels it, then she eats.’ (FH4:27)

The remaining pronouns, *ka=* ‘1SG’, *ma=* ‘2SG’, *ta=* ‘1PL.INCL’ and *ma=* ‘2PL’, are only found in irrealis sentences. Irrealis mood is used to express that there is no evidence for an event or state having actually happened (Payne 1997:244). In Central Lembata, three irrealis contexts are attested: the expression of intentions or future events, negated sentences and imperatives. In the following examples, the speaker expresses the intention to do something, in (55) the time reference is immediate future and in (56) it is distant future.

- (55) *Ma ka=tutu ré bo tentang Jon no Méri.*
 want 1SG=tell now DISC about NAME and NAME
 ‘I want to tell now about John and Mary.’ (N1:1)

- (56) *Sampé minggu jua kia, nepo ka=k-ai duur-a éka.*
 until week two INCEP later 1SG=1SG-go look-3SG garden
 ‘Only in two weeks from now, I will go to look after the garden.’
 (N3:117)

In addition to intentions or future events, proclitics are also used in negated sentences (57), and in imperatives (58). The 3PL suffix in (58) is unexpressed.

- (57) *Ta ka=k-éten-a si.*
 NEG 1SG=1SG-know-3SG NEG
 ‘I don’t know.’ (UM:34)
- (58) *Ma=guté-Ø wé ka=lou-ngi.*
 2SG=take-3PL so.that 1SG=rinse-3PL
 ‘Take them [washed cloths], so that I can rinse them.’ (C3:93)

Since first and second person proclitics (except for 1PL.EXCL) can only be used in irrealis contexts, no first and second proclitics can be used when realis contexts are expressed, instead free pronouns are used in such contexts. The fact that the restriction for irrealis contexts only applies to first and second person might be related to the fact that first and second person are more often used in expressions of intention and in imperatives, which are two of the typical irrealis context in Central Lembata.

3.4.2.3 Pronominal prefixes

The pronominal prefixes listed in Table 3.33 attach to a fixed set of vowel-initial verb roots given in (59). The majority of synchronic vowel-initial verbs in Central Lembata do not take any prefix. In contrast to the productive S/P marking suffixes discussed in the following section, these prefixes index S/A arguments on the verb and they cannot be left out.

Table 3.33: Prefix pronouns for S/A core arguments

1SG	<i>k-</i>
2SG	<i>m-</i>
3SG	<i>n-</i>
1PL.INCL	<i>t-</i>
1PL.EXCL	<i>m-</i>
2PL	<i>m-</i>
3PL	<i>d-</i>

- (59) Verbal roots which take person prefixes
*-a~a*¹² ‘unintentionally’, *-ai* ‘go’, *-al* ‘wear’, *-ao* ‘be like’, *-ar* ‘do’, *-ata* ‘the one that ...’, *-awa* ‘still’, *-énu* ‘drink’, *-era* ‘become; use; wear’, *-éten* ‘know; understand’, *-eti* ‘bring’, *-ewa* ‘catch; reach’, *-ia* ‘stay’, *-o* ‘towards’, *-olu* ‘precede’, *-ora* ‘join; be with’.

For some of the roots listed in (59), a generalisation of the third person singular prefix *n-* to all persons can be observed. This affects the stems *-ao* ‘be like’ in collocations as *nao gané* ‘how’ or *nao wé* ‘like this’, the aspect marker *-awa* ‘still’, the destination marker *-o* ‘towards’ and the commitative marker *-ora* ‘join; be with’. The first person plural inclusive prefix *t-* and the third person plural prefix *d-* are still found on these words but not in all contexts that would semantically demand them. Other person prefixes on these words are not found at all in my corpus.

3.4.2.4 Pronominal suffixes

Central Lembata has two functionally identical suffix sets given in 3.34. Both suffix sets mark S and P arguments. The two suffix sets are named after the initial consonants of their first and second person singular forms. For the 2SG, 3SG and 3PL suffixes in the set of G-suffixes, different allmorphs are used for vowel-final and consonant-final stems, as for example *-ngi* ‘3PL’ for vowel-final stems and *-i* for consonant-final stems.

¹² This is a fully reduplicated form where both parts take the obligatory prefix. An example is *kaka* ‘1SG-unintentionally~1SG-unintentionally’.

Table 3.34: Pronominal S/P suffixes

	G-suffixes	K-suffixes
1SG	-ga	-ka
2SG	-gu / -u	-ku
3SG	-nga / -a	-na
1PL.INCL	-sa	-sa
1PL.EXCL	-mi	-mi
2PL	-mi	-mi
3PL	-ngi / -i	-Ø

The G-suffixes are highly productive and are used with the majority of verbal stems in Central Lembata. In addition to verbal predicates, the G-suffixes also attach to nominal predicates. In (60), the suffix *-i* ‘3PL’ marks the intransitive subject of the verb. The same argument is also marked by a proclitic, this leads to double marking which is frequently found in Central Lembata (cf. §3.6.1).

- (60) *Pa da=panaw-i.*
 then 3PL=walk-3PL
 ‘Then they left.’ (N4:37)

The G-suffixes are also found on recent loans from Indonesian as in (61), where the third singular suffix *-a* ‘3SG’ marks the transitive object (P) of the verb *rékam* ‘record’ which is a loan from Indonesian.

- (61) *Mo rékam-a wé.*
 2SG record-3SG PROX
 ‘You recorded this.’ (C1:176)

The K-suffixes are a minor pattern which is less productive and only found on certain verbal predicates. The K-suffixes have the same functions as the G-suffixes, thus indexing S or P such as in (62) and (63). In (62), the intransitive verb *toba* ‘fall’ indexes the first person singular subject (S) argument using *-ka* ‘1SG’.

- (62) ... *ké toba-ka.*
 thus fall-1SG
 ‘... so, I fell down.’ (N2:69)

In (63), the transitive verb *bélo* ‘cut’ takes the first person singular suffix *-ka* ‘1SG’ to mark the object (P) argument.

- (63) *Ma=bélo-ka.*
 2SG=cut-1SG
 ‘You will kill me.’ (N4:74)

Unlike the G-suffixes, K-suffixes appear on a closed class of verbal stems. The presence of the phoneme *k* in the 1SG and 2SG suggests they are older forms as *k* occurs in the Proto-Malayo-Polynesian (PMP) pronouns **aku* ‘1SG’ and **kahu* ‘2SG’ (Blust and Trussel 2010). This distinction between G- and K-suffixes is also relevant in §3.5.2 on verb classes, as the verb classes are named after the suffix paradigm they take, thus K-suffix verbs and G-suffix verbs.

The use of pronominal *suffixes* is determined by the inherent aspect of the predicate, i. e. event or state. Stative predicates are marked with a suffix, while dynamic predicates, expressing an event, do not take a suffix. This can be observed with intransitive verbs (64) as well as transitive verbs (65). Unlike for the suffixes, the presence or absence of a pronominal *proclitic* does not influence the interpretation of the predicate as a state or event.

- (64) Intransitive verb: state vs. event
- a. *Na=géwé-na la watu lolo.*
 3SG=rise-3SG LOC stone top
 ‘He has climbed onto the rock.’ (NT:5)
- b. *Na=géwi la watu lolo.*
 3SG=rise LOC stone top
 ‘He is climbing onto the rock.’ (NT:5)
- (65) Transitive verb: state vs. event
- a. *Ema-n rio-na na=kopo.*
 mother-3SG.POSS wake.up-3SG 3SG.POSS=child\s
 ‘The mother has woken up her child.’ (NB:7)
- b. *Ema-n rio na=kopo.*
 mother-3SG.POSS wake.up 3SG.POSS=child\s
 ‘The mother wakes up her child.’ (NB:7)

The two sentences in (64) and (65) are only differentiated by the presence or absence of the person suffix *-na* ‘3SG’ on the verb. The verb *géwi* ‘ascend’ undergoes vowel lowering of the final vowel when a suffix is added. This is a regular pattern for final high vowels in K-suffix verbs (cf. §3.5.2.2).

The use of the person suffix *-na* ‘3SG’ in (64a) yields a stative interpretation of the clause, whereby the person has completed the process of climbing onto the rock. Without the suffix, as in (64b), the sentence means that the person is in the process of climbing up, thus expressing a dynamic event. Parallel to this, the suffix *-na* ‘3SG’ in (65a) yields a stative meaning, thus the action of waking up the child is completed and the child is awake now. Whereas in (65b), the event of waking up the child is ongoing.

Stative predicates, marked by S/P suffixes, can be combined with a the perfective aspect marker *kéi*, such as in (66a), and dynamic predicates, without suffixes, can occur with the imperfective marker *nawa* ‘still’, such as in (66b). Although the main verb *lebo* ‘bathe’ cannot take a suffix in the imperfective context, the verb *nawa* ‘still’ itself has a pronominal suffix. The predicate as a whole is clearly expressing a dynamic event but the verb *nawa* ‘still’ as such is non-dynamic. This explains the use of the suffix on this verb.

- (66) a. *Go lebo-ka waak kéi.*
 1SG bathe-1SG finish PFV
 ‘I already took a shower.’ (NT:15)
- b. *Go nawa-ka lebo.*
 1SG still-1SG bathe
 ‘I’m still showering.’ (NB:25)

Stative predicates do not necessarily refer to a past context. In example (63) above, a future state of being killed is expressed. Also in example (67), an immediate future state is expressed with the suffixed verb *lebo-ka* ‘bathe-1SG’. This event is not taking place at the moment of the utterance, therefore the suffix can be used.

- (67) *Go ka=lebo-ka kia.*
 1SG 1SG=bathe-1SG INCEP
 ‘I will take a shower now.’ (NT:14)

Under specific circumstances, S/A suffixes may also mark nominal predicates, which are similar to states in being non-dynamic. In (68), I provide an

example of the noun *user* ‘content’ which takes the verbal person suffix *-a* ‘3SG’ to express the meaning of ‘having content’.

- (68) *Lisor-u* *mulai na=user-a*.
 rice.plant-SPEC start 3SG=content-3SG
 ‘The rice plants start to have grains.’ (T3:109)

Nominal predicates with property nouns, such as in (69), can also be marked with S/P suffixes. In Central Lembata, property concepts are in most contexts expressed by inalienable nouns (cf. §3.6.6 and §3.3.1.1), and these nouns can function predicatively.

- (69) a. *Buku mipiw-a*.
 book thin.one-3SG
 ‘The book gets thin(ner).’ (NB:19)
 b. *Go golok-ga*.
 1SG tall.one-1SG
 ‘I am tall. / I get taller.’ (NB:65)

In most cases, the use of S/P suffixes on the nominal predicate leads to a procedural reading or a change of state, such as for *mipiw* ‘thin.one’ in (69a), where the book becomes thinner. In some cases the semantics of the suffixed predicate are more vague, such as in (69b). The translation of (69b), containing the nominal predicate *golok* ‘tall.one’ with a first person suffix, can either be ‘I am tall’ or ‘I get taller’. Possibly, the suffixes are generally favoured in a third person context, as they are related to uncontrolled, non-dynamic states rather than events, and these are more likely to occur with third persons than with first and second person. This might be a reason why the interpretation of (69b) remains vague. When modified with *bé* ‘rather’, as in (70a), the nominal predicate is obligatorily suffixed. The same sentence without a suffix, such as in (70b) is ungrammatical.

- (70) a. *Da=duu éka* *si* *bé* *lérék-a*.
 3PL=see condition a.bit rather flat.one-3SG
 ‘They saw that the land was rather flat.’ (L2:23)
 b. * *Da=duu éka* *si* *bé* *lérék*.
 3PL=see condition a.bit rather flat.one
 Intended: ‘They saw that the land was rather flat.’ (NB:88)

In addition, S/P suffixes occur in imperatives, such as in (71).

- (71) *Mo boté-na kopo wo nau, wé go boti.*
 2SG carry-3SG child DIST descend so.that 1SG carry
 ‘Bring that child down here, so that I can carry it.’ (NT:5)

Without the S/P suffix *-na* ‘3SG’, the sentence has a declarative reading, as shown in (72), where the verb *boti* ‘carry’ appears without suffix.

- (72) *Mo boti kopo wo nau.*
 2SG carry child DIST descend
 ‘You are carrying the child down here.’ (NB:87)

The pattern of stative predicates with suffixes and event predicates without suffixes could be a remnant of a semantic alignment system that marks S like P (with a suffix) in stative predicates, and S like A (without a suffix) in event predicates. This is a common pattern in the region (Klamer 2008). However, in Central Lembata not only the marking of S appears to be affected by the dynamics of the predicate but also the marking of P. Thus, P is not always suffixed but this is done only in non-dynamic contexts. In all other contexts P is not marked on the verb.

In addition to the general free pronouns introduced in §3.4.2.1, Central Lembata has complex plural pronouns formed from numerals and pronominal suffixes. Examples of this are given in Table 3.35.

Table 3.35: Complex pronouns based on numerals

	Suffix	<i>jua</i> ‘two’	<i>telu</i> ‘three’	<i>paat</i> ‘four’
1PL.INCL	<i>-ta / -sa</i>	<i>jua-ta</i>	<i>telo-sa</i>	<i>s-paat-i-sa</i>
1PL.EXCL	<i>-m(i)</i>	<i>jua-mi</i>	<i>telo-mi</i>	<i>s-paat-i-mi</i>
2PL	<i>-m(i)</i>	<i>jua-mi</i>	<i>telo-mi</i>	<i>s-paat-i-mi</i>
3PL	<i>-Ø</i>	<i>jua-Ø</i>	<i>telo-Ø</i>	<i>s-paat-i-Ø</i>

The complex pronouns can appear in all core argument positions, as *juata* in subject position in (73) and *juam* in object position in (74). The 1PL.EXCL and 2PL suffixes can be reduced to *-m*.

- (73) *Bambé bo jua-ta t-ai lau éka.*
 tomorrow DISC two-1PL.INCL 1PL.INCL-go seawards garden
 ‘Tomorrow, the two of us will go to the garden.’ (N3:72)
- (74) *Mo tutu mio jua-m néné Kara ro wé.*
 2SG tell 2PL two-2PL grandfather NAME FIN PROX
 ‘Tell about the two of you and grandfather Kara!’ (C6:178)

The person suffixes, except for *-ta* ‘1PL.INCL’, used in the complex pronouns above are the same as the verbal K-suffixes (cf. Table 3.34). The suffix *-ta* ‘1PL.INCL’ which only appears with numeral *jua* ‘two’ could be an older form of the first person plural inclusive suffix with the plosive *t* not being fricativised to *s*. For numerals higher than three, the circumfix *s-...-i* is added for all persons. Basically, all one-word-numerals can become complex pronouns by affixation. However, the higher the numeral, the less natural such affixation is.

As numerals share their morpho-syntactic properties with K-suffix verbs, numerals are analysed as a subclass of verbs (cf. §3.5.2.2). Thus the complex numeral pronouns are structurally intransitive verbal predicates.

3.4.2.5 Alignment of core arguments

Free pronouns can be used for all three main argument types, S, A and P. Central Lembata does not have case marking on nouns or pronouns.

However, Central Lembata has a head-marking structure which is manifested in bound pronouns that mark arguments on the verb. As shown in the previous sections, there is one proclitic set (cf. §3.4.2.2), one prefix set (cf. 3.4.2.3) and two suffix sets (cf. §3.4.2.4). When considering grammatical alignment patterns, proclitics and prefixes show nominative-accusative alignment by indexing S and A arguments with the same clitic or prefix on the verb, whereas P arguments remain unmarked. On the other hand, the suffixes follow a ergative-absolutive pattern by marking S and P with the same suffixes on the verb, whereas A remains unmarked.

The great majority of verbs are found in clauses where both alignment patterns operate at the same time. This is shown in the following examples. In (75) and (76), the accusative alignment is visible in the proclitics, as the same proclitic *na=* ‘3SG’ is used in (75) to mark S and in (76), it is used to mark A.

- (75) *Na=gelé-na.*
 3SG=lie-3SG
 ‘She is lying down.’ (N4:21)

- (76) *Nepo na=supeng-u.*
 later 3SG=pick.up-2SG
 ‘Then she will pick you up.’ (Cl:240)

In (76) and (77), the ergative alignment is visible in the suffixes, as the same suffix *-u;gu* ‘2SG’ is used in (76) to mark P, and in (77) to mark S.

- (77) *Nepo mo tué-gu.*
 later 2SG return-2SG
 ‘The you will return home.’ (Cl:229)

Obviously, for the class of suffixless verbs (cf. §3.5.2), only nominative-accusative marking is available. As most of the suffixless verbs are intransitive, they only have an S argument and cannot take a suffix. This pattern suggests a split-S alignment system, where, for some verbs, such as the suffixless verbs described here, S is marked in the same way as A, namely without a suffix, and for other verbs, S is marked in the same way as P, namely with a suffix. Further investigation of this is needed to clearly show that this apparent split-S alignment is based on semantic properties of the S argument.¹³

3.4.3 Possessor pronouns and suffixes

Central Lembata has different sets of possessor pronominals for alienable and for inalienable nouns listed in Table 3.36. For alienable nouns, a prenominal free possessor pronoun or a proclitic refers to the possessor, while for inalienable nouns a bound possessor suffix is obligatory and a free pronoun or a proclitic can be added optionally preceding the noun.

¹³ Grangé (2015b:43) describes semantic alignment for Eastern Adonara, a variety of Western Lamaholot.

Table 3.36: Possessor pronominals

	Alienable		Inalienable		
	Free	Proclitic	Free	Proclitic	Suffix
1SG	<i>goé</i>	<i>go=</i>	<i>goné</i>	<i>go=</i>	<i>-ga</i>
2SG	<i>moé</i>	<i>mo=</i>	<i>moné</i>	<i>mo=</i>	<i>-m(u)</i>
3SG	<i>naé</i>	<i>na= / né=</i>	<i>nané</i>	<i>na=</i>	<i>-n(u); -V; -Ø</i>
1PL.INCL	<i>tité</i>		<i>tité</i>		<i>-sa</i>
1PL.EXCL	<i>kamé</i>	<i>kam=</i>	<i>kamé</i>	<i>kam=</i>	<i>-mi</i>
2PL	<i>mio</i>		<i>mio</i>		<i>-mi</i>
3PL	<i>daé</i>	<i>da= / dé=</i>	<i>dané</i>	<i>da=</i>	<i>-ja; -i</i>

Alienable possessor pronouns precede the possessed noun as in (78).

- (78) *Naé teman oli.*
 3SG.POSS friend arrive
 ‘Her/her friend is coming.’ (FH2:15)

In this example, the pronoun *naé* ‘3SG.POSS’ marks the noun *teman* ‘friend’ as possessed by a third person singular possessor. The free possessor pronouns either appear in their full form as *naé* ‘3SG.POSS’ such as in (78) above or as a proclitic as *go=* ‘1SG.POSS’ in (79).

- (79) *go=unan*
 1SG.POSS=house
 ‘my house’ (C1:187)

Full forms and proclitics are free variants. Note that most proclitics overlap with core argument pronouns or proclitics (cf. §3.4.2.1 and 3.4.2.2). However, as for most persons the full forms are divergent, it is possible to identify whether a proclitic is a possessor or a core argument pronoun by testing the possibility of replacement with the long forms. In example (79) above, the proclitic *go=* ‘1SG.POSS’ can be replaced by *goé* ‘1SG.POSS’ but a replacement with *goné* ‘1SG’ would be ungrammatical.

The bound possessor pronouns in Table 3.36 are attached to inalienable nouns such as in (80), where the inalienable kinship term *ema* ‘mother’ is marked for its first person singular inalienable possessor using the suffix *-ga* ‘1SG.POSS’.

- (80) *Ema-ga noto kérjan ré rai~rai.*
 mother-1SG.POSS give work PROX many~many
 ‘My mother gave me a lot of work.’ (N2:37)

The allomorphs *-ja* and *-i* for the third person plural in Table 3.36 are used for vowel-final stems and consonant-final stems respectively. However, the three possessor suffix variants *-nu*, *-V* and *-Ø* for the third person singular cannot be merely explained by the phontactics of the stem to which they attach. There appears to be a tendency for *-nu* to be used with kinship terms. Vowel lengthening (doubling of the final vowel) and the zero suffix occur with vowel-final and consonant-final body part terms respectively. This variation is illustrated with the inalienable nouns *ina* ‘mother’, *lotor* ‘knee’ and *lima* ‘hand’ in Table 3.37. The inalienable nouns in the table can be optionally preceded by an inalienable possessor pronoun. An example is *goné ina-ga* with a free pronoun or *go=ina-ga* ‘my mother’ with a proclitic. Alternatively, it is also possible to express the possessor by a preposed full noun, such as in *witi ulu-n* ‘goat head-3SG.POSS’ = ‘goat’s head’. The preposed inalienable possessor noun is part of the core NP with the possessed noun and cannot co-occur with a free possessor pronoun or clitic (cf. §3.3.4).

Table 3.37: Inalienably possessed nouns

	‘mother’	‘knee’	‘hand’
1SG.POSS	<i>ina-ga</i>	<i>lotor-ga</i>	<i>lima-ga</i>
2SG.POSS	<i>ina-mu</i>	<i>lotor-mu</i>	<i>lima-mu</i>
3SG.POSS	<i>ina-nu</i>	<i>lotor-Ø</i>	<i>limaa</i>
1PL.INCL.POSS	<i>ina-sa</i>	<i>lotor-sa</i>	<i>lima-sa</i>
1PL.EXCL.POSS	<i>ina-mi</i>	<i>lotor-mi</i>	<i>lima-mi</i>
2PL.POSS	<i>ina-mi</i>	<i>lotor-mi</i>	<i>lima-mi</i>
3PL.POSS	<i>ina-ja</i>	<i>lotor-i</i>	<i>lima-ja</i>

A special type of possessive pronouns are pronouns that replace an entire possessive construction in the same way as the English pronoun *mine* can refer to *my house*. These pronouns are listed in Table 3.38. They can appear in all NP positions, for example in subject position such as in (81) and as a nominal predicate of an equative clause as in (82).

Table 3.38: Possessive pronouns referring to possessed items

Gloss	Singular	Plural
‘mine’	<i>goénu</i>	<i>goésa</i>
‘yours (sg.)’	<i>moénu</i>	<i>moésa</i>
‘his, hers, its’	<i>naénu</i>	<i>naésa</i>
‘ours (incl.)’	<i>titénu</i>	<i>titésa</i>
‘ours (excl.)’	<i>kaménu</i>	<i>kamésa</i>
‘yours (pl.)’	<i>mionu</i>	<i>miosa</i>
‘theirs’	<i>daénu</i>	<i>daésa</i>

- (81) *Bo naésa bo také.*
 DISC 3SG.POSS.PL DISC not.exist
 ‘He does not have any [children].’ Lit. ‘His do not exist.’ (C6:265)
- (82) *Una-ja wo ro titésa.*
 house-PL DIST FIN 1PL.INCL.POSS.PL
 ‘Those houses belong to us.’ Lit. ‘These houses are ours.’ (NB:46)

These possessive pronouns are composed of a free possessor pronoun and the inalienable possessor suffix *-nu* (< PMP **nu* ‘genitive case marker for common nouns’) for singular possessed items and *-sa* for plural possessed items. These inalienable possessor suffixes are homophonous with the possessor suffixes *-nu* ‘3SG.POSS’ and *-sa* ‘1PL.INCL.POSS’. However, in these constructions their semantics is reduced to singular and plural and the person value is lost.

3.4.4 Summary

Central Lembata has five sets of pronominals encoding core arguments: a set of free pronouns, a set of proclitics, a set of prefixes and two sets of suffixes. Free pronouns can refer to all three core argument types of a clause, i. e. S, A and P. The proclitics and the prefixes can only index S and A arguments, while the suffixes only index S and P arguments. A subset of the proclitics only occurs in irrealis contexts. The prefixes are obligatory for a fixed set of vowel-initial verbs. The suffixes are associated with non-dynamic and

uncontrolled predicates. To express a possessor, for inalienable and for alienable nouns different sets of pronominals are used. Alienable nouns are obligatorily marked with a preposed free possessor pronoun or a possessor proclitic from a set specific to alienable nouns. Inalienable nouns are obligatorily marked with a possessor suffix and can be preceded optionally by a free pronoun or proclitic from a set specific to inalienable nouns.

3.5 Verbs

3.5.1 Overview

In this section, I introduce three morpho-syntactic verb classes: G-suffix verbs, K-suffix verbs, and suffixless verbs. In Central Lembata, numerals and locationals are subclasses of verbs. Numerals fall into the class of K-suffix verbs and locationals are suffixless verbs. In §3.5.3, I present two cases of verbs that grammaticalised in several functional directions. These are the verb *-ora* ‘be with; join’ and the verb *maa* ‘say’.

3.5.2 Verb classes

Verbs can be divided into three different classes based on their eligibility to take person suffixes (cf. §3.4.2.4).

1. G-suffix verbs (open)
2. K-suffix verbs (closed)
3. Suffixless verbs (open)

The distribution of verbs over these classes is lexicalised and not determined semantically. The class of G-suffix verbs is an open class that contains the majority of the verbs. Verbs in this class can take suffixes from the G-suffix set. The class of K-suffix verbs is a small closed class of verbs that takes suffixes from the K-suffix set. There are about 30 K-suffix verbs in my corpus out of 360 verbs in total. This class is a relic of an older inflection pattern which is only used with the small set of K-suffix verbs. The class of suffixless verbs contains those verbs that cannot take any suffix, including locationals which synchronically have clear verbal properties. Suffixless verbs are an open class of roughly equal size with the class of K-suffix verbs.

3.5.2.1 G-suffix verbs

G-suffix verbs are verbs that can take the suffixes of the G-set given in Table 3.39 (cf. §3.4.2.4).

Table 3.39: G-Suffixes

1SG	-ga
2SG	-gu / -u
3SG	-nga / -a
1PL.INCL	-sa
1PL.EXCL	-mi
2PL	-mi
3PL	-ngi / -i

With over 300 lexical entries, G-suffix verbs constitute the great majority of verbs in my corpus. Table 3.40 lists a few examples of G-suffix verbs grouped according to their stem type.

Table 3.40: G-suffix verbal stems and their stem types

	V-final	C-final	Coda alternating
C-initial	<i>tula</i> ‘make’	<i>bekat</i> ‘fly’	<i>duu</i> / <i>duur</i> ‘see’
V-initial	<i>odo</i> ‘give’	<i>élaj</i> ‘cut’	<i>era</i> / <i>eraw</i> ‘cry’
Prefixed	<i>-era</i> ‘become’	<i>-éten</i> ‘know’	<i>-énu</i> / <i>-énum</i> ‘drink’

G-suffix verb stems are either vowel-final as *tula* ‘make’, consonant-final as *bekat* ‘fly’ or have two stem forms, that alternate between vowel-final and consonant-final form such as it is found for *duu* / *duur* ‘see’. In this case of alternating coda, the consonant-final form is always employed when the verb has a suffix as in (83). If it appears without a suffix, the vowel-final form is used as in (84). It is different to nominal coda alternation which is based on the syntactic position of the noun (cf. §3.3.1.2).

- (83) *Nepo ka=k-ai duur-a éka.*
 then 1SG=1SG-go see-3SG garden\s
 ‘Then, I will go to have a look at the garden.’ (N3:117)

- (84) *Ka=duu ewa.*
 1SG=see animal\|s
 ‘I will have a look at the animals.’ (C1:195)

G-suffix verb stems are disyllabic with a (C)(C)V.(C)V(C) pattern. Inflected G-suffix verbs can allow a sequence of two consonants in word medial position when a consonant-final stem takes a consonant-initial suffix as in (85). These consonant sequences are only found at morpheme boundaries within a morpho-syntactic word but not within a stem (cf. §3.2.2.2 and 3.2.4). In this example, the verb *duler* ‘visit’ takes the suffix *-ga* ‘1SG’. This results in a word-medial sequence of /r.g/ at a syllable and morpheme boundary.

- (85) *Da=duler-ga.*
 3PL=visit-1SG
 ‘They visited me.’ (C6:80)

This contrasts with the phontactics of nouns, where consonant clusters are resolved when suffixes are added according to specific rules discussed in §3.3.1.2. This could point to a looser connection between verbal stems and pronominal suffixes than between nominal stems and plural and specificity suffixes.

3.5.2.2 K-suffix verbs

K-suffix verbs are verbs that can take the suffixes of the K-set given in Table 3.41 (cf. §3.4.2.4). With around 30 intransitive and transitive content verbs (86) and twelve basic numerals (87), K-suffix verbs are a small closed class.

Table 3.41: K-suffixes

1SG	<i>-ka</i>
2SG	<i>-ku</i>
3SG	<i>-na</i>
1PL.INCL	<i>-sa</i>
1PL.EXCL	<i>-mi</i>
2PL	<i>-mi</i>
3PL	<i>-Ø</i>

(86) Intransitive K-suffix verbs

biti 'execute a certain weaving process', *bura* 'be full (of food)', *diri* 'stand', *gelé* 'lie down; sleep', *geto* 'be broken', *gewa* 'be rotten', *géwi* 'ascend; enter', *iré* 'visit; play; walk about', *kari* 'run', *lebo* 'take a shower; take a bath', *mea* 'be on one's own', *modo* 'fall', *nebu* 'not be thirsty', *nubu* 'grow', *wara* 'burn'

Transitive K-suffix verbs

baju 'pound', *bari* 'pick', *batu* 'throw', *boti* 'carry; hold', *-eti* 'bring', *giki* 'bite', *gliku* 'hold; place a holder', *guti* 'take', *liwu* 'fill in', *rio* 'wake someone up', *sodi* 'catch', *tota* 'burn down', *tuno* 'grill', *widu* 'pull'

(87) Monomorphemic numerals

tune / *tu* 'one', *jua* 'two', *telu* 'three', *paat* 'four', *léém* 'five', *enem* 'six', *pito* 'seven', *buto* 'eight', *siwa* 'nine', *spulo* 'ten', *ratu* 'hundred', *ribu* 'thousand'

The K-suffixes probably represent an older inflection pattern which has been retained with several intransitive as well as transitive verbs (cf. §3.4.2.4). Most of the verbs in (86) are part of the basic vocabulary. Possibly, the older suffix pattern is retained on these verbs because they are frequently used.

The stems taking K-suffixes start with a consonant and end in a vowel. With the K-suffixes which all start with a consonant, these verbs show a very regular C(C)V.(C)V.CV structure when inflected.

Another feature only found with K-suffixes is vowel lowering of the stem-final vowel. A suffix attached to a stem ending in /i/ or /u/ causes this final vowel to be lowered to /e/ and /o/ respectively, such as in *guté-na* 'take-3SG' from *guti* 'take', *karé-na* 'run-3SG' from *kari* 'run' or *wido-ka* 'pull-1SG' from *widu* 'pull'. The suffix can be left unexpressed, while the vowel lowering is retained and therewith the stative meaning of the phrase (cf. §3.4.2.4), such as contrasted in (88).

- (88) a. *Go guti mok.*
 1SG take mug
 'I take the mug.' (NT:5)
- b. *Go guté-na mok kéi.*
 1SG take-3SG mug PFV
 'I already took the mug.' (NT:5)
- c. *Go guté-Ø mok kéi.*
 1SG take-3SG mug PFV
 'I already took the mug.' (NT:5)

In the set of K-suffix verbs there are a few base forms with final mid-vowels, such as *gelé* 'lie down; sleep', *geto* 'be broken', *rio* 'wake s.o. up', *tuno* 'grill' (< PMP *tunu), *pito* 'seven' (< PMP *pitu), *spulo* (<PMP *sa-ŋa-puluq) and *buto* 'eight' (< PFL *butu 'eight; bunch; group'). As none of these final mid-vowels traces back to a historic mid-vowel (see proto-forms given in brackets if available), it can be hypothesised that for all or at least some of these verbs, the mid-vowels were initially caused by suffixation but then generalised to all forms of these verbs.

Numerals with suffixes are predicates but they are not fully verbal anymore as they cannot function as the predicate of a full independent clause. These numeral predicates rather fill the slot of free pronouns (cf. §3.4.2.1), as for example *telo-mi* 'three-1PL.EXCL' in (89).

- (89) *Telo-mi m-énu wai.*
 three-1PL.EXCL 1PL.EXCL-drink water
 'The three of us were having drinks.' (C1:182)

As for other K-suffix verbs, stem-final high vowels in numerals are lowered to mid vowels when a suffix is attached, such as *telu* 'three' becomes *telo-mi* 'three-1PL.EXCL' or *telo-Ø* 'three-3PL'. In the latter example of the third person plural suffix -Ø, the vowel lowering on the stem is evidence for the presence of a zero suffix.

In addition to their ability to take person suffixes, numerals often appear with a core argument clitics preceding them, such as in (90).

- (90) *Da=telo-Ø wo d-ai ...*
 3PL=three-3PL DIST 3PL-go
 'The three of them went there ...' (M2:73)

The ability of numerals to be preceded by core argument clitics is additional evidence for their verbal status.

3.5.2.3 Suffixless verbs

Suffixless verbs are a set of mainly intransitive verbs that cannot occur with a pronominal suffix. These verbs represent a small class with currently about 30 content words in my corpus listed in (91) and a set of 9 locationals with several variants discussed in §3.7.1.

(91) Intransitive suffixless verbs

-ai ‘go’, *belo(k)* ‘appear’, *blélot* ‘look back’, *blonget* ‘look down’, *boko* ‘stand up’, *bowo* ‘bark’, *denger* ‘hear’, *édo* ‘shake’, *géji* ‘ascend; to enter’, *géka* ‘laugh’, *glasa* ‘play’, *gulur* ‘snore’, *gwajo* ‘visit’, *iwol* ‘move the mouth because of food which is too hot’, *jadi* ‘give birth’, *ksemil* ‘smile’, *maa* ‘say; to think’, *péko* ‘turn’, *péngos* ‘turn away’, *sidol* ‘walk/stand on tiptoe’, *suduk* ‘bend’, *swedok* ‘limp’, *tukar* ‘climb’, *wosuk* ‘move backwards’

Transitive suffixless verbs

-aar ‘do’, *mawar* ‘share’, *nanam* ‘plait’, *sigi* ‘prick’

Suffixless verbs can begin and end in either a vowel or a consonant. Their template corresponds to what is generally found in Central Lembata disyllabic stems, namely (C)(C)V.(C)V(C).

Most suffixless verbs are intransitive. As suffixes mark S and P arguments on verbs but not A, intransitive verbs that do not take suffixes treat their S argument like A, namely unmarked or marked by a free pronoun or proclitic. This supports analysing Central Lembata as having a split-S system (cf. §3.4.2.5).

3.5.3 Grammaticalised verbs

In this section, I discuss two verbs that have grammaticalised in several directions. These are the verb *-ora* ‘join; be with’ belonging to the class of G-suffix verbs and the verb *maa* ‘say’ which is a suffixless verb. The verb *maa* ‘say’ probably goes back to Proto-Flores-Lembata (PFL) *ma ‘tongue’ (possibly connected with PMP *həma ‘tongue’) or PFL *madi ‘say’.

3.5.3.1 The verb *-ora* ‘join; be with’

Although not frequent, the verb *-ora* ‘join; be with’ is still found as a fully inflected verb, such as in (92) or (93).

- (92) *Go k-ora k-ai weli éka.*
 1SG 1SG-join 1SG-go sideways garden
 ‘I join in going to the garden.’ (NB:8)

- (93) *Kam=m-ora-gu pnua.*
 1PL.EXCL=1PL.EXCL-be.with-2SG talk
 ‘We are talking with you.’ (Cl:133)

In both examples, the verb *-ora* ‘join; be with’ appears as the first verb in a serial verb construction. It is inflected for subject person and in example (93), the object is also indexed on the verb using the bound pronoun *-gu* ‘2SG’ which shows that this verb is transitive and follows the G-suffix pattern.

Next to this verbal use, we find the verb *-ora* grammaticalised into an existential (cf. §3.6.8), prepositions (94a), (94b) and (94c), and a coordinating conjunction (95).

As a preposition, the word *nora* can introduce different kinds of objects, such as a comitative argument as in (94a) and (94b), or a P argument as in (94c). The preposition *nora* or its short form *no* always appears with the default agreement prefix *n-* ‘3SG’. It does not agree with the S/A argument of the clause. An example is given in (94a) in which the subject is first person plural exclusive and in (94c) in which the subject is third person plural.

- (94) a. *Kam=mojip tali nora liso nora kwaru.*
 1PL.EXCL=live again COM rice COM corn
 ‘We live again of rice and corn.’ (M3:63)
- b. *Ta=minum no tata nona kia, ina.*
 1PL.INCL=drink COM older.sibling young.woman INCEP girl
 ‘Let’s drink with your older sister now, my dear.’ (Cl:38)
- c. *Da=bara-nga nora bala.*
 3PL=carry-3SG COM ivory.tusk
 ‘They are carrying an ivory tusk.’ (Tl:123)

As a coordinating conjunction, *nora* or *no* is translated with ‘and’, as in (95).

- (95) *Bapa tuné nora na=kopo ...*
 father one COM 3SG.POSS=child
 'A father and his child ...' (S2:12)

In contrast to the prepositional use of *nora*, as a coordinator the long form *nora* appears less grammaticalised. The use of the default agreement prefix *n-* '3SG' is not possible for non-3SG NPs as the ungrammatical phrase in (96a) demonstrates. A construction such as in (96b) which is inflected for 1SG must be used instead. Different to the long form *nora*, the coordinator in its short form *no* can be used to coordinate a first person pronoun with another nominal element as can be seen in (96c). In this example, the conjunction *no* does not agree with the preceding first person singular pronoun which leads to the conclusion that *no* is completely grammaticalised into a conjunction here and has lost its verbal properties.

- (96) a. **go nora bapa-ga*
 1SG and father-1SG
 Intended: 'me and my father' (NB:8)
- b. *Go k-ora kopong-u wé ...*
 1SG 1SG-COM child-SPEC PROX
 'Me and this child ...' (NB:8)
- c. *go no bapa-ga*
 1SG and father-1SG
 'me and my father' (NB:8)

The bound verbal stem *-o* 'towards' may come from the verb *-ora* 'join' as well, such as in (97). The fact that this element is inflected clearly points to a verbal origin, however the word is so short that it is difficult to determine where it came from and the connection between *-o* 'towards' and *-ora* 'join' has to remain tentative.

- (97) a. *Biné tuné diro géji n-o ju una*
 female one PROG ascend 3SG-towards downwards house
or n-ai.
 inside 3SG-go.
 'A girl is climbing into the house.' (FH5:5)
- b. *Pindah t-o lau lewu weru t-ai.*
 move 1PL.INCL-towards seawards village new.one 1PL.INCL-go.
 'Let's move to the new village (down there in seaward direction).'
 (L2:138)

In sum, the verb *-ora* 'join; be with' has grammaticalised into comitative prepositions, a coordinator and, perhaps, a destination marker. As a comitative and as a coordinator in the short form *no*, only the 3SG prefix is used. In other uses of *-ora* 'join; be with', the inflectional paradigm is still intact.

3.5.3.2 The verb *maa* 'say'

The suffixless verb *maa* 'say' has grammaticalised into several directions as well. As a main verb, the verb *maa* 'say' is frequently used in stories to introduce direct speech as in (98).

- (98) *Da=maa: kame Ata Wuwur.*
 3PL=say 1PL.EXCL NAME
 'They said: We are Ata Wuwur (clan members).' (L3:367)

Further, the verb *maa* grammaticalised into a speech verb complementiser, a marker of future, intention, volition or ability and finally into a conditional conjunction.

As a complementiser, the verb *maa* 'say' appears as the second verb in a multi-verb sequence together with another speech verb as in (99), where the verb *kai* 'say' is the main verb and the verb *maa* functions as a complementiser.

- (99) *Kai-nga maa na=n-eti tange-ja.*
 say-3SG say 3SG=3SG-bring raw-PL
 'Tell her that she should bring the raw ones.' (C3:20)

As a marker of volition, future, intention or ability the verb *maa* also appears in a multi-verb sequence, but this time as the first element. An example

is given in (100), where the verb *maa* ‘say’ precedes the locational verb *ju* ‘downwards’ and the motion verb *-ai* ‘go’.

- (100) *Go maa ka=ju wosu k-ai.*
 1SG say 1SG=downwards outside 1SG-go
 ‘I want to go outside.’ (FH7:10)

As a conditional conjunction, *maa* appears clause-initially and optionally, it can be combined with the Indonesian conditional conjunction *kalau* ‘if’. Two examples are given in (101). See also §3.6.11 for more details on conditional clauses.

- (101) a. *Ara maa balar také, bo da=ganti d-era witi.*
 but say ivory.tusk not.exist DISC 3PL=replace 3PL-use goat
 ‘But if there is no ivory tusk, they replace it with a goat.’ (T1:124)
- b. *Kalau maa da=jua suka nermoi ...*
 if say 3PL=two like really
 ‘If they really like each other, ...’ (T1:4)

The grammaticalisation pathway of words meaning ‘say’ to quotation markers and complementisers is cross-linguistically and also regionally common (Klamer 2000).

3.5.4 The verbalising prefix *g-*

Central Lembata has a fossilised verbalising prefix *g-*. This prefix is only found on three verbs in my corpus which are listed in Table 3.42.

Table 3.42: Verbalisation with the prefix *g-*

Noun	Gloss	Verb	Gloss
<i>bliko</i>	‘wooden protector’	<i>gliku</i>	‘place a wooden protector’
<i>mekul</i>	‘part’	<i>gmekul</i>	‘split’
<i>mata</i>	‘eye’	<i>gmata</i>	‘speak face to face’

When the prefix is added to a consonant cluster initial word, the initial consonant is dropped and replaced by the prefix, such as in *bliko* ‘wooden protector’ → *gliku* ‘place a wooden protector’.

3.5.5 Summary

Central Lembata verbs can be divided into three subclasses based on their morpho-syntactic properties. G-suffix verbs can take pronominal suffixes of the G-suffix set which is a full set of pronominal suffixes that mark S and P arguments. K-suffix verbs are inflected with pronominal suffixes from the K-suffix set which is a more archaic set to mark S and P arguments. It is only found with a minority of frequently used verbal stems. Lastly, there is the class of suffixless verbs which cannot take any suffix. Most of the suffixless verbs are intransitive. This means that they mark their S argument in the same way as an A argument, namely using free pronouns or proclitics. The verb *-ora* ‘join; be with’ and the verb *maa* ‘say’ both show typical grammaticalisation pathways in Central Lembata. The bound stem *-ora* ‘join; be with’ can still be used as inflected for person as a predicate but in other contexts, its meaning has become more generic and is used as an existential, a comitative preposition or a coordinating conjunction. In those cases, the third person singular inflection is used as a default. The verb *maa* ‘say’ is used as a complementiser with other speech verbs, but also as a marker of intentions and conditions. Central Lembata shows a few examples of verbalisation using the prefix *g-*.

3.6 Basic clausal syntax

3.6.1 Simple verbal clauses

Central Lembata verbal clauses show SV order in intransitive clauses and AVP order in transitive clauses. S stands for an intransitive subject, A for a transitive subject, V stands for a verb and P stands for the most object-like argument of a transitive verb. Example (102) shows an intransitive clause.

- (102) *Kopo lamé tuné na=tobé-nga.*
 child\3 male\3 one 3SG=sit-3SG
 ‘A boy sits.’ (S2:23)

In (102), the intransitive verb *tobé* ‘sit’ is marked twice for the third person singular subject (S). First with a third person singular S/A proclitic *na=*, then with a third singular S/P suffix *-nga*. This double marking appears frequently

in the corpus. However, it is not obligatory. Either the proclitic, the suffix, or both, can be left out, as shown in (103), where the intransitive verb *tobé* ‘sit’ is only marked with the third singular proclitic *na=*.

- (103) *Grapu tu na=tobé.*
 woman\3 one 3SG=sit
 ‘A woman sits.’ (S6:7)

The use of the S/P suffix is most likely related to the aspectual properties of the predicate, i. e. state (suffix) versus event (no suffix), as discussed in §3.4.2.4.

In both examples of intransitive clauses given in (102) and (103) above, a full subject NP is present. This is not obligatory, a full subject NP can be either replaced by a full pronoun or can be left out completely.

The examples in (104) show transitive clauses. In (104a), the verb *tula* ‘make’ is marked for first person singular subject (A) employing the free pronoun *go* ‘1SG’. In (104a), the object (P) is not marked on the verb but rather appears as a bare noun. It is also possible to mark the object on the verb as in (104b) or to express it by a bound pronoun on the verb alone as in (104c). Here neither A or P are expressed in a noun or noun phrase, both are expressed by pronouns. The bound pronoun in (104b) and (104c) yields a stative meaning, thus here meaning completed.

- (104) a. *Go tula nowi.*
 1SG make sarong.male\3
 ‘I make a sarong (for males)’ (I2:4)
- b. *Go tula-nga nowi kéi.*
 1SG make-3SG sarong.male\3 PFV
 ‘I made the sarong already.’ (NB:88)
- c. *Go tula-nga kéi.*
 1SG make-3SG PFV
 ‘I made it.’ (NB:88)

Objects can also be optionally introduced by the comitative preposition *no* ‘COM’, such as in (105).

- (105) a. *Na=soga no kaju majak.*
 3SG=hold COM wood\3 dry.one
 ‘She holds a dried wooden stick.’ (FH3:7)

- b. *Na=soga kaju mekul.*
 3SG=hold wood\’s part.
 ‘She holds a piece of wood.’ (FH2:35)

The following verbs are examples of verbs that typically introduce their object with a comitative: *bawat* ‘bring’, *-eti* ‘bring’, *galew* ‘ask’, *-énu* ‘drink’, *bée* ‘give’, *tula* ‘make’, *perak* ‘see’, *soga* ‘hold’, *guti* ‘take’, *pusaj* ‘step rice (husk)’, *pékat* ‘exchange’, *témuk* ‘receive’, *pata* ‘boil’, *muul* ‘plant’, *-ewa* ‘catch’, *pétén* ‘miss; think of’.

Although semantically, some of these verbs, such as *bée* ‘give’ or *galew* ‘ask’, are typical two object verbs, it is not obligatory to overtly express the theme (T) and the recipient (R) in the same clause; it is possible that one of the objects is omitted and is interpreted from the context. Most of the time, the only object is T and it can optionally be preceded by *no(ra)*, such as illustrated in the two examples with the verb *galew* ‘ask’ in (106).

- (106) a. *Mo galew no né=ke’ada’an lau siné, Bosu.*
 2SG ask COM 3SG.POSS=condition seawards a.bit
 ‘Ask a bit about their conditions there overseas, Bosu.’ (Cl:104)
- b. *Mo tutu galew ke’ada’an tak jelas si.*
 2SG speak ask condition NEG clearly NEG
 ‘You are asking not clearly about the conditions.’ (Cl:104)

Although not frequently found, it is possible to express two objects, denoting a recipient (R) and a theme (T). These constructions are found with the verbs *bée* and *noto* both meaning ‘give’. R is expressed by a bound pronoun (107a), a free pronoun (107b), by a prepositional phrase using *lane* /*la* ‘LOC’ as in (107c) or by full NPs as in (107d). The latter construction with a direct full NP without preposition is less frequent. The constituent order of T and R depends on the formal properties of R. If R is expressed by a pronoun, it precedes T such as in (107a) and (107b), however if R is expressed by a prepositional phrase or a full NP it follows T, as in (107c) and (107d).

- (107) a. *Mo noto-ga glas kédak.*
 2SG give-1SG glass big
 ‘You gave me a huge glass’ (Cl:15)
- b. *Go noto nané na=muku tuné.*
 1SG give 3SG 3SG.POSS=banana\’s one
 ‘I give him a banana.’ (UM:12)

- c. *Na=bée buku la kopo lamé tu.*
 3SG=give book LOC child\ s male\ s one
 ‘She gives a book to a boy.’ (FH1:6)
- d. *Na=noto apél réné né=teman.*
 3SG=give apple PROX 3SG.POSS=friend
 ‘He gives this apple to his friend.’ (FH3:42)

In the theme NP in (107b) the 3SG.POSS proclitic is used although the translation of the NP is ‘a banana’ and not ‘his banana’. This is because the banana given to R is not his banana from the start. The possessor pronoun is used because it becomes his banana due to the process of giving. In Central Lembata theme NPs with a transfer verb such as *bée* ‘give’ frequently contain a possessor pronoun.

T is usually expressed by a simple NP as in the previous examples. It is also possible to express T with a prepositional phrase using the preposition *no* ‘COM’, while R is expressed by a bound pronoun on the verb, as illustrated in (108).

- (108) *Da=langsung bée-nga no bala.*
 3PL=immediately give-3SG COM tusk\ s
 ‘The immediately hand over the ivory tusk.’ (T1:141)

In sum, all arguments can be expressed by full NPs or pronouns with the possibility of omission when an argument is deductible from the context. Objects of monotransitive verbs, as well as T arguments of ditransitives, can be optionally preceded by the preposition *no(ra)* ‘COM’. R arguments are introduced by locationals, such as *la(ne)* ‘LOC’ or *réné* ‘PROX’, when they follow the T argument.

3.6.2 Temporal adverbs

Central Lembata neither marks tense on the verb nor uses any other kind of morphology for this purpose. Instead, the language possesses a set of adverbs to place an event in time. An overview of the temporal adverbs is given in Table 3.43. They are discussed in detail below.

Table 3.43: Temporal adverbs

Time frame	Temporal adverb	Translation
Past	<i>negero jema jua (bé(né))</i>	Lit. yesterday night two PROX Fr. 'in the past'
	<i>jema bé(né)</i>	Lit. night PROX Fr. 'in the past'
	<i>jema jua bé(né)</i>	Lit. night two PROX Fr. 'two days ago'
Immediate past	<i>negero</i>	'yesterday'
	<i>nakété</i>	'just now'
Present	<i>wé(né) / ré(né)</i>	Lit. PROX Fr. 'now'
Non-present	<i>nepo</i>	'later; just now'
Immediate future	<i>béné</i>	'later today'
	<i>bambé</i>	'tomorrow'
Future	<i>bambé jema jua</i>	Lit. tomorrow night two Fr. 'in the future'
	<i>wé jaé n-ai</i>	Lit. PROX hillwards 3SG-go Fr. 'in the future'
	<i>jema jua</i>	Lit. night two Fr. 'in two days'

Fr. = free translation, Lit. = literal translation

Without the use of any temporal adverb, a sentence in Central Lembata is unspecified concerning temporal information. To express the notion 'now', and thus convey present, it is possible to use one of the proximate demonstratives *wé(né)* or *ré(né)* (cf. §3.3.5), as shown in (109).

- (109) *Maa ka=tutu ré bo tentang Jon no Méri.*
 want 1SG=tell PROX DISC about John and Mary
 'I would like to tell now about John and Mary.' (N1:1)

As the form of the temporal expressions for 'now' and the proximate demonstratives are the same, it is not always clear if the speaker intends a tem-

poral or a deictic sense. It may also be questioned whether these two senses should be distinguished at all. The proximate demonstratives *wé(né)* and *ré(né)* can also be followed by words expressing parts of a day, such as *baén* 'morning', *lejo* 'day time' and *buuk* 'night'. This leads to the following collocations: *wé baén* 'now (morning)', *wé(né) lejo* 'now (midday)' and *wé(né) buuk* 'now (evening)'. The words expressing parts of the day, such as *baén* 'morning', are the only nouns in Central Lembata that can be preceded by a demonstrative. In all other cases, demonstratives follow the noun (cf. §3.3.4). In addition to proximate demonstratives, such as *wé(né)*, distal demonstratives can also be combined with parts of the day, such as in *wo(né) lejo* 'DIST day' meaning 'that day'.

Adverbs to express immediate past and immediate future are found in two pairs: *béné* / *nakété* and *bambé* / *negero*. The adverb *béné* 'later today' expresses a point in time which is in the future but still on the same day and the adverb *nakété* 'earlier; just now' expresses a point in time which is in the past but still on the same day or the evening before. The adverb *béné* is also a proximate locational but cannot be a demonstrative. These two immediate non-present expressions can be combined with the parts of the day in the same way as shown for the proximate demonstratives above. This leads to expressions like *nakété lejo* 'earlier at noon', *nakété buuk* 'earlier at night', *béné buuk* 'in the coming evening', *béné baén* 'in the coming morning' or 'tomorrow morning'. In these collocations of *nakété* / *béné* and parts of the day, *nakété* and *béné* can be shortened to *nak* and *bé*, thus for example *bé baén* for 'tomorrow morning' or *nak buuk* 'lat night'.

The other immediate non-present pair is *negero* 'yesterday' and *bambé* 'tomorrow'. The word *bambé* 'tomorrow' can also be combined with the parts of the day as in *bambé baén* 'tomorrow morning' (sometimes contracted to *bangbén*) or *bambé buuk* 'tomorrow evening'.

The remaining future and past expressions are all collocations composed of multiple words. The word *jema* historically means 'night' but synchronically only appears in these fixed expressions. The use of *bé(né)* 'PROX' in these collocations is surprising. As has been shown above, this word is also used as the immediate future adverb denoting 'in a bit; later today' but in collocations, it only appears in expressions pointing to past events. When referring to immediate future, *bé(né)* stands alone or precedes another word as in *béné buuk* 'later tonight', while in the collocations expressing a past point in time, it is the last element, such as in *jema jua bé* 'two days ago'.

The non-present marker *nepo* ‘later; just now’ either denotes a point in time in the immediate past as in (110) or in the future as in (111). There is a crucial syntactic difference between the past and the future reading of *nepo*. To obtain the past reading, the adverb *nepo* has to directly precede the main verb, such as given in (110), no pronoun can interfere.

- (110) *Na=nepo oli.*
 3SG=NON.PRES arrive
 ‘He just arrived.’ (C5:83)

For the future reading however, a pronoun appears between the adverb *nepo* and the main verb as is the case in (111) and (112).

- (111) *Nepo na=supeng moné.*
 NON.PRES 3SG=pick.up 2SG
 ‘Then / Later she will pick you up.’ (C1:249)

The future use of *nepo* leads to the use of a proclitic, such as *ka*= ‘1SG’ in (112), typically found in irrealis contexts (cf. §3.4.2.2).

- (112) *Na=panaw-a nepo ka=guté-Ø.*
 3SG=walk-3SG NON.PRES 1SG=take-3PL
 ‘When he has left, I will take it.’ (C5:103)

In (112), *nepo* can be analysed as a sequential conjunction (cf. §3.6.11) as it connects two events in a sentence. It introduces a clause expressing an event that is going to take place after another event has happened, thus potentially in the future.

3.6.3 Aspectual adverbs

To specify the internal temporal properties of an event, Central Lembata mainly employs the aspectual adverbs given in Table 3.44. Two semantically opposing pairs can be found. First, the use of *kéi* ‘PFV’ or *nawa* ‘still’ defines whether something has taken place already or not yet. Second, the use of *naro* ‘COMPL’ or *kia* ‘INCEP’ defines whether something is beginning or terminating. The progressive marking *diro* ‘PROG’ does not have a clear counterpart.

Table 3.44: Aspectual adverbs

Adverb	Meaning	Aspect	Position
<i>kéi</i>	‘already’	perfective (PFV)	clause-final
<i>nawa</i>	‘still’	imperfective (still)	pre-predicate
<i>naro</i>	‘finished’	completive (COMPL)	clause-final
<i>kia</i>	‘first’	inceptive (INCEP)	clause-final
<i>diro</i>	‘in progress’	progressive (PROG)	pre-verbal

The pre-verbal adverb *diro* ‘PROG’ is used to mark an event as progressive, as illustrated in (113).

- (113) *Kopo biné jua da=diro tobé.*
 child female two 3PL=PROG sit
 ‘Two girls are sitting.’ (FH2:11)

The progressive adverb *diro* is placed directly preceding the main verb and the proclitic pronoun is attached to this aspect marker.

The clause-final *kéi* ‘PFV’, as in (114a), and the pre-predicate *nawa* ‘still’, as in (114b), express perfective (PFV) and imperfective (still) aspect.

- (114) a. *Go=kopo biné lodo kéi.*
 1SG.POSS=child female exit PFV
 ‘My daughter left home already.’ (=is married) (M3:157)
- b. *Awa-ga nawa na=lau Ilé Apé.*
 spouse-1SG.POSS still 3SG=seawards NAME
 ‘My wife is still at Ile Ape.’ (C1:189)

A clause containing *kéi* ‘PFV’ cannot be negated, as shown in (115a), whereas for a clause with *nawa* ‘still’, negation is possible and yields the meaning ‘not yet’ as in (115b).

- (115) a. * *Go=kopo biné ta na=lodo kéi.*
 1SG=child female NEG 3SG=exit PFV
 Intended: ‘My daughter has not left home yet.’ (=not yet married) (NB:87)

- b. *Bapa nawa ta na=lebo si.*
 father still NEG 3SG=bathe NEG
 'Father has not taken a shower yet.' (UM:33)

Usually, the imperfective marker *nawa* precedes the predicate and the perfective marker *kéi* follows it. However, in some cases, *nawa* 'still' can also appear clause-finally as in (116).

- (116) *Reno anak nawa.*
 NAME little still
 'Reno is still little.' (M3:181)

The adverb *nawa* is of verbal origin. It seems to originate from a vowel-initial verb that is inflected for S/A person as the third person plural form *d-awa* '3PL-still' is still frequently found. However, all other persons appear to have generalised to *nawa* which contains a fossilised 3SG prefix. S/P person suffixes of the more archaic K-suffix set (cf. §3.4.2.4) can be added to the stem *nawa* 'still', such as *-ku* '2SG' in (117).

- (117) *Mo gerep nawa-ku.*
 2SG young.woman still-2SG
 'Are you still single?' (UM:98)

It is not entirely clear what the meaning of a verb *-awa* would have been. Possible meanings are 'be incomplete' or 'be ongoing'. Verbal traces such as these cannot be found for *kéi* 'PFV' which always appears clause-finally, which is not a typical position for verbs in Central Lembata.

Compared to the aspectual adverbs *kéi* and *nawa*, the clause-final adverb *naro* 'COMPL' is less frequent but it also occurs in contexts that express an event that is completed or has terminated, such as in (118).

- (118) *Tité ékan we bo slaén-a naro.*
 1PL.INCL.POSS garden\I PROX DISC clean-3SG COMPL
 'Our garden is now cleared.' (N3:62)

The adverb *naro* has a similar function to the perfective adverb *kéi* discussed above. Both express a terminated action.

The adverb *kia* 'INCEP' expresses an event which has to be done or happen first before another event can start as in example (119). This adverb is also often used in imperatives as in (120). I analyse the adverb *kia* as an inceptive (INCEP) as it is related to the starting of an event.

- (119) *Mo bantu ema-m maa pus piri kia.*
 2SG help mother-2SG.POSS want wash dish INCEP
 ‘Help your mother first to wash the dishes.’ (N2:16)
- (120) *Gaa uta kia.*
 eat peanuts INCEP
 ‘Eat peanuts now.’ (Cl:202)

In discourse, the two aspectual adverbs *naro* ‘finished’ and *kia* ‘INCEP’ are often combined to the expression *naro kia* meaning ‘that’s enough for now’, ‘let’s stop now’ or the like.

3.6.4 Negation and prohibition

Central Lembata has six negating lexemes which are summarised in Table 3.45. The non-existential *také* ‘not.exist’ is a predicate, while the other negators negate a predicate.

Table 3.45: Negators in Central Lembata

Function	Subject		Predicate	
Non-existential	...		<i>také</i>	
Negation of N / V predicates	...	<i>tak / ta</i>	...	<i>si(né)</i>
Prohibition of predicates	...	<i>aké</i>	...	<i>baé</i>

The negator *také* ‘not.exist’ only appears in isolation as the answer “no” to a question or as a non-existential (cf. §3.6.8). The three negators *tak*, *ta* and *si(né)* express negation and the two negator *aké* and *baé* express prohibition. The negators *si(né)* and *baé* are clause-final, the others appear before the predicate. The clause-final negators *si* and *siné* are in free variation.

There are three double negation patterns in Central Lembata: 1) *ta ... si(né)* to negate a verbal predicate (121), 2) *tak ... si(né)* to negate a non-verbal predicate (122) and 3) *aké ... baé* to express prohibition (123).

- (121) *Ta da=waak-i si.*
 NEG 3PL=finish-3PL NEG
 ‘It (the food) will never be finished.’ (T3:242)

- (122) *Tak tulor-u siné.*
 NEG frog-SPEC NEG
 'It's not the frog.' (F1:25)
- (123) *Aké péten baé.*
 PROH think PROH
 'Don't worry.' (N3:69)

Although most frequently found in a double pattern, it is possible to leave out one of the negators and thus have single pre-predicate or single clause-final negation.

As Austronesian languages generally show pre-predicate negation, the presence of a clause-final negator in Central Lembata is regarded as an innovation. This hypothesis is discussed in detail in §10.3.

3.6.5 Multi-verb sequences

Multi-verb sequences are verbal clauses that contain two or more verbs that share at least one argument. The verbs can be separated by pronouns or nominal objects but not by conjoining elements. Multi-verb sequences are frequently found in Central Lembata. A full analysis of these constructions that includes the question of which of the Central Lembata verbal sequences qualify formally as serial verb constructions in the sense of Aikhenvald and Dixon (2006) remains to be done.

Central Lembata has the types of multi-verb sequences listed below. V1 stands for the first verb in the sequence, while V2 stands for the second verb in the sequence. There are also sequences with more than three verbs but they are not discussed in detail here.

- (124) Types of multi-verb sequences
1. V1 + V2 = Two events
 - a) simultaneous
 - b) sequential
 2. V1 + V2 = One single event
 - a) two aspects of the same event
 - b) new meaning

- c) two verbs with very similar meanings
- d) V2 locational and deictic motion verb
- e) V2 grammaticalised

Multiverb sequences are divided into two main classes, those that denote two events and those that denote one single event. A multi-verb sequence that expresses two events can express two or more events that are happening at the same time as in (125a) or shortly after each other as in (125b).

- (125) a. *Nepo ta=seli ptuto-sa tali wi.*
 later 1PL.INCL=meet talk-1PL.INCL again DISC
 ‘Later we will meet and talk again.’ (C1:247)
- b. *Da=ola-nga tube-nga pa da=kaa.*
 3PL=cut-3SG cook.in.bamboo-3SG then 3PL=eat:3PL
 ‘They cut it, cooked it in bamboo and then they ate.’ (L3:110)

Multiverb sequences that denote one single event make use of the two verbs in very different ways as listed above. In the following, I exemplify all five subtypes that express one single event.

Motion verbs can occur in multi-verb sequences to express different aspects of one event, such as the way of moving and the direction of the movement. An example is given in (126). There the first verb *ksopel* ‘jump’ expresses the way the dog moves and the second verb *lodor* ‘descend’ expresses the direction of movement.

- (126) *Aor ksopel lodor-a.*
 dog jump descend-3SG
 ‘The dog jumps down.’ (F2:49)

Locational verbs, such as *lau* ‘seawards’ that indicate a location in seaward direction (cf. §3.7.1), can also be part of multi-verb sequences, as illustrated in (127). The locational verb alone does not specify yet that any motion is involved. The first verb *pana* ‘walk’ indicates the way of movement and through that, it becomes clear that the sentence expresses motion. These constructions are discussed in more detail in §3.7.2 on locative expressions.

- (127) *Da=jua pana da=lau éka keja.*
 3PL=two walk 3PL=seawards garden middle
 ‘The two of them went to the middle of the garden.’ (N3:93)

Other aspects of the movement can also be expressed by the second verb in a multi-verb sequence. In the following two examples, the second verb indicates whether the agent of the movement is being followed (128a), or the agent of the movement is following another agent (128b). The first verb *kari* ‘run’ in (128a), expresses the way of motion, while the second verb *-olu* ‘precede’ specifies the semantic aspect of preceding another entity. Whereas in (128b), the first verb *beka* ‘fly’ specifies the way of movement, while the second verb *dori* ‘follow’ expresses that the agent of the multi-verb sequence, the owl, is following someone else, in this case the child.

- (128) a. *Kopong nakété di kari n-olu kéi.*
 child ANAPH also run 3SG-precede PFV
 ‘That child is running ahead already.’ (F2:27)
- b. *Na=beka dori-nga kopo nakété.*
 3SG=fly follow-nga child ANAPH
 ‘It [the owl] flies following the child.’ (F2:36)

In addition to motion events, there are also other events that can be expressed by a multi-verb sequence whose members specify different aspects or parts of the same event. These can be resultative constructions or constructions where the two verbs have related meanings but the second is more specific than the first. These cases are shown in (129) and (130). In the multi-verb sequence in (129), the first verb *pegu* ‘cut’ expresses the action, while the second verb *-era* ‘become’ introduces the result of the action. Both verbs have the same agent subject which is expressed by the free pronoun *go* ‘1SG’ and by the inflection on the second verb which is a vowel-initial prefixing verb (cf. §3.4.2.3).

- (129) *Go pegu k-era mekul jua.*
 1SG cut 1SG-become part two
 ‘I cut it into two pieces.’ (FH7:16)

In (130), the speech verb *tutu* ‘tell’ expresses that the person is speaking, while the second verb *galew* ‘ask’ specifies that the person is not only speaking but asking about something.

- (130) *Mo tutu galew keadaan.*
 2SG tell ask condition
 ‘You are asking about the circumstances.’ (C1:104)

The next subtype of single event multi-verb sequences is the case where two verbs are combined to yield a new meaning. One could also analyse them as verbal compounds. Examples are *bélu bakat* glossed as ‘cut cut’ meaning ‘murder’, *kari pana* ‘run walk’ meaning ‘flee’ and *gelé tobé* ‘lie sit’ meaning ‘live’ or ‘stay’.

A specific form of multi-verb sequences are those that combine verbs which are synonyms. Examples are *diki dongot* ‘carry’, *gawak kpulé* ‘hug’ and *tobé bau* ‘sit’. In all three examples, the individual parts mean the same as both parts together, so no new meaning is generated. This is different from the case of verbal compounds, given in the preceding paragraph.

Finally, there are multi-verb sequences that have a second element which is highly grammaticalised and could be analysed as a preposition. In that case, it is questionable whether synchronically, these constructions should still be seen as multi-verb sequences. But historically they have their origin in multi-verb sequences. The second verb of a grammaticalised multi-verb sequences can be a grammaticalised form of *-ora* ‘be with’ or *maa* ‘say’ which are discussed in §3.5.3, a locational verb which are introduced in §3.7.1 or an aspectual markers of verbal origin, discussed in §3.6.3.

3.6.6 Nominal predicates

Nominal predicates usually contain a noun phrase with only a noun and no other elements. In case of coda alternating nouns, the free form is found (cf. §3.3.1.2). Nominal predicates occur mainly in equative clauses. A subtype of these are clauses containing a noun that expresses a property concept in predicate position. The subject NP in these clauses is usually more complex than the predicate NP. In some cases, the nominal predicate can be marked with a S/P suffix but most nominal predicates are unmarked. For examples and more details on predicative nouns with suffixes see §3.4.2.4.

Equative clauses are built by juxtaposing two NPs; there is no copula or other linking element. Two examples of equative clauses are given in (132) and (131). Noun phrase boundaries are marked with brackets. In both examples, two NPs are combined. The first NP is more complex, whereas the second NP is simply a noun. In (131), the first NP contains an inalienable possessed noun. In (132), the first NP contains the noun *magun* ‘man’, the demonstrative *wo* ‘DIST’ and the discourse marker *bo* ‘DISC’.

- (131) [*Bapa-m*] [*pegawén*].
 father-2SG.POSS official
 ‘Is your father a government official?’ (CI:145)

- (132) [*Magun wo bo*] [*molán*].
 man DIST DISC healer
 ‘That man is a healer.’ (NT:12)

In predicate position, property concepts are expressed by unmarked nouns, as in (133), where the property noun *kédak* ‘big.one’ is not marked for number. Property nouns in predicate position cannot take any suffix, thus are never marked for number.

- (133) [*Una-ja wo*] [*kédak*].
 house-PL DIST big.one
 ‘Those houses are big.’ (NB:74)

Equative clauses are differentiated from attributive property nouns (§3.3.4.2) by the fact that the subject NP contains at least one second element next to the noun that marks the end of the phrase. Examples of eligible elements include a demonstrative, such as *wo* ‘DIST’, as in (133), or the end of phrase marker *ro* ‘FIN’.

The subject of nominal predicates can also be expressed by a full pronoun, as in (134).

- (134) [*Gone ro*] [*golok*].
 1SG FIN tall.one
 ‘I am tall.’ (NB:80)

On the other hand, a short form of a pronoun *go* ‘1SG’ cannot be the subject in this construction shown in (135).

- (135) * [*go*] [*golok*].
 1SG tall.one
 Intended: ‘I am tall.’ (NB:65)

The short form is only possible if the nominal predicate takes a S/P marking suffix as discussed in §3.4.2.4. An example is (69b) in that section.

3.6.7 Adjectival root predicates

Central Lembata has a set of adjectival roots which can be derived into nouns, such as given in Table 3.46. For the derivational morphology involved see §3.3.6.

Table 3.46: Adjectival roots and their nominal derivatives

Gloss	Adjective	Noun
‘big’	<i>kéda</i>	<i>kédak</i>
‘long (time)’	<i>koda</i>	<i>kodak</i>
‘dirty’	<i>mila</i>	<i>milan</i>
‘quiet’	<i>meli</i>	<i>meling</i>
‘bitter’	<i>pait</i>	<i>pnait</i>
‘heavy’	<i>berat</i>	<i>bnerat</i>
‘hot’	<i>kati</i>	<i>knating</i>
‘rotten’	<i>gewa</i>	<i>knewak</i>
‘salty’	<i>peju</i>	<i>pnejuk</i>
‘sour’	<i>gilu</i>	<i>kniluk</i>

This section discusses the use of the adjectival roots. Their nominal derivatives are discussed as nominal predicates in §3.6.6 and as modifiers within NPs in §3.3.4. The only syntactic context where the adjectival roots are used is intensification or negation, using one of the modifying means listed in Table 3.47. An example of an adjectival root in use is *buja* ‘white’ in (136) which is modified by *kéda* ‘big’.

- (136) *Kumi kéda buja.*
 mustache big white
 ‘The mustache was very white.’ (L3:271)

Table 3.47: Modifiers for adjectival predicates

Modifying means	Meaning
<i>ro</i>	‘very’
<i>kéda</i>	‘big’
<i>(tu)ga</i>	‘extremely’
<i>di</i>	‘more’
<i>diro</i>	‘PROG’
<i>kéi</i>	‘PFV’
<i>ta ... si(né)</i>	‘NEG’
full reduplication	‘very very’

In all other contexts, property concepts are expressed as nouns, which are either nominalisations of their adjectival counterparts, such as the nouns in Table 3.13 or they are underived nouns, such as *letes* ‘cold.one’ which does not have an adjectival root.

Property nouns that do not have an adjectival root can also be combined with the modifiers listed above in a nominal predicate, as for example *letes* ‘cold’ in (137).

- (137) *Lau Kalikasa éka buuk kéda letes.*
 seawards NAME condition night very cold
 ‘In Kalikasa, the nights are very cold.’ (NT:1)

Adjectival root predicates are different from nominal predicates. The subject of a nominal predicate, such as *golok* in *goné ro golok* ‘1SG FIN tall.one’ = ‘I am tall.’, always needs a phrase-final element, such as *wo* ‘DIST’ or *ro* ‘FIN’ (cf. §3.6.6). Otherwise, the predicative noun could not always be distinguished from a attributive property noun, such as in a noun phrase as *ata golok* ‘person tall.one’ = ‘a tall person’.

This is not the case for adjectival root predicates, which are in contrast very similar to verbal predicates. They can take short pronouns or proclitics (138) to mark their subjects which is not possible for nominal predicates.

- (138) *Mo diro golo.*
 2SG PROG tall
 ‘You are getting taller.’ (NB:76)

However, I do not analyse adjectival roots as verbs, because they obligatorily appear in intensified or negated contexts, while verbs are only optionally modified by the same modifiers, such as *diro* ‘PROG’ or *ta ... si(né)* ‘tak’.

3.6.8 Existential clauses

Central Lembata has two ways to construct an existential clause, either with the intransitive verb *nong* ‘exist’ or with the existential *nora/no* that has grammaticalised from the verb *-ora* ‘join; be with’ (cf. §3.5.3). The intransitive existential verb *nong* ‘exist’ cannot be inflected. It can either be used with an intransitive subject as in (139) or in isolation as an answer to a question as in (140).

- (139) *Kopi bo nong ara kopi pnait.*
 coffee DISC exist but coffee bitter
 ‘There was (or: they had) coffee but the coffee was bitter.’ (M3:84)

The existential *nong* ‘exist’ also has a negative counterpart which is the intransitive verb *také* ‘not exist’ as in the question in example (140).

- (140) Q: *No mo=kléréng také?*
 exist 2SG.POSS marble not.exist
 ‘Do you have marbles?’
 A: *Nong.*
 exist
 ‘Yes, I do.’ (C2:14-15)

The existential *nong* and its negative counterpart *také* are syntactically distributed in the same way. They can stand in isolation or they can be the intransitive predicate of a subject argument.

The transitive existential is attested in two forms, *nora* and *no*. The existential *nora* can appear with a transitive subject, as in (141). Any coda alternating noun, as in (142), appears in the short form as is the rule for post-verbal object.

- (141) *Guru pegawén nora da=gula.*
 teacher official exist 3PL.POSS=sugar
 ‘The public teachers had their sugar.’ (M3:88)

In (142), the coda alternating noun *rusa/rusar* ‘deer’ appears in the short form when following *nora* ‘exist’. This provides evidence that this NP is an object.

- (142) *Weli watu paap ro bo nora rusa.*
 sideways stone side FIN DISC exist deer\S
 ‘On the other side of the stone, there is a deer.’ (F2:39)

Although taking an object argument, the existential *nora* ‘exist’ is not completely verbal. When negated, it follows the pattern for non-verbal elements (cf. §3.6.4) by using *tak* ‘NEG’ as in (143). It might be that here the noun *wai* ‘water\S’ is negated rather than *nora* ‘exist’ and this is triggering the negation for non-verbal elements.

- (143) *Tak nora wai si.*
 NEG exist water\S NEG
 ‘There is no water.’ (NB:86)

The shorter form *no* has even fewer verbal properties. It cannot have a subject and for coda alternating nouns, the nominal element following the existential *no* does not appear in the vowel-final form as a verbal object would, but in the consonant-final form, as in *apér* ‘fire\L’ in (144) and *waér* ‘water\L’ in (145).

- (144) *Tak no apér si.*
 NEG exist fire\L NEG
 ‘There was no fire.’ (N3:16)

- (145) *No waér na=weli.*
 exist water\L 3SG=sideways
 ‘There is water there.’ (L3:203)

3.6.9 Questions

Questions in Central Lembata have the same word order as declarative sentences. Polar questions are usually unmarked but they can be marked optionally by an interrogative particle such as those listed in Table 3.48.

Table 3.48: Interrogative particles

Type	Particle	Gloss
Native	<i>i</i>	‘Q’
	<i>bo</i>	‘DISC’
Loan (Indonesian)	<i>to</i>	‘Q’
	<i>ka</i>	‘Q’
Negator	<i>take</i>	‘not.exist’

The particle *bo* is not glossed with ‘Q’ for question marker as it can also occur in declarative sentences (cf. §3.6.10). The interrogative particles listed in Table 3.48 appear clause-finally as in (146), except for *bo* ‘DISC’ which is mainly found clause-initially in questions, as in (147).

- (146) *Buuk ua telu i?*
 night CLF three Q
 ‘That’s for three nights, right?’ (C6:22)

- (147) *Bo Jérman né di gésak tali?*
 DISC German 3SG.POSS also different again
 ‘And in German, it is different again?’ (C1:35)

The interrogative particle *take* — originally a non-existential — is often combined with other interrogative markers such as *i* in (148).

- (148) *No mo=kléréng i take?*
 exist 2SG.POSS=marble Q not.exist
 ‘Do you have marbles?’ (C2:14)

Information questions in Central Lembata contain an interrogative word. An overview of the question words is provided in Table 3.49.

Table 3.49: Interrogative words

Simple	Gloss	Word class
<i>ga(né)</i>	‘where’	locational verb
<i>énak / éna</i>	‘who’	coda-alternating(?) noun
<i>aan</i>	‘what; which’	inalienable(?) noun
<i>pira</i>	‘how many’	numeral
Compound	Gloss	Components
<i>nao ga(né)</i>	‘how’	‘be like’ + ‘where’
<i>jema pira</i>	‘when’	‘time unit’ + ‘how many’
<i>bo aan</i>	‘why’	‘DISC’ + ‘what’
<i>(bo) puken aan</i>	‘why’	(‘DISC’) + reason + ‘what’
<i>bi aanenu</i>	‘why’	‘not know’ + ‘what-3SG.POSS’

Question words in Central Lembata do not form a word class on their own. Simple question words belong to the word class of their expected answer. The word to ask for a place *ga(né)* ‘where’ is a locational verb. The word *énak/éna* ‘who’ to ask for a person and the word *aan* ‘what’ to ask for a thing are nouns. And the word *pira* ‘how many’ to ask for an amount is a numeral. This classification is based on the morpho-syntactic properties of the question words that match with those of the word classes they are assigned to.

In (149), the question word *ga* ‘where’ is placed between the pronoun *mo* ‘2SG’ and the clause-final verb *oli* ‘arrive’.

- (149) *Weki-m snebar wo bo mo ga oli?*
body-2SG.POSS wet DIST DISC 2SG where arrive
‘Your body is wet, where are you coming from?’ (L3:202)

This position between pronoun and clause-final verb is the same position found for locationals such as *wé(né)* ‘PROX’ or *wo(né)* ‘DIST’. Another property that the question word *ga(né)* ‘where’ has in common with locationals is the presence of a long and a short form which are synonyms created by the presence or absence of the final syllable *né* (cf. §3.7.1).

The question words *énak* ‘who’ and *aan* ‘what’ appear in core argument positions in the same way as nouns. In (150), the question word *énak* ‘who’

appears in subject position, whereas *aan* ‘what’, in (151), appears in object position.

- (150) *Énak era weli?*
 who cry sideways
 ‘Who is crying over there?’ (C2:70)
- (151) *Mio tula aan tali?*
 2PL do what again
 ‘What else did you do?’ (C6:207)

Apart from their placement in core argument positions, the question words *énak* ‘who’ and *aan* ‘what’ also show morphological properties of nouns. The question word *énak* ‘who’ can take a specificity suffix *-u* ‘SPEC’ (152) as is the case with alienable nouns (cf. §3.3.1.2).

- (152) *Bo mo najan-mu énak-u?*
 DISC 2SG name-2SG.POSS who-SPEC
 ‘And what is your name?’ (UM:57)

The question word *aan* ‘what’ can take an inalienable possessor suffix (153) as is the case with inalienable nouns (cf. §3.3.1.1).

- (153) *Manuk aan(e)-n?*
 chicken what-3SG.POSS
 ‘What kind of chicken?’ (NB:38)

Finally, the question word *pira* ‘how many’ behaves like a numeral as it can be used with a classifier as in (154) and it can take the ordinal number forming circumfix *kes...n* as shown in (155).

- (154) *Go=kléré ua pira bla?*
 1SG.POSS=marble CLF how.many PROX
 ‘How many marbles do I have here?’ (C2:64)
- (155) *Kopo kes-pira-n?*
 child ORD-how.many-ORD
 ‘Which child is it? (first, second..?)’ (C6:93)

3.6.10 Clause and phrase boundary markers

Central Lembata has three discourse particles that appear either at the end or beginning of a clause or at the end of a noun phrase (cf. §3.3.4). Their forms and possible positions are given in Table 3.50.

Table 3.50: Clause and phrase boundary markers

Form	Gloss	Clause-initial	Clause-final	NP-final
<i>bo</i>	‘DISC’	yes	yes	yes
<i>ro</i>	‘FIN’	no	yes	yes
<i>di</i>	‘also’	no	yes	yes

From this table, it can be seen that the particle *bo* can appear in all three positions, whereas the other two particles *ro* and *di* are only found at the end of a clause or a NP. The origin of these particles is not entirely clear but the following are some possible language-internal sources. The particles *bo* and *ro* might go back to the distal locationals *wo(né)* and *ro(né)* (cf. §3.5.2.3). On the other hand, the particle *di* probably goes back to *di* meaning ‘also; additional; more’.

For now, *ro* is glossed as ‘FIN’ meaning final marker because it marks the end of a clause or NP, and *bo* is glossed as ‘DISC’ meaning discourse marker as it has wider discourse functions. The phrase final particle *di* is glossed ‘also’ as this is the primary meaning of the word.

The particles *ro* and *bo* are extremely frequent. In final position of a NP or a clause, they can be combined as *ro bo*. The particle *di* is more marginal. In phrase initial position, *bo* very often appears in questions as in example (156). It is also part of the question word *bo aan* ‘why’ which is composed out of the particle *bo* and the question word *aan* ‘what’ (cf. §3.6.9).

- (156) *Bo ta=muul aan?*
 DISC 1PL.INCL=plant what
 ‘What are we going to plant?’ (N4:34)

In clause- or NP-final position, *bo* as well as *ro* seem to emphasise the phrase

boundary.¹⁴ This can be between a subject and a non-verbal predicate as in (157) or a subject and a verbal predicated as in (158) and (159).

- (157) *Maguja wo bo mola-ja.*
 man.PL DIST DISC healer-PL
 ‘Those men are healers.’ (NT:12)
- (158) *Kamé wé bo gaa kwaru netuk.*
 1PL.EXCL PROX DISC eat:1PL.EXCL corn crushed
 ‘We eat crushed corn here.’ (Cl:6)
- (159) *Kopong ro géwé na=wo.*
 child FIN ascend 3SG=DIST
 ‘The child climbed up there.’ (F2:29)

But both final markers, *bo* and *ro*, can also appear at the end of a clause, thus at the end of a VP, as in (160) and (161).

- (160) *Kuun duang réné toba-nga ro.*
 ficus.species big PROX fall-3SG FIN
 ‘This big ficus tree fell down.’ (C6:60)
- (161) *Kreka bo na=oli bo.*
 startle DISC 3SG=come DISC
 ‘Suddenly, he came.’ (C6:146)

In the case of clause-final *bo*, there is often a connotation of surprise or questioning in the sentence. However, *bo* as a non-final elements is much more frequent.

In final position, the combination of *ro* and a plain verb without pronoun or proclitic preceding it, as in (162), yields an imperative meaning.

- (162) *T-énu ro.*
 1PL.INCL-drink FIN
 ‘Let’s drink.’ (Cl:11)

¹⁴ At first glance, the particle *ro* appears to be similar to the Lewoingu Lamaholot particle *-ke* which Nishiyama and Kelen (2007:129) describe as a focus particle. However, when trying to reproduce the Lewoingu sentences containing *-ké* in the Lamaholot variety of Central Lembata, it can only be replaced by Central Lembata *ro* where Lewoingu *-ké* appears at the end of a determiner phrase.

3.6.11 Clause combining

In Central Lembata coordination is much more frequent than subordination. Subordination can be found for relative clauses using a non-native relative marker discussed in §3.3.4.5, or with the complementiser *ma* discussed in §3.5.3. Table 3.51 lists clause-initial conjunctions for clause coordination.

Table 3.51: Conjunctions

Type	Conjunction	Meaning
Sequential	<i>pa / waak pa</i>	‘then’
Sequential	<i>nepo</i>	‘then’
Contrasting	<i>ara</i>	‘but’
Causal	<i>puken</i>	‘because’
Consequential	<i>ké</i>	‘so’
Purpose	<i>wé</i>	‘so that’
Conditional / Temporal	<i>maa</i>	‘if; when’

To combine two clauses that express a sequence of two events, the conjunction *pa*, *waak pa* or *nepo* can be used. The conjunction *nepo* ‘then’ is a grammaticalisation of the temporal adverb *nepo* that indicates a point in the immediate past (cf. §3.6.2).

Frequently used Indonesian conjunctions in Central Lembata are Indonesian *kalau* ‘if’ and Indonesian *karena* ‘because’. The Indonesian conjunction *kalau* ‘if’ is often directly followed by the native conjunction *maa* ‘if; when’. The conjunction *karena* ‘because’ is never combined with the native conjunction *puken* ‘because’.

In Central Lembata conditional clauses can be expressed using the conjunction *maa* which is grammaticalised from the verb *maa* ‘say’. The Indonesian conjunction *kalau* ‘if’ is also used in conditional clauses, either on its own or combined with *maa*. Table 3.52 shows the amount of instances of both markers, the native *maa* and the Indonesian *kalau*, in my corpus.

Table 3.52: Conditional markers in Central Lembata

Conditional marker	Occurrences in corpus
<i>maa</i>	49
<i>kalau maa</i>	21
<i>kalau</i>	12

From the table, it can be seen that the native construction using *maa* with 49 occurrences in my corpus, such as in (163), is more frequent than the use of the Indonesian conditional conjunction *kalau* which appears only in 12 instances on its own, such as in (164), and in 21 instances combined with *maa*, such as in (165). However, the usage of the Indonesian loan is high enough not to be neglected when discussing conditional constructions in Central Lembata.

- (163) *Nakété maa ta ka=lega kaju si bo, ta ta=onik api*
 ANAPH if NEG 1SG=split wood NEG DISC NEG 1PL.INCL=set fire
si.
 NEG
 ‘If, just now, I did not split wood, we would not make fire [now].’
 (N1:18)
- (164) *Karena, kalau go mea-ka en bo, go sibuk ré bé lewu.*
 because if 1SG self-1SG just DISC 1SG busy PROX PROX village
 ‘Because, if I am on my own, I will be busy here in the village, ...’
 (C1:195)
- (165) *Kalau maa da=bukut-i bo ro, bo kam=dapat no*
 if if 3PL=protect-3PL DISC FIN DISC 1PL.EXCL=get COM
tapu mirek wai.
 coconut candlenut some
 ‘If they protected them [the young plants], they got some coconut
 and candlenut trees.’ (M3:136)

Apart from the conditional conjunctions *kalau* and *maa*, the phrase boundary marker *bo* is an important part of conditional clauses. This marker is found in between the two parts of the sentence. It introduces the clause stating the consequence that results from the condition expressed in the first clause.

3.6.12 Summary

Central Lembata has SV order in intransitive sentences and SVP order in transitive sentences. Arguments can be expressed by NPs, pronominals or both. A recipient argument in a ditransitive can be expressed by a pronominal, a full direct object NP or a prepositional phrase. Central Lembata employs temporal and aspectual adverbs. Predicates are negated by double negation. A range of multi-verb sequences are attested in Central Lembata. Nominal predicates are juxtaposed to the subject NP. Adjectival roots can only be used in intensified or negated contexts. In all other cases, a (derived) property noun is employed. The existential *nong* and the non-existential *také* are predicates. *Nong* has grammaticalised into *no* and lost its verbal qualities. Polar questions are optionally marked by an interrogative particle. Information questions contain question words that belong to the word class of their expected answer. Central Lembata has a few discourse markers that occur at clause or phrase boundaries. A range of conjunctions is attested, of which some can be combined with Indonesian loans that have the same function.

3.7 Location

Location is an important feature in the grammar of Central Lembata. This section discusses locationals, their semantics and their functions. Locationals are a subtype of suffixless verbs (cf. §3.5.2.3) which specify features such as distance and direction from the deictic centre (§3.7.1). In the following sections, I use the terms ‘locationals’ or ‘locational verbs’ as synonyms. A subset of locationals also function as demonstratives. Locational verb phrases can contain a complex object NP, as well as a phrase-final deictic motion verb that specifies the direction of movement away or towards the deictic centre (§3.7.2).

3.7.1 Locationals

Syntactically, locationals or locational verbs appear in predicate position. They share their syntactic features with suffixless verbs as they usually appear with a pro-clitic pronoun but never take a suffix. They can be intransitive, as in (166), or take a nominal object, as in (167).

- (166) *Kopongu na=bété.*
 child.SPEC 3SG=DIST.VIS
 ‘Is the child there?’ (NB:2)

- (167) *Watoru na=ju una.*
 stone.SPEC 3SG=downwards house\s
 ‘The stone is at the house (down there).’ (NB:3)

In (167), the coda alternating noun *una* / *unan* ‘house’ appears in the short form. This is the form which is obligatorily used for object nouns in the VP (cf. §3.3.1.2).

Semantically, locational verbs in Central Lembata express distance, direction, elevation and, to a limited extent, visibility from the deictic centre. Table 3.53 gives an overview of locational terms in Central Lembata.

Table 3.53: Locational verbs in Central Lembata

Semantic distinctions	Gloss	Basic form	<i>b</i> -initial
Default	LOC	<i>la(né)</i>	
	PROX	<i>we(né)</i> <i>re(né)</i>	<i>bla(né)</i> <i>bé(né)</i>
Distance	DIST	<i>wo(né)</i> <i>ro(né)^a</i>	
	DIST.VIS	<i>bété</i>	
Vertical direction	upwards	<i>jé(né)</i>	<i>bjé(né)</i>
	downwards	<i>ju(né)</i>	<i>bju(né)</i>
Horizontal direction	hillwards	<i>jaé</i>	<i>bjaé</i>
	seawards	<i>lau</i>	<i>blau</i>
	sidewards	<i>weli</i>	<i>beli</i>

^a *Ro(né)* only appears in the dialect of Lewaji.

In Table 3.53, the locational verbs are listed according to their semantic properties. There is one default locational *la(né)* that is not semantically specified further.¹⁵ There is a two way distinction in distance: proximate and

¹⁵ The locational *la(né)* is probably etymologically related to a historical preposition. The

Each locational has a basic disyllabic form. The second syllable *né* is optional and not inherited. It has been added to create disyllabic forms in the same way as has been described for free pronouns (cf. §3.4.2.1). In addition, most locationals have a second form with a prefix *b-* of unknown origin. Locationals with initial *w* drop their initial *w* when the prefix *b-* is added, such as *wéné* ‘PROX’ which becomes *béné* and *weli* ‘sideways’ which becomes *beli*. This yields up to four variants for one locational. For example ‘upwards’ can be realised as *jé*, *jéné*, *bjé* or *bjéné*. These long and short forms occur in free variation, usually depending on speech rate and pragmatic features. The proximate locational *blané* ‘PROX’ appears to be historically derived from *lané* ‘LOC’ via prefixation. Synchronically, however, these two locationals have divergent semantics.

(168) a. * *Da=d-ia da=lané.*
 3PL=3PL-stay 3PL=LOC
 Intended: ‘They are staying here.’ (NB:2)

b. *Water-u na=lané una.*
 stone-SPEC 3SG=LOC house\’s
 ‘The stone is at the house.’ (NB:3)

first syllable *la* is used as a locative preposition in several Austronesian languages of the region, as for example in Kambera spoken on Sumba (Klamer 1998:122) or in Roma, a language spoken on the island Roma eastwards of Timor (Steven 1991:67).

a verb that grammaticalised into a locative preposition but the locative preposition *la* that developed into a locational verb to adapt to the system of locational verbs found in the language.

Apart from a location, *la(né)* can also mark a benefactive as in (169). In this example, *lané* precedes the recipient argument *kopo lamé* ‘a boy’ of the main verb *béé* ‘give’.

- (169) *Na=béé-nga lané kopo lamé.*
 3SG=give-3SG LOC child\s male\s
 ‘She gives it to a boy.’ (FHI:42)

A subset of the Central Lembata locationals discussed above can be used as demonstratives. Table 3.54 shows all demonstratives in Central Lembata, the subset of locationals and the temporal adverb *nakété* ‘just now’ (cf. §3.6.2) that can be used as an anaphoric demonstrative.

Table 3.54: Demonstratives

Gloss	Full	Short
PROX	<i>wéné</i>	<i>wé</i>
	<i>réné</i>	<i>ré</i>
DIST	<i>woné</i>	<i>wo</i>
	<i>(roné)</i>	<i>(ro)^a</i>
ANAPH	<i>nakété</i>	<i>nak</i>

^a *Ro(né)* only appears in the dialect of Lewaji.

Each demonstrative has a full form and a short form. There is no semantic difference between the full and the short form, they can freely replace each other. Pragmatically, the full form puts more emphasis on the demonstrative. The distal and the proximate demonstratives come in synonym pairs. The use of proximate /w/-initial or /r/-initial demonstratives is idiolectal or free variation. Compare §3.3.5 for the use of demonstratives in noun phrases.

3.7.2 Locational verb phrases

Locational verbs can be the only verb of a predicate, such as in (170) or the final verb in a multi-verb sequence, such as in (171). In (170), the distal locational verb *wo(né)* ‘DIST’ specifies the location of the object noun *bnelak* ‘field’ in terms of distance from the speaker. In the free English translation, the distance aspect is expressed with the demonstrative ‘that’.

- (170) *Na=woné bnelak.*
 3SG=DIST field
 ‘It [the stone] was on that open field.’ (L3:259)

The sentence (171) shows the use of *la(né)* ‘LOC’ as the last verb in a multi-verb sequence.

- (171) *Pa na=toko lané kaju wikil.*
 then 3SG=hang LOC wood branch
 ‘Then he hangs (it) on a tree branch.’ (FH3:31)

Locational verb phrases can have complex complements that specify the relation of two elements in space, such as being on top of something else or moving towards a certain place. In (172), I present a template for Central Lembata locational phrase.

- (172) Locational phrase template
 [LOC.V + [N + (LOC.N)]_{NP} + (DMV)]

The first two elements, the locational verb (LOC.V) and a noun (N) denoting a place are the only obligatory elements in this template. Locative nouns (LOC.N) and deictic motion verbs (DMV) are optional elements, discussed in the subsections hereafter. An example of a basic locational phrase is (173) with the locational verb *lau* ‘seawards’ and a nominal object *Watuwawer*, a place name.

- (173) *D-ia lau Watuwawer.*
 3PL-stay seawards NAME
 ‘They stayed in Watuwawer (in seaward direction).’ (L2:48)

In example (173), the locational *lau* ‘seawards’ specifies that Watuwawer is located in seaward direction from the speaker’s position. This information

could only be incorporated into the free English translation by adding a relative clause such as ‘which is located in seaward direction from here’.

The wide range of locationals in Central Lembata is used to specify several semantic aspects of the subject referent in relation to the deictic centre, which often is the speaker. Depending on the locational verb used, the distance, direction or elevation of the subject referent can be specified further from the deictic centre’s point of view.

In example (174), the locational *jé* ‘upwards’ specifies the location of the garden *éka* as being located higher than the position of the speaker.

- (174) *Go bolé muku luam jé éka.*
 1SG pick banana heart upwards garden
 ‘I’m picking banana hearts at the garden.’ (NT:18)

In addition to the locational verb containing deictic information, there are two optional slots available in a locational phrase. The slot for locative nouns (LOC.N) is discussed in §3.7.2.1 and the slot for deictic motion verbs (DMV) is introduced in §3.7.2.2.

3.7.2.1 Locative nouns

The locative noun (LOC.N) slot can be filled with a locative noun that conveys information on the topological relation of the subject referent and another object or place. Table 3.55 gives an overview of the most frequent locative nouns in Central Lembata.

Table 3.55: Locative nouns

Locative noun	Gloss
<i>oor</i>	‘inside’
<i>wutu</i>	‘end’
<i>lolo</i>	‘top’
<i>leeng</i>	‘lower part’
<i>keja</i>	‘middle’
<i>weja</i>	‘side’

In both examples (175) and (176) the locational noun *oor* ‘inside’ indicates that the subject referent is located inside another object, which is in this

case a storage house. Through the context in which the sentences appear, it is known that the subject refers to harvested rice and maize from the fields.

- (175) *Da=doka la wétak oor.*
 3PL=store LOC storage inside
 ‘They are storing it in the storage house.’ (T3:167)
- (176) *D-er waak-i doka-ngi d-ewa jé wétak oor.*
 3PL-do finish-3PL lift-3PL 3PL-reach upwards storage inside
 ‘They lifted all of it into the storage house.’ (T3:183)

3.7.2.2 Deictic motion verbs

Whenever motion is involved and the locational phrase expresses a direction (goal) or an origin (source) of the subject referent, a deictic motion verb (DMV) is added in the final slot of the locational phrase. The deictic verbs listed in Table 3.56 semantically express the direction of the movement from the deictic centre’s perspective.

Table 3.56: Phrase-final deictic verbs ‘go’ and ‘come’

	Motion verb	Gloss
go	<i>-ai</i>	‘go’
	<i>géji / géwi</i>	‘go.up’
	<i>lodo</i>	‘go.down’
come	<i>méné(k-)</i>	‘come’
	<i>aka(j-)</i>	‘come.up’
	<i>nau(n-)</i>	‘come.down’

The phrase-final deictic verbs can be classified into ‘come’ and ‘go’ verbs. ‘Come’ verbs express a movement that goes towards the deictic centre, whereas ‘go’ verbs express a movement away from the deictic centre. In addition to the aspect of coming or going, the speaker can specify whether the movement is in upward or downward direction by using deictic verbs that incorporate elevation. The ‘go’ verbs *géji/géwi* ‘go.up’ and *lodo* ‘go.down’ can also be used as medial verbs with the meanings ‘climb/ascend/enter’ and ‘descend/exit’. In medial position, they can be inflected using person suffixes

but the remain uninflected in final position. The ‘go’ verbs *géji* ‘go.up’ and *géwi* ‘go.up’ are free variants. The ‘come’ verbs only appear in final position and can be inflected in that position.

In the following, I illustrate the use of the DMV in come-go pairs according to the level of elevation. The first pair of DMV are the verbs *-ai* ‘go’ and *méné* ‘come’ which are unspecified for elevation. However, in combination with locational verbs (cf. Section 3.7.1), elevation can be expressed, as shown in the following examples. In (177a), the phrase with the verb *-ai* ‘go’ expresses a downward movement ‘to the ground’, as it is combined with the locational verb *juné* ‘downwards’, whereas in (177b), another phrase with the same verb expresses an upward movement ‘up to your house’ because it is used in a clause that contains the locational verb *jé* ‘upwards’.

- (177) a. *Kopo anak kerka-nga, pa péal-a no juné ena n-ai.*
 child little startle-3SG then fall with downwards soil 3SG-go
 ‘The little child is startled, then it falls to the ground.’ (F1:37)
- b. ... *mo jé una m-ai* ...
 2SG upwards house 2SG-go
 ‘... you go up to your house, ...’ (N3:111)

The same two options can be found with the verb *méné* ‘come’. In (178a), the verb *méné* ‘come’ is preceded by the locational verb *jéné* ‘upwards’, expressing a downward movement from a higher place. Whereas in (178b), the verb *méné* ‘come’ expresses an upward movement from a lower place, as it is preceded by the locational verb *ju* ‘downwards’.

- (178) a. *Korar-u di na=beka-nga jéné méné.*
 owl-SPEC also 3SG=fly-3SG upwards come
 ‘The owl also flies (down) from above.’ (F1:41)
- b. *Da=ju lala méné.*
 3PL=downwards road come
 ‘They are approaching from the road (down there).’ (NB:12)

The two DMV *aka* ‘come.up’ and *nau* ‘come.down’ both encode a movement in the direction towards the deictic centre, as well as information on elevation, as illustrated in (179) and (181a). In (179) the verb *aka* ‘come.up’ is expressing a movement towards the deictic centre which is located at a higher

place than the origin of the movement. The DMV is preceded by the locational verb *lau* 'seawards' with its object the placename *Ilé Apé*. The locational *lau* 'seawards' expresses that the place Ile Ape is located in seaward direction from the deictic centre, the speaker, thus located at a lower position than the speaker.

- (179) *Na=tué lau Ilé Apé aka, hari Sabtu.*
 3SG=return seawards Ile Ape come.up day Saturday
 'She will return from Ile Ape on Saturday.' (CI:193)

The DMV *aka* 'come.up' is usually preceded by a locational verb that expresses a lower location and this location is then the place of origin of the movement, such as shown in (179) above. However, the verb *aka* 'come.up' can also be preceded by a locational unspecified for elevation, such as in (180). In this case, the locational *bé* preceding the deictic motion verb *aka* 'come.up' expresses the destination place of the movement, *Lewaji*, which is identical to the deictic centre and located higher than the origin of the movement.

- (180) *Da=kari d-o bé Lewaji aka.*
 3PL=run 3PL-with PROX NAME come.up
 'They fled up here to Lewaji.' (L2:75)

A similar pattern is found with the verb *nau* 'come.down'. This verb can be either preceded by a locational unspecified for elevation, and then expressing the destination of the movement, such as in (181a). Or it can be preceded by a locational expressing a higher place which encodes the origin of the movement, such as in (181b). In (181a), the verb *nau* 'come.up' is preceded by the proximate locational *bé* which denotes the destination of the movement expressed by the deictic motion verb. In (181b), the deictic motion verb *nau* 'come.down' is preceded by the locational *je* 'upwards' which denotes a higher place than the deictic centre and thus the origin of the movement. A stem-final consonant *n* is added when the verb takes a suffix.

- (181) a. *Wawi [...] tu na=bé nau.*
 pig one 3SG=PROX come.down
 'a pig is coming down here' (L2:57)
 b. *Da=je lala naun-i.*
 3PL=upwards road come.down-3PL
 'They are coming down from the road (up there).' (NB:11)

The third pair of motion verbs are *lodo* ‘go.down’ and *géji* or *géwi* ‘go.up’. These verbs encode a downward or upward movement and they are placed clause-finally. In most cases, the motion verb *lodo* ‘go.down’ is preceded by a upwards locational verb, such as *jé* ‘upwards’ in (182a) and (182b).

- (182) a. *Da=jé lala lodo.*
 3PL=upwards road go.down
 ‘They are going down from the road (up there).’ (NB:12)
- b. *Bo korar-u na=du jé kaju lolo lodo.*
 DISC owl-SPEC 3SG=see upwards tree top go.down
 ‘The owl looks down from the top of the tree.’ (FI:44)

In parallel to this pattern, the synonymic motion verbs *géji* and *géwi*, both glossed as ‘go.up’, are often preceded by a downwards locational verb, such as in (183).

- (183) *Da=belo ju ena géji.*
 3PL=appear downwards soil go.up
 ‘They appeared out of the ground.’ (NT:19)

The DMV verbs *lodo* and *géji* / *géwi* can also express an outward versus inward movement. In that case, *lodo* ‘go.up’ corresponds to an outward movement and can also be read as meaning ‘exit’, while *géji* / *géwi* ‘go.up’ corresponds to an inward movement and can also be read as ‘enter’. An example is given in the elicited sentence in (184).

- (184) a. *Aor-u kari la kam=una lodo.*
 dog-SPEC run LOC 1PL.EXCL.POSS=house exit
 ‘The dog runs out of our house.’ (NT:19)
- b. *Aor-u kari la kam=una géji.*
 dog-SPEC run LOC 1PL.EXCL.POSS=house enter
 ‘The dog runs into our house.’ (NT:19)

In addition to the clause-final deictic motion verbs just discussed, Central Lembata has adapted the clause-final position for an additional set of verbs appearing with locational phrases. These clause-final verbs are mainly posture verbs and a few other intransitive verbs, listed in 3.57.

Table 3.57: Additional clause-final verbs in Central Lembata

Posture verbs	
<i>tobé</i>	‘sit’
<i>diri</i>	‘stand’
<i>gelé</i>	‘lie (human)’
<i>pawa</i>	‘lie (thing)’
<i>swakat</i>	‘crouch’
<i>toko</i>	‘hang’
<i>suduk</i>	‘bend’
Other verbs	
<i>oli</i>	‘arrive’
<i>géka</i>	‘laugh’
<i>irak</i>	‘shout’
<i>sédu</i>	‘dance’

In (185), the verb *diri* ‘stand’ is placed clause-finally. The clause also contains the locational verb *ju* ‘to be at a lower place’ glossed as ‘downwards’ which is placed clause-medially, directly after the subject. The two verbs are separated by the object NP *watu puk* ‘foot of the stone’. The object NP is not obligatory, as a locational verb, such as *wo* ‘DIST’ in (186), can be directly followed by a clause-final verb.

- (185) *Aoru na=ju watu puk diri.*
 dogsPEC 3SG=downwards stone stem stand
 ‘The dog is standing at the foot of the stone.’ (F1:47)

- (186) *Ké ata dikeja, go denger, da=wo irak.*
 so person person.PL 1SG hear 3PL=DIST shout
 ‘So, I hear people shouting there.’ (S2:31)

Clause-final verbs are an innovation in Central Lembata. The canonical word order is verb-medial and only a small subset of verbs is phrase-final, i. e. the deictic motion verbs and a couple of other verbs, for example posture verbs, as has been shown in this section. §10.2 investigates the emergence of these clause-final verbs in Central Lembata and other Flores-Lembata languages.

3.7.3 Summary

Central Lembata has a set of locational verbs that express distance and direction from the deictic centre. A subset of these locationals functions also as demonstratives. A locational verb phrase can contain the following elements: a locational verb, a noun denoting the location as a point of reference, and optionally a locative noun and a deictic motion verb. Locative nouns convey information on the topological relation between the subject and the location. Deictic verbs are used to indicate movement by differentiating between ‘come’ and ‘go’, as well as elevation.

3.8 Summary and conclusions

Central Lembata has an average-sized phoneme inventory with little distributional restrictions for phonemes. Syllables are open in medial position and open or closed in final position. All stems are generally disyllabic and stressed at the penultimate syllable, except for a penultimate syllable with a schwa which results in ultimate stress. Central Lembata is the only language in the Flores-Lembata family for which a plural suffix as well as a specificity suffix is attested for nouns. The phenomenon of coda alternating nouns is also unique to Central Lembata within the FL family. The noun phrase has a core part for nouns and numerals, while demonstratives and discourse markers occur on the right edge of the NP located outside of the core NP. In addition to free pronouns, Central Lembata has a set of S/A marking proclitics of which a subset only appears in irrealis contexts and two sets of S/P marking suffixes that are predominantly found in non-dynamic predicates. Free possessor pronouns are used for possessed alienable nouns, also if a possessor NP is present. Possessor suffixes are only allowed on inalienable nouns. Verbs can be divided into three morpho-syntactic classes based on which suffix set they take. The verb *-ora* ‘join; be with’ and the verb *maa* ‘say’ both show typical grammaticalisation pathways. These are from the verb meaning ‘be with; join’ to an existential, a comitative, a coordinator and maybe a destination marker; and from the verb meaning ‘say’ to a complementiser, a marker of future, intention, volition or ability and to a conditional conjunction. Central Lembata has SVP basic word order. Nominal predicates are juxtaposed to their subject NP and optionally marked with

S/P suffixes. Tense and aspect are expressed through adverbs. Double negation is the most common negation strategy for clauses. Location marking is a crucial aspect in the Central Lembata grammar. Locational verbs express direction or distance from the deictic centre. In a locational phrase, detailed information on location and movement can be conveyed by locative nouns and clause-final deictic motion verbs.

Part II

Historical phonology and lexical innovations

CHAPTER 4

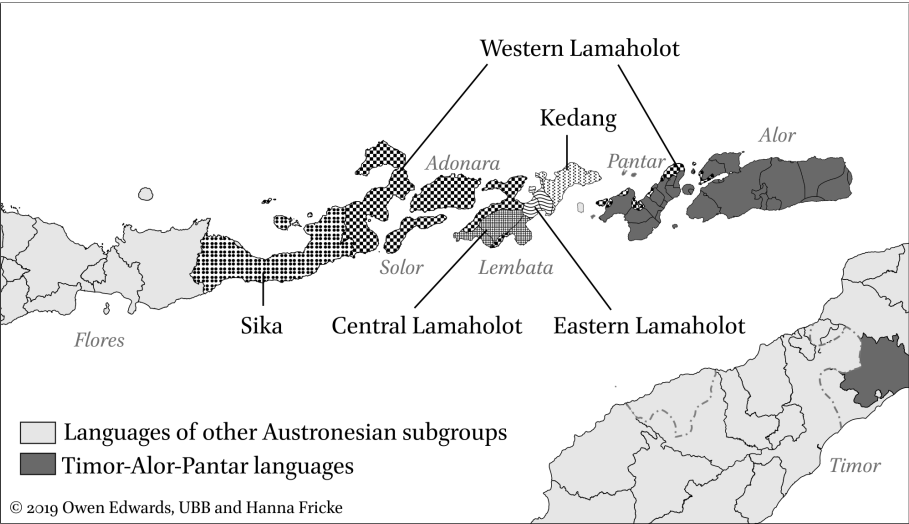
Introduction to Part II

4.1 The Flores-Lembata languages

4.1.1 Overview

The local languages spoken in the eastern part of Flores and in the Solor Archipelago are grouped together as the Flores-Lembata languages, displayed on the map in Figure 4.1. The Flores-Lembata languages are a subgroup within the Malayo-Polynesian branch of Austronesian. All neighbouring languages are also Austronesian, except for the non-Austronesian languages on the islands of Alor and Pantar to the east which belong to the Timor-Alor-Pantar family (cf. §1.3.1). In coastal areas of the islands of Alor and Pantar, an Austronesian language is spoken: the Western Lamaholot language Alorese (cf. §1.2.4.2). Towards the west of the Flores-Lembata languages, the Austronesian Central Flores languages are spoken (Elias 2018). On the island of Timor, southeast of the Solor Archipelago, the Austronesian Timor-Babar languages and Central Timor languages are found (cf. §1.3.1).

Figure 4.1: The Flores-Lembata subgroups and their linguistic context



Based on exclusively shared sound changes (cf. §5.3.1), I distinguish five Flores-Lembata subgroups which are listed in Table 4.1 in geographical order from west to east along with their abbreviations as used throughout this dissertation. In Chapter 5, I provide more details on these subgroups, the internal structure of the Flores-Lembata family, as well as on the classification of Flores-Lembata within Malayo-Polynesian.

Table 4.1: The Flores-Lembata subgroups

Subgroup	Abbreviation
Sika	SK
Western Lamaholot	WL
Central Lamaholot	CL
Eastern Lamaholot	EL
Kedang	KD

Based on lexical similarity or mutual intelligibility many more than five languages can be established.¹ Thus, each of the subgroups contain one or more

¹ Keraf (1978a) contains a lexicostatistical approach on the Lamaholot subgroups which distinguishes 17 languages based on the criterion of a lexical similarity of 80 % or higher.

languages.

Previous studies on the Flores-Lembata languages have not considered Central Lamaholot and Eastern Lamaholot as independent subgroups or branches of Flores-Lembata. The geographic areas of Central Lamaholot and Eastern Lamaholot were usually included in the so-called Lamaholot dialect chain or cluster (cf. Grangé 2015b) but remained linguistically undescribed. Most published research on individual varieties of the Lamaholot dialect chain so far has been conducted on varieties of the Western Lamaholot group.

In this dissertation, I use the term Lamaholot to refer to the three subgroups Western Lamaholot, Eastern Lamaholot and Central Lamaholot as a unit of closely-related subgroups that have been in close contact.² However, there is no evidence that Lamaholot, thus the three subgroups, forms an innovation-defined group within Flores-Lembata (cf. §5.3.2). The reasons for not abandoning the label and concept of Lamaholot encompassing the three subgroups as a whole is (i) the fact that the speakers of the three Lamaholot subgroups see themselves as belong to one socio-cultural unit opposed to their neighbours Kedang in the east and Sika in the west and (ii) the three subgroups have been in contact until today and share certain structural features that are not attested in Sika and Kedang, such as clause-final negation and an alienability distinction in the possessive construction (cf. Part III). Also lexically, they are more similar to each other than to Sika and Kedang. In the following section, I introduce each of the Flores-Lembata subgroups in detail.

² The term Lamaholot itself probably goes back to a place name. The first part *lama* means 'place' (Keraf 1978a:7). It is found in many place names throughout the Solor archipelago. The second term *holot* means 'stick together' (Arndt 1937:3; Keraf 1978a:7) and according to Arndt (1937:3) in older forms of the language it also means 'human'. The word *holot* is cognate with the name Solor which refers to an island between Flores and Lembata and to the whole Solor archipelago covering Adonara, Solor and Lembata. The sound change from PFL *s > WL h is regular (cf. §5.2.3). So, the island name Solor represents the more archaic form of the name. The final *t* in *holot* replaced the *r* in the place name Solor. It could be related to the suffix *-t* found on attributes in Sika (cf. §8.3.3.1). So Lamaholot could mean something like 'the place of the Solorese' (cf. Hägerdal 2012:22). However, earlier sources, i. e. the Majapahit chronicles Nagarakertagama and the Pigafetti records from the Portuguese ship Victoria mention the name *Solot* with final *t* referring to the area west of Pantar and Lembata. Barnes (1982:409) suggests that the version of the name with final *t* is the original.

4.1.2 Sika

Sika (SK) is the westernmost language of the Flores-Lembata family, spoken by around 175,000 people spreading over an area of about 80 km from west to east in the eastern half of the island of Flores (Lewis and Grimes 1995:601), see map in Figure 4.1. The language is spoken in the whole Sikka regency, with its capital Maumere, as well as in the villages Hewa, Kokang, Wodong and parts of Pantai 'Oa in the southwestern edge of the Flores Timur regency.³ According to the literature, there are three main dialects of Sika (Rosen 1986:40-41; Lewis 1988:9-10; Lewis and Grimes 1995:601): the dialect from the village "Sikka Natar" on the south coast, the dialect of "Krowe" (central mountain area of the Sikka regency) and the dialect of "Tana 'Ai" in the very east of the Sikka regency.

The oldest and perhaps most detailed description of the language is a grammar by the missionary Paul Arndt (1931) written in German. It covers a wide range of fields, it is however rather difficult to read for linguists today as example sentences are usually not translated and no wordlist is given. Other descriptions are Rosen (1986) on phonology, verb classes and verbal suffixes in Sika, Lewis and Grimes (1995), a very short sketch of Sika, and Bolscher (1982), an introduction to the Sika language written by a missionary in Indonesian for foreign priests coming to the area. Pareira and Lewis (1998) is a Sika-Indonesian dictionary. Lewis also published several anthropological works and articles on the Sika culture, especially on the Tana 'Ai area (Lewis 1988; Lewis 1998; Lewis 2006). The most recent publication is Fricke (2014a) on phonology, verbal inflection, possessive constructions and spatial language in the Hewa variety of Sika.

4.1.3 Western Lamaholot

Western Lamaholot (WL) is the biggest Flores-Lembata subgroup in terms of geographic space and number of speakers, covering an area of about 90 km from west to east, see map in Figure 4.1. Western Lamaholot varieties are spoken all over the Flores Timur regency, on the eastern tip of Flores with its capital Larantuka, the islands Solor and Adonara, as well as in coastal areas

³ In the linguistic literature, the language Sika is usually written with a single *k*, while in official documents as well as in the name of the regency, Sikka is spelled with double *k* reflecting Dutch spelling from colonial times.

of the Lembata regency including its capital Lewoleba, and in a few villages, such as the village Ojang in the northeastern part of the Sikka regency. The Western Lamaholot varieties in the Sikka regency are called Muhang and are largely undescribed.⁴ In addition, Alorese spoken in coastal areas of the islands of Alor and Pantar is linguistically part of Western Lamaholot although this group of speakers has been separated from their homeland for several hundreds of years and the language has diverged considerably since then (cf. §1.2.4.2). I come back to the history of Alorese at the end of this section. The total number of speakers of Western Lamaholot varieties must amount to at least 300,000, as the Flores Timur regency alone has about 240,000 inhabitants according to census data, not including the speakers on Lembata, Alor and Pantar.

Western Lamaholot varieties have a long history of linguistic and ethnographic research. The first linguistic description of Western Lamaholot varieties is Arndt (1937), a grammatical description by a German missionary using information from several western varieties but excluding those on Lembata (Arndt 1937:3). The grammar is written in German and has a structure based on a grammar of a European language. For modern-day use, it is unfortunate that the language examples are not always translated and never glossed on a word-by-word basis. However, it represents an early record of the language which, in some cases, gives information on older stages of the language. In the 1970s, two linguists, native to different parts of the Lamaholot area, published grammatical descriptions of Western Lamaholot varieties in Indonesian. Fernandez (1977) is a short description of the Western Lamaholot variety spoken around the mountain Ile Mandiri close to the town of Larantuka on Flores. Keraf (1978a) describes the morphology of the Western Lamaholot variety spoken in the village of Lamalera on the south coast of Lembata. The exceptional aspect of Keraf's work is his lexical survey on 33 Lamaholot dialects of all three subgroups, which he includes in the appendix of his work. Contemporary research on Lamaholot has involved linguists of various origins. Nishiyama and Kelen (2007) is a brief description of the variety of the villages Lewoingu and Lewolaga located in the eastern part of Flores. Nagaya (2011) represents a grammatical description of one of the westernmost Lamaholot varieties spoken around the

⁴ Most likely the variety Keraf (1978a) labels as Pukaunu, located in the western edge of the Flores Timur regency, is similar or identical to Muhang.

mountain Lewotobi located on Flores. Klamer (2011) is a sketch grammar of Alorese, the WL offspring spoken on Alor and Pantar. Grangé (2015b) describes split intransitivity in East Adonara Lamaholot and compares several Lamaholot varieties in terms of their lexicon. Grangé (2015b:47) proposes eastern Adonara as the homeland of Lamaholot, which most likely means the homeland of the Western Lamaholot subgroup because his hypothesis is based only on data from Western Lamaholot varieties.⁵ Kroon (2016) is a descriptive grammar of the variety of Lamaholot spoken on the island of Solor. Akoli (2010) and Michels (2017) are Master theses comparing different aspects in several Western Lamaholot varieties. Pampus (1999) is an extensive dictionary of the Lewolema variety of WL with translations into Indonesian and German.

Vatter (1932), Arndt (1938) and Arndt (1940) are early records of cultural aspects in the Western Lamaholot area. Vatter (1932) is not restricted to the Western Lamaholot area but also contains information on the wider region of the whole Solor and Alor Archipelago. Barnes (1987) is a study on weaving practices among the Lamaholot people, which also takes the Central and Eastern Lamaholot areas into account. Barnes (1996) is an ethnographic study on the whale hunting community of Lamalera on Lembata. Kohl (1998) is an extensive ethnographic description of the Lewolema community in the eastern tip of Flores. Rappoport (2014) is a musicological study on songs in Tanjung Bunga, the northeastern tip of Flores.

A more extensive note on Alorese is in order here. According to historical and ethnographic sources, speakers of Alorese emigrated from the Lamaholot area to Pantar around the year 1300 (cf. §1.2.4.2), and later spread further to Alor (Klamer 2011:16). It has been suggested that Alorese is a dialect of Lamaholot by Stokhof (1975:9) but based on lexical divergence — only 55% lexical similarity between Alorese and WL varieties could be attested — and morphological simplification compared to WL varieties, Alorese should be considered a separate language (Klamer 2011:24).⁶ This percentage of lexical similarity between Alorese and WL indeed shows that Alorese is lexically as different from other WL varieties as these are differ-

⁵ Grangé (2015b:48) himself suggests that more research on the Lamaholot varieties on Lembata is needed to confirm or reject his proposal of Eastern Adonara as the Lamaholot homeland.

⁶ The WL varieties used for comparison to Alorese are Lewoingu, Solor and Lamalera (Klamer 2011:118-127).

ent from Central Lamaholot and Eastern Lamaholot. Keraf (1978a:Appendix VI) calculates 55% lexical similarity between Western Lamaholot, Central Lamaholot and Eastern Lamaholot respectively. In contrast, according to Keraf (1978a), Sika and Kedang only show 30% of lexical similarity to all Lamaholot varieties. But the shared sound change of Proto-Flores-Lembata (PFL) *r > ʔ in Western Lamaholot and Alorese (cf. §5.2.5 and §5.3.1.2) and the shared innovation of the clause-final negator Proto-Western Lamaholot (PWL) *hala (< PMP *salaq ‘wrong’) > Alorese *lahé* (cf. §10.3.4.4) points to a common ancestor for WL and Alorese: Proto-Western Lamaholot (PWL). Therefore, Alorese can be considered an offspring of PWL which has undergone independent and more drastic changes compared with other varieties of WL. This divergence can be explained by the fact that Alorese has been geographically separated from other varieties of WL for about 700 years, as well as by longstanding contacts between speakers of Alorese and speakers of Alor-Pantar languages.

4.1.4 Central Lamaholot

Central Lamaholot (CL) varieties are spoken in the central and southwestern parts of the island Lembata, see map in Figure 4.1. Most parts of this area are mountainous and roads to smaller villages are not always in good conditions. This makes it hard to reach the villages. The estimated number of speakers amounts to around 17,000 people based on census data of the Lembata regency. The Atadei district in the central mountains of Lembata is the only district which is completely Central Lamaholot speaking. Other districts that are partly CL speaking are Wulandoni, Nagawutung and Nubatukan.

Central Lamaholot has been largely undescribed. Until recently, the 200 items wordlists of eight Central Lamaholot varieties in Keraf (1978a) were the only published information on Central Lamaholot. The grammar sketch of the CL variety Central Lembata in Chapter 3 of this dissertation is the first detailed description of a Central Lamaholot variety. There are two recent works of Masters students that discuss varieties of Central Lamaholot. Akoli (2010) describes aspects of the Lewokukung variety, based on a wordlist of 200 basic items and a transcribed and translated folk story. Krauße (2016) is a brief grammar sketch on the CL variety spoken on the Atadei peninsular of Lembata, named Atadei Painara or Eastern Atadei.

4.1.5 Eastern Lamaholot

Eastern Lamaholot (EL) is spoken in a small area of about 20 km in the eastern part of the island Lembata, see map in Figure 4.1. Speakers of Eastern Lamaholot are mainly found in the Lebatukan district, but probably also partly in the Nubatukan district to the west or the Omesuri district to the east. In the west of this area Western Lamaholot and Central Lamaholot varieties are spoken. Towards the east the language is adjacent to Kedang. The number of speakers amounts up to 8,000 people based on census data. Until today, Eastern Lamaholot remains largely undescribed. The only materials available on Eastern Lamaholot are two 200 items wordlists collected by Keraf (1978a) of the varieties Lewoeleng and Lamatuka.

4.1.6 Kedang

Kedang (KD) is spoken by approximately 29,000 people at the easternmost tip of Lembata in 28 villages surrounding the volcano Ujolewung (Samely 1991a:1), see map in Figure 4.1. The area is part of the Lembata regency which includes the whole island of Lembata where, besides Kedang, also several different Lamaholot varieties are also spoken. Samely (1991a) is the only published grammatical description of Kedang. This grammar is focused on the phonetics and phonology of the language. Samely and Barnes (2013) is a comprehensive dictionary resulting from Samely's data collection and many decades of anthropological fieldwork carried out by Barnes. Several anthropological publications about Kedang by Robert Barnes and Ruth Barnes are available, such as Barnes (1974), an ethnographic study of the Kedang people, or Barnes (2005) about the innovation of weaving practices in Kedang.

4.2 Language sample and data sources

The historical reconstruction and lexical comparison carried out in Part II of this dissertation is based on lexical data from 46 Flores-Lembata varieties, listed in Table 4.2. The lexical items are taken from published dictionaries and, to large extents, from the LexiRumah database which collects comparative wordlists from various sources at <https://lexirumah.model-ling.eu/lexirumah/> (Kaiping and Klamer 2018; Kaiping et al. 2019). In the list below, I

first give the abbreviation of the variety I use in this dissertation, then the ISO 369-3 code, the name of the variety and in the last column the sources I used for the lexical data. The varieties are sorted by subgroup. The names of the varieties are taken from the data sources. No claims are made as to whether these varieties represent definable languages or dialects within the given subgroup. In most cases, the names are identical with village names, other place names or mountain names. As some larger units of several varieties have been defined as languages, I indicate those language names in brackets after some of the varieties in the list.

Table 4.2: The Flores-Lembata varieties in the language sample of Part II

Abbr.	ISO 369-3	Variety	Sources
Sika (SK)			
HEW	ski	Hewa	Keraf 1978b; Fricke 2014b, Klamer 2015a
KR	ski	Krowe	Pereira and Lewis 1998
MM	ski	Maumere	Fricke 2014b
TA	ski	Tana 'Ai	Lewis 1995
Western Lamaholot (WL)			
AB / AL	aol	Alor Besar (Alorese)	Moro 2016a
BN / AL	aol	Baranusa (Alorese)	Moro 2016b
MS / AL	aol	Munaseli (Alorese)	Moro 2016c
PD / AL	aol	Pandai (Alorese)	Moro 2016d
AD	adr	Adonara	Klamer 2015c
BT	adr	Botun	Keraf 1978b
DL	adr	Dulhi	Keraf 1978b
HRW	adr	Horowura	Keraf 1978b
KO	adr	Kiwangona	Keraf 1978b
LMK	adr	Lamakera	Keraf 1978b
WT	adr	Watan	Keraf 1978b
WW	adr	Waiwadan	Keraf 1978b
BL	lmr	Belang	Keraf 1978b
LH	lmr	Lamahora	Keraf 1978b

Abbr.	ISO 369-3	Variety	Sources
LML	lmr	Lamalera	Keraf 1978b
MD	lmr	Merdeka	Keraf 1978b
ML	lmr	Mulan	Keraf 1978b
WK	lmr	Wuakerong	Keraf 1978b
BM	slp	Bama	Keraf 1978b
BP	slp	Baipito	Keraf 1978b
LG	slp	Lewolaga	Keraf 1978b
LWI	slp	Lewoingu	Klamer 2015d
LWL	slp	Lewolema	Keraf 1978b; Pampus 1999
RE	slp	Ritaebang	Keraf 1978b
TJ	slp	Tanjung	Keraf 1978b
WB	slp	Waibalun	Keraf 1978b
IA	ila	Ile Ape	Keraf 1978b
LTB	lwt	Lewotobi	Keraf 1978b
PK	-	Pukaunu	Keraf 1978b

Central Lamaholot (CL)

KK	lvu	Kalikasa (Central Lembata)	Fricke 2015a; Fricke 2019
LK	lvu	Lewuka	Keraf 1978b
LWK	lvu	Lewokukung	Keraf 1978b
IL	lmj	Imulolo	Keraf 1978b
LP	lmj	Lewopenutu	Keraf 1978b
LWT	lmj	Lewotala	Keraf 1978b
MR	lmj	Mingar	Keraf 1978b
LR	lmf	Lerek (Atadei Painara)	Fricke 2015b
PN	lmf	Painara (Atadei Painara)	Keraf 1978b

Eastern Lamaholot (EL)

LMT	lmq	Lamatuka	Keraf 1978b
LE	lwe	Lewoeleng	Keraf 1978b

Kedang (KD)

LB	ksx	Leubatang	Klamer 2015b

Abbr.	ISO 369-3	Variety	Sources
LW	ksx	Leuwayang	Samely 1991b; Samely and Barnes 2013

In parts of Chapter 5, I also use lexical data from other Austronesian languages outside of the Flores-Lembata subgroup. These languages and their sources are provided in Table 4.3 hereafter. For these languages no abbreviations are used in this dissertation.

Table 4.3: Other Austronesian languages used for comparison in Part II

ISO 369-3	Language	Sources
Central Flores and Palu'e		
-	Proto-Central Flores (PCF)	Elias 2018
ljl	Lio	Elias 2017b
end	Ende	Aoki and Nakagawa 1993
end	Nga'o	Elias 2017e; Elias 2017d
xxk	Keo	Baird 2002
nxe	Nage	Elias 2017c
nxg	Ngada	Arndt 1961
ror	Rongga	Arka et al. 2007
ple	Palu'e	Donohue 2003
Western Flores		
mgy	Manggarai	Verheijen 1967
kvh	Komodo	Verheijen 1982
Bima and Sumba-Hawu		
bhp	Bima	Ismail 1985
xbr	Kambera	Onvlee 1984; Klamer 1998
hvn	Hawu	Grimes et al. 2008; Jonker 1908
Timor (AN)		
-	Proto-Rote-Meto (PRM)	Edwards in prep

The Proto-Malayo-Polynesian (PMP) sounds and reconstructions in Part II are from Blust and Trussel (2010). For the purpose of this dissertation, I take the PMP forms by Blust and Trussel as they are without evaluation their evidence as this is not feasible within the scope of this work. However, it has to be noted that some of the PMP forms are based on comparably small sets of data with sometimes a geographical bias towards the western part of the Austronesian language family. Therefore, it is possible that with further research in this field and increasing availability of lexical data, some PMP form may have to be revised. It also needs to be taken into account when reading the PMP forms in this dissertation that they are based on a much larger set of languages than the languages in the subgroup of Flores-Lembata. Blust and Trussel (2010) also reconstruct lower-level forms, such as Proto-Central-Malayo-Polynesian (PCMP) and Proto-Central-Eastern-Malayo-Polynesian (PCEMP). However, as the existence of these subgroups is still disputed (cf. §1.3.1), these forms are not taken into account systematically. Nevertheless, if such a PC(E)MP reconstruction is known for a cognate set discussed here, I indicate this.

4.3 Methodology

To answer the research questions for Part II, 46 wordlists of Flores-Lembata varieties (cf. §4.2) were analysed using the online tool EDICTOR (etymological dictionary editor) at <http://edictor.digling.org>. The wordlists are accessible through the lexical database LexiRumah but originally come from various sources as indicated in the database and in the tables in §4.2. The wordlists contain between 200 and 600 lexical items each. In total they cover 607 different concepts of basic as well as special vocabulary. For some concepts, additional information from dictionaries, also listed in §4.2, was added.

From these wordlists, I collected over 400 lexeme sets. I define a lexeme set as a set of related words that appear across languages. These sets are often cognate sets which means that they trace back to a reconstructible proto-form in a proto-language, such as Proto-Flores-Lembata (PFL) or/and Proto-Malayo-Polynesian (PMP). An example of a cognate set is the set for the concept ‘seven’ in Table 4.4. However, a lexeme set can also be a set of related forms that cannot (yet) be reconstructed to a common proto-form,

such as in the set for ‘tongue’ in Table 4.4. Such an unreconstructible form is marked by a hashtag (#) (cf. §4.4). The unreconstructibility of sets could be due to missing cognates, unexplained irregularities or borrowing.

Table 4.4: Examples of lexeme sets

	‘seven’	‘tongue’
PMP	*pitu	-
PFL	*pitu	#ebel
SK	<i>pitu</i>	-
CL (KK)	<i>pito</i>	<i>evel</i>
WL (LWI)	<i>pito</i>	<i>veveləŋ</i>
EL	[...]	<i>eblə</i>
KD	<i>pitu</i>	<i>ebel</i>

I establish a set when a lexeme occurs in at least two of the Flores-Lembata subgroups (cf. §4.1). Occasionally, my database contains lexemes that are only found in one Flores-Lembata subgroup but these are only considered if they go back to a PMP form.

The lexeme sets used for Proto-Flores-Lembata (PFL) reconstructions and those sets of similar forms that cannot be reconstructed to PFL, are listed in Appendix B. In the appendix, the sets are sorted alphabetically according to concept. Each set contains a PFL reconstruction or a potential reconstruction marked by a hashtag, as well as all related forms in individual languages that are attested in the LexiRumah database and occasionally in dictionaries. The language varieties in the appendix are marked with the abbreviations listed in Table 4.2.

The lexeme sets, potentially cognate sets, are analysed in several ways to answer the research questions in §1.4, replicated here. The first two questions are addressed in Chapter 5 and the last three questions are answered in Chapter 6.

- (1) How are the Proto-Malayo-Polynesian (PMP) proto-sounds reflected in the Flores-Lembata languages?
- (2) What is the evidence for (i) subgroups within Flores-Lembata, (ii) Flores-Lembata as a subgroup, and (iii) Bima-Lembata as a higher-

level subgroup including Flores-Lembata and other Austronesian languages?

- (3) Which Flores-Lembata lexical items are inherited from an Austronesian ancestor?
- (4) Which Flores-Lembata lexical items are not inherited from an Austronesian ancestor?
- (5) Which Flores-Lembata lexical items can be reconstructed to Proto-Flores-Lembata?

To answer question (1) on how the Proto-Malayo-Polynesian (PMP) sounds are reflected in the Flores-Lembata languages, I used the comparative method (cf. Campbell and Poser 2008) to establish regular sound correspondences in lexeme sets that go back to PMP forms. The results of this analysis are shown in the first part of Chapter 5.

To answer question (2) on evidence for subgroups, I identified exclusively shared sound changes from the regular correspondences established before. On the basis of regular sound changes as shared innovations, five Flores-Lembata subgroups could be established. To provide evidence for Flores-Lembata and higher level subgroups, other Austronesian languages in the region but outside of the Flores-Lembata group had to be taken into account. As the LexiRumah database also contains lexemes from Austronesian languages of Timor and Flores, for the same concepts as the Flores-Lembata wordlists, I could use this database for the purpose of finding evidence for exclusively shared sound changes in Flores-Lembata and in Bima-Lembata, with the addition of a few other sources (cf. Table 4.3). The evidence for subgroups is provided in the second part of Chapter 5.

Chapter 6 examines the Flores-Lembata lexicon according to PMP and non-PMP origin of lexeme sets and their reconstructability to PFL. To answer questions (3) and (4) which concern the origin of the Flores-Lembata lexemes, I evaluated each set according to the presence or absence of a related PMP form. Lexeme sets with regular sound correspondences and a PMP form were considered to be of PMP origin, thus going back to an Austronesian ancestor. Lexeme sets without related PMP form were considered of non-PMP origin.

For a cognate set to be reconstructible to Proto-Flores-Lembata (PFL), as asked for in question (5), the following criteria must be fulfilled. First, the sound correspondences between the reflexes in different subgroups have to

be regular. The regularity is determined based on the regular correspondences established in Chapter 5. Second, there are two possible conditions that lead to a PFL reconstruction. (i) If the lexeme set can be traced back to a PMP form (attested in Blust and Trussel 2010), then the set is always reconstructed to PFL, or (ii) if no related PMP form is known, the set must be attested in at least Sika and Kedang, then the set can also be reconstructed to PFL. This means that a form that appears in only one or two Flores-Lembata subgroup and has a PMP form is reconstructed to PFL. However, if no PMP form exists, only items that are attested in Sika and Kedang are reconstructed to PFL. Sika and Kedang are the two Flores-Lembata subgroups that are geographically the furthest apart and therefore, the occurrence of related forms in these two languages points to inheritance from Proto-Flores-Lembata (PFL) rather than to diffusion after the split of the family. The possibility of Sika and Kedang forming a subgroup within Flores-Lembata (and therefore sharing vocabulary) has been ruled out (cf. §5.3.2). In the majority of non-PMP sets, the lexeme is not only attested in Sika and Kedang but also in at least one Lamaholot variety. In the sample, there are only 6 sets of related lexical items that are attested in Sika and Kedang but no known reflexes in the Lamaholot variety. Five out of these six go back to Proto-Malayo-Polynesian which suggests that the Lamaholot varieties have replaced these concepts with new words.⁷ A lexeme shared by several subgroups, but not fulfilling the criteria just set out, is not reconstructed to PFL. For example, if a set of forms with regular correspondences is currently only known to be found in Kedang and Lamaholot, and does not trace back to PMP, I do not consider it reconstructible to PFL because I cannot rule out diffusion between adjacent subgroups. For such a set, a potential reconstruction is established and marked with a hashtag (#) and a subgroup abbreviation, such as LH-KD for Lamaholot-Kedang.

The following limitations apply to the analysis in Chapter 6. For reconstructions and regularity judgements in this chapter, I focus on consonants, as in the scope of this study, an equally systematic analysis of the vowel reflexes is not feasible. Nevertheless, I take into account that a sporadic change

⁷ The Flores-Lembata cognate sets in my sample that only have reflexes in Sika and Kedang but in none of the Lamaholot varieties are: PFL *m-tida 'sharp' < PMP *tazim, PFL *tave 'laugh' < PMP *tawa, PFL *udu 'grass; bush' < PMP *udu, PFL *vivir 'lips' < PMP *biRbiR 'lower lip', PFL *posok / *blosok 'rub; wipe' < PMP *usuq (?) and PFL *ipoho? 'breathe' with no known PMP source.

between phonetically more similar vowels, such as *i* > *e*, is more likely than a change between very different vowels, such as *a* > *u*. Thus, a possible lexeme set with a very unlikely vowel change would not be considered reconstructible to PFL. In this chapter, a systematic research on whether the reconstructions of the non-Austronesian Flores-Lembata lexicon are also attested in other Austronesian languages of the area or in the non-Austronesian Timor-Alor-Pantar (TAP) languages was not within the scope of this dissertation. Nevertheless, connections to forms in other languages were indicated when known. It is also possible that some of them turn out to be actually of Austronesian origin if more comparative research is done on Austronesian vocabulary.

4.4 Data representation

4.4.1 Transcription conventions

PMP phonemes, as well as occasional Proto-Austronesian (PAN) forms, are transcribed as in Blust and Trussel (2010). Most of the symbols used by Blust and Trussel (2010) are equivalents to IPA. However, the grapheme <j> in the PMP reconstructions represents [g], [ɣ] or [gʲ], the grapheme <z> represents a voiced palatal affricate [dʒ] or [jʒ], the grapheme <R> represents [r], the grapheme <y> represents [j] and the grapheme <e> represents schwa [ə] (Ross 1992; Wolff 2010; Blust 2013:245, 554, 588, 601). To allow a differentiation between schwa [ə] and the unrounded front vowel [e] on lower levels, I retranscribe all PMP *e with *ə. In all other cases, I keep the transcriptions in Blust and Trussel (2010). The non-IPA symbols in the PMP forms are listed in Table 4.5.

My own reconstructions of PFL are given using IPA symbols. Only for the palatal approximant [j], I use the symbol <y> instead of its IPA symbol [j] to avoid confusion with a voiced palatal affricate [dʒ]. All reconstructions are marked with an asterisk <*> in front of the word and given in normal font. Individual phonemes are highlighted in bold to show correspondences. If no PFL form can be reconstructed but a set of similar forms is attested, I give a form that could potentially be a PFL reconstruction based on reversing the sound changes they undergo in the subgroups. I mark these unreconstructible forms with a hash tag (#), instead of an asterisk (*). A form marked by

thus refers to a lexeme shared by several subgroups but which, based on current knowledge, is not reconstructible to a proto-language.

Table 4.5: Non-IPA symbols in PMP forms in Blust and Trussel (2010)

Blust and Trussel 2010	IPA symbol	This dissertation
<j>	[g] / [ɣ] / [gʲ]	<j>
<z>	[dʒ] / [ʝ]	<z>
<R>	[r]	<R>
<e>	[ə]	<ə>
<y>	[j]	<y>

Reflexes of entire words or phonemes that are attested in present-day languages are given in *italic* and transcribed in phonemic IPA. Again with the exception of the palatal approximant [j], which I represent as <y> to avoid confusion with a voiced palatal affricate [dʒ]. The symbol *w* is retranscribed as *v* for the Flores-Lembata languages, as it is realised as a voiced fricative [v] or approximant [ʋ] in all languages of Flores-Lembata. The vowels *ɛ* and *æ* in some Lamaholot sources are retranscribed as *e*, as they are not phonemic. The same is done for *ɔ* and *ə* which are retranscribed as *o* and schwa *ə* respectively for the same reason. To make the transcriptions of Kedang lexical data from two different sources consistent, I retranscribe Samely (1991a)'s <e> [ɛ] and <è> [æ] as *e* and *ɛ* respectively and I also retranscribe Klammer (2015b)'s [w] as *v*. In addition, all Kedang [ɔ] are retranscribed as *o* as there is no phonemic contrast between [o] and [ɔ]. Note that the words in the lexeme sets in Appendix B are given according to the sources and are not retranscribed.

4.4.2 Organisation of tables

Tables with sets of related lexemes in this part are composed according to the following conventions. For an example of a lexeme table, see Table 5.4 in Chapter 5 or Table 4.4 in §4.3. Each column lists reconstructions and their reflexes for one concept. The first rows show reconstructions. Then in the subsequent rows, the respective reflexes in the five Flores-Lembata subgroups are provided. The concepts given in the last row of the tables are the

concepts associated with the synchronic meanings of the reflexes. If these do not match with the reconstructed meanings or if the meaning of one lexical item in the set has changed, this is indicated. The subgroups are given in the left-most column. The varieties of one subgroup sometimes differ considerably in their lexicon and for some items, PMP reflexes are only found in a subset of varieties. Therefore, I indicate the variety in brackets if necessary. See §4.2 for a list of varieties and their abbreviations. A dash in the table means that the concept is expressed by a new lexical item that is not cognate. The symbol [...] means that no data is available on this concept for the given language. A vertical line (|) marks a (historic) morpheme boundary in a reflex, mostly to show fossilized morphology or that a certain part of the word is added later and not reconstructible to PFL or PMP. In contrast, a hyphen marks a morpheme boundary signalling active morphology or a morpheme boundary in reconstructions. Angle brackets < > mark infixes. Morpheme boundaries of an infix are marked differently to other morpheme boundaries to avoid confusion with two affixes. A dash in stem-final position means that person suffixes are usually attached to this word. A question mark (?) before a form means that it is not entirely clear if this form belongs to the set. Lexical items given in square brackets [] have been identified as lexical replacements or borrowings.

Chapter 6 contains tables with lists of reconstructions based on cognate sets. The reconstructions are given alphabetically according to the Flores-Lembata meaning and with a PMP source if applicable. The meaning of the PMP source is only indicated when diverging from the PFL meaning. If no PFL form can be reconstructed following the criteria set out in 4.3, I give a form with a hash tag (#) that could potentially be a PFL reconstruction (cf. §4.4.1). For some lexemes two forms are reconstructed. If they are separated by a slash (/), I assume that there was variation between the two forms in the proto-language. This variation is often related to (historic) affixation. If the two forms are separated by 'or', I mean that it is not possible to decide on the actual proto-form based on the current data. Parts of reconstructions may be represented in brackets to indicate that there was probably variation between two forms. An example is PFL *(v)uvuŋ 'ridge' which was probably realised as PFL *vuvuŋ and PFL *uvuŋ.

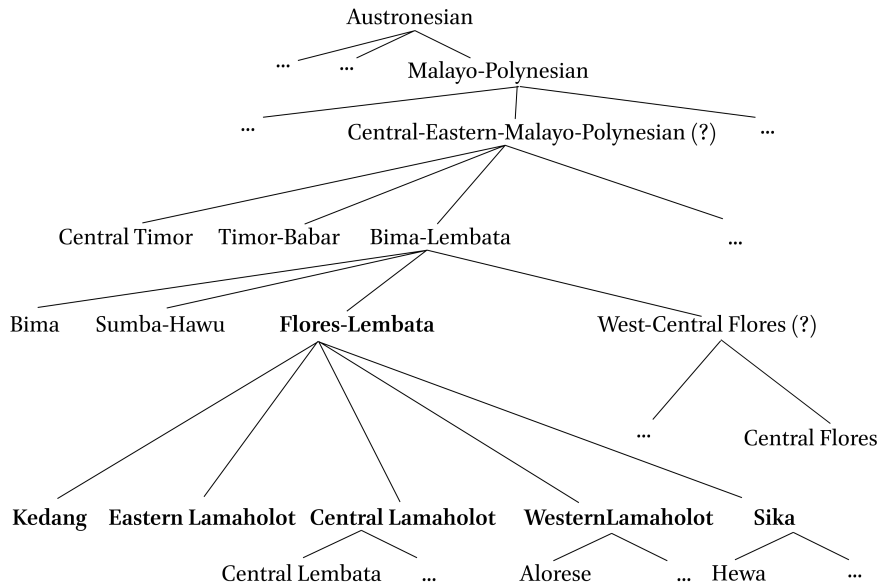
CHAPTER 5

Historical phonology and subgroupings

5.1 Introduction

This chapter presents a historical perspective of the Flores-Lembata phonology with the aim of reconstructing PFL phonemes and the establishment of shared innovations that support the five Flores-Lembata subgroups, as shown in Figure 5.1. The present study is the first attempt to establish the internal structure of the Flores-Lembata family based on exclusively shared sound changes, taking into account languages of all five subgroups in this family. As higher-level subgroups within Malayo-Polynesian are still debated (cf. §1.3.1 for more details), only a thorough bottom-up reconstruction can finally reveal the higher-level subgroups within the Malayo-Polynesian family. This chapter on historical phonology in the Flores-Lembata languages is a contribution to this endeavour.

Figure 5.1: Genealogical classification of the Flores-Lembata languages



I show in this chapter that the five Flores-Lembata subgroups are supported by exclusively shared sound changes and lexical innovations (cf. §5.3) and that there is no clear evidence for mid-level subgroups within the Flores-Lembata family (cf. §5.3.2). Evidence for Flores-Lembata as a subgroup is provided in §5.4 and evidence for a Bima-Lembata subgroup as an immediate ancestor of Flores-Lembata in §5.5. Evidence for the potential sister subgroups of Flores-Lembata and Bima-Lembata are not within the scope of this dissertation.

All three Lamaholot subgroups are internally diverse, and within each of these groups we can identify several languages. Some of these languages undergo further regular sound changes, such as PFL *s > h. However, these sound changes are not subgroup-defining, as they are very common changes that appear to have diffused between the languages. Sika and Kedang, in contrast, do not contain as much internal diversity. No further regular sound changes are attested on lower levels, with the exception of possible merger of final *n* and *ŋ* in western Sika varieties.

The internal structure of Flores-Lembata presented in Figure 5.1 above is in line with previous research by Fernandez (1996:174) who establishes a

Flores-Lembata subgroup (named *Flores Timur*) with three primary branches Sika, Lamaholot (i. e. Western Lamaholot) and Kedang. In contrast to Fernandez and the proposal here, Doyle (2010:30) groups Sika and Lamaholot (i. e. Western Lamaholot) together as a subgroup within Flores-Lembata. Previous subgroupings neither included Eastern Lamaholot nor Central Lamaholot varieties.

In addition to the subgrouping works by Fernandez and Doyle, a lexicostatistic analysis has been carried out by Keraf (1978a) with a focus on Lamaholot, but also including one Sika and one Kedang variety. Thus this work also takes into account varieties from all subgroups of Flores-Lembata but uses a different method, namely lexicostatistics, while the present study is based on shared innovations in form of sound changes. Based on 33 Lamaholot wordlists with 200 basic vocabulary items, Keraf's work distinguishes the three main groups of Lamaholot, which each share 55% of lexical similarity: Western Lamaholot, Central Lamaholot and Eastern Lamaholot (Keraf 1978a:Appendix VI). Keraf's work also shows that the Lamaholot subgroups are lexically closer to each other (55% lexical similarity) than to the other languages of Flores-Lembata, Sika and Kedang which only share about 30% of their vocabulary with the Lamaholot varieties. Elias (2017a) uses Keraf's lexical data to apply the comparative method (Campbell and Poser 2008) and historical glottometry (François and Kalyan 2018), with the aim of examining the internal subgrouping of the Lamaholot dialect chain. His findings of shared innovations also confirmed the main groupings into Western, Central and Eastern Lamaholot.

Table 5.1 and 5.2 show all sets of non-identical consonantal sound correspondences that are attested in the family of Flores-Lembata, following a top-down approach from PMP consonants. Other PMP consonants are retained unchanged in the Flores-Lembata languages. Some exceptions occur in final position. Two reflexes separated by a slash point to an unconditioned split, a split determined by the phonological environment or a split between two different varieties within the subgroup. Details on these splits are provided in the respective subsections.

Table 5.1: PMP obstruents and their non-identical reflexes

PMP	*k	*q	*h	*s		*b		*d/*z/*j		*z	*g
PFL	*k	*ʔ/∅	*ʔ/∅	*s	*h	*b	*v	*d	*l	*g	
Env.	#_ V_V #_VL										
SK	ʔ	ʔ/∅	∅	<i>h</i>	<i>h</i>	<i>b</i>	<i>v</i>	<i>r</i>	<i>r</i>	<i>l</i>	<i>g</i>
WL	<i>k</i>	ʔ/∅	ʔ/∅	<i>h/∅</i>	<i>h/∅</i>	<i>b</i>	<i>v</i>	<i>d</i>	<i>r</i>	<i>l/r</i>	<i>g</i>
CL	<i>k</i>	∅	∅	<i>s</i>	∅	<i>b</i>	<i>v</i>	<i>d</i>	<i>dʒ</i>	<i>l/r</i>	<i>g</i>
EL	ʔ/∅	ʔ/∅	ʔ/∅	<i>h</i>	<i>h</i>	<i>b</i>	<i>v</i>	<i>d</i>	<i>r</i>	<i>l/r</i>	<i>g</i>
KD	ʔ	ʔ/∅	ʔ/∅	<i>h/∅</i>	<i>h/∅</i>	<i>b</i>	<i>v</i>	<i>d</i>	<i>r/y/∅</i>	<i>l</i>	<i>k</i>

PMP *k [k], *q [q], *h [h], *s [s], *b [b], *d [d], *z [dʒ]/[ʝ], *j [g]/[ɣ]/[gʲ]

Table 5.2: PMP non-obstruents and their non-identical reflexes

PMP	*ŋ		*R	*y
Env.	#_	V_V		
PFL	*n	*ŋ	*r	*y
SK	<i>n</i>	<i>n</i>	<i>r</i>	<i>y/i</i>
KD	<i>n</i>	<i>ŋ</i>	<i>r</i>	<i>y/i/e</i>
CL	<i>n</i>	<i>ŋ</i>	<i>r</i>	<i>y/dʒ</i>
WL	<i>n</i>	<i>ŋ</i>	ʔ	<i>y/dʒ</i>
EL	<i>n</i>	<i>ŋ</i>	<i>r</i>	[...]

PMP *ŋ [ŋ], *R [r], *y [j]

The chapter is structured as follows. In §5.2, I establish regular sound correspondences between Flores-Lembata lexical items that go back to Proto-Malayo-Polynesian (PMP) reconstructions. That the sound correspondences in these lexical items are mostly regular is a sign for inheritance from a common ancestor (Campbell and Poser 2008:172). This means that the Flores-Lembata languages descend from PMP and ultimately from Proto-Austronesian. PMP is taken as a point of reference as this is the ancestor of all Austronesian languages that are found outside of the island Taiwan and because a large amount of reconstructions are available for PMP in Blust

and Trussel (2010). All PMP reconstructions in this dissertation are taken from this source (cf. §4.2). In §5.3, I provide evidence for the lowest level subgroups within the Flores-Lembata family (§5.3.1), but I also show that there is little evidence for mid-level groups within the Flores-Lembata family (§5.3.2). In §5.4, I provide evidence for Flores-Lembata as an innovation-defined subgroup. In addition, in §5.5, I show that Flores-Lembata is part of a bigger Bima-Lembata subgroup encompassing also languages of Bima, Sumba and Western and Central Flores.

5.2 Reflexes of PMP sounds

5.2.1 PMP voiceless stops *p, *t, *k and *q

5.2.1.1 Initial and intervocalic position

The Proto-Malayo-Polynesian (PMP) root-initial and intervocalic voiceless stops are relatively well-preserved in all languages of Flores-Lembata, except for PMP *q which changed into a glottal stop in Proto-Flores-Lembata (PFL) or earlier. For a discussion on root-final stops see §5.2.1.3. Table 5.3 shows the PMP voiceless stops with their reflexes in Flores-Lembata. Central Lamaholot and Western Lamaholot prove to be most conservative in this respect because these languages not only retain PMP *p and *t but also *k. In Sika, Kedang and Eastern Lamaholot, there is a sound change from PMP *k > ʔ/Ø. PMP *q is reflected as ʔ or Ø in all Flores-Lembata languages. Central Lamaholot is the only language that subsequently lost ʔ completely, thus does not show any reflex of PMP *q.

Table 5.3: Reflexes of PMP voiceless stops *p, *t, *k and *q

Env.	#_	V_V	#_	V_V	#_	V_V	#_	V_V
PMP	*p-	*-p-	*t-	*-t-	*k-	*-k-	*q-	*-q-
PFL	*p-	*-p-	*t-	*-t-	*k-	*-k-	*ʔ-	*-ʔ-
SK	<i>p</i>	<i>p</i>	<i>t</i>	<i>t</i>	<i>ʔ/∅</i>	<i>ʔ</i>	<i>ʔ/∅</i>	<i>ʔ</i>
WL	<i>p</i>	<i>p</i>	<i>t</i>	<i>t</i>	<i>k</i>	<i>k</i>	<i>ʔ/∅</i>	<i>∅</i>
CL	<i>p</i>	<i>p</i>	<i>t</i>	<i>t</i>	<i>k</i>	<i>k</i>	<i>∅</i>	<i>∅</i>
EL	[...]	<i>p</i>	<i>t</i>	<i>t</i>	<i>∅</i>	<i>ʔ</i>	<i>∅</i>	<i>∅</i>
KD	<i>p</i>	<i>p</i>	<i>t</i>	<i>t</i>	<i>ʔ/∅</i>	<i>ʔ/∅</i>	<i>ʔ/∅</i>	<i>∅</i>

Below I present cognate sets with regular reflexes of PMP voiceless stops in the languages of Flores-Lembata. Table 5.4 and 5.5 provide examples of reflexes of PMP *p in the languages of Flores-Lembata. The PMP stop *p is regularly reflected as *p* in all languages of Flores-Lembata for which data is attested. For Eastern Lamaholot, there is no data for reflexes of initial PMP *p-.

Table 5.4: Reflexes of initial PMP *p-

PMP	*pitu	*palu	*pusəj	*sa-ŋa-puluq	*piliq
PFL	*pitu	*palu-k	*pusər	*s-pulu	*piliʔ
SK	<i>pitu</i>	-	<i>puher</i>	<i>pulu</i>	(?) <i>liʔi</i>
WL (LWI)	<i>pito</i>	-	<i>kə puhur</i>	<i>pulo</i>	<i>pileʔ</i>
CL (KK)	<i>pito</i>	<i>palu k</i>	<i>kə pusər</i>	<i>s pulo</i>	<i>pili</i>
EL	[...]	[...]	[...]	[...]	[...]
KD	<i>pitu</i>	<i>palu ʔ</i>	<i>puhɛː</i>	<i>pulu</i>	<i>pil</i>
	‘seven’	‘hit’	‘navel’	‘ten’	‘choose’

Table 5.5: Reflexes of intervocalic PMP *-p-

PMP	*hapuy	*sapu	*apa	*ipən	*ma-kapal	*kiput
PFL	*api	*hapu	*apa	*ipə-	*m-kapal	*kiput
SK	<i>api</i>	<i>hapus</i>	<i>apa</i>	-	<i>apar</i>	-
WL (LML)	<i>ape</i>	<i>hapu</i>	-	<i>ipã</i>	-	-
WL (DL)	<i>ape</i>	-	-	<i>ipə</i>	-	<i>kipu?</i>
WL (MS)	<i>ape</i>	<i>hapo</i>	<i>apai</i>	-	<i>gapal</i>	<i>kupo?</i>
CL (LT)	<i>ape ru</i>	<i>apu dza</i>	-	-	-	-
CL (LR)	<i>ape r</i>	-	<i>apoi</i>	<i>pai</i>	-	<i>k<ən>ipot ən</i>
EL	<i>ape</i>	<i>hapu</i>	<i>ape</i>	<i>ipe</i>	-	-
KD	<i>api</i>	-	<i>ape</i>	-	<i>kapal</i>	<i>ipe?</i>
	‘fire’	‘wipe’	‘what’	‘teeth’	‘thick’	‘narrow’

Table 5.6 provides cognate sets that show the regular reflexes of PMP *t. In initial and intervocalic position, PMP *t is regularly reflected as *t* in all Flores-Lembata languages. In Kedang *va?* ‘stone’ < PMP *batu in Table 5.6, the final syllable is lost. Therefore, the reflex of intervocalic PMP *t appears now in final position which causes the change of PMP *t > ? which is regular in Kedang in final position (cf. §5.2.1.3).

Table 5.6: Reflexes of initial and intervocalic PMP *t

PMP	*talih	*taŋis	*qa-təluR	*batu	*m-atay	*qutin
PFL	*tali	*tani	*təlur	*vatu	*matay	*uti
SK	<i>tali</i>	<i>tani</i>	<i>təlo n</i>	<i>vatu</i>	<i>mate</i>	<i>uti</i>
WL (LI)	<i>tale</i>	<i>taniŋ</i>	<i>telu</i>	<i>vato</i>	<i>mata</i>	-
WL (AD)	<i>tale?</i>	<i>tani</i>	<i>telu k</i>	<i>vato</i>	<i>mata:</i>	<i>ute</i>
CL (KK)	<i>tali</i>	-	<i>təlu k</i>	<i>vatu</i>	<i>matadz</i>	<i>uti</i>
EL	<i>tale</i>	<i>tani</i>	<i>təlũ</i>	<i>vato</i>	<i>mata</i>	[...]
KD	-	-	<i>tolor</i>	<i>va?</i>	<i>mate</i>	<i>vuti</i>
	‘rope’	‘cry’	‘egg’	‘stone’	‘die’	‘penis’

Table 5.7 and 5.8 show cognate sets containing reflexes of PMP *k. Sika has changed PMP *k > ? in all positions. In initial position, a few instances of

loss of PMP *k are observed but they may relate to transcriptions issues as not all researchers consistently transcribed initial glottal stops. Also in Kedang and Eastern Lamaholot, PMP *k > ʔ or Ø. This change of PFL *k > ʔ in Sika, Kedang and Eastern Lamaholot is not entirely complete, there are sporadic retentions of PFL *k with the reflex *k*. Central Lamaholot and Western Lamaholot always retain PMP *k as *k*. The Lamaholot reflexes of PMP *aku in Table 5.8 have an irregular initial *g* and also the Kedang form contains irregular vowel changes. For a discussion of these forms see §6.2.4.

Table 5.7: Reflexes of initial PMP *k-

PMP	*kutu	*kahiw	*kaən ^a	*kami	*kulit
PFL	*kutu	*kayu	*kaan	*kami	*kulit
SK	ʔutu	ʔai	ʔaa	ʔami	ʔulit
WL (LWL)	kuto	kadzɔ	kãã	kame	kulit
CL (KK)	kutu	kadzɔ	kaa	kame	-
EL	[...]	-	aa	ame	ulit ã
KD	ʔutu	ʔai	aa	(k)e:	-
	‘headlice’	‘tree; wood’	‘3PL.eat’	‘1PL.EXCL’	‘skin’

^a The original transcriptions of the reflexes in this set contained single vowels.

Table 5.8: Reflexes of intervocalic PMP *-k-

PMP	*ikuR	*hikan	*takaw	*aku
PFL	*ikur	*ikan	*t<əm>akav	*aku
SK	iʔur	iʔan	? toʔi	aʔu
WL (LI)	ikuʔ uŋ	ikaŋ	təmaKa	? go:
WL (LML)	iku	ikã	[...]	? go e
CL	ikur	ikan	takav	? go ne
EL	iʔũ	iʔã	[...]	? go ʔe
KD	-	iʔa	maʔo	? εʔi
	‘tail’	‘fish’	‘steal’	‘1SG’

Table 5.9 and 5.10 provide cognate sets with reflexes of PMP *q. The reflex

of PMP *q is either zero or a glottal stop in the languages of Flores-Lembata. I suggest that PMP *q regularly changed into glottal stop in PFL or even earlier. Subsequently, this glottal stop was lost on an irregular basis in the daughter languages. The irregular pattern between zero and glottal stop may to a certain extent result from transcription issues, especially in initial and final position. Researchers may not always have distinguished an onsetless syllable from a syllable with an initial glottal stop. Central Lamaholot has lost glottal stops completely, so all reflexes of PFL *ʔ < PMP *q are zero.

Table 5.9: Reflexes of initial PMP *q-

PMP	*qapuR	*quzan	*qabu	*quma	*quay
PFL	*ʔapur	*udan	*ka-ʔavu	*uma	*uay
SK	<i>ʔapur</i>	<i>uran</i>	<i>avu</i>	<i>uma</i>	<i>ue</i>
WL (AD)	<i>ʔapuʔ</i>	<i>ʔuraŋ</i>	<i>kə ʔavu k</i>	<i>mã</i>	<i>ʔua</i>
WL (LWI)	<i>apuʔ</i>	<i>uraŋ</i>	<i>kə ʔavu</i>	<i>maŋ</i>	<i>uvay əŋ</i>
CL	<i>apur</i>	<i>udʒan</i>	<i>k avo k</i>	-	<i>uadz</i>
EL	[...]	<i>ura</i>	[...]	[...]	[...]
KD	<i>apur</i>	<i>uya</i>	<i>ava ʔ</i>	<i>ʔ lumar</i>	<i>vua vɛi</i>
	‘lime’	‘rain’	‘ash’	‘gardenfield’	‘rattan’

Table 5.10: Reflexes of intervocalic PMP *-q-

PMP	*ma-paqit	*taqi	*waqay
PFL	*m-paʔit	*taʔi	*vaʔi
SK	<i>baʔit</i>	<i>taʔi</i>	<i>vaʔi</i>
WL (AD)	<i>pait</i>	<i>tae</i>	-
WL (MS)	<i>paiʔ</i>	<i>taɪŋ</i>	-
CL	<i>pait</i>	<i>tai</i>	-
EL	[...]	[...]	-
KD	<i>pɛiʔ</i>	-	-
	‘bitter’	‘excrements’	‘foot; leg’

The PMP stop *q is reflected as ʔ or zero in the languages of Flores-Lembata.

Therefore, I reconstruct PFL *ʔ in all positions as a reflex of PMP *q. However, if only WL-Adonara shows a ʔ as a potential reflex of PMP *q, I do not reconstruct it to PFL because WL-Adonara seems to insert glottal stops before every initial vowel. The alternation between zero and glottal stop in the reflexes in Sika, Western Lamaholot, Eastern Lamaholot and Kedang appears to be irregular with zero being the norm and glottal stop the exception. For Sika, there is an apparent regularity which favours glottal stop in intervocalic position and zero in initial position for reflexes of PMP *q.

5.2.1.2 Initial stops with PMP *ma-

There are instances of PFL *m directly preceding root-initial stops. In most cases this initial *m goes back to the PMP stative prefix *ma-. As all words that reflect PMP *ma- are stative concepts in the languages of Flores-Lembata, it is likely that the function of the initial *m in PFL was still retained as marking stative. In the modern languages this initial *m has merged with the following consonant in different ways. In the Sika reflex *nauʔ* ‘tinea’, the prefix appears to be lost. Table 5.11 provides examples of reflexes of PFL *m-p.

Table 5.11: Reflexes of initial PMP *ma-p

PMP	*ma-paqit	*ma-pənuq	*ma-pəju	*panaw
PFL	*m-paʔit	*m-pənu	*m-pədu	*m-panau
SK	<i>baʔit</i>	<i>bənu</i>	<i>bəru</i>	<i>nauʔ</i>
WL (AD)	<i>pait</i>	<i>pəno</i>	<i>pəro</i>	<i>mənao</i>
WL (LWI)	<i>pai k</i>	<i>mənu ŋ</i>	<i>pəro</i>	<i>mau</i>
CL	<i>pait</i>	<i>mənu k</i>	<i>pədʒu</i>	-
EL	[...]	[...]	[...]	[...]
KD	<i>pɛiʔ</i>	<i>pənu</i>	<i>pɛyu</i>	-
	‘bitter’	‘full’	‘salty’	‘tinea’

In Sika, the merging of PFL *m-p resulted in voicing of the initial consonant and dropping of initial *m, thus PFL *m-p > Sika *b*. In the Central and Western Lamaholot data, either the nasal or the stop was kept and the other was dropped completely. This apparent irregularity of keeping either the

nasal or the stop in Central and Western Lamaholot is most likely related to two functionally different forms of many property nouns that are attested in Lamaholot varieties: a base form and a derived form (cf. §8.3.3.2). In the table, the reflexes of words denoting property ('adjectival') concepts, such as 'bitter', 'full' and 'salty', are either given in their base form, with the initial *p* going back to a form without prefix, or in their derived form, with initial *m* going back to a form with prefix **m*- that replaced the initial *p*, thus PFL **m*-*p* > *m*. The change of PFL **m*-*p* > *m* appears to be regular, as it is also attested in the reflexes of PFL **m*-panaw where PFL **m*-*p* > WL *m*. In addition, to interpret the CL and WL forms with initial *m* as derived property nouns is supported by the initial presence of additional final consonants that are related to derivational processes -*k* and -*ŋ* (cf. §3.3.6.1).

The same development is observed in reflexes of PFL **m*-*t* given in Table 5.12. In Sika **m*-*t* is reflected as *d*. In the other languages either *t* or *m* is retained. In Kedang, sporadic voicing of the consonant occasionally occurs, such as in PFL **m*-tidəm > Kedang *dɛyeʔ* 'sharp'.

Table 5.12: Reflexes of initial PMP **ma*-*t*

	PMP * <i>ma</i> - <i>tasak</i>	* <i>mantalaq</i> 'Venus'	* <i>ma</i> - <i>tuquR</i>	* <i>ma</i> - <i>tuqah</i>	* <i>tazim</i> 'whet'
PFL	* <i>m</i> - <i>tasak</i>	* <i>mətala</i>	* <i>m</i> - <i>tuʔur</i>	* <i>m</i> - <i>tuʔa</i>	* <i>m</i> - <i>tidəm</i>
SK	<i>dahaʔ</i>	<i>dala</i>	<i>duʔur</i>	<i>duʔa n</i>	<i>diran</i>
WL	<i>tahak</i>	<i>pə tala</i>	<i>tuʔu k</i>	<i>t<en>uʔe</i>	-
CL	<i>tasak</i>	-	-	<i>tua</i>	-
EL	[...]	<i>malã</i>	-	[...]	-
KD	<i>taʔɛn</i>	<i>malɛ tala</i>	? <i>turi</i> [<i>roka</i>]	<i>tua</i>	<i>dɛyeʔ</i>
	'ripe'	'star'	'dry'	'old'	'sharp'

Reflexes of PMP **ma*-*k* and **ma*-*q* are given in Table 5.13. For PMP **ma*-*k*, the loss of the prefix is observed in Sika, while in WL-Alorese, it is merged with the initial consonant resulting in *g*. The Kedang form *kapal* could be a result of PMP **ma*-*k* > *g* > *k*. The first change is the same as attested in Alorese and the change of **g* > *k* is regular in Kedang (cf. §5.2.1.1). In reflexes of **ma*-*q*, the *m* of the prefix is retained while the stop is lost. There is no evidence to reconstruct a reflex of initial PMP **q* to PFL for the words starting with a reflex of **ma*-*q*.

Table 5.13: Reflexes of initial PMP *ma-k and *ma-q

PMP	*ma-kapal	*ma-qitəm	*ma-qudip
PFL	*m-kapal	*mitəm	*modip
SK	<i>ʔapar</i>	<i>mitan</i>	<i>more t</i>
WL (AL)	<i>gapal</i>	<i>mitəŋ</i>	<i>mori k</i>
WL (LWI)	-	<i>mitəŋ</i>	<i>mori</i>
CL	-	<i>mitəm</i>	<i>modʒip</i>
EL	-	<i>mitã</i>	<i>mori</i>
KD	<i>kapal</i>	<i>mitəŋ</i> ‘black; dirty’	-
	‘thick’	‘black’	‘alive’

5.2.1.3 Final stops PMP *p, *t, *k and *q

Sika and Kedang change PMP *-k > -ʔ and PMP *-q becomes ʔ or Ø in PFL. The conditioning of the split of PMP *-q > ʔ/Ø remains unclear. Further, PFL *-ʔ (< PMP *-q) is sporadically lost in most FL languages and lost completely in Central Lamaholot. In addition to the sound changes already observed for initial and medial position, Kedang changes final PMP *t > ʔ, while initial and intervocalic PMP *t is retained in Kedang. Alorese loses all final consonants, except for the sporadic retention of *k*. Central Lamaholot retains all final consonants unchanged except for PMP *-q which is lost. Further, sporadic loss of final consonants is attested in all Flores-Lembata languages. In Western Lamaholot, the consonant is also sporadically changed into glottal stop instead of being lost completely.

Table 5.14: Reflexes of final PMP *-p, *-t, *-k and *-q

PMP	*-p	*-t	*-k	*-q
PFL	*-p	*-t	*-k	*-ʔ/∅
SK	∅	t/∅	ʔ/∅	ʔ/∅
WL	p/∅	t/ʔ/∅	k/ʔ/∅	ʔ/∅
WL (AL)	∅	∅	k/ʔ/∅	∅
CL	p	t	k	∅
EL	∅	t/∅	ʔ/∅	[...]
KD	[...]	ʔ/∅	ʔ	ʔ/∅

The evidence for reflexes of final PMP *-p is scarce as only one cognate set, given in Table 5.15 could be found. The Sika reflex *more|t* ‘alive’ most likely contains an attributive suffix (cf. §8.3.3.1).

Table 5.15: Reflexes of final PMP *-p

PMP	*ma-quḍip
PFL	*modip
SK	<i>more t</i>
WL (LWI)	<i>mori</i>
WL (KO)	<i>morip</i>
CL	<i>modzip</i>
EL	<i>mori</i>
KD	-
	‘alive’

Table 5.16 lists examples of cognate sets reflecting final PMP *-t. There is a regular change of final PMP *-t > ʔ in Kedang.

Table 5.16: Reflexes of final PMP *-t

PMP	*Ramut	*lipət	*ma-paɣit	*ma-bəRəqat	*kʉlit
PFL	*ramut	*ləpət	*m-paʔit	*bərat	*kʉlit
SK	<i>ramut</i>	<i>ləpet</i>	<i>baʔit</i>	<i>bərat</i>	<i>ʔulit</i>
WL (LWI)	<i>ramu</i>	<i>ləpə k</i>	<i>pai k</i>	<i>baʔa</i>	<i>kuli</i>
WL (AD)	<i>ʔamut</i>	<i>ləpət</i>	<i>pait</i>	<i>baʔat</i>	-
WL (MS)	<i>ramu k</i>	<i>ləpeʔ</i>	<i>paiʔ</i>	<i>baʔ</i>	<i>kuli k</i>
CL	<i>ramut</i>	<i>ləpət</i>	<i>pait</i>	<i>bərat</i>	-
EL	<i>ramut ã</i>	[...]	[...]	<i>bəra</i>	<i>ulitã</i>
KD	<i>ramuʔ</i>	<i>ləpiʔ</i>	<i>peiʔ</i>	<i>baraʔ</i>	-
	‘root’	‘fold’	‘bitter’	‘heavy’	‘skin’

Table 5.17 lists reflexes of PMP final *k. PMP *k regularly changes into glottal stop in Sika, Kedang and Eastern Lamaholot. Sporadic PMP *k > ʔ is attested in final position in Western Lamaholot and sporadic loss of final *k is attested in all languages, except for Central Lamaholot. The Sika form *anak* ‘small’ appears to retain an irregular final *k*. However, the final *k* could also be a suffix found on property nouns (cf. §8.3.3).

Table 5.17: Reflexes of final PMP *-k

PMP	*manuk	*anak	? *təbək
PFL	*manuk	*anak	*tubak
SK	<i>manuʔ</i>	<i>anak</i>	-
WL (AD)	<i>manuk</i>	<i>ʔanaʔ</i>	-
WL (LWI)	<i>manu</i>	<i>anaʔ</i>	-
CL	<i>manuk</i>	<i>anak</i>	<i>tubak</i>
EL	<i>manuʔ</i>	<i>ana</i>	<i>tuba</i>
KD	<i>manuʔ</i>	<i>anaʔ</i>	<i>tubaʔ</i>
	‘chicken; bird’	‘child; small’	‘stab’

Table 5.18 provides reflexes of PMP *q in final position. For some words, especially those expressing properties, no reflex of final PMP *q can be reconstructed to PFL. An example is PFL *doa ‘far; long’ (< PMP *zauq). For other

words, final PMP *q is reflected as ʔ in PFL and retained as such in Sika and Kedang.

Table 5.18: Reflexes of final PMP *-q

PMP	*zauq	*mamaq	*budaq	*budaq	*salaq
PFL	*doa	*mamaʔ	*budaʔ	*vuda	*sala
SK	-	maʔmaʔ	buraʔ	vura n	hala
WL (AD)	doã	-	buraʔ	vurhã	n alã
WL (MS)	doa ŋ	mame	bura k	? bura ŋ	hala ŋ
CL	doa	mamo t	budʒa k	pə vudʒa	s<n>ala k
EL	doa	[...]	burõ	[...]	[...]
KD	doa	mameʔ	buyaʔ	vura n	-
	‘far; long’	‘chew’	‘white’	‘foam’	‘wrong’

In many cases, the final consonant appears to be replaced by a suffix, such as *-k*, *-n* or *-ŋ* (cf. §3.3.6.2 and §8.3.3). However, synchronically these suffixes are often interpreted as part of the root and thus can be regarded as fossilised. The final *t* in the Central Lamaholot verb *mamot* ‘chew’ in Table 5.18 remains unexplained. There is a suffix *-t* attested on property nouns, such as in Sika *moret* ‘alive’ (cf. Table 5.15) but as ‘chew’ is a verb, no such suffix is expected.

5.2.2 PMP voiced obstruents *b, *d, *z, *j and *g

5.2.2.1 Overview

In this section, I discuss the Flores-Lembata reflexes of PMP *b, *d, *z [dʒ/jj], *j [g/gʲ/ɣ] and *g. The three voiced PMP stops *b, *d and *g have the phonetic values of [b, d, g]. PMP *b and PMP *d appear in all positions, while PMP *g is rather rare in general and does not occur in final position. PMP *j [g/gʲ/ɣ] is only found in intervocalic and final positions. PMP *z [dʒ/jj] is only found in initial and intervocalic positions. For more information on the phonetic values of the PMP sounds *j and *z see §4.4.

Table 5.19 is an overview of sound correspondences going back to the PMP voiced obstruents. PMP *b unconditionally splits into PFL *b and *v

in initial position and in intervocalic position, all instances of final PMP *-b > PFL *-v. An unconditioned split here means that the conditioning of this split remains unknown.

PMP *d, *z and *j have merged into PFL *d for initial and intervocalic position. In final position PMP *d > PFL *r and PMP *j > PFL *y. Initial PMP *z followed by an intervocalic liquid becomes *l* or in some Lamaholot varieties also *r*. Only very few reflexes of PMP *g are found in the languages of Flores-Lembata. These cognates show a regular change of PMP *g > *k* in Kedang.

Table 5.19: Reflexes of PMP *b, *d, *g, *j [g/gʲ/ɣ] and *z [dʒ/jj]

Env.	#	#	V_V	#	#_VL	#	V_V	#	#	#	V_V
PMP		*b		*b	*z	*d/*z	*d/*z/*j	*d	*j	*g	*g
PFL		*b	*v	*b	*l	*d	*d	*r	*y	*g	*g
SK	<i>b</i>	<i>v</i>	<i>b</i>	<i>v</i>	<i>l</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>Ø</i>	<i>g</i>	<i>g</i>
WL	<i>b</i>	<i>v</i>	<i>b</i>	<i>v</i>	<i>l/r</i>	<i>d</i>	<i>r</i>	<i>r/Ø</i>	<i>Ø</i>	<i>g</i>	<i>g</i>
CL	<i>b</i>	<i>v</i>	<i>b</i>	<i>v</i>	<i>l/r</i>	<i>d</i>	<i>dʒ</i>	<i>r</i>	<i>dʒ/Ø</i>	<i>g</i>	<i>g</i>
EL	<i>b</i>	<i>v</i>	<i>b</i>	<i>v</i>	<i>l/r</i>	<i>d</i>	<i>r</i>	[...]	[...]	<i>g</i>	<i>g</i>
KD	<i>b</i>	<i>v/Ø</i>	<i>b</i>	<i>v</i>	<i>l</i>	<i>d</i>	<i>y/r/Ø</i>	<i>r</i>	<i>Ø</i>	<i>k</i>	<i>k</i>

In the following, the split of PMP *b, the merger of PMP *d/*z/*j and the retention of PMP *g are discussed in more detail.

5.2.2.2 Split: PMP *b > PFL *b/*v

Unconditional PMP *b > PFL *b/*v is attested in initial position, as shown in Table 5.20. In intervocalic position, PMP *b is less frequent and all intervocalic PMP *b > PFL *v, as shown in Table 5.21.

There is evidence that PMP *b > *b/*v already occurred on a higher level than PFL. PMP *b > *b/*v is evidence for a Bima-Lembata subgroup including Flores-Lembata, Flores, Bima and Sumba-Hawu. This is discussed in more detail in §5.5.

PFL *b remains *b* in all daughter languages and PFL *v remains *v* in all daughter languages, except for Kedang where PFL *v > *Ø* before *u*. An example is Kedang *ua|n* ‘fruit’ (< PFL *vua < PMP *buaq). The change appears

to be incomplete. Thus, for example, KD-Leubatang has *vura|n* ‘foam’ (< PFL *vuda < PMP *bujəq) with retention of *ʋ*, but also *ua|n* ‘fruit’ (< PFL *vua < PMP *buaq) with loss of *ʋ*. There is one apparent instance of intervocalic PFL *v > h in Kedang: Kedang *tehu* ‘sugarcane’ < PFL *təvu < PMP *təbuh.

Table 5.20: Reflexes of initial PMP *b-

PMP	*b-	*babuy	*batu	*buaq	*b-	*bayu	*budaq
PFL	*v-	*vavi	*vatu	*vua-n	*b-	*bayu	*buda?
SK	ʋ	<i>vavi</i>	<i>vatu</i>	<i>vua n</i>	<i>b</i>	<i>bai</i>	<i>bura?</i>
WL (LWI)	ʋ	<i>vave</i>	<i>vato</i>	<i>vua ŋ</i>	<i>b</i>	<i>bayo</i>	<i>bura</i>
CL	ʋ	<i>vavi</i>	<i>vatu</i>	<i>vua k</i>	<i>b</i>	<i>badzu</i>	<i>budza k</i>
EL	ʋ	[...]	<i>vato</i>	<i>vuʔã</i>	<i>b</i>	[...]	<i>burʔã</i>
KD	ʋ	<i>vavi</i>	<i>va?</i>	<i>ua n</i>	<i>b</i>	<i>bae</i>	<i>buya?</i>
		‘pig’	‘stone’	‘fruit’		‘pound’	‘white’

Table 5.21: Reflexes of intervocalic PMP *-b-

PMP	*-b-	*qabu	*təbuh
PFL	*-v-	*ka-ʔavu	*təvu
SK	ʋ	<i>avu</i>	<i>təvu</i>
WL (LWI)	ʋ	<i>kə ʔavu</i>	<i>təvo</i>
CL (KK)	ʋ	<i>k avo k</i>	<i>təvu</i>
CL (LR)	ʋ	<i>k avo k</i>	<i>təvo r</i>
EL	?	[...]	[...]
KD	ʋ	<i>ava?</i>	<i>təhu</i>
		‘ash’	‘sugarcane’

In the southern Lembata varieties of Central Lamaholot, such as CL-Imulolo, and Western Lamaholot, such as WL-Lamalera, a consecutive sound change of PFL *v > f is found. The same change is found in the Alorese varieties Alor Besar and Baranusa.

Evidence for final PMP *-b is rare, only one cognate set given in Table 5.22 is attested in my database where PMP *-b > PFL *-v.

Table 5.22: Reflexes of final PMP *-b

	‘yawn’
PMP	*ma-huab
PFL	*muav
SK	<i>moaʔ</i>
KD	<i>moa</i>
CL	<i>kə muav</i>
WL	<i>pə moa</i>
EL	[...]

Most of the FL languages, except for Central Lamaholot, lose PFL *v in final position. In Sika, final PFL *v is reflected as ʔ instead of being lost.

5.2.2.3 Merger: Initial and intervocalic PMP *d/*j/*z > PFL *d

Initial and intervocalic PMP *d, *j and *z have merged in PFL as they have the same reflexes in all Flores-Lembata languages. As initial PMP *d- is retained unchanged in Kedang and Lamaholot, I suggest that PMP *d, *j and *z merged into PFL *d. In §5.2.2.4, I show why I propose PFL *d not only for initial position but also for intervocalic position, although in the modern languages of Flores-Lembata no intervocalic *d* is attested. Final PMP *-d and *-j did not merge, but PMP *-d > PFL *-r and PMP *-j > PFL *-y.

Table 5.23 provides cognate sets that contain reflexes of initial PMP *d- and PMP *z- that merged into PFL *d-. Initial PFL *d- is reflected as *r* in Sika, whereas it is retained unchanged in Lamaholot and Kedang.

Table 5.23: Reflexes of initial PMP *d- and *z- [dʒ/ʝ]

PMP	*dəŋəR	*diRi	*dahun	*zəkət	*zauq	*zaqit	*zaRum	‘needle’
PFL	*dəŋər	*diri	*doun	*deket	*doa	*daʔit	*daru	
SK	<i>rəna</i>	-	<i>roun</i>	-	-	<i>raʔit</i>	-	
WL (LWI)	<i>veŋe</i>	<i>deʔiŋ</i>	-	-	<i>doa ŋ</i>	-	<i>dau ŋ</i>	
WL (AD)	<i>deŋeʔ</i>	<i>deʔi</i>	-	-	<i>doã</i>	-	-	
CL	<i>dəŋər</i>	<i>diri</i>	-	-	<i>doa</i>	-	-	
EL	<i>dəŋe</i>	<i>diri</i>	-	<i>deʔe</i>	<i>doa</i>	-	-	
KD	<i>deŋər</i>	<i>? ma dər</i>	-	<i>? duʔ</i>	<i>doa</i>	-	-	
	‘hear’	‘stand’	‘leaf’	‘burn’	‘far; (fields)’long’	‘sew’	‘sew’	

The word for ‘sew’ comes from two PMP sources: *zaqit ‘sew’ > Sika *raʔit* and *zaRum ‘needle’ > WL-Lewoingu *dau|ŋ* ‘sew’. This is a case of semantic change from ‘needle’ to ‘sew’. The vowel change PMP *au > PFL *oa in the reflexes of PMP *zauq ‘far; long’ cannot be explained so far. Also the apparent change of PFL *d > ʋ in WL-Lewoingu *veŋe* ‘hear’ is unexpected and cannot be explained further at the current stage.

Reflexes of initial PMP *z- or *d- in words that contain an intervocalic liquid PMP *-l- or PMP *-R- behave exceptionally. In these words PMP *z or *d is not reflected as PFL *d but as PFL *l due to the assimilation to the following liquid in the word. Table 5.24 provides the three cases found in my data set.

Table 5.24: Reflexes of initial PMP *z-/*d- before intervocalic liquid

PMP	*zalan	*duRi ‘thorn; fish bone’	*daləm
PFL	*lalan	*luri	*laləm
SK (MM)	<i>lala</i>	<i>luri n</i>	-
SK (HW)	<i>lara</i>	<i>luri n</i>	-
WL (AD)	<i>rarã</i>	<i>riʔũ:</i>	-
WL (LWI)	<i>lara ŋ</i>	<i>riʔu</i>	-
WL (LWT)	<i>larã</i>	<i>riʔu</i>	-
WL (BN)	-	<i>ruʔi ŋ</i>	-
CL	<i>lalan</i>	<i>riu k</i>	-
EL	<i>lara</i>	<i>riʔũ</i>	-
KD	<i>lala</i>	<i>luri n</i>	<i>laɛ ŋ</i>
	‘road’	‘bone’	‘inside’

For PFL *lalan (> PMP *zalan), Sika, Kedang and Central Lamaholot show regular reflexes. In the WL varieties, the medial *l* has been changed into *r* and for some varieties, this even effected the initial *l* to change into *r*. In SK-Hewa, this change is also found. I suggest that this happened due to influence of the neighboring WL variety Lewotobi which has *larã* ‘road’.

The reflexes of PMP *duRi ‘thorn, splinter, fish bone’ in Sika and Kedang are regular. The Lamaholot forms need more explanation. WL-Baranusa, a variety of Alorrese, is most conservative in this form. It retains the order of the vowels, first *u* then *i*. The change of PFL *r > ʔ is regular. After the sound change of PFL *r > ʔ in Western Lamaholot, the initial *l* is changed into *r*, as has also been observed in WL-Adonara *rarã* ‘road’. In all other Western Lamaholot varieties, metathesis of the two vowels has occurred in addition to these changes. The Central Lamaholot form *riuk* ‘bone’ (< PFL *luri < PMP *duRi) is irregular as usually PFL *-r- = *r* in Central Lamaholot. In the word *riuk*, the intervocalic PFL *-r- appears to be deleted. I propose that CL *riuk* is a loan from WL *riʔu* with the addition of the suffix *-k* and the deleting of the glottal stop which is regular in Central Lamaholot.

Intervocalic PFL *-d- (< PMP *d *z *j) undergoes several different changes in individual languages as shown in Table 5.25. Central Lamaholot reflects PMP *-d- as *dʒ* intervocalically, whereas Sika, WL and EL reflect PMP

*-d- as *r* in intervocalic position. In Kedang, intervocalic PMP *-d- > *y* which is then sometimes reduced to zero. It is very likely that Kedang went through the intermediate stage of PFL *-d- > *-dʒ- before > *y* or Ø. There is evidence from loanwords, such as *yendela* ‘window’ from Indonesian *dʒendela* and *yadi* ‘become; happen’ from Indonesian *dʒadi*, that Kedang *y* in initial position comes from an earlier *dʒ* (Samely and Barnes 2013:712). So, it could be that initial *dʒ* and intermediate *dʒ* (from PFL *-d) became *y*. This is the same change as attested in Central Lamaholot which underwent PFL *-d- > PCL *-dʒ-. Also, in some varieties of Central Lembata, intervocalic *dʒ* is weakened sporadically to *y*, such as in CL-Lerek *nayan* ‘name’ from Proto-Central-Lamaholot *-nadʒan ‘name’.

Table 5.25: Reflexes of intervocalic PMP *-d-, *-z- and *-j-

PMP	*ma-qudip	*budaq	*quzan	*tazim ‘whet’	*ŋajan	ŋijun
PFL	*modip	*budaʔ	*udan	*m-tidəm	*nadan	*(n)iduŋ
SK	<i>moret</i>	<i>buraʔ</i>	<i>uran</i>	<i>diran</i>	<i>naran</i>	<i>iru</i>
WL (LWI)	<i>mori</i>	<i>bura</i>	<i>uraŋ</i>	-	<i>naraŋ</i>	<i>iru ŋ</i>
WL (AD)	<i>mori t</i>	<i>buraʔ</i>	<i>ʔuraŋ</i>	-	<i>narã</i>	<i>iru net</i>
CL	<i>modʒip</i>	<i>budʒa k</i>	<i>udʒan</i>	-	<i>nadʒan</i>	<i>nidʒu</i>
EL	<i>mori</i>	<i>burɔ</i>	<i>ura</i>	-	<i>nara</i>	<i>nirũ</i>
KD	-	<i>buyaʔ</i>	<i>uya</i>	<i>dɛyeʔ</i>	<i>naya</i>	<i>ni ŋ</i>
	‘alive’	‘white’	‘rain’	‘sharp’	‘name’	‘nose’

In the examples in Table 5.25, Kedang reflects PFL *-d (> *-dʒ) > *y*/Ø. However, there are also cases where PMP *-d > *r*/*s* in Kedang. Table 5.26 shows the cognate sets in my dataset that contain possible words with intervocalic PFL *-d > *r*/*s* in Kedang. The reflexes in the other Flores-Lembata languages are regular, following the intervocalic pattern in Table 5.19.

Table 5.26: Irregular Kedang reflexes of intervocalic PMP *-d-

PMP	* <i>udu</i>	* <i>huaji</i>	* <i>budaq</i>	* <i>si ida</i>
PFL	* <i>udu</i>	* <i>vadi</i>	* <i>vuda</i>	* <i>hida</i>
SK	<i>uru ŋ</i>	<i>vari</i>	<i>vura n</i>	<i>r imu</i>
WL (LWI)	-	<i>ari ŋ</i>	<i>vura haŋ</i>	<i>raː</i>
WL (PD)	-	<i>ariŋ</i>	<i>wuraŋ</i>	<i>hire</i>
CL	-	<i>vadzɪ</i>	<i>pə wudʒa</i>	<i>da ne</i>
EL	[...]	(<i>v</i>) <i>ari</i>	[...]	? <i>ro ʔe</i>
KD	(<i>v</i>) <i>uru</i>	? <i>ari</i> ?	<i>vura n</i>	? <i>suo</i>
	‘grass; bush’	‘younger sibling’	‘foam’	‘3PL’ ^a

^a The Central Lamaholot form *dane* ‘3PL’ (< PFL **sida*/**hida*) follows the pattern of initial PMP **d*. The initial syllable of the pronoun was lost before the intervocalic change of PFL **d* > *dʒ* took place (cf. §5.2.2.4). The Eastern Lamaholot reflex could be related to the WL enclitic pronoun =*ro* ‘3PL’ only used for objects (Michels 2017:42). The Sika pronoun *rimu* ‘3PL’ is the result of the inherited third person plural pronoun merging with a reflex of the Central Flores reconstruction **imu* ‘friend; companion’ (Elias 2018:118). In several Central Flores languages, the word has undergone a semantic shift from ‘friend; companion’ to ‘3SG’.

In the first three examples in the table above, PFL *-d- > *r* in Kedang. It remains unclear what causes this irregularity as usually PFL *-d- > *y* in Kedang as shown above.

The reflex *s* < PFL **d* appears only in one example: Kedang *suo* ‘3PL’. As I have proposed intervocalic PFL *-d- (> **dʒ*) > *y*/Ø for Kedang, an intermediate stage of PFL **hida* > **hidʒa* > *dʒa* with loss of the initial syllable could be imagined. Alternation between *dʒ* and *s* is also attested in other Kedang words, such as in *yadi* (from *dʒadi*) / *sadi* ‘become, happen’ from Indonesian *dʒadi* (Samely and Barnes 2013:712). Therefore, it is possible that *dʒa* ‘3PL’ became *sa* but in other words, the initial *dʒ* was kept and later changed into *y*. However, the diphthongisation from *a* > *uo* which would be necessary to gain the present-day form *suo* ‘3PL’ remains obscure. Therefore, it is also possible that *suo* is not cognate with the other forms of the set.

5.2.2.4 Alternative options for the PFL reflex of PMP *d/*j/*z

In §5.2.2.3, I have shown, by providing evidence for identical correspondence sets, that the reflexes in the Flores-Lembata languages indicate a merger of initial and intervocalic PMP *d, *j and *z. I propose that these three PMP consonants merged into PFL *d in initial and intervocalic position despite the fact that synchronically no intervocalic *-d-* tracing back to PMP *d/*j/*z is attested (cf. §5.2.2.3). Therefore, reconstructing the value of the merged Proto-Flores-Lembata (PFL) sound is not straightforward. In this section, I motivate the reconstruction of PFL *d < PMP *d/*j/*z in initial and intervocalic position.

Based on the reflexes attested in the individual subgroups, there are three options for the PFL reflexes of PMP *d/*j/*z in initial and intervocalic position respectively:

1. PFL *d- and *-dʒ-
2. PFL *d- and *-d-
3. PFL *dʒ- and *-dʒ-

I argue for the second option which proposes PFL *d in both positions. In the following, I lay out the consequences for each of the three reconstruction options.

Option 1 Proposing PFL *d- in initial position and PFL *-dʒ- in intervocalic position would require the following subsequent sound changes in SK, KD, WL and EL. No sound change would be required in CL.

- PFL *d- > r / #_ in SK
- PFL *-dʒ- > r / V_V in SK
- PFL *-dʒ- > r / V_V in WL
- PFL *-dʒ- > r / V_V in EL
- PFL *-dʒ- > y / V_V in KD

A downside of this option is that proposing the allophones *d- and *-dʒ- in PFL cannot easily explain the Central Lamaholot reflexes *da* ‘3PL’ from a putative PFL *hidʒa ‘3PL’ listed in Table 5.27 below.¹ Assuming PFL *hidʒa

¹ The evidence for reconstructing the initial syllable *hi in PFL *hidʒa/*hida comes from the WL variety Alorese which retains the form *hire* ‘PL’ (< PMP *si ida) as a plural word. In other WL varieties, the 3PL pronoun *ra* is attested. Thus, the reflex of PFL *-dʒ- is *r* as proposed above. For the WL reflexes, there are no issue of putative PFL *-dʒ- as long as assuming that the change of *-dʒ- > r occurred before the loss of the initial syllable.

would require an exceptional change of intervocalic $*-d\zeta-$ > d in the word *da* as a free 3PL pronoun in most CL varieties. This change does not appear in any other word.

Table 5.27: Reflexes of PMP $*si\ ida$ ‘3PL’ reconstructions of Option 1

PMP	$*si\ ida$	‘3PL’
PFL	$*hid\zeta a$	‘3PL’
PCL	$*id\zeta a$	
CL (CLB, PN, LK)	<i>da </i> (<i>ne</i>)	‘3PL’
CL (IL, LP, MR)	<i>da </i> (<i>ro</i>)	‘3PL’
CL (CLB, PN)	<i>-i</i> / <i>-d\zeta a</i>	‘3PL.POSS’
CL (CLB)	<i>-d\zeta a</i>	‘PL’
CL (LWK)	<i>ya ne</i>	‘3PL’
WL	<i>ra </i> (<i>?e</i>)	‘3PL’
WL-Alorese	<i>hire</i>	‘PL’

PMP=Proto-Malayo-Polynesian
PFL=Proto-Flores-Lembata
PCL=Proto-Central Lamaholot

A possible solution would be to propose that there was variation between PFL $*da$ / $*hid\zeta a$ ‘3PL’. Consequently, most CL free pronouns would come from PFL $*da$, while the CL suffixes, as well as the LWK free pronoun would come from PFL $*hid\zeta a$. In this scenario, Proto-Western Lamaholot (PWL) would have lost PFL $*da$ and only retains reflexes of PFL $*hid\zeta a$.

The other forms do not cause problems when assuming PFL $*hid\zeta a$. The Central Lamaholot suffixes *-i* ‘3PL.POSS’ and *d\zeta a* ‘3PL.POSS/PL’ going back to PMP $*si\ ida$ can be explained by PFL $*hid\zeta a$ as they contain intervocalic $-d\zeta-$. Also the CL-Lewokukung (LWK) form can be explained, as *y* in this variety comes from an earlier *d\zeta*.

Additional support for Option 1 — PFL $*d-$ in initial and $*-d\zeta-$ in intervocalic position — could be the reconstruction of PFL $*d\zeta-$ in initial position as a marginal phoneme (cf. §5.2.2.6).

Option 2 Proposing, PFL *d- in initial position and PFL *-d- in intervocalic position would require the following subsequent sound changes in the FL subgroups.

- PFL *d- > r / #_ in SK
- PFL *-d- > r / V_V in SK
- PFL *d- > r / V_V in WL
- PFL *-d- > r / V_V in EL
- PFL *-d- > (*dʒ >) y / V_V in KD
- PFL *-d- > dʒ / V_V in CL

In contrast to the first option, this option would more easily explain the Central Lamaholot reflexes *da* ‘3PL’ in Table 5.27. The CL varieties that have *da* ‘3PL’ (all except for CL-LWK) have lost the initial syllable of Proto-Central Lamaholot (PCL) *ida (< PFL *hida) before the change of PFL *-d- > dʒ in Central Lamaholot. In contrast, CL-LWK lost the initial syllable after the change of PFL *-d- > *-dʒ-, and subsequently underwent the change of *dʒ > y. As initial PFL *d is retained in CL, the form remains *da*.

Option 3 Proposing, PFL *dʒ- in initial position and PFL *-dʒ- in intervocalic position would require the following subsequent sound changes in the Flores-Lembata subgroups.

- PFL *dʒ- > r / #_ in SK
- PFL *dʒ- > d / #_ in KD
- PFL *dʒ- > d / #_ in CL
- PFL *dʒ- > d / #_ in WL
- PFL *dʒ- > d / #_ in EL
- PFL *-dʒ- > r / V_V in SK
- PFL *-dʒ- > r / V_V in WL
- PFL *-dʒ- > r / V_V in EL
- PFL *-dʒ- > y / V_V in KD

Like Option 2, Option 3 would easily explain the Central Lamaholot reflexes *da* ‘3PL’ in Table 5.27. The CL varieties that have *da* ‘3PL’ would have lost the initial syllable of their reflexes from PCL *idʒa (< PFL *hidʒa) before the change of PFL *dʒ- > d in Central Lamaholot. As initial PFL *dʒ- > d, the form became *da*.

However, assuming this sound change of PFL *dʒ- > d- causes an additional problem. There are a few CL words that have initial *dʒ* which does not come from PFL *y but has to be reconstructed to PFL *dʒ- (cf. §5.2.2.6). Therefore, assuming PFL *dʒ- > d- would require explaining why these words did not change their initial consonant while all others did.

Weighing these three options against each other, Option 2 — reconstructing initial and medial *d — is the most likely. The main problem with Option 1 is that it cannot explain the CL pronoun *da* ‘3PL’. The main problem with Option 3 is that it cannot explain the CL words with initial *dʒ*. Both problems are avoided when choosing Option 2.

5.2.2.5 Final PMP *-d and *-j

Table 5.28 displays reflexes of final PMP *-d and PMP *-j. In final position, these two consonants do not merge. Final PMP *-d > PFL *r > r/Ø in the present-day languages and final PMP *-j [g/ɣ/gj] > PFL *y [j]. PFL *y is later lost in most varieties. In Central Lamaholot, PFL *y either becomes *dʒ* or is lost (c.f. §5.2.6). The reflexes of PMP *pusəj ‘navel’ have an irregular final *r* in the reflexes, as usually final PMP *-j > y.

Table 5.28: Reflexes of final PMP *-d and *-j

PMP	*qulu tuhud	*batad ‘millet; sorghum’	*pusəj	*qənaj	*qunəj	*sakaj
PFL	*lotur	*vatar	*pusər	*ənay	*una	*hakay
SK	<i>tur</i>	<i>vatar</i>	<i>puher</i>	<i>ne</i>	<i>une</i>	<i>haʔe</i>
WL	<i>lotor</i>	<i>vata</i>	<i>kə puhur</i>	-	<i>ono</i>	<i>haka</i>
CL	<i>lotor</i>	-	<i>kə pusər</i>	<i>ənadʒ</i> ‘soil’	<i>una</i> ‘house’	<i>aka</i> ~ <i>akadʒ</i>
EL	[...]	[...]	[...]	-	-	[...]
KD	-	<i>vatar</i>	<i>puhe:</i>	<i>ene</i>	-	<i>aʔ</i>
	‘knee’	‘corn’	‘navel’	‘sand; soil’	‘inside’	‘ascend; climb’

5.2.2.6 Evidence for PFL *dʒ with no PMP source

Following a bottom-up approach, there is some evidence provided in Table 5.29 to reconstruct PFL *dʒ in initial position. However, only one set, PFL *dʒua ‘two’, can clearly be reconstructed to PFL, all other sets have no reflexes in Kedang and Eastern Lamaholot. PFL *dʒua ‘two’ is probably an irregular reflex of PMP *duha ‘two’, as regularly PMP *d- > PFL *d-, but for the other words no PMP sources could be found.

Table 5.29: Reflexes of initial PFL *dʒ- without regular PMP source

PMP	-	? *duha	-	-	-	-
PFL	*dʒ-	*dʒua	#dʒəma	#dʒae	#dʒe(ta)	#dʒu
SK	<i>r</i>	<i>rua</i>	<i>rema</i> ‘time unit’	-	<i>re ta</i> ‘hillwards’	-
WL	<i>r</i>	<i>rua</i>	<i>rəmaʔ</i>	<i>rae</i>	-	-
CL	<i>dʒ</i>	<i>dʒua</i>	<i>dʒəma</i> ‘time unit’	<i>dʒae</i>	<i>dʒe</i>	<i>dʒu</i>
EL	?	[...]	[...]	[...]	[...]	[...]
KD	<i>s</i>	<i>sue</i>	-	-	-	-
		‘two’	‘night’	‘hillwards’	‘upwards’	‘downwards’

The sound correspondences in the table are different from reflexes of initial PFL *d-, as CL has *dʒ*-, KD has *s*- and WL has *r*-, while all three reflect PFL *d- as *d*-. Note that the correspondence set that leads to the reconstruction of PFL *dʒ- is similar to the reflexes of intervocalic PFL *-d- (cf. Table 5.25).

5.2.2.7 The retention of PMP *g

Table 5.30 provides reflexes of PMP *g in initial and intervocalic position. In the data I used for this study, there is not much evidence for reflexes of PMP *g in the languages of Flores-Lembata. Only four cognate sets could be found. In Kedang, the reflex of PMP *g is *k*. Given that Kedang also has PFL *k > ʔ (cf. §5.2.1), PFL *g > *k* must have happened after PFL *k > ʔ had been completed.

Table 5.30: Reflexes of PMP *g

PMP	*gaRaŋ	*gaRuʔ	*gatəl	*baqagi
PFL	*garaŋ	*garu	*gatər	*bagi
SK	-	<i>garu</i>	<i>gatar</i>	<i>bige</i>
WL (LWI)	-	<i>raguʔ</i>	<i>gatə</i>	-
WL (AD)	-	<i>raguʔ</i>	<i>gatə k</i>	-
WL (MS)	-	<i>gau</i>	<i>gate ʔ</i>	<i>bage</i>
CL	-	<i>kə ragu</i>	<i>gətə k</i>	-
EL	-	<i>ragu</i>	[...]	[...]
KD	<i>karaŋ</i>	<i>karo</i>	-	<i>boʔ</i>
	‘rough’	‘scratch’	‘itchy’	‘divide’

In the Kedang word *boʔ* ‘divide’, the intervocalic PFL *g has become the final consonant of the word and is therefore changed into glottal stop. Kedang only allows glottal stop, nasals and liquids in final position (Samely 1991a:46-47).

For the words going back to PFL *garu ‘scratch’, the Lamaholot varieties have undergone metathesis of the initial and medial consonant which led to *ragu* instead of *garu*, while Alorese retains the unmetathesised form *gau*.

5.2.3 PMP fricatives *h and *s

5.2.3.1 Loss of PMP *h

Table 5.31 show the reflexes of PMP *h in the languages of Flores-Lembata. PMP *h is almost always lost in the Flores-Lembata languages. However, there are a few potential cases of PMP *h retained as *h* or *ʔ* which are discussed further below.

Table 5.31: Reflexes of PMP *h

Env.	#_	V_V	_#
PMP	*h	*h	*h
PFL	*Ø	*h/Ø	*ʔ
SK	Ø	Ø	Ø
WL	Ø	h/ʔ/Ø	ʔ/Ø
CL	Ø	Ø	Ø
EL	ʔ/Ø	Ø	Ø
KD	ʔ/Ø	h/Ø	Ø

Table 5.32, 5.33 and 5.34 provide cognate sets with reflexes of PMP *h in the languages of Flores-Lembata. In initial position, PMP *h- was most likely already lost in Proto-Flores-Lembata or earlier. In Eastern Lamaholot and Kedang, a glottal stop is found sporadically at the position of initial PMP *h-. However, as both languages occasionally also insert glottal stops before initial vowels, such as in EL-Lewoeleng *ʔiʔu* ‘tail’ (< PMP *ikuR), it cannot be decided without doubt whether the initial glottal stop has been added at a later stage or is a reflex of PMP *h. For WL-Adonara a regular pattern of inserting a glottal stop before every initial vowel is attested.

Table 5.32: Reflexes of initial PMP *h-

PMP	*hapuy	*hikan	*haŋin	*huaji	*hular
PFL	*api	*ikan	*aŋin	*vadi	*ular
SK	<i>api</i>	<i>iʔan</i>	<i>anin</i>	<i>vari</i>	<i>ular</i>
WL	<i>ape</i>	<i>ikaŋ</i>	<i>aŋin</i>	<i>ari ŋ</i>	<i>ulaʔ</i>
CL	<i>api</i>	<i>ikan</i>	<i>aŋin</i>	<i>vadzɪ</i>	<i>ular</i>
EL	<i>ape</i>	<i>ʔiʔa</i>	<i>aŋin</i>	<i>vari</i>	<i>ula</i>
KD	<i>api</i>	<i>ika</i>	<i>aŋin</i>	<i>ʔariʔ</i>	<i>ular</i>
	‘fire’	‘fish’	‘wind’	‘younger sibling’	‘snake’

In intervocalic position, PMP *-h- is also generally lost as shown in Table 5.33. In WL, there are a few instances of PMP *h reflected as glottal stop in

between vowels, such as in WL-Lamahora *vaʔi* ‘water’. Similar to the case of initial PMP *h-, the glottal stop between two vowels could also be an insertion rather than a reflex of PMP *h. However, this appears unlikely as insertion of glottal stop intervocalically is not otherwise found in the phonologies of the Flores-Lembata languages.

Table 5.33: Reflexes of intervocalic PMP *-h-

PMP	*kahu	*luhəq	*wahiR	? *duha	*dahun	*buhək
PFL	*kau	*luu	*vaʔir	*dʒua	*down	*vuhak
SK	<i>ʔau</i>	<i>lu</i>	<i>vair</i>	<i>rua</i>	<i>roun</i>	-
WL (LWI)	-	<i>lou ŋ</i>	<i>vai</i>	<i>rua</i>	-	-
WL (LH)	-	[...]	<i>vaʔi</i>	[...]	-	-
WL (LTB)	-	[...]	<i>vaiʔ</i>	<i>rua</i>	-	-
CL	-	-	<i>vai</i>	<i>dʒua</i>	-	-
EL	-	-	<i>vae</i>	[...]	-	-
KD	-	<i>lu n</i>	<i>vei</i>	<i>sue</i>	-	<i>uha</i>
	2SG	‘tear’	‘water’	‘two’	‘leaf’	‘hair’

Kedang *uha* ‘hair’ (< PMP *buhək) provides possible evidence for the retention of intervocalic PMP *-h-. This would be exceptional as intervocalic PMP *-h- has been lost in all known languages of the area. In the word *uha* ‘hair’ (< PMP *buhək), the loss of initial PFL *w (< PMP *b) before *u* is regular in Kedang. PMP *ə > *a* before *k is also attested in PMP *təbək ‘stab’ > Kedang *tubaʔ*. Due to the regularity of the other reflexes, the Kedang example *uha* ‘hair’ provides possible evidence for occasional retention of intervocalic PMP *h in Kedang. Based on current data, this would be the only case of PMP *h = h in Kedang.

In final position, PMP *-h is mainly lost or sporadically retained as glottal stop, such as in WL-Adonara *taleʔ* ‘rope’ in Table 5.34. However, as there is just one example with a glottal stop at the position of PMP *-h, this could also be an irregular insertion.

Table 5.34: Reflexes of final PMP *-h

PMP	*tali ^h	*təbuh	*qilih	*tumah	*ma-tuqah
PFL	*tali	*təvu	*ili	*tuma	*m-tuʔa
SK	<i>tali</i>	<i>təvu</i>	<i>ili n</i>	-	<i>dua n</i>
WL (AD)	<i>taleʔ</i>	<i>tevo</i>	<i>ʔile</i>	<i>kə tumã</i>	<i>t<en>uʔe</i>
WL (LWI)	<i>tale</i>	<i>təvo</i>	<i>ile</i>	<i>kə tuma</i>	<i>t<ən>ue ŋ</i>
CL	<i>tali</i>	<i>təvu</i>	<i>ili</i>	<i>kə tumav</i>	<i>tua na</i>
EL	<i>tale</i>	-	<i>ile</i>	[...]	[...]
KD	-	<i>təhu</i>	<i>ili</i>	-	-
‘rope’ ‘sugarcane’ ‘mountain’ ‘cloths louse’ ‘old’					

5.2.3.2 Split: PMP *s > PFL *s/*h

Table 5.35 shows the reflexes of PMP *s in the languages of Flores-Lembata. PMP *s splits unconditionally into PFL *s and PFL *h in initial and intervocalic position. All Flores-Lembata languages, except for Central Lamaholot, later complete this change by changing the reflexes of PFL *s into *h*. In Central Lamaholot PFL *s = *s* while PFL *h > $\emptyset$.

Table 5.35: Reflexes of PMP *s

Env.	#_		V_V		_#
PMP	*s		*s		*s
PFL	*s	*h	*s	*h	* $\emptyset$
SK	<i>h</i>	<i>h</i>	<i>h</i>	<i>h</i>	$\emptyset$
WL	<i>h</i> / $\emptyset$	<i>h</i>	<i>h</i>	<i>h</i> / $\emptyset$	$\emptyset$
CL	<i>s</i>	$\emptyset$	<i>s</i>	$\emptyset$	$\emptyset$
EL	[...]	<i>h</i>	<i>h</i>	<i>h</i>	$\emptyset$
KD	<i>h</i> / $\emptyset$	<i>h</i> / $\emptyset$	<i>h</i> / $\emptyset$	<i>h</i> / $\emptyset$	$\emptyset$

Tables 5.36, 5.37 and 5.38 provide cognate sets with reflexes of PMP *s in the languages of Flores-Lembata. In initial and medial position, there is an unconditioned split of PMP *s into PFL *s / *h. No conditioning for the split

of PMP *s in Proto-Flores-Lembata could be determined and the instances of PMP *s = PFL *s and PMP *s > PFL *h in my data set are roughly equal in number. It is probably a result of an incomplete sound change of *s > *h in Proto-Flores-Lembata.

Most Central Lamaholot varieties retain PFL *s = s, with the exception of CL-Painara which has PFL *s > h. All other FL languages have undergone PFL *s > h in the vast majority of lexemes. However, this change is not entirely complete, as there are sporadic retentions of PFL *s = s in Sika and Western Lamaholot. PFL *h is lost in Central Lamaholot. In Kedang initial PFL *h > Ø while in other positions PFL *s/*h are retained as *h*, *ʔ* or lost. Table 5.36 shows reflexes of initial PMP *s.

Table 5.36: Reflexes of initial PMP *s

PMP	*s	*siwa	*sama	*salaq	*s	*sakay	*qasiRa
PFL	*s	*siva	*sama	*sala	*h	*hakay	*hira
SK	<i>h</i>	<i>hiva</i>	<i>hama</i>	<i>hala</i>	<i>h</i>	<i>haʔe</i>	-
WL (AD)	<i>h</i>	<i>hiva</i>	<i>hama</i>	<i>n alã</i>	<i>h</i>	<i>haka</i>	[<i>siʔa</i>]
CL	<i>s</i>	<i>siva</i>	<i>s<n>ama ŋ</i>	<i>s<n>ala k</i>	Ø	<i>aka(dʒ)</i>	<i>ira</i>
EL	?	[...]	[...]	[...]	<i>h</i>	[...]	<i>hira</i>
KD	<i>h</i>	-	<i>hama</i>	<i>ke he</i>	<i>h/Ø</i>	<i>aʔ</i>	-
		'nine'	'same'	'wrong'		'climb'	'salt'

In Table 5.36, the WL word *siʔa* 'salt' contains an irregular retention of initial *s*. Possibly WL lost the inherited cognate for the word for 'salt' and reintroduced it through borrowing from a language that still retained the *s*. As salt is a trade commodity, it is highly borrowable. A possible donor languages for WL *siʔa* 'salt' would be a Central Flores language, such as Lio which has *siʔe* 'salt' (cf. §6.2.4).

Table 5.37 lists reflexes of intervocalic PMP *s. Kedang shows sporadic PFL *s/*h > ʔ/Ø. PFL *(t)usu 'breast' (< PMP *susu) has an irregular change of PMP *s > PFL *t initially, as well as irregular loss of the initial consonant in Sika.

Table 5.37: Reflexes of intervocalic PMP *s

PMP	*s	*tasak	*pusəj	*susu	*s	*asu	*əsa	*pusuŋ
PFL	*s	*m-tasak	*pusər	*(t)usu	*h	*ahu	*əha	*puhuŋ
SK	<i>h</i>	<i>dahaʔ</i>	<i>puher</i>	<i>uhu</i>	<i>h</i>	<i>ahu</i>	<i>ha</i>	<i>puhuŋ</i>
WL (AD)	<i>h</i>	<i>tahak</i>	<i>ke puhur et</i>	<i>tuho</i>	<i>h</i>	<i>aho</i>	[...]	<i>puho</i>
CL	<i>s</i>	<i>tasak</i>	<i>kə pusər</i>	<i>tusu</i>	Ø	<i>au</i>	<i>m ea</i>	<i>puo</i>
EL	?	[...]	-	[...]	<i>h</i>	<i>aho</i>	[...]	<i>puho</i>
KD	<i>h/ʔ</i>	<i>taʔen</i>	<i>puhɛː</i>	<i>tuʔu</i>	<i>h/Ø</i>	<i>au</i>	<i>ehaʔ</i>	-
		‘ripe’	‘navel’	‘breast’		‘dog’	‘one; alone’	‘heart’

Table 5.38 shows reflexes of final PMP *s. In PFL *tani ‘cry’ (< PMP *taŋis) final *s is lost. In PFL *bukat ‘open’ (< PMP *buŋkas) final *s is irregularly replaced by *t*. Final PFL *t is regularly reflected as *ʔ* in Kedang (cf. §5.2.1.3). In the PFL reflex *menipihi ‘thin’ (< PMP *ma-nipis), the final *s has become intervocalic. Therefore, it is retained as *h* in Western Lamaholot and Eastern Lamaholot. In Central Lamaholot PFL *h > Ø. The insertion of *ʋ* in CL varieties may be explained by the avoidance of two adjacent vowels in final position.

Table 5.38: Reflexes of final PMP *s

PMP	*ma-nipis	*taŋis	*buŋkas
PFL	*m-nipih-i	*tani	*bukat
SK	-	<i>tani</i>	-
WL (KO)	<i>məniphi</i>	<i>tani</i>	[...]
WL (AD)	<i>menipi</i>	<i>tani</i>	<i>buka</i>
CL (LWK)	<i>mipivu</i>	-	[...]
CL (KK)	<i>mipiv</i>	-	<i>bukat</i>
EL	<i>mipihi</i>	<i>tani</i>	[...]
KD	<i>mipi</i>	-	<i>bukaʔ</i>
	‘thin’	‘cry’	‘open’

5.2.4 PMP nasals *m, *n and *ŋ

The PMP nasals *n, *m and *ŋ are mainly retained as such in the languages of Flores-Lembata as shown in Table 5.39. The nasal which undergoes the most changes is *ŋ. Word initially PMP *ŋ- > n/Ø for all FL languages. In Sika medial *-ŋ- > n.

Table 5.39: Reflexes of PMP nasals *m *n and *ŋ

Env.	#_	V_V	_#	#_	V_V	_#	#_	V_V	_#
PMP	*m-	*-m-	*-m	*n-	*-n-	*-n	*ŋ-	*-ŋ-	*-ŋ
PFL	*m	*m	*m	*n	*n	*n	*n	*ŋ	*ŋ
SK	<i>m</i>	<i>m</i>	Ø	<i>n</i>	<i>n</i>	<i>n/ŋ/Ø</i>	Ø	<i>n</i>	<i>ŋ/n/Ø</i>
WL	<i>m</i>	<i>m</i>	Ø	[...]	<i>n</i>	<i>n/ŋ/Ø</i>	<i>n/Ø</i>	<i>ŋ</i>	<i>ŋ</i>
CL	<i>m</i>	<i>m</i>	<i>m</i>	[...]	<i>n</i>	<i>n</i>	<i>n</i>	<i>ŋ</i>	<i>ŋ/Ø</i>
EL	<i>m</i>	<i>m</i>	Ø	[...]	<i>n</i>	<i>n/Ø</i>	<i>n/Ø</i>	<i>ŋ</i>	Ø
KD	<i>m</i>	<i>m</i>	Ø	[...]	<i>n</i>	<i>n/Ø</i>	<i>n</i>	<i>ŋ</i>	<i>ŋ/n</i>

Table 5.40 gives an example of PMP *m reflexes for each position in the word. In initial and medial position, PMP *m is retained in all FL languages, while in final position only Central Lamaholot keeps PMP *m in final position.

Table 5.40: Reflexes of PMP *m

PMP	*manuk	*Ramut	*ma-qitəm	*tazim ‘whet’
PFL	*manuk	*ramut	*mitəm	*m-tidəm
SK	<i>manu?</i>	<i>ramut</i>	<i>mita n</i>	<i>dira n</i>
WL (AD)	<i>manuk</i>	<i>ʔamut</i>	<i>mitə</i>	-
WL (LWI)		<i>ramu</i>	<i>mitə ŋ</i>	-
CL	<i>manuk</i>	<i>ramut</i>	<i>mitəm</i>	-
EL	<i>manu?</i>	<i>ramu</i>	<i>mitã</i>	-
KD	<i>manu?</i>	<i>ramu?</i>	<i>mitə ŋ</i> ‘black; dirty’	<i>dɛyɛ ?</i>
	‘chicken’	‘root’	‘black’	‘sharp’

Table 5.41 provides cognate sets illustrating the development of PMP *n for each position. Initial PMP *n is scarce in my data and there are no examples of initial *n* (< PMP *n) in Lamaholot and Kedang. In Sika, there is also the word *niur* ‘coconut’ (< PMP *niuR). Intervocalic PMP *n is regularly reflected as *n*. Final PMP *n is mainly reflected as *n*. However, sporadically, the reflex of final PMP *-n > Ø in Western Lamaholot and Kedang. In some varieties, such as SK-Krowe or WL-Munaseli (a variety of Alorese), final PMP *-n > *ŋ*. This change is also sporadically attested in other Western Lamaholot varieties.

Table 5.41: Reflexes of PMP *n

PMP	*niuR	*ina	*anak	*haŋin	*bulan
PFL	*niur	*ina	*anak	*aŋin	*vulan
SK (HW)	<i>niur</i>	<i>ina</i>	<i>anak</i>	<i>anin</i>	<i>vulan</i>
SK (KW)	<i>niur</i>	<i>ina</i>	-	<i>aniŋ</i>	<i>vulaŋ</i>
WL (LWI)	-	-	<i>ana?</i>	<i>aŋin</i>	<i>vula</i>
WL (AD)	-	<i>?ina</i>	<i>?ana?</i>	<i>?aŋi</i>	<i>vulã</i>
WL (MS)	-	<i>ina</i>	<i>anaŋ</i>	<i>aŋiŋ</i>	<i>vulaŋ</i>
CL	-	<i>ina</i>	<i>ana k</i>	<i>aŋin</i>	<i>vulan</i>
EL	-	<i>ina</i>	<i>ana</i>	<i>aŋin</i>	[...]
KD	-	<i>?ine</i>	<i>ana?</i>	<i>aŋin</i>	<i>vula</i>
	‘coconut’	‘mother’	‘child; small’	‘wind’	‘moon’

Table 5.42 shows cognate sets that contain reflexes of initial and intervocalic PMP *ŋ. Initial PMP *ŋ > *n* in PFL. Intervocalic PMP *ŋ is retained as *ŋ* in Kedang and Lamaholot, while Sika shows PMP *-ŋ- > *n*. In Sika, this leads to a merger of PMP *n/*ŋ > *n* in intervocalic position. The Western Lamaholot and Eastern Lamaholot reflexes of PMP *taŋis ‘cry’ have irregular reflexes containing intervocalic *n* (< PMP *ŋ) most likely going back to PFL, therefore reconstructed as PFL *tani ‘cry’.

Table 5.42: Reflexes of initial and intervocalic PMP *ŋ

PMP	*ŋusu	*ŋajan	*ŋijuŋ	*dəŋəR	*taŋis	*naŋuy
PFL	*nusu	*nadan	*(n)iduŋ	*dəŋər	*tani	*naŋi
SK	-	<i>naran</i>	<i>iru</i>	<i>rəna</i>	<i>tani</i>	<i>nani</i>
WL (LWI)	-	<i>naraŋ</i>	<i>iruŋ</i>	<i>weŋe</i>	<i>taniŋ</i>	<i>naŋe</i>
WL (LML)	-	<i>naraŋ</i>	<i>niruŋ</i>	<i>dəŋa</i>	<i>tani</i>	<i>naŋe</i>
WL (MS)	<i>nuhu ŋ</i>	<i>naraŋ</i>	<i>iruŋ</i>	<i>dəŋa</i>	<i>taniŋ</i>	<i>naŋge</i>
CL	<i>nus</i>	<i>nadʒan</i>	<i>nidʒu</i>	<i>dəŋər</i>	-	<i>naŋe</i>
EL	<i>nuhe</i>	<i>nara</i>	<i>nirũ</i>	<i>dəŋe</i>	<i>tani</i>	<i>naŋi</i>
KD	? <i>nunu</i>	<i>naya</i>	<i>niŋ</i>	<i>dəŋər</i>	-	<i>naŋi</i>
	‘mouth’	‘name’	‘nose’	‘hear’	‘cry’	‘swim’

Final PMP *-ŋ appears to be easily lost in the languages of Flores-Lembata. For most lexical items it can still be reconstructed to PFL but is not found in all reflexes. Sometimes, such as in CL *gaʔa|k* ‘rough’ or WL *iru|net* ‘nose’, the final nasal is replaced by a suffix.

Table 5.43: Reflexes of final PMP *-ŋ

PMP	*pusuŋ	*bubuŋ	*garaŋ	*kədəŋ	*ŋijuŋ
PFL	*puhun	*(v)uvuŋ	*garaŋ	*kəda	*(n)iduŋ
SK	<i>puhuŋ</i>	-	-	<i>ʔəra</i>	<i>iru</i>
WL (AD)	<i>puho</i>	<i>uvuŋ</i>	<i>gaʔa k</i>	-	<i>iru nət</i>
WL (AB)	-	<i>fufuŋ</i>	-	-	<i>niruŋ</i>
CL	<i>puo</i>	<i>uvuŋ</i>	-	-	<i>nidʒu</i>
EL	<i>puho</i>	[...]	[...]	-	<i>nirũ</i>
KD	-	-	<i>karaŋ</i>	-	? <i>niŋ</i>
	‘heart’	‘ridge’	‘rough’	‘stand’	‘nose’

In sum, while the nasal *m* is clearly separate in all FL languages, the nasals *n* and *ŋ* have partly merged in initial and final position. In Sika this merger also occurs intervocalically. In some Western and Eastern Lamaholot varieties, final *n* or *ŋ* has often been deleted with the preceding vowel being nasalised, such as in WL-Lewolema *puhũ* ‘flower’ (< PMP *pusuŋ).

5.2.5 PMP liquids *l and *R

The PMP liquids *l and *R [r] are generally reflected in a regular manner. PMP *l and *R [r] are usually retained unchanged with the exception of PMP *R > WL ʔ. This sound change is complete in intervocalic and final position but possibly incomplete initially. The reflexes of PMP *l and *R are given in Table 5.44.

Table 5.44: Reflexes of PMP *l and *R [r]

Env.	#_	V_V	_#	#_	V_V	_#
PMP	*l-	*-l-	*-l	*R-	*-R-	*-R
PFL	*l	*l	*l	*r	*r	*r
SK	<i>l</i>	<i>l</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
WL	<i>l</i>	<i>l</i>	<i>l</i>	ʔ/ <i>r</i>	ʔ	ʔ
CL	<i>l</i>	<i>l</i>	[...]	<i>r</i>	<i>r</i>	<i>r</i>
EL	<i>l</i>	<i>l</i>	[...]	<i>r</i>	<i>r</i>	[...]
KD	<i>l</i>	<i>l</i>	<i>l</i>	<i>r</i>	<i>r</i>	<i>r</i>

Table 5.45 shows cognate sets that contain reflexes of PMP *l. Most instances of PMP *l are retained as *l*. However, there are occasional irregular changes of PFL *l > *r* in Western and Eastern Lamaholot, mainly influenced by the presence of a second liquid in the word, such as in WL-Adonara *rəra* ‘day, sun’ (< PMP *qaləjaw). Other examples of sporadic *l > *r* are reflexes of PFL *lalan ‘road’ in some Western Lamaholot varieties and reflexes of the unreconstructible lexeme set #latar ‘hair’ in Western and Eastern Lamaholot. Unreconstructible lexeme sets will be discussed in §6.3.

Only one cognate set is found which contains reflexes of PMP *l in final position. In this set PMP *l becomes *r* in Sika.

Table 5.45: Reflexes of PMP *l

PMP	*qaləjaw	*qalima	*təlu	*hulaR	*ma-kapal
PFL	*lədav	*lima	*təlu	*ular	*m-kapal
SK	<i>ləro</i>	<i>lima</i>	<i>təlu</i>	<i>ular</i>	<i>ʔapar</i>
WL (LWI)	<i>ləraː</i>	<i>lima ŋ</i>	<i>telo</i>	<i>ulaʔ</i>	-
WL (AD)	<i>rəra</i>	-	<i>telo</i>	<i>ʔulaʔ</i>	-
WL (MS)	<i>ləra</i>	<i>lima ŋ</i>	<i>təlo</i>	<i>ula</i>	<i>gapal</i>
CL (KK)	-	<i>lima</i>	<i>təlu</i>	<i>ular</i>	-
CL (LWT)	<i>lədʒaf</i>	<i>lima ha</i>	[...]	<i>ular u</i>	-
EL	<i>ləra</i>	<i>lima</i>	[...]	<i>ula</i>	[...]
KD	<i>loyo</i>	<i>liŋ</i>	<i>təlu</i>	<i>ular</i>	<i>kapal</i>
	‘sun’	‘hand’	‘three’	‘snake’	‘thick’

Table 5.46 provides cognate sets that contain reflexes of initial PMP *R. From that set it is clear that the change of *R > ʔ in WL is not complete in all varieties in word initial position. An example is PMP *Ramut > WL-Lamalera *ramut* ‘root’ in which initial PFL *r > r.

Table 5.46: Reflexes of initial PMP *R

PMP	*Ramut	*Rumaq	*Raya
PFL	*ramut	*ruma	*raya
SK	<i>ramut</i>	-	-
WL (LML)	<i>ramut</i>	-	-
WL (AD)	<i>ʔamut</i>	-	-
WL (BN)	<i>ramu k</i>	<i>ʔuma</i>	-
CL (KK)	<i>ramut</i>	-	-
CL (LR)	<i>ramut</i>	-	<i>rayan</i>
EL	<i>ramu</i>	-	-
KD	<i>ramuʔ</i>	-	<i>ria</i>
	‘root’	‘house’	‘big’

Table 5.47 shows cognate sets containing reflexes of intervocalic and final PMP *R. In intervocalic and final position the sound change of PMP *R > ʔ/Ø

has been completed in Western Lamaholot. Final PMP *-R is also mainly lost in Sika. Central Lamaholot and in Eastern Lamaholot sometimes added a suffix.

Table 5.47: Reflexes of intervocalic and final PMP *R

PMP	*baqəRu	*laRiw	*bəRəqat	*qapuR	*dəŋəR	*təluR
PFL	*vəru	*plari/*kari	*bərat	*ʔapur	*dəŋər	*təlur
SK	<i>vərun</i>	<i>p lari</i>	<i>bərat</i>	<i>ʔapur</i>	<i>rəna</i>	<i>təlo n</i>
WL (LWI)	<i>vuʔu ŋ</i>	<i>pə laʔe</i>	<i>baʔa</i>	<i>apuʔ</i>	<i>weŋe</i>	<i>telu</i>
WL (LWL)	<i>vuʔũ</i>	<i>pə laʔe</i>	<i>baʔat</i>	<i>apuʔ</i>	<i>weŋe</i>	<i>təlu k</i>
WL (MS)	<i>vunoŋ</i>	<i>p laeŋ</i>	<i>baʔ</i>	<i>apu</i>	<i>dəŋa</i>	<i>təlu k</i>
CL	<i>vərun</i>	<i>k ari</i>	<i>bərat</i>	<i>apur</i>	<i>dəŋər</i>	<i>təlu k</i>
EL	<i>vəru</i>	[...]	<i>bəra</i>	[...]	<i>dəŋe</i>	<i>təlũ</i>
KD	<i>verun</i>	-	<i>baraʔ</i>	<i>apur</i>	<i>dəŋər</i>	<i>tolor</i>
	‘new’	‘run’	‘heavy’	‘lime’	‘hear’	‘egg’

5.2.6 PMP glides *w *y

In this section, I discuss reflexes of initial and intervocalic PMP glides *w and intervocalic PMP *y. PMP *y does not appear word-initially. Reflexes of final glides are analysed together with their preceding vowels in §5.2.8. The PMP glide *w merges with some instances of PMP *b as voiced fricative *v* in Proto-Flores-Lembata (cf. Section 5.2.2). In initial and intervocalic position, there are no further changes, except for PMP *w > PFL *v. PMP *y is weakened to a high vowel or zero in some cases, but strengthened to *dʒ* in others.

Table 5.48: Reflexes of PMP initial and intervocalic *w and *y

Env.	#_	V_V	V_V
PMP	*w-	*-w-	*-y-
PFL	*v	*v	*y
SK	v	[...]	y
WL	v	v	y/dʒ
CL	v	v	y/dʒ
EL	v	v	[...]
KD	v	v	y

Table 5.49 provides cognate sets with reflexes of PMP *w. PMP *w is generally retained as *v* in initial and intervocalic position.

Table 5.49: Reflexes of initial and intervocalic PMP *w

PMP	*wahiR	*ka-wanan	*ka-wiri	*ma-tawa	*qasawa
PFL	*vaʔir	*vanan	*viri	*tave	*hava
SK	<i>vair</i>	<i>vanan</i>	<i>viri</i>	<i>to</i>	<i>vai</i>
WL (LWI)	<i>vai</i>	<i>vanaj</i>	-	-	<i>kə vae</i>
CL	<i>vai</i>	<i>vana</i>	-	-	<i>ava</i>
EL	<i>vae</i>	<i>vana</i>	<i>viri</i>	-	<i>hava</i>
KD	<i>vei</i>	<i>vana</i>	<i>veri</i>	<i>tave</i>	<i>veʔ</i>
	‘water’	‘right side’	‘left side’	‘laugh’	‘spouse’

Table 5.50 shows intervocalic reflexes of the PMP glide *y [j] in the languages of Flores-Lembata. PMP *y only appears in intervocalic and final position. Kedang and Sika retain *y in intervocalic position. In Lamaholot, reflexes of PMP *y are either retained as *y*, or strengthened to *dʒ*. The conditioning factors are unknown. In CL-Central Lembata the change of PMP *y > *dʒ* is completed for final and intervocalic position. In the WL varieties Alorese and Lewolema, the change is also complete. CL-Lerek and WL-Adonara show an incomplete change of PMP *y > *dʒ*. WL-Lewoingu regularly retains *y* as a reflex of PMP *y in intervocalic position. In Sika and Kedang, the glide *y* often becomes *i* or *e*. For EL, not enough data is available.

Table 5.50: Reflexes of intervocalic PMP *-y-

PMP	*bayu	*layaR	*kahiW	*Raya	*ma həyaq
PFL	*bayu	*layar	*kayu	*raya	*məya
SK	<i>bai</i>	<i>layar</i>	<i>ʔai</i>	-	<i>mea ŋ</i>
WL (MS)	<i>badʒo</i>	<i>ladʒa</i>	<i>kadʒu</i>	-	<i>m<n>ia ŋ</i>
WL (LWL)	[...]	[...]	<i>kadʒo</i>	-	[...]
WL (AD)	<i>badʒo</i>	<i>layaʔ</i>	<i>kayo</i>	-	<i>mia</i>
WL (LWI)	<i>bayo</i>	<i>layaʔ</i>	<i>kayo</i>	-	<i>mia</i>
CL (LR)	<i>badʒo</i>	<i>layar</i>	<i>kayo</i>	<i>raya n</i>	-
CL (KK)	<i>badʒu</i>	<i>ladʒar</i>	<i>kadʒu</i>	-	-
EL	[...]	[...]	-	-	[...]
KD	<i>bae</i>	<i>layar</i>	<i>ʔai</i>	<i>ria:</i>	-
	‘pound’	‘sail’	‘wood; tree’	‘big’	‘shy; ashamed’

5.2.7 PMP vowels

The PMP vowels *a, *i and *u are unchanged in non-final position. In final-position, Western Lamaholot undergoes vowel lowering of high vowels: PMP *-i > e and PMP *-u > o. In Eastern Lamaholot and Central Lamaholot, vowel lowering is found sporadically for reflexes of final PMP *-u and final PMP *-i. In Kedang, final PMP *-a > e/ɛ/o. In most cases, the change PMP *a > e or ɛ is attested, with only one example of final PMP *a > o; Kedang *mato* ‘eye’ (< PMP *mata).

Table 5.51: Reflexes of PMP vowels *a, *i, *u

	Non-final			Final		
PMP	*a	*i	*u	*-a	*-i	*-u
PFL	*a	*i	*u	*-a	*-i	*-u
SK	<i>a</i>	<i>i</i>	<i>u</i>	<i>a</i>	<i>i</i>	<i>u</i>
WL	<i>a</i>	<i>i</i>	<i>u</i>	<i>a</i>	<i>e</i>	<i>o</i>
CL	<i>a</i>	<i>i</i>	<i>u</i>	<i>a</i>	<i>i/e</i>	<i>u/i</i>
EL	<i>a</i>	<i>i</i>	<i>u</i>	<i>a</i>	<i>i/e</i>	<i>u/o</i>
KD	<i>a/ε</i>	<i>i</i>	<i>u</i>	<i>e/ε/o</i>	<i>i</i>	<i>u</i>

Table 5.52 and Table 5.53 provides examples which contain reflexes of PMP *a, *i and *u. Final and non-final refers to the position of the proto phoneme. Sometimes the synchronic reflex of a non-final proto phoneme can be final, such as Western Lamaholot *ramu* < PMP *ramut ‘root’.

Table 5.52: Reflexes of non-final PMP *a, *i and *u

PMP	*ŋajan	*hikan	*pitu	*ma-paqit	*kutu	*Ramut
PFL	*nadan	*ikan	*pitu	*m-paʔit	*kutu	*ramut
SK	<i>naran</i>	<i>iʔan</i>	<i>pitu</i>	<i>baʔit</i>	<i>ʔutu</i>	<i>ramut</i>
WL (LWI)	<i>naraŋ</i>	<i>ikaŋ</i>	<i>pito</i>	<i>pait</i>	<i>kuto</i>	<i>ramu</i>
CL	<i>nadʒan</i>	<i>ikan</i>	<i>pito</i>	<i>p<n>ait</i>	<i>kutu</i>	<i>ramut</i>
EL	<i>nara</i>	<i>iʔã</i>	[...]	[...]	[...]	<i>ramu</i>
KD	<i>naya</i>	<i>iʔa</i>	<i>pitu</i>	<i>pɛiʔ</i>	<i>ʔutu</i>	<i>ramuʔ</i>
	‘name’	‘fish’	‘seven’	‘bitter’	‘headlice’	‘root’

Table 5.53: Reflexes of final PMP *a, *i and *u

PMP	*mata	*ina	*kami	*diRi	*təlu	*batu
PFL	*mata	*ina	*kami	*diri	*təlu	*vatu
SK	<i>mata</i>	<i>ina</i>	<i>ʔami</i>	-	<i>təlu</i>	<i>vatu</i>
WL (LWI)	<i>mata</i>	-	<i>kame</i>	<i>deʔi ŋ</i>	<i>telo</i>	<i>vato</i>
WL (AD)	<i>mata k</i>	<i>ʔina</i>	<i>kame</i>	<i>deʔi</i>	<i>telo</i>	<i>vato</i>
WL (MS)	<i>mata ŋ</i>	<i>ina ŋ</i>	<i>kame</i>	-	<i>təlo</i>	<i>vato</i>
CL	<i>mata</i>	<i>ina</i>	<i>kame</i>	<i>diri</i>	<i>təlu</i>	<i>vatu</i>
EL	-	<i>ina</i>	<i>ame</i>	<i>diri</i>	[...]	<i>vato</i>
KD	<i>mato</i>	<i>ʔine</i>	<i>(k)e:</i>	-	<i>telu</i>	<i>vaʔ</i>
	‘eye’	‘mother’	‘1PL.EXCL’	‘stand’	‘three’	‘stone’

The Western Lamaholot reflexes of PMP *diRi ‘stand’ have undergone metathesis of their vowels. The final vowel *e* (< PMP *i) that had been lowered is moved to the penultimate syllable, while the vowel *i* that had been in the penultimate has been moved to the end. The nasal *ŋ* in WL-Lewoingu is a later insertion.

The reflexes of PMP *ə are more complex than those of the PMP vowels *a, *i and *u discussed above. Reflexes of PMP *ə are summarised in Table 5.54. While Sika, CL and EL show regular reflexes, WL and KD have unconditioned splits of PFL *ə in both positions. With more data, a conditioning environment for these splits could possibly be found. In the final syllable Sika has completed a regular change of PMP *ə > *a* and EL shows regular PMP *ə > *e*. CL is most conservative and retains PMP *ə = *ə* in all positions.

Table 5.54: Reflexes of PMP *ə

	Penultimate	Ultimate
PMP	*ə	*ə
PFL	*ə	*ə
SK	ə	a
WL (LWL)	ə/e	ə
WL (LWI)	ə/e	ə/e
WL (AD)	ə/e	ə/a
WL (MS)	ə/a	ə/e/a
WL (AB)	e/a	e/a
CL	ə	ə
EL	ə	e
KD	e/ɛ	e/ɛ

Table 5.55 provides examples with reflexes of penultimate PMP *ə. In Kedang, a *o* in the ultimate syllable leads to the assimilation of the penultimate vowel reflecting PMP *ə, such as in *loyo* ‘day’ (< PMP *qaləjaw) and *tolor* ‘egg’ (< PMP *qatəluR). The *o* in the ultimate syllable of *tolor* ‘egg’ is an irregular reflex of PMP *u, while the final *o* in *loyo* ‘day’ is a regular reflex of PMP *aw. Table 5.56 lists examples containing reflexes of PMP *ə in the ultimate syllable of the word.

Table 5.55: Reflexes of penultimate PMP *ə

PMP	*qaləjaw	*qatəluR	*dəŋəR	*qatiməla	*təlu
PFL	*lədav	*təlur	*dəŋər	*təməla	*təlu
SK	<i>ləro</i>	<i>təlo</i>	<i>rəna</i>	<i>məla</i>	<i>təlu</i>
WL (LWL)	<i>rəra</i>	<i>təlu k</i>	<i>dəŋəʔ</i>	<i>təməla</i>	<i>telo</i>
WL (LWI)	<i>ləraː</i>	<i>telu</i>	<i>weŋe</i>	-	<i>telo</i>
WL (AD)	<i>rəra</i>	<i>telu k</i>	<i>dəŋəʔ</i>	<i>təməla</i>	<i>telo</i>
WL (MS)	<i>lara</i>	<i>təlu k</i>	<i>dəŋa</i>	<i>məre</i>	<i>təlo</i>
WL (AB)	<i>lara</i>	<i>talukuj</i>	<i>daŋa</i>	<i>taməla</i>	<i>telo</i>
CL	<i>lədʒa</i>	<i>təlu k</i>	<i>dəŋər</i>	<i>təməla</i>	<i>təlu</i>
EL	<i>ləra</i>	<i>təlu</i>	<i>dəŋe</i>	[...]	[...]
KD	<i>loyo</i>	<i>tolor</i>	<i>dəŋər</i>	<i>mələ</i>	<i>telu</i>
	‘day; sun’	‘egg’	‘hear’	‘flea’	‘three’

Table 5.56: Reflexes of ultimate PMP *ə

PMP	*ənəm	*gatəl	*dəŋəR	*kədəŋ
PFL	*ənəm	*gatər	*dəŋər	*kəda
SK	<i>əna</i>	<i>gatar</i>	<i>rəna</i>	<i>ʔəra</i>
WL (LWL)	<i>nəm (ə)</i>	<i>gatə k</i>	<i>dəŋəʔ</i>	-
WL (LWI)	<i>nəm uj</i>	<i>gatə</i>	<i>weŋe</i>	-
WL (AD)	<i>nam u</i>	<i>gatə k</i>	<i>dəŋəʔ</i>	-
WL (MS)	<i>nəm u</i>	<i>gateʔ</i>	<i>dəŋa</i>	-
WL (AB)	<i>nam uj</i>	<i>gate</i>	<i>daŋa</i>	-
CL	<i>enəm</i>	<i>gatə k</i>	<i>dəŋər</i>	-
EL	[...]	[...]	<i>dəŋe</i>	-
KD	<i>enəŋ</i>	-	<i>dəŋər</i>	-
	‘six’	‘itchy’	‘hear’	‘stand’

Apart from the four PFL vowels *a, *i, *u and *ə that go back to PMP vowels, there is evidence for PFL *e without a PMP source. PFL *e is attested in intervocalic and final position. Table 5.57 provides cognate sets reflexes

tracing back to PFL *e. The word #ale ‘waist’ is not reconstructible to PFL because there is no reflex of it in Sika. Therefore, it is marked with # instead of * (cf. 6.3).

Table 5.57: Cognate sets containing PFL *e without PMP source

PFL	*tena	*kə-melu	#ale	*kera ^a
SK	<i>tena</i>	<i>melur</i>	-	<i>ʔera</i>
WL (AD)	<i>təna</i>	<i>kəmelut</i>	-	<i>keʔa</i>
WL (MS)	<i>təna</i>	<i>mɛluk</i>	<i>alɛ ŋ</i>	<i>kea</i>
CL	<i>tena</i>	<i>kəmelut ən</i>	<i>ale</i>	<i>kera</i>
EL	[...]	[...]	[...]	[...]
KD	<i>tɛnɛ</i>	<i>mɛluʔ</i>	<i>alɛ n</i>	<i>ere</i>
	‘canoe’	‘smooth’	‘waist’	‘turtle’

^a This word for ‘turtle’ has been reconstructed to PCEMP *kera ‘turtle’.

5.2.8 PMP vowel-glide sequences in final position

Table 5.58 shows the reflexes of PMP final vowel-glide sequences. Central Lamaholot retains the final glides after *a*. In all other instances the final glide is lost. The loss of the final glide can influence the quality of the preceding vowel. The last row in the table indicates the number of examples found for this pattern in my dataset.

Table 5.58: Reflexes of PMP *-aw, *-ay, *-iw, *-uy

PMP	*-aw	*-ay	*-iw	*-uy
PFL	*-av	*-ay	*-i/*-yu	*-i
SK	<i>o</i>	<i>e/i</i>	<i>i</i>	<i>i</i>
WL	<i>a</i>	<i>ay/a/e/i</i>	<i>e/yo</i>	<i>e</i>
CL	<i>av/a</i>	<i>ay/adʒ</i>	<i>i/dʒu</i>	<i>i/e</i>
EL	<i>a</i>	<i>a</i>	[...]	<i>i/e</i>
KD	<i>o</i>	<i>e</i>	<i>i</i>	<i>i/e</i>
Number of examples in database	3	8	2	3

As shown in the examples in Table 5.59, the final vowel-glide sequence PMP *-aw > av in Central Lamaholot with retention of the final glide, while in Sika and Kedang PMP *-aw > o and in Western and Eastern Lamaholot PMP *-aw > a. In CL-Imulolo, final *v > f.

Table 5.59: Reflexes of final PMP *-aw

PMP	*panaw	*takaw	*qaləjaw ‘day’
PFL	*panav	*t<əm>akav	*lədav ‘sun; day’ ^a
SK	<i>pano</i>	? <i>toʔi</i>	<i>ləro</i> ‘sun’
WL (LWI)	<i>pana</i>	<i>təmaʔa</i>	<i>rəra:</i> ‘sun’
CL (KK)	<i>pana(v)</i>	<i>takav</i>	-
CL (IL)	-	[...]	<i>lədʒaf</i> ‘sun’
EL	<i>pana</i>	[...]	<i>ləra</i> ‘sun’
KD	<i>pan</i>	<i>maʔo</i>	<i>loyo</i> ‘sun; day’
	‘walk’	‘steal’	‘sun; day’

^a All Flores-Lembata languages have a word meaning ‘day’ that is derived from the words for ‘sun’ listed here. In Kedang, both words have the same shape. CL-Kalikasa has replaced the word for ‘sun’ but retains a reflex in the word *lədʒon* ‘day’ which is derived with a suffix -n.

Central Lamaholot is most conservative in the retention of final PMP *w. Nevertheless, a partial loss of the final PMP glide *w is observed. The Central Lamaholot form *panav* ‘walk’ only appears with suffixes, such as in *da=panav-i* ‘3PL-walk-3PL’ = ‘they went’, otherwise *pana* is used. However, the CL form *takav* ‘steal’ never appears without the final consonant.

The final PMP sequence *-ay undergoes fortition to *adʒ* in Central Lamaholot, as well as sporadically in Western Lamaholot. Final *-ay is thus reconstructed to PFL, with the exception of body part nouns that take a nasal suffix. In these words, the final glide is deleted. In Sika and Kedang, PMP *-ay > e. The Sika reflex *vaʔi-n* ‘leg’ (< PMP *waqay) is an exception, as PMP *-ay > i. The Kedang reflex *vua* ‘rattan’ (< PMP *quay) also appears to be an exception, as PMP *-ay > a in this word. For Eastern Lamaholot little evidence is available. Nevertheless, it appears from EL *mata* ‘die’ (< PMP *matay) that PMP *-ay > EL a in final position.

Table 5.60: Reflexes of final PMP *-ay

PMP	*quay	*qatay	*matay	*qənay	*waqay	*sakay
PFL	*uay	*ate-n	*matay	*ənay	*vaʔi	*hakay
SK	<i>ue</i>	<i>vate n</i>	<i>mate</i>	<i>ne</i>	<i>vaʔi n</i>	<i>haʔe</i>
WL (LWI)	<i>uvay əŋ</i>	<i>ate ŋ</i>	<i>mata</i>	-	-	<i>haka</i>
WL (LH)	[...]	-	<i>matadʒ</i>	[...]	-	[...]
WL (MS)	<i>uve</i>	<i>ate ŋ</i>	<i>mate</i>	<i>əni</i>	-	-
CL (KK)	<i>uadʒ</i>	-	<i>matadʒ</i>	<i>ənadʒ</i> ‘soil’	-	<i>akadʒ</i>
EL	[...]	-	<i>mata</i>	[...]	-	[...]
KD	<i>vua</i>	<i>hate n</i>	<i>mate</i>	<i>ene</i>	-	<i>aʔ</i>
	‘rattan’	‘liver’	‘die’	‘sand’	‘leg’	‘ascend’

My dataset only contains two cognate sets that go back to PMP forms with the final vowel-glide sequence *-iw given in Table 5.61. Most likely due to the quality of the PMP consonant preceding the final sequence, the two sets develop in different ways. PMP *kahiw ‘wood; tree’ becomes PFL *kayu due to loss of medial *h and PMP *laRiw becomes PFL *p-lari/*kari. The initial syllable *pə is an innovation. The reflexes of PMP *kahiw > PFL *kayu follow the regular pattern of intervocalic PFL *y. In Sika and Kedang, ultimate PFL *yu > *i*. The reflexes of PMP *laRiw > PFL *p-lari/*kari follow the regular pattern of final PFL *i, which is lowered to *e* in Western Lamaholot.

Table 5.61: Reflexes of final PMP *-iw

PMP	*kahiw	*laRiw
PFL	*kayu	*p-lari/*k-ari
SK	<i>ʔai</i>	<i>p lari</i>
WL (LWI)	<i>kayo</i>	<i>pə laʔe</i>
CL (KK)	<i>kadʒu</i>	<i>k ari</i>
EL	-	[...]
KD	<i>ʔai</i>	-
	‘wood; tree’	‘run’

Final PMP *-uy is reflected regularly as PFL *i. The reflexes in the Flores-

Lembata languages follow the pattern of PFL *-i.

Table 5.62: Reflexes of final PMP *-uy

PMP	*hapuy	*babuy	*naŋuy
PFL	*api	*vavi	*naŋi
SK	<i>api</i>	<i>vavi</i>	<i>nani</i>
WL (LWI)	<i>ape?</i>	<i>vave</i>	<i>naŋe</i>
CL (KK)	<i>api</i>	<i>vavi</i>	<i>naŋe</i>
EL	<i>ape</i>	[...]	<i>naŋi</i>
KD	<i>api</i>	<i>vavi</i>	<i>naŋi</i>
	‘fire’	‘pig’	‘swim’

5.2.9 The Proto-Flores-Lembata phoneme inventory

This section summarises the PFL phonemes reconstructed based on regular sound correspondences in the Flores-Lembata languages. In addition, I summarise the PMP sources for the PFL sounds. For details on the reflexes please consult the individual sections above. Table 5.63 presents the Proto-Flores-Lembata (PFL) vowel inventory and Table 5.64 the PFL consonant inventory.

Table 5.63: Vowel inventory of Proto-Flores-Lembata

	Front	Central	Back
High	*i		*u
Mid	*e	*ə	*o
Low		*a	

The PFL vowels are retained from their PMP sources as such, with the addition of PFL *e which does not have a PMP source.

Table 5.64: Consonant inventory of Proto-Flores-Lembata

	Labial	Coronal	Dorsal	Glottal
Voiceless stops	*p	*t	*k	*ʔ
Voiced stops	*b	*d	*g	
Affricate		*dʒ		
Fricative	*v	*s		*h
Nasal	*m	*n	*ŋ	
Rhotic		*r		
Lateral		*l		
Approximant			*y [j]	

All reconstructed PFL consonants have regular PMP sources listed in 5.65, except for PFL *dʒ. Nevertheless, there is evidence for PFL *dʒ as a marginal phoneme of PFL (cf. §5.2.2.6). In Table 5.65, initial, intervocalic and final phonemes are only listed separately when different changes apply.

Table 5.65: PMP sources for PFL phonemes

PMP source	PFL	Position	Type of change
*p	*p	all	no change
*t	*t	all	no change
*k	*k	all	no change
*q	*ʔ	all	lenition
*b	*b	all	no change
*d / *j / *z	*d	all	merger
*g	*g	all	no change
*m	*m	all	no change
*n/*ŋ	*n-	initial	merger
*n	*-n-	intervocalic	no change
*ŋ	*-ŋ-	intervocalic	no change
*b/*w	*v	all	merger
*s	*s	all	no change
*s	*h	all	lenition
*R	*r	all	no change

PMP source	PFL	Position	Type of change
*R/ *d	*-r	final	merger
*l	*l	all	no change
*y	*y	all	no change
*j	*-y	final	lenition
-	*dʒ	initial	-
*a	*a	all	no change
*i	*i	all	no change
*u	*u	all	no change
*ə	*ə	all	no change
-	*e	all	-

5.2.10 Reconstructed initial clusters in PFL

Table 5.66 provides examples of cognates sets with reconstructed initial consonant clusters that most likely alternated with a single consonant form or another cluster. Three of the five sets are of PMP origin but the clusters cannot be traced back to PMP. No data on Eastern Lamaholot is available for these concepts.

The table shows different patterns in each set but there are similarities in the fact that a simple onset consonant gains a complex variant. For those sets with a PMP source, the original form and the innovation can be easily identified, such as PFL *vani ‘bee’ is the original, the base, and PFL *blani is the innovated form, as they are traced back to PMP *wani ‘bee’. In the synchronic forms, it appears that Sika and Kedang avoid complex onsets. Nevertheless, some of the Kedang forms, such as *lani* ‘bee’ and *nəbi?* ‘wall’ most likely go back to forms with complex onsets, similar to the forms still found in some Lamaholot varieties.

The processes behind this variation in onset cannot be entirely explained at the current stage. However, for the set denoting the concept ‘wall’, nominalising morphology can be recovered. The reconstruction *gəbi / *gnəbin ‘wall’ shows a process of nominalisation that is attested in CL-Central Lem-bata (cf. §3.3.6) and is most likely more wide-spread. Base forms starting with *g* are nominalised by the infix *-n-* and a suffix *-k* or *-n*. In some varieties, this process transforms the voiced *g* into voiceless *k*.

Table 5.66: PFL initial consonant clusters

	‘bee’	‘cheek’	‘run’	‘wall’	‘shoulder’
PMP	*wani	*pipi	*lariw	-	-
PFL	*vani / *blani	*pipi / *klipi	*kari / *plari	*gəbi / *gnəbi n	*kpali k / *kwali k
SK (HEW)	<i>vani</i>	<i>pipi</i>	<i>plari</i>	<i>gəbi</i>	<i>pali k</i>
KD	<i>lani</i>	<i>pipi n</i>	-	<i>nəbi?</i>	<i>vali</i>
CL (KK)	<i>blani</i>	<i>pipi</i>	<i>kari</i>	<i>kənəbin</i>	<i>kwale k</i>
CL (LR)	[...]	<i>klipi</i>	<i>kar</i>	<i>kənabe r</i>	<i>kwale k</i>
WL (LWL)	<i>vane</i>	<i>kəlipi?</i>	<i>pəlaʔe</i>	<i>kənəbi?</i>	<i>kpali k</i>
WL (AD)	<i>vane</i>	<i>pipi kət</i>	<i>palaʔe</i>	<i>kənəbi?</i>	-
WL (LWI)	<i>vane</i>	<i>kəlipi</i>	<i>pəlaʔe</i>	<i>kənəbi</i>	-
WL (MS)	<i>blane</i>	<i>pipi ŋ</i>	<i>plaeŋ</i>	<i>gəbe</i>	-
WL (PD)	<i>bəlane</i>	<i>pipi ŋ</i>	<i>plae</i>	<i>gnabeŋ</i>	-

From the data, which mainly comes from wordlists, it is not clear whether only one form is retained in the language, either the base or the derived form, or whether both forms are still in use but only one was given. In Pampus (1999:631) for WL-Lewolema, *gəbi?* is given as a verb meaning ‘construct a bamboo wall for a house’², while the derived form *kənəbi?* means ‘wall’. However, as varieties such as SK-Hewa with *gəbi* ‘wall’ and WL-Munaseli with *gəbe* ‘wall’ appear to use the base form to refer to the nominal concept ‘wall’, the nominalising morphology is most likely not functional anymore in these varieties.

5.3 Subgroups within Flores-Lembata

In this section, I summarise evidence for the subgroups established within the family of Flores-Lembata. First, in §5.3.1, I provide evidence for the lowest levels which have been considered individual languages or dialect clusters in previous work. These are Sika, Kedang, Central Lamaholot, Western Lamaholot and Eastern Lamaholot. Second, in §5.3.2, I show that there is little

² Original definition: ‘mit einer Wand (aus gespaltenem Bambus) versehen’

evidence to group these languages further into mid-level groups. Some groupings are more likely than others but none of them shows very convincing evidence. For examples of the sound changes given in the following section, see §5.2.

5.3.1 Evidence for low-level subgroups

5.3.1.1 Sika

The following regular sound changes only occur in Sika and thus define Sika as an independent branch of Flores-Lembata.

1. PFL *d > r in all positions
2. PFL *-ŋ- > n in intervocalic position
3. PFL *mp- > b in initial position
4. PFL *mt- > d in initial position

Most of these changes could have been active at the same point in time. Only Change 4 must have occurred after Change 1 was completed. Otherwise the phonemes *d* resulting from PFL *m-t word initially would have become *r* as well.

In addition, Sika also undergoes PFL *k > ʔ in all positions, PFL *d > r in all positions and PFL *s > h in all positions. However, these changes are not unique to Sika but also occur in other Flores-Lembata languages.

Examples of lexical innovations in Sika are (i) *gahar* ‘tall’, (ii) *heret* ‘yellow’ (≠ PFL *kumas-ən ‘yellow’ < ? Malay *kuning* ‘yellow’ + *mas* ‘gold’) and (iii) *lapo* ‘house’ (≠ PFL *ruma ‘house’ < PMP *Rumaq ‘house’). For the concept ‘tall’ no PFL form is known. All subbranches appear to have innovated different words. PFL *ruma ‘house’ is only retained in Alorese with the form *uma*, all other Flores-Lembata languages have innovated a word for house.³

Possible cognates of *lapo* ‘house’ are found on Rote and Timor, such as Termanu *lopo* ‘shelter’ and Meto *lopo* ‘Timorese round houses for social activities’ (Owen Edwards, pers. comm.). Sika could have borrowed the word from these languages and changed *o* > *a*. Nonetheless, it remains unknown how the contact between the Sika speakers and the Rote-Meto speakers could have taken place. For the other two innovations, no cognates or similar forms could be found in other languages so far.

³ WL (except for Alorese) has *layo* ‘house’, CL has *una* ‘house’ and KD has *huna* ‘house’. The CL and KD words probably go back to PMP *qunəj ‘inside’.

5.3.1.2 Western Lamaholot

In Western Lamaholot varieties one exclusively shared sound change is attested: PFL *r > ʔ.

Western Lamaholot also undergoes intervocalic PFL *-d- > r and PFL *s > h in all positions. However, these are changes that are also attested in other subgroups. Therefore, they are not subgroup-defining for Western Lamaholot. Nevertheless, the change of PFL *-d- > r in Western Lamaholot must have occurred after the change of PFL *r > ʔ, as otherwise PFL *r and PFL *-d- would have merged to ʔ. Also for *s > h it is likely that it is a rather recent change in Western Lamaholot because s remains in ritual speech and in some fossilised derivatives (Pampus 1999:28).

There are two subsequent diffused changes attested in individual varieties of Western Lamaholot. These are Proto-Western Lamaholot (PWL) *w > f and PWL *y > dʒ. PWL *w > f is mainly found in the varieties of southern Lembata and in Alorese on Alor. Fortition of PWL *y > dʒ is attested scattered in several areas throughout the whole Lamaholot area (Elias 2017a).

Examples of lexical innovations in the Western Lamaholot subgroup are *blaha* ‘long’ and the clause-final negator PWL *hala ‘NEG’ (< PMP *salaq ‘wrong’). The innovation of the negator PWL *hala is not only a semantic change but also a syntactic change. PFL had pre-predicate negation but this new negator is placed clause-finally (cf. §10.3). Another exclusively shared innovation of the Western Lamaholot subgroup is the loss of initial schwa and the addition of final u(ɣ) to the numeral ‘six’ from PFL *ənəm ‘six’, which is now realised as *namu* or *nəmuɣ* in the WL varieties.

Alorese spoken on the islands of Alor and Pantar has been identified as an independent language (Klamer 2011). Based on the shared sound change of PFL *r > ʔ and the other shared innovations just listed, Alorese is part of the Western Lamaholot.

5.3.1.3 Central Lamaholot

The following sound changes define the subgroup of Central Lamaholot. They go back to the level of Proto-Central Lamaholot (PCL).

1. PFL *-d- > PCL *-dʒ- in intervocalic position
2. PFL *h > Ø in all positions
3. PFL *ʔ > Ø in all positions

There are a few subsequent changes attested in individual varieties of Central Lamaholot. For example, the variety of Painara changes all PCL $*s > h$ and PCL $*dʒ$ ($< \text{PFL } *-d- < \text{PMP } *d/*j/*z$) undergoes an unconditioned split of $*dʒ$ to y and $dʒ$. Also in Lewokukung, PCL $*dʒ$ sporadically changes into y but this change affects a different selection of items as in Painara. In Central Lembata, the reflex of PCL $*y$ is $dʒ$.

The change of PCL $*s > h$ and PCL $*dʒ > y$ in Painara and the retention of PCL $*s = s$ and the change of PCL $*y > dʒ$ in Central Lembata creates a salient distinction in these two adjacent varieties. Painara has the phonemes h and y but no s and almost no $dʒ$, while Central Lembata has the phonemes s and $dʒ$ but no h nor y . This clear linguistic border coincides with the socio-cultural division of Paji and Demon villages that is found throughout the Lamaholot area (cf. §1.2.3.4). In the past, Paji and Demon were enemies involving a lot of mistrust and violent acts of killing. In the case of the adjacent varieties of Painara and Central Lembata the phonemic distinctiveness of the varieties appears to coincide with their socio-cultural distinctiveness. It could be that the enmity between the two areas has contributed to the phonological diversification of the two varieties.

In the varieties of Central Lamaholot and beyond, cases of sound change diffusion can be observed. Glide fortition of $*y > dʒ$ is also attested in all other Central Lamaholot varieties, except for Painara, as well as in several varieties of Western Lamaholot (Elias 2017a). In the western varieties of Central Lamaholot, the changes of $*s > h$ and $*w > f$ have probably diffused from neighboring Western Lamaholot varieties (Elias 2017a).

Overall, Central Lamaholot is phonologically the most conservative subgroup of Flores-Lembata. Central Lamaholot is the only FL group that retains PFL $*s = \text{PCL } *s$ and PFL $*-av = av$. In addition, final consonants are retained more frequently in Central Lamaholot than in other Flores-Lembata languages.

Examples of lexical innovations of Central Lamaholot are *keda* $|k$ 'big' ($\neq$ PFL $*raya < \text{PMP } *Raya$), *s<n>əgur* 'smoke' and *luvak* 'sun' ($\neq$ PFL $*lədav$ 'sun' $< \text{PMP } *qaləjaw$ 'day'). A reflex of PFL $*raya$ 'big' is only retained in Kedang *ria*: 'big', all other subgroups have innovated a new form for this concept.⁴ No PFL form for 'smoke' can be reconstructed.

In addition to lexical innovations, there are two morphological innova-

⁴ WL and EL have *bela* 'big' or similar forms, and SK has *gəte* 'big'.

tions that are exclusive to Central Lamaholot: plural marking with the suffix *-dʒa* and coda alternation in alienable nouns which results in nominal lexemes with two surface forms, such as *au/aor* ‘dog’ (cf. §8.4.2.1 and §3.3.3).

5.3.1.4 Eastern Lamaholot

No exclusively shared sound change is attested in the Eastern Lamaholot varieties. Eastern Lamaholot undergoes PFL *-d- > *r*, PFL *s > *h* and PFL *k > ʔ. These changes have also occurred in other subgroups and may have diffused to Eastern Lamaholot. As all three changes are very common sound changes, it may also be pure chance that they occur in more than one subgroup.

Examples of lexical innovations in Eastern Lamaholot are *ʔuhu* ‘narrow’ (≠ PFL *kiput < PMP *kiput ‘narrow’) and *əso* ‘tree’ (≠ PFL *kayu ‘tree; wood’ < PMP *kahiw ‘wood; tree’). The latter could be related to forms in Alor-Pantar languages, such as Kula *asaka* ‘tree’ or Sawila *asəkə* ‘tree’.

5.3.1.5 Kedang

The following sound changes only occur in Kedang and thus define Kedang as an independent subgroup within Flores-Lembata.

1. PFL *g > *k* in all positions
2. PFL *-d- > (*dʒ >) *y*/∅ in intervocalic position; PFL *d > *r* before *u(a)*

Kedang also undergoes PFL *s > *h and PFL *k > ʔ in all positions which is not listed as an exclusive innovation because it also occurs in Sika and Eastern Lamaholot. However, the sound change PFL *g > *k* must have happened after the change of PFL *k > ʔ had occurred. Otherwise, PFL *k and PFL *g would have merged as ʔ.

Examples of lexical innovations in Kedang are *ʔebo* ‘tail’ (≠ PFL *ikur < PMP *ikuR) and *uben* ‘night’. For the concept night, no PFL form is known. All subgroups appear to have innovated different words for ‘night’.

5.3.2 No good evidence for mid-level subgroups

At the current stage of research, there is no good evidence for mid-level subgroups in the Flores-Lembata family. Mid-level subgroups would unite two

or more of the low-level groups defined in §5.3 by exclusively shared innovations.

Nevertheless, there are three sound changes that are attested in more than one subgroup:

- PFL *k > ʔ in SK, KD and EL
- PFL *-d- > r in SK, WL and EL
- PFL *s > h in SK, KD, WL and EL
- PFL *-aw > -o in SK and KD

There are two main reasons for *not* basing mid-level subgroups on these changes. (1) The changes just listed are all cross-linguistically very common, thus they would only be weak subgrouping evidence. If all three sound changes would be exclusive to the same set of languages, one may be able to take them as evidence for a subgroup. However, this is not the case for the Flores-Lembata languages. (2) For at least some of these changes, I have pointed out in the section above that they must have happened after another sound change in that subgroup to explain the synchronic forms in the subgroup. For example, the subgroup-defining change for WL PFL *r > ʔ must have happened before PFL *-d- > r, as -r- < PFL *-d- is retained unchanged in the WL varieties and not changed into ʔ. For these reasons, I suggest that the shared changes are either independent developments in the individual subgroups or spread through vertical diffusion after the split-up of Proto-Flores-Lembata. For Sika, an individual development of PFL *-d- > r is likely as Sika undergoes the same change also in initial position.

Apart from the shared sound changes just discussed, there is another potentially shared sound change of PFL *-d- > dʒ in CL and KD. In Kedang, there is only indirect evidence for this change, as nowadays all PFL *-d- have become y in Kedang. While the change of PFL *-d- > *dʒ could provide evidence for a CL-Kedang subgroup, it must be borne in mind that reconstruction of medial *-d- for PFL is somewhat problematic (§5.2.2.4) and an alternate reconstruction is PFL *-dʒ-. However, in this alternative case of reconstruction, intervocalic dʒ would be a shared retention in CL and Kedang and provide no evidence for subgrouping. Given this possibility, I prefer not to subgroup CL and Kedang together on the basis of *-d- > dʒ.

Despite the absence of shared sound changes, there is lexical evidence for treating the Lamaholot varieties as a single unit apart from Sika and Kedang. This is based on lexical statistics by Keraf (1978a) who calculates

about 50% shared basic vocabulary between WL, EL and CL, whereas only about 20-30% is shared by these three groups with Kedang or Sika. However, this evidence is not particularly strong subgrouping evidence because lexical items could have easily diffused between these adjacent languages. Furthermore, their speakers are in intensive contact with each other. Only in combination with exclusively shared sound changes, could the lexical evidence could provide additional support. For the three Lamaholot subgroups, no shared sound change is attested. Therefore, there is no good basis for proposing a proto-language, such as Proto-Lamaholot. Nevertheless, due to the relative lexical similarity and the more recent shared social history it is justifiable to speak of Lamaholot as a unit when keeping in mind that this language group historically encompasses three independent subgroups of Flores-Lembata.

There might be a slight evidence for Kedang-Lamaholot as a subgroup including Kedang, Western Lamaholot, Eastern Lamaholot and Central Lamaholot based on shared lexical innovations. Kedang and the Lamaholot subgroups have 73 shared lexical innovations compared with 41 lexical items shared between Lamaholot and Sika (cf. §6.3.2 and §6.3.3). However, given the lack of any supporting evidence I do not consider this evidence alone strong enough to assume a Kedang-Lamaholot subgroup.

5.4 Evidence for a Flores-Lembata subgroup

The subgroup of Flores-Lembata encompasses exclusively all five innovation-defined groups discussed in §5.3.1: Sika, Kedang, Central Lamaholot, Western Lamaholot and Eastern Lamaholot. The Flores-Lembata subgroup is defined by the following three sound changes. Examples of cognate sets containing the sounds affected by the PFL sound changes have been listed earlier in the tables in §5.2.

1. PMP *ŋ-/*n- > PFL *n- / #_
2. PMP *z/*d/*j > PFL *d
3. PMP *s > PFL *s/*h

While each of these individual changes may be found in other groups, such as for example *ŋ > n in Timor-Babar, no other languages share the exact set of these three changes.

5.5 Evidence for a Bima-Lembata subgroup

After having shown that Flores-Lembata is an innovation-defined subgroup, I propose that Flores-Lembata can be grouped together with the languages of Central Flores, Western Flores, Bima, Sumba-Hawu based on the unconditioned split of initial PMP *b- > *b-/*w- in the same lexical items, as in Flores-Lembata (cf. §5.2.2.2). I name this subgroup Bima-Lembata as these two islands are the westernmost island (Bima) and the easternmost island (Lembata) of this subgroup. All languages in between these islands, including the Sumba-Hawu languages in the south, are part of Bima-Lembata. Examples of initial PMP *b- > w in a representative selection of languages of this area are listed in Table 5.67.

Table 5.67: Initial PMP *b- > PBL *w- in the Bima-Lembata languages

	‘pig’	‘stone’	‘fruit’	‘moon’	‘woman’	‘foam’
PMP	*babuy	*batu	*buaq	*bulan	*bahi	*bujəq
PFL	*vavi	*vatu	*vua	*vulan	*vai	*vuda
SK	<i>vavi</i>	<i>vatu</i>	<i>vua</i>	<i>vulan</i>	<i>vai</i>	<i>vura n</i>
KD	<i>vavi</i>	<i>vaʔ</i>	<i>(v)ua</i>	<i>vula</i>	-	<i>vura n</i>
CL	<i>vavi</i>	<i>vatu</i>	<i>vua k</i>	<i>vulan</i>	-	<i>pə vudʒa</i>
WL (LWI)	<i>vave</i>	<i>vato</i>	<i>vua ŋ</i>	<i>vula</i>	<i>vai</i>	<i>vura haŋ</i>
EL	[...]	<i>vato</i>	<i>vuʔã</i>	[...]	-	[...]
PCF	*wawi	[...]	[...]	*wula	*fai	*woda
Lio	<i>vavi</i>	<i>vatu</i>	-	<i>vula</i>	<i>fai</i>	<i>vora</i>
Ende	<i>wawi</i>	<i>watu</i>	-	<i>wurha</i>	<i>hai</i>	<i>wora</i>
Ngao	<i>vavi</i>	<i>vatu</i>	-	<i>wuʒa</i>	-	<i>vo.ia</i>
Keo	<i>wawi</i>	<i>watu</i>	-	<i>wuda</i>	<i>fai</i>	-
Nage	<i>vavi</i>	<i>vatu</i>	-	<i>vuda</i>	<i>fai</i>	<i>vo.ia</i>
Ngada	<i>vavi</i>	<i>vatu</i>	<i>vua</i>	<i>vula</i>	<i>fai</i>	<i>voza</i>
Rongga	<i>wawi</i>	<i>watu</i>	-	<i>wula</i>	<i>fai</i>	<i>vo.ia</i>
Palu'e	<i>wawi</i>	<i>watu</i>	-	<i>wula</i>	<i>wai</i>	-
Manggarai	-	<i>watu</i>	<i>wua</i>	<i>wulaŋ</i>	<i>wai</i>	<i>wusa</i>

	‘pig’	‘stone’	‘fruit’	‘moon’	‘woman’	‘foam’
Komodo	-	<i>batu</i>	<i>wua</i>	<i>wulan</i>	-	-
Bima	<i>vavi</i>	<i>vadu</i>	<i>vua</i>	<i>vura</i>	-	[...]
Kambera	<i>wei</i>	<i>watu</i>	<i>wua</i>	<i>wulan</i>	-	<i>wura</i>
Hawu	<i>vavi</i>	<i>vo vadu</i>	<i>vue</i>	<i>vəru</i>	-	<i>voro</i>
PRM	*bafi	*batu	*bua-k	*bulan	*fee	*fudzə

PMP = Proto-Malayo-Polynesian, PRM = Proto-Rote-Meto (Owen Edwards, pers. comm.)

PCF = Proto-Central Flores (Elias 2018), PFL = Proto-Flores-Lembata

I propose that the change of *b- > w in relevant lexical items occurred at the level of Proto-Bima-Lembata, and those probably merged with reflexes of PMP *w. Proto-Bima-Lembata (PBL) *w then further develops into *v*, *ν* or *f* in individual languages. The reflex of PBL *w is PFL *v. The change of PBL *w > f only occurs in individual lexemes of Proto-Central Flores (PCF), such as PCF *fai ‘woman’ (< PMP *bahi). A similar change of PBL *w > PFL *v > f is regular in some varieties of Western and Central Lamaholot (cf. 5.3.1). Komodo *batu* ‘stone’ is the only irregular reflex in these cognate sets as it shows PMP *b > *ḃ*.

The last line of the table provides Proto-Rote-Meto (PRM) reconstructions to show that the Rote-Meto languages do not follow the same pattern. The Austronesian Rote-Meto languages are the geographically closest eastern neighbour of Bima-Lembata.

Words which do not undergo PMP *b > w have various reflexes in the languages of Flores-Lembata, Flores, Bima and Sumba, as shown in Table 5.68. In Flores-Lembata *b is retained as *b*, while in the other languages different patterns emerge. Some extend the shift to *w* further, such as some Central Flores languages, while others keep *b* or change it into *ḃ*. The most regular is Palu'e where all words with PMP *b that do not undergo *b > w/ν underwent *b > *p* instead. As the same lexical items undergo PMP *b > w initially in all languages of Bima-Lembata, while the other lexemes reflect PMP *b in different ways, the split of PMP *b- > *b-/*w- in initial position is strong evidence for an innovation-defined subgroup. The Austronesian languages of Timor show a different pattern. Thus for example, while the Rote-Meto languages show a split of PMP *b > f/b, the distribution does not affect the same lexemes as for the languages of Flores, Bima and Sumba.

Table 5.68: Initial PMP *b- > b/ɓ/p/w in the Bima-Lembata languages

	‘return’	‘divide’	‘pound’	‘heavy’	‘white’	‘flower’
PMP	*balik	*baqagi	*bayu	*bərəqat	*budaq	*buŋa
PFL	*balo(ŋ)	*bagi	*bayu	*bərat	*budaʔ	*buŋa
SK	<i>balon</i>	<i>bige</i>	<i>bai</i>	<i>bərat</i>	<i>buraʔ</i>	-
KD	-	<i>boʔ</i>	<i>bae</i>	<i>baraʔ</i>	<i>buyaʔ</i>	-
CL	-	-	<i>badzu</i>	<i>bərat</i>	<i>budza k</i>	<i>buŋa</i>
WL (LWI)	[...]	-	<i>bayo</i>	<i>baʔa</i>	<i>bura</i>	-
EL	-	[...]	[...]	<i>bəra</i>	<i>burō</i>	-
PCF	*ɓale	[...]	*wayu	-	-	[...]
Lio	<i>ɓale</i>	-	<i>wadzu</i>	-	-	<i>voŋa</i>
Ende	<i>ɓaɛ</i>	<i>bagi</i>	<i>wadzu</i>	-	-	<i>woŋa</i>
Ngao	-	-	<i>vadzu</i>	-	-	<i>voŋa</i>
Keo	<i>ɓade</i>	<i>bagi</i>	<i>wadzu</i>	-	-	<i>woŋa</i>
Nage	-	-	<i>watfu</i>	-	-	<i>voŋa</i>
Ngada	<i>ɓale</i>	<i>bagi</i>	<i>wadzu</i>	-	-	<i>voŋa</i>
Rongga	<i>ɓale</i>	<i>bayi</i>	<i>wadzu</i>	-	-	-
Palu'e	<i>palu</i>	-	<i>padzu</i>	<i>pədza</i>	<i>pura</i>	-
Manggarai	-	<i>bahi</i>	-	-	-	-
Komodo	<i>waleʔ</i>	[...]	<i>wadzu</i>	-	-	-
Bima	<i>mbali</i>	<i>bage</i>	<i>mbadzu</i>	<i>bara</i>	<i>bura</i>	<i>buŋa</i>
Kambera	<i>beli</i>	-	<i>bai</i>	-	<i>burahu</i>	-
Hawu	[...]	<i>bəke</i>	-	-	<i>? vo pudi</i>	[...]
PRM	*bali/*ɓali	-	*mbau	*berat	-	*ɓuna-k

PMP = Proto-Malayo-Polynesian, PRM = Proto-Rote-Meto (Owen Edwards, pers. comm.),

PCF = Proto-Central Flores (Elias 2018), PFL = Proto-Flores-Lembata

Therefore, I conclude that the split of initial PMP *b- > *b-/*w- is a shared innovation and which can be posited as evidence for subgrouping the languages of Flores-Lembata, Central Flores, Western Flores, Bima, and Sumba-Hawu. This subgroup can be called Bima-Lembata based on its geographic extension from west to east. Such a subgroup has been suggested by Blust (2008:48) who proposed that the Sumba-Hawu group may include languages

of Western and Central Flores, but only can include Bima if also languages further east are included. These languages further east are the Flores-Lembata languages. The internal division of Bima-Lembata remains to be investigated further. At the present stage of research, Sumba-Hawu (Blust 2008), Central Flores (Elias 2017b) and Flores-Lembata (§5.4) are innovation-defined subgroups within the Bima-Lembata languages.

5.6 Conclusions

In this chapter, I have shown regular reflexes of Proto-Malayo-Polynesian (PMP) consonants and vowels in all languages of Flores-Lembata. I defined five subgroups within Flores-Lembata which are each supported by shared sound changes and lexical innovations.

Sika is defined by the exclusive sound changes of PFL *d > r, PFL *-ŋ- > n, PFL *mp- > b and PFL *mt- > d, as well as lexical innovations, such as *ga-har* ‘tall’, *heret* ‘yellow’ and *lapo* ‘house’. Kedang is defined by the exclusively shared sound changes of PFL *g > k and PFL *-d- > j/Ø, as well as lexical innovations, such as *?ebo* ‘tail’ and *uben* ‘night’. Central Lamaholot is defined by the exclusively shared sound changes of PFL *-d- > dʒ, PFL *h > Ø, and PFL *ʔ > Ø, as well as lexical innovations, such as *kedak* ‘big’, *snəgur* ‘smoke’, *luvak* ‘sun’. Western Lamaholot is defined by the exclusively shared sound change of PFL *r > ʔ and lexical innovations, such as *blaha* ‘long’, and the semantic change of *hala* ‘wrong; mistake’ > *hala* ‘NEG’. Eastern Lamaholot does not undergo any exclusively shared sound change, but there are lexical innovations, such as *?uhu* ‘narrow’ and *əso* ‘tree’.

This work is the first to include Eastern Lamaholot and Central Lamaholot varieties and establish them as independent subgroups of the Flores-Lembata family. There is no evidence to group any of the five subgroups together as a mid-level subgroup within Flores-Lembata.

Further, I provided evidence for Flores-Lembata as a subgroup based on three shared sound changes: PMP *ŋ- > PFL *n in initial position, PMP *j/*z/*d > PFL *d and PMP *s > PFL *s/*h.

Finally, I provided evidence for a larger Bima-Lembata subgroup including Flores-Lembata, Bima, Sumba-Hawu and West-Central Flores based on the lenition of initial PMP *b- > *w- in a specific set of lexical items which do not show this lenition in other languages of the region.

CHAPTER 6

Inheritance and innovation in the lexicon

6.1 Introduction

This chapter is about the lexicon of the Flores-Lembata languages and their ancestor Proto-Flores-Lembata (PFL). I show that PFL has a largely Austronesian vocabulary, as about 80% of my PFL reconstructions have a Proto-Malayo-Polynesian (PMP) source. However, the individual subgroups of Flores-Lembata underwent a considerable addition of vocabulary that cannot be traced back to PMP. The amount of additional vocabulary varies per subgroup. Larger amounts of additional vocabulary is found in Lamaholot, followed by Kedang, and then Sika with the smallest amount of lexical additions. In the Central Lamaholot subgroup, the non-Austronesian component amounts to more than 50%.

I propose that this new vocabulary is a lexical substrate that entered the subgroups due to contact with now extinct non-Austronesian languages. A lexical substrate is a layer of vocabulary from one or more substrate languages that are not spoken any more. Typically, the speakers of the substrate languages have shifted to new languages — here the proto-languages of the Flores-Lembata subgroups — and retained part of the lexicon of their original language. An alternative proposal could be to identify the newly added

vocabulary in the Flores-Lembata subgroups as (large scale) borrowing. In this case, one would not assume that non-Austronesian speakers shifted to the Flores-Lembata languages but that the speakers of the Flores-Lembata languages borrowed large amounts of vocabulary from a non-Austronesian source. There are two arguments that point to substrate rather than large scale borrowing. First, due to the genetic diversity of the population of the region that shows an almost equal Austronesian / non-Austronesian mix (cf. §1.3.2), there must have been a considerable amount of non-Austronesian speakers all over this region in the past. Nowadays all languages spoken in the region, except for the Timor-Alor-Pantar languages, are Austronesian. More likely than the extinction of a large amount of non-Austronesian speakers, which would have caused the genetic signal to be much weaker, is that the speakers of non-Austronesian languages shifted to Austronesian languages. Second, not only the lexicon but also the morpho-syntax of the Flores-Lembata languages shows non-Austronesian features as I will show in Part III of this dissertation. The transfer of morpho-syntactic features is a typical outcome of language shift preceded by a period of bilingualism (Muysken 2010:272). Borrowing grammatical features from an unrelated language is less likely.

This chapter is based on a systematic analysis of 422 lexeme sets collected from the Flores-Lembata wordlists stored in the Lexirumah database. A lexeme set is a set of related forms based on cognacy or borrowing. See §4.3 for more details on the methodology applied for this chapter and §4.4 for conventions in data representation.

Table 6.1 gives an overview of the results of this chapter. The lexeme sets are sorted by reconstructability to PFL according to three categories: (i) 210 sets that can be reconstructed to Proto-Flores-Lembata (PFL), (ii) 185 sets that cannot be reconstructed to PFL but have regular correspondences and (iii) 27 sets with irregular correspondences. Further, the three types of sets are assessed according to the availability of a PMP source. Most of the PFL reconstructions have a PMP origin. The lexeme sets of non-PMP origin are largely regular and with 248 sets in total, they outnumber those of PMP origin with 174 sets in total.

Table 6.1: Lexeme sets analysed for this chapter

	Total	PMP	no PMP source	Section
PFL	210	173	37	§6.2
Unreconstructible regular sets	185	-	185	§6.3
Unreconstructible irregular sets	27	1	26	§6.4
Total	422	174 (40%)	248 (60%)	

The lexeme sets that show regular sound correspondences but are classified as unreconstructible may ultimately be reconstructible to PFL but reflexes in some of the subgroups were lost or are missing from current data sources. More comprehensive data from the Flores-Lembata languages will likely show that a number of these lexeme sets are indeed reconstructible to PFL. However, it is also possible that these words are (early) borrowings that entered the Flores-Lembata languages after the break-up into subgroups but before the subgroup-defining sound changes occurred. These sets cannot be identified as late loans as they underwent the expected regular sound changes. The scenario of early borrowings would presuppose geographical separation of the subgroups, then contact resulting in lexical borrowing and only after new vocabulary had entered the subgroups, the regular sound changes occurred.

This chapter is divided into four sections. §6.2 presents the 210 lexeme sets that can be reconstructed to PFL according to the criteria explained in §4.3. §6.3 discussed the 185 unreconstructible lexeme sets which, nevertheless, show regular sound correspondences between subgroups. §6.4 discusses cases of lexeme sets with irregular correspondences. Some of these may be explicable by borrowing. §6.5 summarises the findings of the chapter and draws conclusions.

6.2 Proto-Flores-Lembata reconstructions

6.2.1 Overview

This section presents 210 cognate sets that can be reconstructed to Proto-Flores-Lembata (PFL). Most of these reconstructions (around 80%) have a

PMP source (§6.2.2), while a subset (around 20%) does not match to any known PMP form (§6.2.3).

6.2.2 PFL reconstructions with PMP sources

This section lists and discusses the 173 PFL reconstructions that have a PMP source. Out of these most have reflexes in all Flores-Lembata subgroups, meaning in Sika, Kedang and in at least one of the Lamaholot varieties (n=113). But there are also a few PFL reconstructions that only have reflexes in a subset of the Flores-Lembata subgroups. These are presented in separate tables. The PFL reconstructions in this section reflect the PMP form in a largely regular way. However, this does not exclude some minor irregularities at the level of individual phonemes, such as insertions of an additional phoneme (marked by a vertical line probably reflecting historical affixation), the loss of a single phoneme, sporadic metathesis or sporadic vowel changes.

Table 6.2 lists 113 PFL forms that are of PMP origin and are reflected with largely regular sound correspondences in all Flores-Lembata subgroups.

Table 6.2: PFL reconstructions and their PMP sources (n=113)

PFL	PFL meaning	PMP
*aku	'1SG'	*i aku
*kami	'1PL.EXCL'	*kami
*kita	'1PL.INCL'	*kita
*hida	'3PL'	*si ida
*tudu	'accuse'	*tuzuq
*pəniki	'bat'	*paniki
*vani/*blani	'bee'	*wani
*manuk	'bird; chicken'	*manuk
*m-paʔit	'bitter'	*paqit
*mitəm	'black'	*ma-qitəm
*puhun	'blossom; flower'	*pusuŋ 'heart; heart of banana'
*prupi/plupi	'blow'	*upi
*vulu-k	'body hair'	*bulu
*luri	'bone'	*duRi
*vuhur	'bow'	*busuR

PFL	PFL meaning	PMP
*(t)usu	'breast'	*susu
*mamaʔ	'chew'	*mamaq
*pipi/*klipi	'cheek'	*pipi
*ana(k)	'child; small'	*anak
*piliʔ	'choose'	*piliq
*hakay	'climb'	*sakay
*mai	'come'	*mai
*vatar	'corn; maize'	*batad 'millet; sorghum'
*lədav	'day; sun'	*qaləjaw 'sun'
*matay	'die'	*m-atay
*gali	'dig'	*kali
*bagi	'divide'	*baqagi
*ahu	'dog'	*asu
*-inu	'drink'	*inum
*mada	'dry; thirsty'	*maja
*pa-vari	'dry in sun'	*waRi
*kVan	'eat'	*kaən
*təlur	'egg'	*qatəluR
*mata	'eye'	*mata
*ama	'father'	*ama
*api	'fire'	*hapuy
*ikan	'fish'	*hikan
*təməla	'flea'	*qatiməla
*vuda	'foam'	*bujəq
*ləpət	'fold'	*lipət
*tuʔan	'forest'	*tuqan
*vua-n	'fruit; betelnut'	*buaq
*m-pənu-k	'full'	*pənuq
*bəli	'give'	*bəRay
*udu	'grass; bush'	*udu
*lima	'hand, arm, five'	*qalima
*kutu	'headlice'	*kutu
*dəŋər	'hear'	*dəŋəR
*bərat	'heavy'	*(ma)bəRəqat

PFL	PFL meaning	PMP
*pida	‘how many’	*pija
*bə-ləma ^a	‘inside; deep’	*daləm
*una	‘inside; house’	*qunəj ‘pith of plant; core’
*viri	‘left side’	*kawiri
*tave	‘laugh’	*tawa
*ʔapur	‘lime’	*qapur
*vivir	‘lips’	*biRbiR ‘lower lip’
*isi-k or *ihi-k	‘meat’	*isi
*vulan	‘moon’	*bulan
*ina	‘mother’	*ina
*ili	‘mountain’	*qilih
*vava	‘mouth’	*baqbaq
*nadan	‘name’	*ŋajan
*pusər	‘navel’	*pusəj
*vəru	‘new’	*baqəRu
*niduŋ/iduŋ	‘nose’	*ŋijuŋ/*ijuŋ
*m-tuʔa	‘old (people)’	*ma-tuqah
*əha	‘one; alone’	*əsa
*uti	‘penis’	*qutin
*ata	‘person’	*qaRta ‘outsider, alien people’
*vavi	‘pig’	*babuy
*bayu	‘pound’	*bayu
*veli	‘price; bride price; expensive; buy’	*bəli
*udan	‘rain’	*quzan
*uay	‘rattan’	*quay
*vanan	‘right side’	*ka-wanan
*m-tasak	‘ripe’	*ma-tasak
*lalan	‘road’	*zalan
*ramut	‘root’	*Ramut
*layar	‘sail’	*layaR
*m-pədu	‘salty’	*qapəju ‘gall’ > *ma-pəju
*sama	‘same’	*sama
*ənay	‘sand’	*qənay
*garu	‘scratch’	*garut

PFL	PFL meaning	PMP
*tahik	'sea'	*tasik
*pitu	'seven'	*pitu
*iu	'shark'	*qihu
*m-tidəm	'sharp'	*tazim 'whet'
*meya	'shy; ashamed'	*ma-həyaq
*ənəm	'six'	*ənəm
*ular	'snake'	*hulaR
*mətala	'star'	*mantalaq 'Venus'
*t<m>akav	'steal'	*takaw
*tai	'stomach; belly'	*tian
*vatu	'stone'	*batu
*mulur	'straight'	*lurus
*təvu	'sugarcane'	*təbuh
*naŋi	'swim'	*naŋuy
*luu	'tear'	*luhəq
*pulu	'ten'	*sa-ŋa-puluq
*m-kapal	'thick'	*ma-kapal
*rivu/*ribu	'thousand'	*Ribu
*təlu	'three'	*təlu
*panav	'walk'	*panaw
*kayu	'tree; wood'	*kahiw
*dʒua ^b	'two'	*duha
*uta	'vegetable; bean'	*qutan
*vaʔir	'water'	*wahiR
*apa	'what'	*apa
*budaʔ	'white'	*budaq
*aŋin	'wind'	*haŋin
*binay	'woman; sister'	*binay 'woman'
*sala	'wrong'	*salaq
*vadi	'younger sibling'	*huaji

^a The prefix *b-* is a nominaliser in CL-Central Lembata (cf. §3.3.6.1).

^b PFL *dʒ- < PMP *d- is an irregular reflex.

Table 6.3 lists the 24 PFL reconstructions that have a PMP form and show largely regular correspondences in the three Lamaholot subgroups and in

Kedang. Despite the absence of a reflex in Sika, this is enough evidence to classify these sets as cognate sets and reconstruct these words to PFL with the assumption that Sika has replaced these concepts with new words or a reflex is not attested in my dataset.

Table 6.3: PFL reconstructions without Sika reflex (n=24)

PFL	PFL Meaning	PMP source
*hakay	‘ascend’	*sakay
*raya	‘big’	*Raya
*tuno	‘burn; grill’	*tunu
*tanem	‘bury’	*tanəm
*doa ^a	‘far; long’	*zauq
*pukət	‘fishnet, fish trap’	*pukət
*kavil ^b	‘fishhook’	*kawil
*əpat	‘four’	*əpat
*paluk	‘hit’	*palu
*k-silap	‘lightning’	*silap ‘sparkle; drizzle’
*təkek	‘lizard’	*təktək
*a(m)pu	‘mother’s brother’	*əmpu ‘grandparent/grandchild’
*nusu	‘mouth’	*ŋusu
*kiput	‘narrow’	*kiput
*garanj	‘rough’	*garanj
*takut	‘scared’	*takut
*kələm	‘sky’	*kələm ‘dark, overcast, obscure’
*diri	‘stand’	*diRi
*lahe-k	‘testicles’	*lasəR
*m-nipih-i	‘thin’	*ma-nipis
*basa	‘wash’	*basəq
*tani ^c	‘weave’	*tənun
*kapik ^d	‘wing’	*kapak
*tuun ^e	‘year’	*taqun

^a PMP *-au- > PFL *-oa- is an irregular change.

^b Sika *kavir* ‘fishhook’ is related but has irregular initial *k =k rather than expected *k > ʔ/Ø.

^c The vowel changes from PMP to PFL are irregular.

^d (i) Sika *kapik* ‘wing’ is related but has irregular initial *k =k rather than expected *k > ʔ/Ø. (ii) PMP *a > PFL *i is an irregular change.

^e PMP *-aqu- > PFL *-uu- is an irregular change.

The 22 PFL reconstructions listed in Table 6.4 have reflexes in Sika and at least one Lamaholot variety but no reflex in Kedang. Nevertheless, these are cognate sets that can be traced back to PMP with regular correspondences. Therefore, these lexemes are reconstructed to PFL and it is assumed that Kedang replaced the respective concepts with new words or my database is missing a Kedang reflex for these words.

Table 6.4: PFL reconstructions without Kedang (n=22)

PFL	PFL meaning	PMP source
*modip	‘alive, live’	*ma-qudip
*ʔavu	‘ash, dust’	*qabu
*uma ^a	‘garden’	*quma
*leba	‘burden stick’	*lemba
*tani ^b	‘cry’	*taŋis
*taʔi	‘excrement’	*taqi
*puhun	‘heart’	*pusuŋ ‘heart; heart of banana’
*laki	‘husband; male’	*laki
*gatər	‘itchy’	*gatəl
*lotur	‘knee’	*qulu tuhud
*siva	‘nine’	*siwa
*meran	‘red’	*ma-iRaq
*gəvalik ^c	‘return’	*balik
*padi	‘rice plant’	*pajay
*tali	‘rope’	*talih
*plari/*kari	‘run’	*lariw
*kulit	‘skin’	*kulit
*g-nilu-k ^d	‘sour’	*ŋilu
*ikur	‘tail’	*ikuR
*m-panau	‘tinea’	*panaw
*puki	‘vagina’	*puki
*hapu	‘wipe’	*sapu

^a Kedang *lumar* ‘garden’ could be related.

^b Intervocalic PFL *-n- < PMP *-ŋ- is irregular.

^c PMP *balik > PFL *gəvalik is most likely PMP *b > *w > *v with the addition of a verbalising prefix *g-* (cf. §3.5.4).

^d Kedang *kiru* ‘sour’ could be related.

The 14 PFL reconstructions listed in Table 6.5 have reflexes in two or more Lamaholot varieties but neither a reflex in Sika nor in Kedang. Nevertheless, these are cognate forms that can be traced back to PMP with regular correspondences. Therefore, these lexemes are reconstructed to PFL and it is assumed that Kedang and Sika replaced the respective concepts with new words or the reflexes have not been attested in my dataset.

Table 6.5: PFL forms without reflexes in Sika and Kedang (n=14)

PFL	PFL meaning	PMP source
*sika	‘chase away’	*sika
*buja/*puja	‘flower’	*buja
*(kə)namuk	‘fly’ (n.)	*ñamuk ‘mosquito’
*tuma	‘louse on clothing’	*tumah
*ta(ke)	no; not	*taq
*bukat	‘open’	*bu(ŋ)kas
*mula	‘plant’	*mula
*(v)uvuŋ ^a	‘ridge’	*bubuŋ
*hira	‘salt’	*qasiRa
*tudu	‘sleep’	*tuduR
*ipe	‘teeth’	*(n)ipən
*baŋun	‘wake up’	*baŋun
*an	‘what’	*anu
*muav	‘yawn’	*ma-huab

^a Sika *puvun* ‘ridge’ could be related.

6.2.3 PFL reconstructions without PMP sources

Table 6.6 lists 37 regular PFL reconstructions that, based on the current stage of knowledge, do not go back to a PMP form. If a related or resemblant form is known to also occur in regional Austronesian languages outside of the Flores-Lembata family, this is indicated in the last column with “Flores” meaning the Austronesian languages of Flores, “Timor (AN)” meaning the in the Austronesian languages of Timor, “Timor (TAP)” meaning in the Timor-Alor-Pantar languages of Timor, and “Alor-Pantar” meaning in the Alor-Pantar languages on the islands of Alor and Pantar. I do not consider

the possible occurrences of the lexemes in languages outside of the East Nusa Tenggara and Timor-Leste region. Further research on the lexicon of the languages in this area and beyond will probably increase the number of these regionally spread items. Currently, 14 out of 37 lexeme sets listed here are also found outside of the Flores-Lembata family. The remaining 23 reconstructions may be considered as exclusive innovations of PFL.

Table 6.6: PFL reconstructions without PMP sources (n=37)

PFL	PFL meaning	Regional spread
*təmisi	‘ant’	
*dasan	‘ask; report’	
*muku	‘banana’	Flores, Timor (AN), Timor (TAP), Alor-Pantar
*təmayuŋ	‘bedbug’	Flores, Timor (AN)
*giki	‘bite’	Flores, Timor (AN), Timor (TAP), Alor-Pantar
*vəki	‘body’	Flores
*tena	‘canoe’	
*laku	‘civet cat’	Flores, Timor (AN), Alor-Pantar
*rusu or *ruhu	‘coral reef’	
*pati	‘cut’	Flores, Timor (AN)
*gurit	‘dig’	
*bao	‘float’	
*lodoŋ	‘fall down; descend’	
*voda-k	‘fat’	Flores
*pə-vunu	‘fight’	
*napu-k	‘flat; stream; river’	
*pau ^a	‘mango’	Flores, Timor (AN)
*motoŋ	‘marungga’	Alor-Pantar
*osan	‘mat’	
*k<n>əpuŋ/*həpuŋ	‘mosquito’	
*kəmeruŋ	‘rice ear bug’	Timor (AN)
*(n)ubak	‘stream; river’	
*vura	‘sand’	
*labur	‘shirt’	Flores, Maluku

PFL	PFL meaning	Regional spread
*kpali-k/*kwali-k	‘shoulder’	
*kamak	‘skin; bark of tree’	
*kə-melu	‘smooth’	
*m-potaŋ	‘spit’ (v.)	
*(k)rəvun	‘sweat’	
*səru-k	‘sweet’	
*alis	‘tendon’	Flores
*kera	‘turtle’	Flores, Timor (AN), Alor-Pantar ^b
*ale	‘waist’	
*hogo	‘wake up’	
*gəbi/*gnəbin	‘wall’	Flores
*(l)oyor	‘wave; sea’	
*nora	‘with’	Flores, Timor (AN)

^a Could be related to PWMP *qambawaŋ ‘manggo’.

^b PCEMP *kəRa or *keRa ‘turtle’.

6.2.4 Irregular reflexes in individual subgroups

In the following, I discuss instances of regular PFL reconstructions that are linked to cognate sets which contain unexpected changes in individual subgroups. These sets have been listed and counted already in the tables above because they can be reconstructed to PFL. Most irregularities appear in Kedang and the Lamaholot varieties. The Sika reflexes are largely regular. This observation is in line with the fact that there is more additional non-PMP vocabulary attested in Lamaholot and Kedang than in Sika as will be shown in §6.3 below. Both, irregularities in inherited words, such as in discussed in this section, and the additional non-PMP vocabulary, as discussed below, may both point to the historic presence of speakers of unrelated languages, especially in the Lamaholot and Kedang areas, that ultimately switched to Lamaholot and Kedang varieties.

Table 6.7 show the reflexes of the first person pronouns in the FL languages. While the Sika reflexes are completely regular, the Kedang and Lamaholot reflexes show several irregularities, highlighted in bold.

Table 6.7: Irregularities in the first person pronouns

PMP	*i aku	*kami	*kita
PFL	*aku	*kami	*kita
SK	<i>aʔu</i>	<i>ʔami</i>	<i>ʔita</i>
WL	<i>go ʔe</i>	<i>kame</i>	<i>tite</i>
CL	<i>go ne</i>	<i>kame</i>	<i>tite</i>
EL	<i>go ʔe</i>	<i>ame</i>	<i>gite</i>
KD	<i>ko / ɛʔi</i>	<i>e / ke</i>	<i>te</i>
	'1SG'	'1PL.EXCL'	'1PL.INCL'

For the 1SG pronouns, it appears that the Lamaholot varieties and Kedang underwent irregular $*k > g$ and lost the initial vowel *a*. The lowering of the final vowel $*u > o$ is regular in WL and EL but not in CL and Kedang (cf. §5.2.7). The subsequent change of $*g > k$ in Kedang to gain *ko* '1SG' is a regular change (cf. §5.2.1). For Kedang two variants for 1SG are given: $\epsilonʔi$ as a general pronoun and *ko* as an emphatic pronoun (Samely 1991a:69). It is unclear if the general pronoun $\epsilonʔi$ comes from PFL $*aku$ as two irregular vowel changes would remain unexplained.

The reflexes of PFL $*kami$ appear largely regular. The Kedang pronoun *e* '1PL.EXCL' can be explained by the loss of the second syllable and the regular changes of PFL $*k > ʔ/\emptyset$ and $*-a > e$. However, the Kedang variant *ke* has an irregular retention of PFL $*k = k$.

The Sika reflex of PFL $*kita$ is regular. The Kedang reflex is also regular when assuming the loss of the initial syllable. However, in the Lamaholot varieties, an irregular change of the initial consonant is observed. CL and WL undergo $*k > t$ - and EL undergoes $*k > g$ - in their reflexes of PMP $*kita$. The initial *t* in CL and WL could be explained by sporadic assimilation of the initial consonant to that of the medial consonant.

Table 6.8 lists examples of sporadic lenition of PFL $*b > v$ attested in the languages of Flores-Lembata. The same type of lenition in different lexical items has been attested already on a higher level in Proto-Bima-Lembata (cf. §5.5). Here again a few lexemes appear to start a new wave of lenition. Two of the three sets have a PMP source. I also include one set without PMP source because it shows the same pattern of lenition.

Table 6.8: Sporadic lenition of PFL *b > v

PMP	*Ribu	*binay	-
PFL	*ribu/ *rivu	*binay 'female; sister'	#ebel
SK	<i>rivu</i>	<i>wine ŋ</i>	-
WL (AD)	<i>ribu</i>	-	<i>əvər ət</i>
WL (LWI)	<i>ribu</i>	-	<i>vevel əŋ</i>
WL (LWL)	<i>ribu</i>	<i>bine?</i>	<i>veve</i>
CL (KK)	<i>ribu</i>	<i>binadz</i> <i>k winadz</i>	<i>evel</i>
CL (LR)	<i>rib</i>	<i>bine</i> [...]	<i>evel</i>
EL	[...]	[...]	<i>eblə</i>
KD	<i>ribu</i>	<i>bine n</i>	<i>ebel</i>
	'thousand'	'woman'	'tongue'

For the concept 'thousand', it is also possible that PFL had *rivu, thus lenition going back to an earlier stage, and the Kedang and Lamaholot varieties borrowed the Malay word *ribu* 'thousand'.

For the concept 'tongue', no PFL form can be reconstructed because this set cannot be traced back to PMP and Sika does not have a related word. It only appears in Kedang and the Lamaholot varieties. As a change of *b* > *v* is more likely than *v* > *b*, as the same change is attested in other sets, I assume that the original form was *ebel* 'tongue' and the the Western and Central Lamaholot varieties underwent lenition. In some varieties, an additional *v* is added before the initial vowel. Sporadic insertion of *v* before vowels is not uncommon in the Flores-Lembata languages. Another example is the Central Lamaholot form for 'bow' from PMP *busuR > PFL *vuhur 'bow'. As PFL *h is lost in Central Lamaholot, *v* is inserted in between the two medial vowels resulting in CL-Kalikasa *vuvor* 'bow'. Also in Sika *vaten* 'liver' from PFL *ate-n < PMP *qatay 'liver', the initial *v* is inserted before the vowel.

The two sets 'salt' and 'how much' in Table 6.9 have irregular forms in some varieties that can possibly be explained as borrowings, marked with an arrow (→). Both concepts are related to the market place. From the Cent-

ral Lamaholot and Eastern Lamaholot forms, PFL *hira ‘salt’ can be reconstructed. The unexpected Western Lamaholot reflexes containing *s* are most likely loans from Central Flores languages which have *siʔe* ‘salt’. PFL *pida (< PMP *pija) ‘how much’ is regularly reflected in most subgroups, as the correspondences of SK *r* - KD *y/Ø* - WL *r* are regular reflexes of PFL *-d- < PMP *-j-. The Central Lamaholot forms are irregular, as PFL *-d- < PMP *-j- is normally reflected as Central Lamaholot *-dʒ-*. It is possible that Central Lamaholot speakers borrowed *pira* ‘how much’ from their Western Lamaholot neighbours.

Table 6.9: Possible borrowings (→) in individual subgroups

PMP PFL	*qasiRa *hira	*pija *pida
SK	-	<i>pira</i>
WL (AD)	→ <i>siʔa</i>	<i>pira</i>
WL (LWL)	→ <i>siʔa</i>	<i>pira</i>
WL (BL)	→ <i>sia</i>	[...]
WL (PD)	→ <i>sia</i>	<i>pira</i>
CL (KK)	<i>ira</i>	→ <i>pira</i>
CL (LR)	<i>ira r</i>	→ <i>pira</i>
EL	<i>hira</i>	[...]
KD	-	<i>pie</i>
	‘salt’	‘how much’

Table 6.10 shows reflexes of PFL *tuʔan (< PMP *tuqan) ‘forest’ and instances of resemblant forms in the Lamaholot varieties that cannot be explained by regular sound changes. The Lamaholot forms underwent unexpected voicing of the initial plosive *t.

Table 6.10: Irregular forms for ‘forest’ in Lamaholot

PMP	*tuqan
PFL	*tuʔan
SK	<i>tuʔan</i>
WL (LWL)	<i>duã</i>
CL (KK)	<i>duan</i>
EL	-
KD	<i>tuen</i>
	‘forest’

6.2.5 Discussion

Table 6.11 summarises the features of the PFL vocabulary reconstructed for this dissertation. The left-most column categorises the spread of the lexeme sets. The Lamaholot subgroups are grouped together as LH here as they are located in the centre of the Flores-Lembata family. Kedang and Sika are at the edges. LH thus means one or more Lamaholot subgroups. For the last category of PFL reconstructions that only contain reflexes in Lamaholot varieties, this means that reflexes are attested in at least two Lamaholot subgroups. Lexeme sets that neither have related forms in Sika and Kedang nor a PMP source are not reconstructible to PFL. These are discussed in §6.3.

Table 6.11: PFL reconstructions (n=210)

	PMP source	no PMP source	Total
SK - (LH) - KD	113	37	150
LH - KD	24	-	24
SK - LH	22	-	22
LH	14	-	14
Total	173 (81%)	37 (19%)	210 (100%)

The 210 PFL reconstructions are to a great extent of Austronesian origin, for 81% of them there is a known PMP source. About one fifth of the PFL

vocabulary remains of unknown origin. PFL, as a descendant of PMP, has thus replaced about 20% of the vocabulary for the concepts in this study since PMP times. PMP was spoken around 4000 years ago (Pawley 2005).

When selecting only basic vocabulary forms from the sample, around 124 PFL forms remain.¹ Out of these only 13% are not of PMP origin. This lower percentage of non-PMP vocabulary in PFL basic vocabulary compared to the whole database confirms that lexical replacement in basic vocabulary is less likely to occur than in other parts of the vocabulary.

The PFL vocabulary which is not of PMP origin could be regarded as a non-Austronesian lexical substrate in PFL. However, at the current stage of research, it is not entirely clear if the set of lexical items in PFL that do not trace back to PMP (listed in §6.2.3) can be part of a substrate because it is unknown how much of this vocabulary traces further back to an earlier ancestor of PFL. In §6.2.3, I have shown that about 30% of the non-PMP vocabulary in PFL has related forms elsewhere in the region. As this number is based on an initial survey, more in-depth systematic investigation into the

¹ About half of the lexeme sets in the database have been classified as denoting basic concepts. Often one basic concept is expressed by two or more lexeme sets. The classification as basic concepts is based on the Leipzig-Jakarta Basic Vocabulary list (Tadmor et al. 2010:238-241) with my own extensions, concerning in particular regionally relevant concepts. In total, the following 192 concepts have been classified as basic for the purpose of this study: 1pl excl; 1pl incl; 1sg; 2sg; 3pl; 3sg; all; ant; ash, dust; back; banana; bathe; betel vine; big; bird, chicken; bite; bitter; black, dirty; blood; blow; body, self; body hair; bone, seed; breast, milk; burn, shine; child, small; cloud, fog; coconut; come; cry; cut, kill; day, sun; deaf; die; dog; dream; drink; drop, fall from above; dry, thirsty; ear; eat; egg; eight; excrements; eye; fall from above, descend; fall over; far, long; fat; fingernail; finished; fire; fish; flat, below, river; flower, blossom; fly; fly (n.); flying fox; foot, leg; forehead; forest; four; fruit, betelnut; full; give; go; good; grass, bush; hair; hand, arm, five; head; headlice; hear; heart; heavy; here; hide; hillwards, above; hit; horn; hot; house; how much, how many; how?; hungry; inside, deep; inside, liver, house; itchy; knee; knife; know; laugh; leaf; lie down (non-human); liver; man; many; meat, flesh; meeting house; moon, market; mosquito; mother; mountain; mouth; name; narrow; navel; near; neck; needle; new; night; nine; no, not; nose; old; one, alone; person; pound; price, bride price, expensive, buy; rain; rat; rattan; red; rice; road; roof rafter; root; rope; round; run; salt; sand, soil; say; sea, wave; see; seven; short; sick, painful; sit; six; skin, bark of tree; sky; sleep, lie down; smoke; snake; soil; spit; stand; star; stomach, belly; stone; storage house, barn; suck; sugar palm; sugarcane; sun; sweet; swim; tail; teeth; ten; that; thatch for roofing; thatched roof; thick; this; thousand; three; tie; tongue, say; tree, wood; two; vomit; wake someone up; wake up; walk; wash, bathe; water; what; where; white; who; wide; wife, husband; wind; wing; woman, sister; yellow; yesterday; younger sibling.

lexicon of the languages of the region and even beyond may shed light on how far this vocabulary can be traced back. Some of it may even ultimately go back to PMP. It is possible that with further research, the number of PFL reconstructions without PMP source becomes so small that one could account for it by lexical replacement that naturally occurs in any language for different reasons, such as avoidance of homophony, semantic change, borrowing and invention of new words.

6.3 Unreconstructible regular lexeme sets

6.3.1 Overview

This section discusses 185 lexeme sets that cannot be reconstructed to PFL but which show regular correspondences between the subgroups in which they occur. These sets cannot be reconstructed to PFL because related forms neither occur in Sika and Kedang, nor is there a PMP source which could justify a PFL reconstruction. The regularity of the related forms in the set makes it possible to reconstruct a hypothetical form that could be a PFL form if the missing forms in Sika and/or Kedang would be found. These hypothetical reconstructions are marked with a hash tag (#). The lexeme sets discussed here may ultimately be reconstructible to PFL when more data becomes available but they could also be early borrowings that entered a subset of the Flores-Lembata subgroups (cf. §6.1).

The lexeme sets are organised in three categories: sets without a reflex in Sika (§6.3.2), sets without a reflex in Kedang (§6.3.3) and sets with neither a reflex in Sika nor in Kedang (§6.3.4). For each category, I provide a list of hypothetical reconstructions with hash tag (#) and their regional spread as far as this information is available to me. The same categories of regional spread as for the PFL reconstructions without PMP source in §6.2.3 apply.

6.3.2 Lamaholot-Kedang lexeme sets

Table 6.12 contains 73 Lamaholot-Kedang (LH-KD) regular but unreconstructible lexeme sets. Out of these, 19 sets are regionally more wide-spread, thus are attested in at least one language outside of the FL subgroup. The great

majority (54 out of 73 sets) is only attested in Kedang and the Lamaholot varieties.

Table 6.12: Shared lexemes in LH and KD without PMP source (n=73)

#Lamaholot-Kedang	Meaning	Regional spread
#soloi	'answer' (v.)	
#gəter	'ask question'	
#bovoŋ	'bark'	
#həbu ^a	'bathe'	
#malu	'betel vine'	Timor (AN), Timor (TAP)
#puur	'blow'	Flores, Timor (AN), AP
#papi	'burn; clear land'	
#letu?	'close' (v.)	
#kova ^b	'cloud; fog'	
#korok	'chest'	
#tapu	'coconut'	
#hekan	'condition; time; garden'	
#mudəŋ	'correct; the following'	
#bəpap	'crocodile'	Alor-Pantar
#belu	'cut; kill'	Flores
#sedu	'dance'	
#klebit	'deaf'	
#butu	'eight; bunch; group'	Flores, Timor (AN), AP ^c
#gokal	'fall over'	
#bəka	'fly'	
#lei	'foot, leg'	
#{kəne) breuŋ ^d	'friend'	
#ne?i	'give'	Timor (AN)
#gedi	'go up; ascend'	
#dikə-n ^e	'good; person'	
#vurek	'gravel'	
#tava ^f	'grow; stem'	
#pohiŋ	'help'	
#vuok	'hole'	
#vetak	'house; barn'	

#Lamaholot-Kedang	Meaning	Regional spread
#nara bone gaku	'how'	
#kverak	'jackfruit'	Alor-Pantar
#kudul	'knee'	
#lolo	'leaf'	
#ləpa	'leaf; sheet; lontar leaf'	
#benehik	'light (not dark)'	
#(kutu) kihan	'louse eggs'	
#kabe	'man; husband; person'	
#rai-k ^g	'many'	
#tudak	'narrow'	
#dahe-k	'near'	
#vuli	'neck'	Alor-Pantar
#batul	'needle'	Alor-Pantar
#payam	'papaya'	
#volar	'ridge'	
#vadək	'rope'	
#doru ^h	'rub; wipe'	Alor-Pantar
#taʔu	'salt'	
#bota(n)	'sand'	
#kəburak	'scabies'	Flores
#kuluk	'seed'	Alor-Pantar
#durum	'sell'	
#saur	'sew'	Timor (AN), Alor-Pantar
#məkul	'short'	
#tobe	'sit'	
#təguʔ	'skewer'	
#molān	'sorcerer'	
#gala(r)	'spear'	Flores
#təmidu ⁱ	'spit'	Timor (AN)
#bəta	'split'	
#tubak	'stab'	
#(kə)boti	'stomach; belly'	
#kebaŋ	'storage house; barn'	Alor-Pantar
#pola	'sugar palm'	

#Lamaholot-Kedang	Meaning	Regional spread
#soŋa	'tie'	
#ebel	'tongue'	
#(bela) bayan	'treaty'	Alor-Pantar
#deko	'trousers'	Flores, Timor (AN), AP
#ləvu	'village'	
#luaŋ	'vomit'	
#hamu	'wipe; sweep'	Timor (AN)
#kumas	'yellow'	
#evian	'yesterday'	

^a Central Lamaholot *ləbo* 'bathe' could be related.

^b Sika *kova* 'cloud' could be related but would involve an irregular retention of PMP *k = Sika *k*. This lexeme set might trace back to PMP *away 'atmosphere, space between earth and sky' with an insertion of initial *k-* and an irregular change of PMP *a > PFL *o.

^c PCEMP *butu 'group, crowd, flock, school, bunch, cluster'

^d Sika *deuy* 'friend' could be related but would involve an irregular correspondence of Lamaholot/Kedang *br-* and Sika *d-*.

^e The set #dikə-n could derive from PMP *diqɑq 'good' with an irregular change of PMP *-q- > PFL *-k- before *ə*. However, as also the change of PMP *-aq > PFL *-ə in this word remains unexplained, PFL *dikə 'good; correct' might also be unrelated to PMP *diqɑq. The original meaning of this set is probably 'good; correct'. The word 'good' is combined with another word for 'person', i.e. PFL *ata, such as still in used for example in Central Lembata *ata dikən* 'person'. This was probably done as an opposition of members of another group that were enemies. Over time, also the second part of the compound acquires the meaning 'person'. However, in some subgroups, such as for example in Kedang and Eastern Lamaholot, both meanings 'good' and 'person' are retained. In Alorese, a reflex of PFL *dikə means 'right side'.

^f Eastern Lamaholot *nava* 'stem' could be related.

^g #rai 'many' could trace back to PMP *Raya 'big'.

^h Western Lamaholot *doruk* 'rub; wipe' could be related but would involve an irregular retention of PFL *r = WL *r*.

ⁱ This could be related to PWMP *qizuR 'saliva; spittle'.

6.3.3 Lamaholot-Sika lexeme sets

Table 6.13 lists 41 lexeme sets with regular reflexes in at least one Lamaholot (LH) variety and in Sika (SK). 18 out of 41 sets have related forms in other languages of the region. 23 out of 41 lexeme sets are only attested in Sika

and the Lamaholot subgroups.

Table 6.13: Shared lexemes in LH and SK without PMP source (n=41)

#Lamaholot-Sika	Meaning	Regional spread
#supel	'arrow'	Flores, Alor-Pantar (?)
#baka	'bite'	Flores
#(sə)mei	'blood'	
#nahi	'breath'	Flores
#ihere	'close' (v.)	
#kobu	'crocodile'	
#gasik	'count'	Timor (AN)
#kəbehar	'cuscus'	
#baŋjak	'flow'	Flores
#-ai	'go'	
#volon	'hill; ridge'	Flores
#tara	'horn'	
#(raʔi) etan	'know'	Timor (AN)
#blavir	'long; far'	
#koli	'lontar palm'	Flores, Alor-Pantar
#(meiŋ)ʔətan	'meat'	
#təker	'narrow'	Flores
#lusir	'needle'	
#guman	'night'	Timor (AN), Alor-Pantar
#dzəma	'night, time unit'	
#pehan ^a	'other'	Flores
#likat	'oven'	Flores
#əpak	'palm of hand; footprint'	
#pahat	'plant yam'	Flores
#tubu	'pull'	
#gide	'pull'	
#gualok	'round'	
#madi	'say'	Flores
#kəmekot	'scorpion'	
#buʔu	'short'	Flores
#blara	'sick; painful'	

#Lamaholot-Sika	Meaning	Regional spread
#tuʔay ^b	‘sleep’	
#nuhi	‘smoke’	Flores, Timor (AN)
#pemek	‘squeeze’	Alor-Pantar
#robak	‘stab’	
#hukut	‘think; remember; miss’	
#kleka ^c	‘thunder’	
#papa lele	‘trade’	
#puʔu	‘wash’	Flores
#kəsako	‘whisper’	
#ledan	‘wide’	

^a Kedang *palan* ‘other’ could be related.

^b Kedang *təʔel* ‘sleep’ could be related.

^c CL-Kalikasa *kəlagor* ‘thunder’ could be related but would require an irregular change of the last syllable #ka to Kalikasa *gor*.

6.3.4 Lamaholot lexeme sets

Table 6.14 lists 71 lexeme sets with regular reflexes in at least two Lamaholot (LH) subgroups but no reflexes in Sika and Kedang. 16 out of 71 sets have a possibly related forms in other languages of the region. 55 out of 71 lexeme sets are only attested in the Lamaholot subgroups.

Table 6.14: Shared lexemes in LH varieties without PMP source (n=71)

#WL-CL-EL	Meaning	Regional spread
#əvan	‘accuse’	
#tapan	‘answer’	Timor (TAP)
#svaol	‘all’	
#knaru	‘back’	
#navak	‘body’	
#ravuk	‘body hair’	Timor (AN)
#esari nai	‘breathe’ (v.)	
#hopi	‘buy’	
#kiri	‘comb’	Alor-Pantar (PAP *kir (Robinson 2015))

#WL-CL-EL	Meaning	Regional spread
#oli	'come; arrive'	
#suda	'command; order' (v.)	
#bisu	'cook'	
#kluok	'cooked rice; uncooked rice'	
#vekan	'divide'	
#knavi	'door'	Alor-Pantar (?)
#ləŋat	'fall from above'	
#gəni	'fight'	
#vahak	'finished'	
#lərek	'flat; below'	
#kənito	'forehead'	
#alus	'good'	
#pehen	'grasp; hold'	
#madu	'grasshopper'	
#latar	'hair'	
#kote	'head'	Timor (AN)
#soronj	'hide'	
#dani	'hit (drum)'	
#umaŋ	'hole'	
#plati/kati	'hot'	
#maluv	'hungry'	
#bati	'hunt'	
#gekay	'laugh'	
#səmekiŋ	'left side'	
#loit	'let go'	
#pavaŋ	'lie' (position for things)	
#kleak ^a	'light (weight)'	
#kmoruŋ	'locust'	
#vuda	'lungs'	Alor-Pantar
#elam	'meat; flesh'	
#vətəm ^b	'millet'	Flores
#vala	'mud'	Alor-Pantar
#niləŋ	'necklace'	
#magun	'old people'	

#WL-CL-EL	Meaning	Regional spread
#toʔu	‘one’	
#gesak	‘other’	
#glasa	‘play’	
#nakiŋ	‘promise’	Alor-Pantar
#vidu	‘pull’	Flores
#magar	‘rack above hearth’	
#tue	‘return’	
#(a)luŋu	‘river; stream’	
#bua	‘sail’ (v.)	
#sodam	‘smell’	Timor (AN)
#m<an>akap	‘sorcerer’	
#pəriŋo	‘spit’	
#piʔuk	‘squeeze’	
#puka	‘stem’	Flores
#mopa	‘straight’	
#kebol	‘sugar palm’	
#luvak	‘sun’	
#blolo/golo	‘tall’	Alor-Pantar
#luʔo	‘thatch for roofing’	
#tnakar	‘thatched roof’	
#pənəŋe	thick	
#prəvak	thick	
#petən	‘think; miss’	
#məna	‘vagina’	Flores
#rio	‘wake someone up’	
#ga(ne)	‘where’	Alor-Pantar, Timor (TAP)
#henaku	‘who’	Timor (AN)
#ugadak	‘wound’	

^a Sika *heak* ‘light (weight)’ and Kedang *ʔahaʔ* ‘light (weight)’ could be related to #kleak.

^b Kedang *vereʔ* ‘millet’ could be related to #vətəm.

6.3.5 Discussion

In the previous section, 185 sets of related lexemes that cannot be reconstructed to Proto-Flores-Lembata (PFL) have been presented. These lexeme

sets have forms in at least two subgroups of Flores-Lembata and the sound correspondences between the lexical items are regular. The sets cannot be reconstructed to PFL because they do not fulfill the criteria set out in §6.1. In short, the criteria for reconstructibility to PFL are the occurrence of reflexes at least in Sika and Kedang, or alternatively, having a PMP source. The 185 lexeme sets examined here only occur in a subset of the Flores-Lembata languages and do not have a known PMP source.

Table 6.15 provides an overview of the numbers of unreconstructible vocabulary sets that are found. There is a set of vocabulary that is attested in Lamaholot varieties and in Kedang (LH-KD), a set that is attested in Lamaholot varieties and in Sika (LH-SK) and a set of vocabulary that is only attested in the Lamaholot varieties (LH). For the three Lamaholot varieties this means that a related form is attested in at least one of the three Lamaholot subgroups when also shared with Sika or Kedang (category LH-SK or LH-KD), but attested in at least two subgroups when not shared with Sika or Kedang (category LH).

Table 6.15: Unreconstructible regular lexeme sets

	no PMP source
LH - KD	73
LH - SK	41
LH	71
Total	185

There are two ways to explain this unreconstructible vocabulary. First, it is possible that the remaining subgroups lost the reflexes so that ultimately the lexemes could be reconstructed to PFL. However, the ultimate origin of the words would still remain unclear, as they do not appear to be of Austro-nesian origin (because they lack a PMP ancestor form). Second, it is possible that the words entered the subgroups after the split of Flores-Lembata into subgroups but before the regular sound changes occurred in the individual subgroups. It is not possible that these words were added to the lexicon of the Flores-Lembata subgroups after all subgroup defining sound changes had occurred because then, the regular correspondences between the words in the subgroups could not be explained.

The second option appears more realistic because for the first option one has to assume that a huge amount of vocabulary, especially in Sika and Kedang, has been lost or is not attested in the dataset. This is very unlikely. I thus argue that the new vocabulary entered the subgroups before the subgroup defining sound changes occurred. New lexical items of unknown origin can indicate that the speakers of the language invented new words or that they borrowed the words from an unknown source. Considering the large amount of new lexical items appearing in the Flores-Lembata subgroups, the invention of such a large amount of new words appears very unlikely. So my hypothesis is that the vocabulary must have come from at least one substrate language. Of course, this does not exclude the possibility that for some of the sets the missing forms in either Sika or Kedang will still be found and the set will thus be reconstructible to PFL.

One may propose that the shared lexical items in several Flores-Lembata subgroups are evidence for mid-level subgroups within the Flores-Lembata family. However, as no phonological evidence for such mid-level subgroups could be found (cf. §5.3.2), shared lexical items alone are very weak evidence for subgrouping. I rather suggest that the shared lexical items in Kedang and Lamaholot, Sika and Lamaholot and among the Lamaholot varieties result from contact with the same substrate language(s).

In fact, three lexical substrates can be proposed. A western substrate that affected Sika and the Lamaholot varieties (§6.3.3), a central substrate that only affected the Lamaholot varieties (§6.3.4) and an eastern substrate that affected Kedang and the Lamaholot varieties (§6.3.2). As Lamaholot is located in the middle, it has been affected by all three substrates. Whether these three substrates were actually three different languages or just represent three different selections of vocabulary from the same language cannot be decided from the present data.

6.4 Irregular lexeme sets

In my dataset, there are 27 irregular lexeme sets that cannot be reconstructed to PFL because they do not show regular sound correspondences between the forms in the different subgroups. Nevertheless, the sets contain forms that are resemblant and their similarity cannot be ignored. Some of these sets can possibly be explained as late loans that entered the subgroups after

the sound changes had occurred but others remain unclear. I discuss possible loans in §6.4.1 and unclear sets in §6.4.2.

6.4.1 Possible loans

This section discusses 13 lexeme sets with irregular sound correspondences that can most likely be explained by borrowing. Table 6.16 lists three lexemes that are most likely loans from Malay.² These words are considered loans because there is a clear Malay source and they did not undergo the regular sound change of PFL *r > ʔ in Western Lamaholot. However, in the sets #ritik ‘drizzle’ and #rusa ‘deer’, the sound changes PFL *s > h in SK and WL, as well as the change of PFL *k > ʔ in Kedang can be seen. These sound changes are not subgroup defining and occur in more than one subgroup. In the set #sorɔŋ ‘give’, the change of PFL *s > h has not occurred. It has been observed that PFL *s > h is sometimes incomplete (§5.2.3). However, as the unexpected s occurs in all subgroups here, it is more likely that #sorɔŋ ‘give’ is a more recent loan from a time when *s > h did not occur any more. In case of the incomplete sound change, some variation between varieties would be expected.

Table 6.16: Malay loans in the Flores-Lembata languages

Malay	<i>rintik</i> ‘speckle’ #ritik	<i>rusa</i> ‘deer’ #rusa	<i>sorɔŋ</i> ‘push; shove’ #sorɔŋ
SK	-	<i>ruha</i>	<i>sorɔŋ</i> ‘serve; stretch out hands’
WL (LWI)	(<i>kite</i>) <i>rite</i> ŋ	<i>ruha</i>	<i>sorɔŋ</i>
WL (MS)	<i>rik</i>	<i>ruha</i>	-
CL (KK)	<i>rətik</i>	<i>rusa</i>	-
EL	[...]	[...]	<i>sorõ</i>
KD	<i>riti</i> ʔ	<i>ruha</i>	<i>sorɔŋ</i>
	‘drizzle’	‘deer’	‘give’

There are four other instances of a missing sound change of PFL *r > ʔ in Western Lamaholot listed in Table 6.17. For these cases, the loan hypothesis

² The Malay meanings are taken from Stevens and Schmidgall-Tellings (2004).

is less easy to prove because there are no known sources for these potential loans. All four lexeme sets are not reconstructible to PFL but they occur in several Lamaholot subgroups, similar to the substrate sets discussed in §6.3.4. However, they are missing regular correspondences in Western Lamaholot. Therefore, I propose that they entered the Flores-Lembata languages, or at least Western Lamaholot, after the sound change of PFL *r > ʔ.

Table 6.17: Lexeme sets that did not undergo PFL *r > WL ʔ

	#turu	#kromi	#raaŋ	#krogoŋ
SK (KR)	-	-	-	ʔruguy
WL (AD)	təʔuru	krome	rã:	-
WL (LWL)	nurəʔ	kərome	ra	krogoŋ
WL (PD)	nuroŋ	kmore	[...]	-
CL (KK)	turən	kromi	ra:ŋ	krogoŋ
CL (LR)	turən	[...]	raŋa	-
EL	[...]	[...]	[...]	[...]
KD	-	-	-	-
	‘dream’ (v.)	‘rat’	‘voice’	‘skinny’

Similar to #soroŋ ‘give’ from Malay *soroŋ* ‘push; shove’, there are six other lexeme sets with unexpected *s* occurring in four subgroups Sika, Western Lamaholot, Kedang and Eastern Lamaholot as listed in Table 6.18. These could be loans that came into the languages after the sound change of *s > *h* in these subgroups. Central Lamaholot varieties regularly retain *s = *s* with the exception of CL-Lerek, as can be seen in the data in the table below. This sound change of Proto-Central-Lamaholot (PCL) *s > *h* in Lerek must be of a much more recent date than the change of PFL *s > *h* in the other Flores-Lembata subgroups. For some of these words in Sika and Kedang, there is also an unexpected *k*. This suggests that the words entered the language also after that sound change had occurred.

Table 6.18: Lexeme sets that did not undergo PFL *s > WL/KD/SK h

	#soka	#pasak	#sadok	#si(n)oŋ	#səgat	#soga
SK (HEW)	<i>soka</i>	<i>pasak</i>	<i>sadok</i>	-	-	-
SK (KR)	<i>soka</i>	<i>pasak</i>	<i>sadok</i> ‘kick ball’	<i>sinonŋ</i>	-	-
WL (LWI)	<i>soka</i>	<i>pasak</i>	-	<i>sionŋ</i>	-	-
WL (LWL)	<i>soka</i>	<i>pasak</i>	<i>sadok</i>	<i>siõ</i>	<i>səgat</i>	<i>soga</i>
CL (KK)	<i>soka</i>	<i>pasak</i>	<i>sadok</i>	-	<i>səgat</i>	<i>soga</i>
CL (LR)	-	<i>pahak</i>	[...]	-	<i>həgat əŋ</i>	-
EL	<i>sokə</i>	<i>pasa</i>	<i>sado</i>	-	-	-
KD	<i>soka</i>	<i>pasa?</i>	-	-	-	-
	‘dance’	‘shoot’	‘fist’	‘smell’	‘stab’	‘hold’

Related forms to the set #soka also occurs in languages further west, such as in Palu'e with the word *tfoka* ‘dance’ and in Bima with *soka* ‘dance’. But no language of origin can be determined. The set #pasak is related to the word *pasa* ‘shoot’ in several Central Flores language and Proto-Central Flores (PCF) **pasa* ‘shoot’ can possibly be reconstructed. Thus, a Central Flores language could be the donor for this lexeme set. The set #si(n)oŋ ‘smell’ could be related to Malay *cium* ‘smell’ but the intervocalic *n* would then remain unexplained. The remaining three sets (#sadok, #səgat, and #soga) are more problematic and it remains unclear if they can be explained by borrowing. No related forms are known and they only appear in a small subset of the Flores-Lembata languages.

6.4.2 Lexeme sets with unexplained correspondences

In this section, I discuss 14 lexeme sets that have obvious similarities but no regular correspondences across the subgroups. Only one out of the sets can be traced back to PMP, namely PMP **susu* ‘breast’. All others are of unknown origin and only occur in a subset of the Flores-Lembata subgroups. The lexeme sets of unknown origin could possibly be part of the lexical substrate discussed in §6.3 but for some reason, they did not undergo regular sound changes after entering the subgroups. Table 6.19 lists two examples of irregular initial correspondences.

Table 6.19: Unexpected initial correspondences

PMP	- #bərəkət/dəkət	*susu ? PFL *(t)usu
SK	-	<i>uhu</i>
WL (AD)	<i>dəkət</i>	<i>tuho</i>
WL (LWI)	<i>bərəkə</i>	<i>tuho</i>
WL (BL)	<i>rəkət</i>	[...]
WL (PD)	<i>dəkek</i>	<i>tuho</i>
WL (MS)	<i>dəkek</i>	<i>tuho</i>
CL (KK)	<i>rəkət</i>	<i>tusu</i>
CL (LWT)	<i>bərəkət</i>	[...]
CL (LR)	<i>rəkət ən</i>	<i>tuho r</i>
EL	<i>raʔe</i>	[...]
KD	-	<i>tuʔu</i>
	‘sharp’	‘breast’

The lexeme set #bərəkət/dəkət ‘sharp’ is of unknown origin and only occurs in the Lamaholot varieties. There is alternation of *(bə)r* and *d* in on-set position. The alternation does not align with the subgroups. A possible explanation for this set is to analyse *dəkət* ‘sharp’ as the base form which is nominalised with the prefix *b-*, attested as a nominaliser in CL-Central Lembata (cf. §3.3.6). As a following step one would assume the change of *b(ə)d* > *b(ə)r* ~ *r*.

The lexeme set meaning ‘breast’ traces back to PMP *susu ‘breast’. In the reflexes, the initial PMP *s is lost in Sika, while in Kedang and the Lamaholot varieties, it is replaced by *t*. The replacement of *s by *t* is sporadically found in other words as well, such as for example WL-Adonara *təratu* ‘one hundred’ < PMP *sa-ŋa-Ratus ‘one hundred’. This does not provide enough evidence for this form to be reconstructed to PFL with either *usu or *tusu for instance.

Table 6.20 lists examples of unexpected irregular correspondences of consonants in intervocalic position. All the sets are of unknown origin.

Table 6.20: Unexpected intervocalic correspondences

	#bəCo(l)	#uduk/uruk	#loka	#dihe	#nudəp
SK	<i>boʔu</i>	-	-	-	-
WL (AD)	<i>beto</i>	-	<i>loʔok</i>	-	<i>nudərət</i>
WL (LWI)	-	-	<i>loʔo</i>	-	<i>ude</i>
WL (LWL)	<i>beʔo / bəso / beto</i>	-	<i>loʔok</i>	-	<i>udet</i>
WL (LML)	<i>bəso</i>	<i>uruk</i>	[...]	<i>die</i>	[...]
WL (WB)	<i>beto</i>	-	[...]	-	[...]
WL (BL)	<i>bəsol</i>	[...]	[...]	<i>didi</i>	[...]
WL (PD)	<i>beta</i>	[...]	-	-	<i>ude k</i>
CL (LR)	<i>bedʒo</i>	<i>uduk</i>	<i>lokaŋ</i>	-	<i>udəm</i>
CL (KK)	-	<i>uduk</i>	<i>loka</i>	<i>didʒi</i>	<i>nudəp</i>
CL (LWT)	<i>bəsol</i>	-	[...]	<i>dihe</i>	[...]
EL	<i>bəso</i>	<i>udu</i>	[...]	-	[...]
KD	-	<i>uruʔ</i>	-	-	-
	'come'	'push'	'let go'	'blow'	'heel of foot'

Table 6.21 shows instances of similar looking pairs in two different subgroups. The initial or medial consonant alternates with zero. Table 6.22 shows three sets with words that obviously look related but no clear pattern can be determined. For the set #dʒeta 'hillwards', I also consider the possibility that these are not all cognates of the same set. There could be #reta to account for the words in Sika and WL-Waibalun (WB) and #dʒ(a)e to account for the forms in CL-Kalikasa and WL-Lewolema.³

³ Proto-Central-Flores *d(ʒ)eta 'above' could be related. (Elias 2018) suggests that this form comes from PMP *i atas 'above'. However, the change of PMP *i a > PCF *d(ʒ)e does not appear to be regular.

Table 6.21: Instances of consonant alternating with zero

LH-KD	#(g)iu	#me(r)it	#(d)anen	#(m)are
WL (LWI)	-	-	-	-
WL (LML)	-	[...]	-	-
WL (BL)	-	[...]	-	-
WL (PD)		-	-	<i>mare</i>
CL (KK)	<i>giu</i>	<i>mərit</i>	<i>danen</i>	-
EL	-	[...]	[...]	<i>are</i>
KD	<i>iu</i>	<i>mɛiʔ</i>	<i>anen</i>	-
	‘cook’	‘knife’	‘(uncooked) rice’	‘smell’

Table 6.22: Resemblant sets with unclear pattern

PMP	(?) *kalawaq	-	-
PFL	*kalaka	#(kə)bukal	#dʒe(ta)
SK	(<i>kuku</i>) <i>raka</i>	<i>bulak</i>	<i>reta</i>
WL (AD)	-	<i>kəbukare</i>	[...]
WL (LWL)	-	-	<i>rae</i>
WL (WB)	[...]	[...]	<i>reti</i> ^b
WL (PD)	<i>klake</i>	-	[...]
CL (LR)	<i>laka (borit)</i>	<i>kəbukal</i>	<i>re dʒe</i> ^a
CL (KK)	<i>lak (borit)</i>	<i>kəbukal</i>	<i>dʒae</i> ‘hillwards’ / <i>dʒe</i> ‘upwards’
EL	[...]	[...]	[...]
KD	-	<i>ʔebal</i>	-
	‘spider’	‘butterfly’	‘hillwards; above’

^a Krauß 2016:126^b Akoli 2010:59

6.5 Conclusions

In this chapter, I have shown that while Proto-Flores-Lembata (PFL) has a largely Austronesian vocabulary, the lexicon of the individual Flores-Lemba-

ta subgroups has been considerably influenced by non-Austronesian substrate languages. In the following, I first discuss the PFL lexicon and then the lexicon of the individual subgroups.

About 81% of the Proto-Flores-Lembata (PFL) vocabulary that I reconstructed has an Austronesian, i. e. Proto-Malayo-Polynesian (PMP), source (§6.2). PMP was presumably spoken 4,500-3,500 years ago in the northern Philippines, and the first Austronesian speakers arrived about 3,500 years ago in the area where the FL languages are spoken today (Klamer 2019; Pawley 2005). This suggests a very fast spread of PMP speakers and languages through the whole Indonesian archipelago. As PFL is not a direct descendant of PMP, and its closest relatives of the Bima-Lembata family (cf. §5.5) are spoken in the area, Proto-Bima-Lembata (PBL) could have been spoken by the first Austronesian speakers on Flores and beyond. It is unknown how much time may have passed between the times of PBL (possibly 3,500 years ago) and the times of PFL. It may have been one or two millennia. If this estimation is correct, a time span of about 1,500 to 2,000 years between PMP and PFL can be proposed. As only about 19% of the PFL reconstructions are of non-PMP origin, one can propose a lexical replacement rate of 19% for the time span between PMP and PFL. In contrast to a stronger substrate hypothesis for some of the individual Flores-Lembata subgroups — discussed in more detail in the following paragraph —, it remains unclear if these 19% of non-PMP vocabulary in PFL (13% when considering only basic vocabulary) can be attributed to substrate influence. Further research into the regional spread of this vocabulary is needed.

As illustrated in Table 6.23, in the individual Flores-Lembata languages that are spoken today, the percentage of Austronesian vocabulary drops further to between 62% in the Sika variety of Hewa and 47% in the Central Lembata variety of Central Lamaholot. When only examining basic vocabulary, the PMP percentages are about 10% higher for each language. See footnote (1) earlier in this chapter on the selection of basic vocabulary in this study. In the table, each subgroup is represented by one variety as I have not observed significant variation between the varieties of one subgroup regarding the distribution of PMP versus non-PMP vocabulary.

As the time span between PFL and the present-day Flores-Lembata languages is the same for all, it can be concluded that since PFL times the Lamaholot subgroups underwent the biggest increase in lexical replacement, followed by Kedang and then Sika. Considering the whole database, Central

Lamaholot with 53%, Western Lamaholot with 51% and Eastern Lamaholot with 46% have the highest percentages of non-PMP vocabulary, followed by Kedang with 43%, and then Sika with 38%. This suggests more non-Austronesian influence in the Lamaholot subgroups than in Sika and Kedang.

Table 6.23: Lexemes of PMP / non-PMP origin in individual languages

	Whole database			Basic vocabulary		
	PMP	non-PMP	Total	PMP	non-PMP	total
SK (HEW)	62%	38%		75%	25%	
	136	84	220	91	31	122
WL (LWI)	49%	51%		61%	39%	
	134	142	276	89	58	147
CL (KK)	47%	53%		57%	43%	
	158	175	333	97	73	170
EL (LMT)	54%	46%		62%	38%	
	69	59	128	57	35	92
KD (LB)	57%	43%		64%	36%	
	131	97	228	78	44	122

The comparably low percentage on non-PMP vocabulary in Eastern Lamaholot is most likely influenced by the small number of non-basic vocabulary that is known for this subgroup. Less than one third of the 128 EL lexical items are non-basic vocabulary, while out of the 333 CL lexical items almost half is non-basic. Due to the different absolute numbers of lexical items available per subgroup, the percentages are not entirely comparable.

About 19% of the non-PMP vocabulary was already present in PFL and is inherited into the individual languages. To this non-PMP vocabulary of PFL, Sika has added about 18% of non-PMP vocabulary after having split from PFL, while in Central Lamaholot, an addition of more than 30% of new vocabulary is attested. Much of the non-PMP vocabulary is shared between Kedang and the Lamaholot varieties (cf. §6.3.2), and another part of it only among the Lamaholot varieties (cf. §6.3.4).

In §6.3.5, I have proposed that the Flores-Lembata languages have been influenced by a lexical substrate after the split of PFL. The Lamaholot sub-

groups located in the centre were affected most and thus have the largest amount of lexical replacement. A larger amount of lexical replacement can also indicate that the contact was more intense or of longer duration. I have pointed out that a number of the lexical items that do not go back to PMP are also found in other Austronesian and non-Austronesian languages of the region, but this is still the minority.

Part III

Morpho-syntactic innovations

CHAPTER 7

Introduction to Part III

7.1 Language sample and data sources

In this part, I compare structural features across languages in the area of the Indonesian province Nusa Tenggara Timur (NTT) and the country of Timor-Leste shown on the map in Figure 7.1. The Flores-Lembata languages are located in the middle of this area, surrounded by other Austronesian languages and by the languages of the Timor-Alor-Pantar family (cf. 1.3.1).

Table 7.1 gives an overview of the Flores-Lembata varieties and the data sources used in Part III and Table 7.2 provides a list of other languages in the area of study that are used for morpho-syntactic comparison. If applicable, dialectal varieties of the languages are given in brackets after the source to indicate that the source takes data from a particular variety. The location of these languages can be found on the map in Figure 7.1.

Table 7.1: The Flores-Lembata varieties in the language sample of Part III

Variety	ISO 639-3	Sources
SK-Hewa	ski	Fricke 2014a
SK-Hewokloang	ski	Rosen 1986
SK-Krowe	ski	Lewis and Grimes 1995
SK-Nita	ski	Rosen 1986
KD-Leubatang	ksx	Klamer 2015b (only lexical)
KD-Leuwayang	ksx	Samely 1991a
CL-Atadei Painara	lmf	Krauß 2016
CL-Central Lembata	lvu	Fricke 2019; Chapter 3
CL-Lewokukung	lvu	Keraf 1978b (only lexical)
WL-Adonara	adr	Grangé 2015a; Klamer 2015c (only lexical)
WL-Alorese	aol	Klamer 2011 (Baranusa, Alor Kecil); Moro 2016a (Alor Besar) (only lexical)
WL-Lamalera	lmr	Keraf 1978a
WL-Lewoingu	slp	Nishiyama and Kelen 2007
WL-Lewotobi	lwt	Nagaya 2011
WL-Solor	adr	Kroon 2016
EL-Lewoeleng	lwe	my own fieldnotes 2017

Table 7.2: Languages outside of Flores-Lembata in the language sample of Part III

Language	ISO 639-3	Island	Family	Subgroup	Sources (variety in source)
Abui	abz	Alor	TAP	Alor-Pantar	Kratochvíl 2007
Adang	adn	Alor	TAP	Alor-Pantar	Robinson and Haan 2014
Amarasi	aaz	Timor	AN	Timor-Babar	Edwards 2016a (Kotos dialect)
Blagar	beu	Pantar	TAP	Alor-Pantar	Steinhauer 2014
Bunaq	bfu	Timor	TAP	Timor	Schapper 2009
Ende	end	Central Flores	AN	Bima-Lembata	Aoki and Nakagawa 1993 (only lexical data)
Fataluku	ddg	Timor	TAP	Timor	van Engelenhoven 2009; Heston 2015
Hawu	hvn	Sabu	AN	Bima-Lembata	Walker 1982
Helong	heg	Timor	AN	Timor-Babar	Edwards 2018a (Funai dialect); Jacob and Grimes 2011; Balle 2017 (Semau)
Kaera	jka	Alor	TAP	Alor-Pantar	Klamer 2014b
Kamang	woi	Alor	TAP	Alor-Pantar	Schapper 2014a
Kambera	xbr	Sumba	AN	Bima-Lembata	Klamer 1998
Keo	xxk	Central Flores	AN	Bima-Lembata	Baird 2002
Komodo	kvh	West Flores	AN	Bima-Lembata	Verheijen 1982 (only lexical data)
Makalero	mkz	Timor	TAP	Timor	Huber 2011
Makasae	mkz	Timor	TAP	Timor	Huber 2008
Mambae	mgm	Timor	AN	Central Timor	Fogaça 2017 (Same dialect); Grimes et al. 2014 (Same dialect)

Language	ISO 639-3	Island	Family	Subgroup	Sources (variety in source)
Manggarai	mqy	West Flores	AN	Bima-Lembata	Semiun 1993; Verheijen 1967 (only lexical data)
Naueti	nxn	Timor	AN	Timor-Babar	Velo 2016
Rote languages	several	Rote	AN	Timor-Babar	Jonker 1908 (Termanu); Jonker 1915; Jacob and Grimes 2011 (Lole); Tamelan 2007 (Dela)
Sawila	swt	Alor	TAP	Alor-Pantar	Kratochvíl 2014
Teiwa	twe	Pantar	TAP	Alor-Pantar	Klamer 2010
Tetun	tet	Timor	AN	Timor-Babar	van Klanken 1999 (Fehan dialect)
Waima'a	wmh	Timor	AN	Timor-Babar	Himmelman et al. 2006
Welaun	-	Timor	AN	Central Timor	Edwards 2019
Wersing	kvw	Alor	TAP	Alor-Pantar	Schapper and Hendery 2014
Western Pantar	lev	Pantar	TAP	Alor-Pantar	Holton 2014

For each linguistic example from a Flores-Lembata language in Part III, I provide a heading containing the subgroup abbreviation and the spelled out variety, as listed in Table 7.1. For example WL-Lewoingu meaning the Lewoingu variety in the Western Lamaholot subgroup. If relevant, the dialectal variety can be specified in brackets after the language name. For linguistic examples from language outside of the Flores-Lembata subgroup, only the language name with an optional indication of the dialectal variety is given, but there is no subgrouping prefix used as is the case for the Flores-Lembata languages.

7.2 Methodology

To answer the research questions (6) and (7) in §1.4, replicated here, about structural features in the Flores-Lembata languages that can be attributed to contact-induced change, I compare morpho-syntactic features in the area of study laid out in §7.1 above.

- (6) Which structural features in the Flores-Lembata languages are innovations?
- (7) Which structural innovations can be attributed to contact-induced change?

The morphological and syntactic features discussed in the chapters of Part III are chosen on the basis of potentially being contact-induced innovations in the languages of Flores-Lembata. Each feature is investigated following three steps.

Firstly, it is shown that the feature is an innovation and not a retention from an Austronesian (AN) ancestor language, such as Proto-Malayo-Polynesian (PMP). The feature is considered innovated if the feature can neither be reconstructed to Proto-Austronesian (PAN) nor Proto-Malayo-Polynesian, nor is it typically found in Austronesian languages as a whole, in particular not in those spoken further west. For this study, the area further west of the Flores-Lembata languages covers the AN languages of Sumba, Central and Western Flores which do not show traces of contact with the non-Austronesian languages of the Papuan area.

Secondly, I provide data on the presence of the respective feature in the eastern part of the area of study, which covers the languages of Flores-Lembata, of Timor and of the Timor-Alor-Pantar (TAP) family. As the Flores-Lembata languages and the AN languages of Timor are both Austronesian but show the features that I have classified as not inherited from an Austronesian ancestor, I conclude that the Flores-Lembata languages and the AN languages of Timor innovated these features.

Lastly, evidence for contact-induced change as well as for internally-driven change is investigated for the innovated features in the Flores-Lembata languages. Potential evidence for contact-induced change are (i) the presence of the structural pattern in the neighbouring non-Austronesian languages of the TAP family which could be related to the unknown contact language and are the only non-AN languages in the area, and (ii) the documentation of similar cases in other areas where it has been shown that non-Austronesian languages have caused the same pattern to arise in Austronesian languages. Potential evidence for internally-driven changes are (i) universal tendencies in language change, such as typical grammaticalisation pathways that are cross-linguistically frequent, and (ii) universal markedness of features because marked features (more difficult to learn) are less likely to survive languages shift and to be taken over by the target language community (Thomason 2001:76).

I do not discuss the emergence of these features in the Austronesian languages of Timor as this is not within the scope of this dissertation. However, based on the very similar observations concerning the occurrence of the features discussed here, one may hypothesize similar developments for the Austronesian languages of Timor. However, the linguistic situation of the Austronesian languages of Timor appears to be more complex than the situation of the Flores-Lembata languages, including more languages of at least two higher-level subgroups (cf. §1.3.1). Therefore, a more fine-grained analysis is needed to reconstruct the raise of the features in these languages.

7.3 Transcription and glossing conventions

In Part III, I transcribe all language data according to a standardised orthography. The same orthography is also used in the Central Lembata grammar sketch in Chapter 3 of Part I. The orthography is based on the Indone-

sian standard orthography as much as possible. Most orthographic characters correspond to IPA symbols. Table 7.3 shows those phonemes that have orthography characters which are different from the corresponding IPA symbols. The IPA symbols approximately represent the realisation of these phonemes but they are not necessarily the exact phonetic realisation which may vary across languages and speakers, sometimes the phonetic environment also plays a role.

Table 7.3: Phonemes with non-IPA orthography

Orthographic character	IPA symbol	Sound
<a>	/ɐ/	near-open central vowel
<é>	/e/	close-mid front vowel
<è>	/æ/	near-open front vowel
<e>	/ə/	central vowel (schwa)
<j>	/dʒ/	voiced post-alveolar affricate
<ng>	/ŋ/	velar nasal
<w>	/v/	voiced labiodental fricative
<y>	/j/	palatal approximant
<'>	/ʔ/	glottal stop

For reasons of comparability, examples from other sources are retranscribed into the orthographic conventions just presented. In Appendix C.1, I provide a table that shows a list of languages from other sources and the original transcription conventions used in these sources. This makes it possible to retrieve the original transcription of the examples from other sources used in this thesis.

Place names and languages names are not retranscribed using the orthographic conventions discussed here. If no English equivalent exists, place names are spelled following local conventions or the published source used.

Glosses in examples from other sources are adapted to the conventions of this thesis. In Appendix C.2, a table with glosses in this thesis and the original glosses from other sources is given. This makes it possible to retrieve the original way in which the examples were glossed.

CHAPTER 8

Innovations in the nominal domain

8.1 Introduction

In this chapter, I discuss three morpho-syntactic features of nouns that have been innovated in the languages of Flores-Lembata or in a subset of these languages. These features are (i) a formal distinction between alienable and inalienable possession attested in Western Lamaholot and Central Lamaholot, (ii) properties (adjectival concepts) expressed by nouns in all Flores-Lembata languages, and (iii) general plural marking on nouns by means of suffixes or plural words attested in Central Lamaholot and Alorese.

Table 8.1 lists each of these features and provides information on their spread in Nusa Tenggara Timur and Timor-Leste (cf. map in §7.1). A minus (-) means that the feature is not attested. A plus (+) means that the feature is attested in all or nearly all languages of this group. A minus/plus (-/+) means that the feature is attested in at least one language or up to half of the languages of the group but not in all. In this table, Flores, Sumba and Timor are geographical labels referring to the Austronesian languages spoken on these islands. Flores refers to western and central Flores, Sumba also includes the island of Sabu, and Timor also includes the island of Rote. Sika (SK), Kendang (KD), Western Lamaholot (WL) and Central Lamaholot (CL) are sub-

groups of the Flores-Lembata language family (cf. §4.1). The Flores-Lembata subgroup of Eastern Lamaholot (EL) had to be excluded because virtually no data on the morpho-syntax of this language group is available. The only non-Austronesian (non-AN) language family in the area are the Timor-Alor-Pantar languages (TAP).

Table 8.1: The spread of innovations in the nominal domain

Feature	AUSTRONESIAN							NON-AN TAP
	Sumba	Flores	SK	KD	WL	CL	Timor	
Alienability	-	-	-	-	+	+	-/+	+
Property n.	-	-	+	+	+	+	-/+	-/+
Plural	-	-	-	-	-/+	+	-/+	+

The three innovations listed in the table are possibly the result of contact-induced change. They have in common that they are attested in the eastern part of the area of study but not in the west, roughly distinguishing east and west by drawing a line between the island of Timor in the east and the island of Sumba in the west. In the east, these features occur in the Austronesian as well as in the non-Austronesian languages.

None of the features discussed here is attested on Sumba and in western and central Flores. The only exception is a minor alienability distinction in the Central Flores language Keo (cf. §8.2.2). When looking at the Flores-Lembata family, property nouns are the only feature shared by all of its members. The other features are only attested in WL and CL. In the Austronesian languages of Timor, a scattered distribution of the three features emerges. For the TAP languages, an alienability distinction and plural marking on nouns is attested throughout the family, while property nouns are only attested in some of the TAP languages.

This chapter follows the methodology explained in §7.2 to demonstrate that each of the three features is an innovation in Flores-Lembata that is likely to have arisen due to contact with non-Austronesian languages. The alienability distinction is discussed in §8.2. Property nouns are examined in §8.3 and plural marking on nouns is taken up in §8.4. In §8.5, I summarise the findings and conclude that property nouns are an areal feature which developed due to the pre-posed possessor word order (cf. §9.2), while the

alienability distinction and plural marking on nouns are new categories that were innovated in the Flores-Lembata languages and the AN languages of Timor due to contact with non-Austronesian languages.

Glosses and transcriptions from other sources are adapted to the conventions of this dissertation (cf. §7.3). A list of adapted glosses and re-transcribed sounds with their original representations is found in Appendix C.

8.2 Alienability

8.2.1 Overview

In the context of this thesis, I define alienability as a semantic property of a possessive construction. According to Payne (1997:105), an alienable possessive construction expresses a relation between two entities that is not permanent, thus can be resolved. An inalienable possessive construction expresses a relation between two entities that are inherently connected and cannot normally be separated. Typical nouns that are used in inalienable possessive constructions are body part nouns, kinship terms or parts of wholes. I refer to these nouns as inalienable nouns, while nouns that are found in alienable possessive constructions are referred to as alienable nouns (Payne 1997:105; Shopen 2007:185). In the languages discussed here, a typical alienable possessive construction could be a more analytical construction with free pronouns, while an inalienable construction may use a bound pronoun or no linking element between the two nouns. On a symbolic level, the grammar represents the semantic nature of these possessive constructions, an alienable possessive relation being more detached and less fixed, while an inalienable possessive relationship is closer and less easy to resolve. This is parallel to the relation between free possessor pronoun and possessed noun which can be separated, thus occur on their own, and bound pronoun and possessed noun which cannot be detached so easily. In some languages, inalienable nouns take an obligatory bound pronoun, which means that the noun and the pronoun cannot be separated at all.

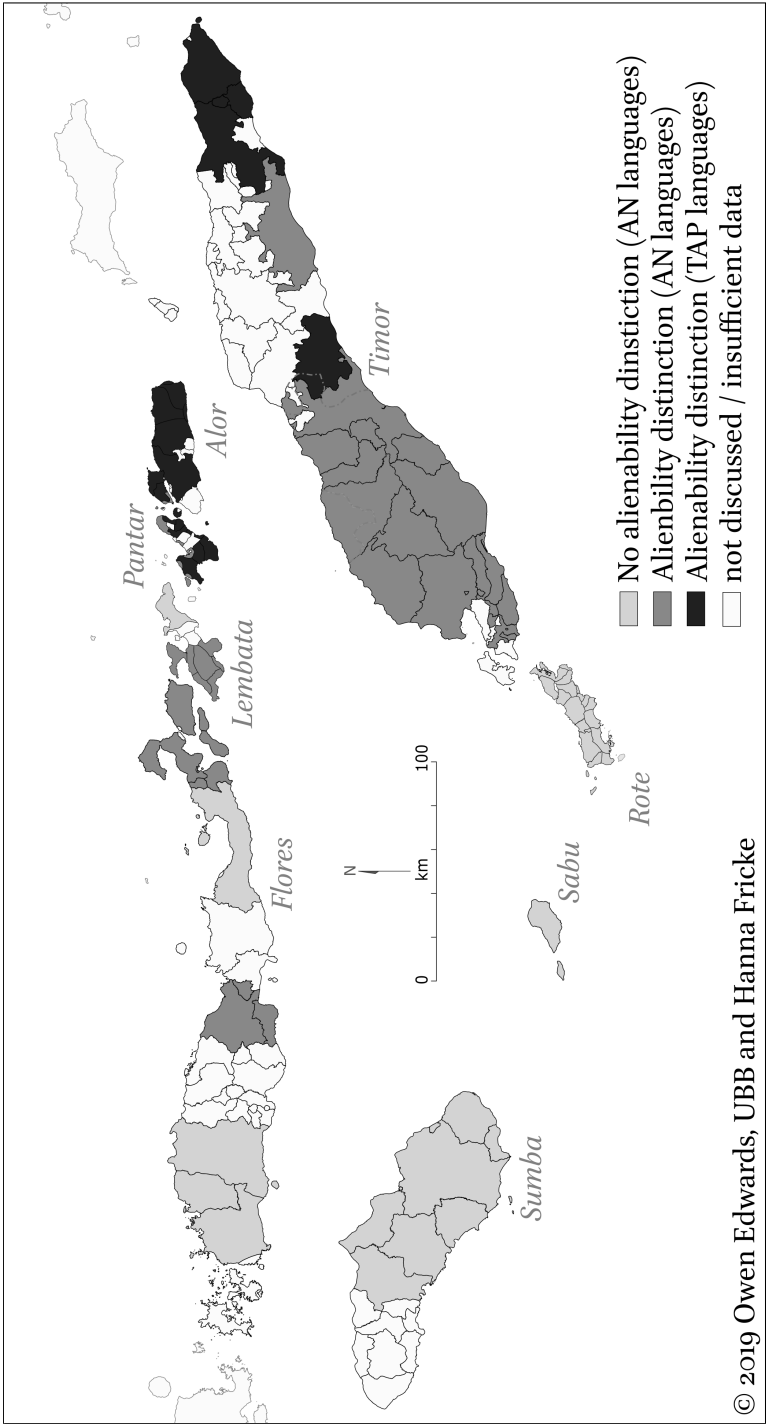
The distinction between alienable and inalienable possession is an innovation in the Austronesian languages, as it is not inherited from Proto-Malayo-Polynesian (PMP). Klammer (2002:373) lists the alienability distinction as a typical feature of eastern Indonesian languages distinguishing them

from those further west. Blust (2013:482) states that possessive constructions in the Austronesian languages of Taiwan, the Philippines and western Indonesia tend to be simple, while further east different constructions are used to express alienable or inalienable possessive relations. The details in how these Austronesian languages distinguish the two constructions vary (Himmelman 2005:165). An overview of different strategies in languages of eastern Indonesia is given in Klamer et al. (2008:116-122). The fact that an alienability distinction is neither attested in all major subgroups of Austronesian nor of Malayo-Polynesian, but only in its eastern branches, makes it unlikely to reconstruct this feature to PMP. This is in line with the result of Reesink and Dunn's study using phylogenetic methods to reconstruct ancestral states for PMP. Reesink and Dunn (2018:946) reconstruct no alienability distinction for PMP with a high degree of confidence. Thus, this feature is very unlikely to be inherited. Klamer et al. (2008:116, 122) have listed the alienability distinction as a non-Austronesian feature in many of the Austronesian language of East Nusantara that arose as a result of contact with non-Austronesian languages.

In East Nusa Tenggara and Timor-Leste, there are languages that just have one type of possessive construction. All nouns can be used in this construction with no morpho-syntactic distinction between alienable and inalienable. But there are also languages that have two different types of constructions, one for inalienable nouns and another one for alienable nouns. The map in Figure 8.1 displays the distribution of languages that grammatically distinguish alienability and those that do not. Alienability is grammaticalised in Lamaholot, in the Austronesian languages of Timor and in the non-Austronesian Timor-Alor-Pantar languages. Further west, in the Austronesian languages of Sabu, of Sumba, in Central and Western Flores, as well as in Kedang and Sika, this distinction is not relevant in the choice of the possessive construction.

In this section on alienability, I argue that Western Lamaholot and Central Lamaholot have innovated an alienability distinction and not inherited this feature from any common ancestor. After having proposed that the alienability distinction is not inherited from PMP above, I provide evidence in §8.2.2 that this feature is neither inherited from a more recent ancestor of the Flores-Lembata family, such as Proto-Bima-Lembata or Proto-Flores-Lembata. To do so, I demonstrate the absence of the feature in the languages of Sumba, in Western and Central Flores, and in Kedang as well as Sika.

Figure 8.1: The alienability distinction in possessive constructions



In §8.2.3, I provide evidence for the presence of the alienability distinction in Western Lamaholot, Central Lamaholot and in the Austronesian languages of Timor as well as the TAP languages. Finally, §8.2.4 provides a discussion of the emergence of the alienability distinction in the Austronesian languages of the area suggesting that Western Lamaholot and Central Lamaholot acquired the alienability distinction due to contact with non-Austronesian languages, possibly belonging to the TAP family or at least being typologically similar to the TAP languages.

8.2.2 No alienability distinction

I show in this section that there is no alienability distinction attested in the closest relatives of Central Lamaholot and Western Lamaholot, i. e. Sika and Kedang, neither in the Austronesian languages of Flores and Sumba which are the languages closest related to the Flores-Lembata family (cf. §5.5). Therefore, I propose that Central Lamaholot and Western Lamaholot innovated the alienability distinction.

For the Flores-Lembata languages Sika and Kedang, no grammatical distinction between alienable and inalienable possessive construction is known (Arndt 1931; Fricke 2014a; Samely 1991a). In Sika, possessive constructions are formed by juxtaposing the possessor NP and the possessed noun without any inserted element regardless of the alienability status of the possessive relation. The same construction is used for an inalienable possessive relations (1a), where the possessed noun *kikir* ‘finger’ is a body part, as well as for an alienable possessive construction (1b), where the possessed noun *labur* ‘shirt’ is an object that may easily switch its possessor.

(1) SK-Hewa

a.	<i>bi'an kikir</i>	
	human finger	
	‘someone’s finger’	(Fricke 2014a:40)
b.	<i>bi'an labur</i>	
	human shirt	
	‘somebody’s shirt’	(Fricke 2014a:40)

In Sika, vowel-final possessed nouns take a possessive suffix in alienable and inalienable adnominal possession (Lewis and Grimes 1995:607; Fricke

2014a:38-39). In contrast to pre-posed nominal possessors, all pronominal possessors in Sika are postposed to the possessed noun. The same pronouns are used for inalienable (2a), as well as alienable possessive relations (2b).

- (2) SK-Hewa
- a. *mé nimu-n*
child 3SG-POSS
'her/his children' (Fricke 2014a:39)
 - b. *lepo nimu-n*
house 3SG-POSS
'his/her house' (Fricke 2014a:42)

Kedang has preposed possessors, both pronominal and nominal. The same construction is used for inalienable, shown in (3a) and (4a), as well as alienable possessive relations, illustrated in (3b) and (4b). The pronominal as well as the adnominal possessive construction are analytical: free possessor pronouns are used, and no morphology.

- (3) a. KD-Leuwayang
sétang né' ning
satan 3SG.POSS nose
'the devil's nose' (Samely 1991a:157)
- b. Kedang
bapa né'é méi'
father 3SG.POSS knife
'father's knife' (Kedang speaker in Malang, pers. comm.)
- (4) KD-Leuwayang
- a. *mo' ana'*
2SG.POSS child
'your children' (Samely 1991a:76)
 - b. *ko' lumar*
1SG.POSS field
'my field' (Samely 1991a:75)

In contrast, for Keo spoken in central Flores, a minor distinction between alienable and inalienable constructions has been described (Baird 2002:208).

In Keo, a possessive construction by means of juxtaposition is more frequently used for inalienable possessive relations (5a), and a construction with the particle *ko'o* is used more often for possessive relations that are alienable (5b).

- (5) Keo
- a. *'udu wawi*
 head pig
 'pig's head' (Baird 2002:210)
 - b. *'aé ko'o kami*
 water POSS IPL.EXCL
 'our water' (Baird 2002:204)

However in Keo, also the nature of the possessor in terms of animacy and humanness plays a role in the choice of construction. While inanimate objects or animals are more likely to appear in a juxtaposed possessive construction, possessors expressed by kinship terms or proper names are preferably used in a possessive construction with particle.

For Kambara on the island of Sumba, for Hawu on Sabu and for Mangarai in western Flores, no distinction between an alienable and an inalienable possessive construction is reported (Klamer 1998; Semiun 1993; Walker 1982).

8.2.3 Alienability distinction

8.2.3.1 Lamaholot

In the varieties of Lamaholot, the distinction between inalienable and alienable possessive construction is grammaticalised, while in the other Flores-Lembata subgroups Kedang and Sika it is not (cf. §8.2.2). All known varieties of Lamaholot make a grammatical distinction between inalienable and alienable possession, though there is variation in formal details. In this section, I lay out these details for varieties of Central Lamaholot (CL) and of Western Lamaholot (WL) varieties.

In the majority of cases described here, which are CL-Central Lembata, WL-Solor and WL-Adonara, the distinction is in construction type: morphological with a suffix for inalienable constructions and analytical with free

pronoun for alienable constructions. However, in the WL varieties, the alienability distinction is losing ground to varying degrees. Both the analytical as well as the morphological construction can be employed for alienable and inalienable nouns. In WL-Lewoingu, one construction is used for inalienable and alienable, while in WL-Solor and WL-Adonara, there is a tendency for free possessor pronouns to occur with alienable nouns and bound possessor pronouns with inalienable nouns. In fewer varieties, which are WL-Lewotobi and WL-Alorese, the distinction lays in the pronominal suffixes or free pronouns of which two variants exist, one for inalienable nouns and one for alienable nouns.

First, I discuss varieties where the construction type differs. In Central Lamaholot this distinction is intact, while in the Western Lamaholot varieties it appears to be disappearing. In a strict sense, it is unknown whether the distinction is disappearing or whether it has never been completely implemented in some of the WL varieties. CL-Central Lembata has two sets of possessor pronouns, one set of bound pronouns that are suffixed to the possessed noun in an inalienable possessive construction, such as *-ga* '1SG.POSS' in (6a), and one set of free pronouns that is obligatory and only found in alienable possessive constructions, such as *goé* '1SG.POSS' in (6b). In the inalienable construction, there is an optional free pronoun, here *go* '1SG.POSS', indexing the possessor and preceding the possessed noun. The full sets of CL-Central Lembata possessive pronouns are listed in §3.4.3.

(6) CL-Central Lembata

a.	(<i>go</i>) <i>najan-ga</i> 1SG.POSS name-1SG.POSS 'my name'	(C1:105)
b.	<i>goé</i> <i>kajor</i> 1SG.POSS wood 'my wood'	(C5:161)

Both possessive constructions can have a nominal possessor in addition to the possessor pronoun. The nominal possessor always precedes the possessed noun. In an inalienable possessive construction, the possessor noun immediately precedes the possessed (7a). In an alienable possessive construction with a nominal possessor (7b), the possessor pronoun, here *da* = '3PL', is obligatorily placed in between the nominal possessor and the possessed noun.

(7) CL-Central Lembata

- a. *witi ulu-n*
 goat\|s head-3SG.POSS
 'the goat's head' (NT:2)
- b. *Jon no Méri da=unan-u*
 NAME and NAME 3PL.POSS=house-SPEC
 'John and Mary's house' (N1:2)

No free possessor pronoun can be inserted in between inalienable possessor and possessed noun and the possessor noun *witi* 'goat\|s' must appear in its short form, marked with \|s in the gloss, while the possessed noun is marked with a possessor suffix (cf. §3.3.1.2). In sum, CL-Central Lembata expresses inalienable possession through morphological means, whereas alienable possession is expressed by an analytic construction.

In WL-Solor, a clear distinction between inalienable and alienable possessive constructions is attested (Kroon 2016:67,94,125). There is a full set of inalienable possessor enclitics, one of them being used in (8a), and postposed free possessive pronouns used with alienable nouns, as shown in example (8b).

(8) WL-Solor

- a. *bapa kote='ẽ*
 father head=3SG.POSS
 'father's head' (Kroon 2016:166)
- b. *labu bapa na'ẽ*
 shirt father 3SG.POSS
 'father's shirt' (Kroon 2016:166)

Not only the pronoun but also the word order differs in the two constructions. In the inalienable construction, the possessor noun is preposed, while in the alienable construction, the possessor is postposed. Kroon (2016:68,166) states that there is a tendency to generalise the inalienable construction to all nouns. Thus, the alienability distinction appears to be losing ground.

WL-Adonara also has a set of postposed free possessor pronouns and a set of enclitics. Both can be used with inalienable and alienable nouns. However, for inalienable nouns, the enclitic construction is generally preferred and it is even obligatory (Grangé 2015a:40).

WL-Lewoingu has a set of post-nominal possessor pronouns which are derived from core argument pronouns (Nishiyama and Kelen 2007:15,23,31) and a full set of suffixes which can index the possessor on the possessed noun. Both constructions, the one with free pronouns and the one with suffixes, can be used with inalienable as well as alienable nouns. There is no alienability distinction. In the data presented by Nishiyama and Kelen (2007:23-27), repeated in (9) and (10), no distinction between inalienable and alienable possessive construction is observed, as body part nouns, kinship terms and nouns referring to transferable objects are all used in the same two constructions.

(9) WL-Lewoingu: Free possessor pronoun

- a. *lima Bala na'én*
hand NAME 3SG.POSS
'Bala's hand' (Nishiyama and Kelen 2007:25)
- b. *oto Lado na'én*
car NAME 3SG.POSS
'Lado's car' (Nishiyama and Kelen 2007:25)

(10) WL-Lewoingu: Possessor suffix

- a. *bapa'-nen*
father-3SG.POSS
'his father' (Nishiyama and Kelen 2007:23)
- b. *lango-nen*
house-3SG.POSS
'his house' (Nishiyama and Kelen 2007:24)

In WL-Lewotobi and in the WL offspring Alorese, not the construction type but the shape of the pronominal denoting the possessor determines the alienability status. The alienability distinction is in these cases only found in the third person. In WL-Lewotobi, inalienably possessed nouns are marked by the suffix *-N*, meaning nasalisation of the preceding vowel, whereas alienably possessed nouns are marked by an enclitic *-kẽ*, such as shown in (11). Both suffixes are analysed as nominalisers.¹

¹ In WL-Lewotobi, the distinction is in the construction, not in the properties of the noun itself. This is suggested because a noun can, according to context, take the alienable or inalienable marker (Nagaya 2011:237-238).

- (11) WL-Lewotobi
- a. *Ika lei-N*
NAME foot-NMLZ
'Tka's foot' (Nagaya 2011:33)
 - b. *Ika doi-kẽ*
NAME money-NMLZ
'Tka's money' (Nagaya 2011:33)

The nominalising suffixes above can only be used as possessive marker for possessed nouns with a third person singular possessor. For other persons, a nominalised person pronoun has to be used. The use of such a free pronoun neutralises the alienability distinction. Thus, alienability can only be distinguished for third person singular possessors (Nagaya 2011:117,233).

In Alorese, there is an alienability distinction in the third person singular possessor pronoun. There are two variants of this pronoun, *no* for inalienable nouns and *ni* or *né* for alienable nouns (Klamer 2011:54). Historically, this seems to be the same system as found in Central Lamaholot. The preposed free possessor pronoun for inalienable nouns is identical with the core argument pronoun. This is *no* in Alorese and *na* in Central Lamaholot. The preposed free possessor pronoun for alienable nouns is different from the core argument pronoun, in Alorese it is *ni* or *né* and in Central Lamaholot it is *naé*.

In sum, in most Lamaholot varieties, including Central as well as Western Lamaholot, the distinction is in construction type: morphological with a suffix for inalienable constructions and analytical with free pronoun for alienable constructions. Different to the distinction in construction, in WL-Lewotobi, the distinction lays in the suffixes of which two variants exist: *-N* for inalienable nouns and *-kẽ* for alienable nouns. In Alorese, the distinction is made in preposed free pronoun of which two variants exist: *no* for inalienable nouns and *ni/né* for alienable nouns.

8.2.3.2 The Austronesian languages of Timor

To place the Flores-Lembata languages into a larger regional context, I show that in many Austronesian languages of Timor the distinction between inalienable and alienable possession is grammatically expressed. In most cases, inalienable nouns take possessor suffixes and free pronouns, while alienable

nouns only take free pronouns, such as in the example (12) from Amarasi, a Meto language in western Timor.

- (12) Amarasi
- a. *au fuuf-k=i*
 1SG fontanelle-1SG.POSS=1DET
 ‘my fontanelle’ (Edwards 2016a:258)
 - b. *au tas=i*
 1SG bag=1DET
 ‘my bag’ (Edwards 2016a:258)

Also in the east Timorese language Tetun, alienability plays a role in the choice of possessive constructions. Inalienable possessive constructions tend to be constructions with a preposed possessor and a possessor enclitic on the possessum, while alienable possessive constructions are constructions using a free pronoun and no enclitic with a pre- or postposed possessor (van Klinken 1999:144). The same is most likely true for many other Timorese languages but published descriptions are still insufficient.

In contrast, for the languages of Rote, which are subgrouped together with the Meto cluster on Timor (cf. Edwards 2018b), no alienability distinction has been documented (Jonker 1915). As can be seen in the examples in (13) from the language Termanu on Rote, inalienable referents such as a feet and alienable referents such as a buffaloes appear in the same possessive construction.

- (13) Termanu
- a. *au ei=ng=ala*
 1SG foot-1SG.POSS-PL
 ‘my feet’ (Jonker 1908:152)
 - b. *au kapa=ng=ala*
 1SG buffalo-1SG.POSS-PL
 ‘my buffaloes’ (Jonker 1908:372)

8.2.3.3 The Timor-Alor-Pantar languages

In all known and described Timor-Alor-Pantar (TAP) languages, an alienability distinction is attested. The AN languages of Timor (§8.2.3.2) and the

Lamaholot varieties (§8.2.3.1) thus typologically pattern with the TAP languages and not with their closest Austronesian relatives (§8.2.2).

The alienability distinction strategy in the TAP languages varies. There are languages that encode inalienable possession with a prefix and alienable possession with a free pronoun, such as Western Pantar, Blagar (Klamer 2017:25-27) or Bunaq (Schapper 2009:311-312). Languages, such as Kaera, Abui or Kamang, have two distinct prefixes for this purpose (Klamer 2017:25-27). Again another strategy is attested in Teiwa and some languages of the Timor branch of TAP (Huber 2011:124; Klamer 2017:27), where possessor prefixes are obligatory in inalienable constructions while optional in alienable constructions.

Especially the strategy of using a free pronoun versus a prefix and the use of two different affixes to express an alienability distinction is structurally similar to what has been described for Lamaholot and the Austronesian languages of Timor (cf. §8.2.3.1 and 8.2.3.2). An example of free pronoun versus prefix strategy is the inalienable construction in Western Pantar which uses a possessor prefix and an optional free pronoun (14a), while the alienable possessive construction uses a free possessor pronoun (14b).

- (14) Western Pantar
- | | | | |
|----|-------------------|-------------------------|------------------|
| a. | (<i>gai</i>) | <i>ga-wasing</i> | |
| | | 3SG.POSS 3SG.POSS-tooth | |
| | | ‘its/his/her tooth’ | (Holton 2014:62) |
| b. | <i>wénang gai</i> | <i>bla</i> | |
| | man | 3SG.POSS house | |
| | ‘the man’s house’ | | (Holton 2014:61) |

These two constructions and their functioning in distinguishing alienability is very similar to what I have described for Lamaholot in §8.2.3.1, especially to the constructions of CL-Central Lembata that also have an optional alienable free possessor pronoun in the inalienable construction, such as *gai* ‘3SG.POSS’ in the Western Pantar example (14a).

8.2.4 The emergence of the alienability distinction

In §8.2.3.1 and §8.2.3.2, I have shown that Lamaholot and the Austronesian languages of Timor grammaticalised the distinction between alienable

and inalienable possession. There are various ways in which this distinction is realised. In most varieties of Lamaholot, an analytical construction with a free pronoun is used for alienable possession, while a bound pronoun is found in constructions expressing inalienable possession. Central Lamaholot uses preposed free pronouns and suffixes, while among Western Lamaholot varieties two main strategies are found: postposed free pronouns and suffixes (WL-Solor and WL-Adonara) or two variants of the same pronoun, either a free element or a bound element, that distinguish alienable and inalienable only in the third person (WL-Lewotobi and Alorese). Similarly, in the Austronesian languages of mainland Timor, the presence of a bound pronoun on the possessed noun is associated with inalienable possession.²

As shown in §8.2.3.3, very similar patterns to distinguish alienable and inalienable possession are attested in the Timor-Alor-Pantar languages. First, the grammaticalisation of a semantic distinction, namely the alienability distinction, is shared between some of the Austronesian and all the non-Austronesian languages of this area. Second, for some languages, also the construction strategies are similar: free pronoun for alienable and bound pronoun for inalienable.

The main difference between the possessive constructions in the Austronesian languages and in the Timor-Alor-Pantar languages lays in the position of the possessive affix which is a prefix in the TAP languages and a suffix in the Austronesian languages. This difference can be explained when considering that, historically, possessor affixes are grammaticalised free possessor pronouns. As the word order in TAP languages is POSS-N, thus a possessor is preceding the noun, this naturally leads to possessor prefixes. In the Austronesian languages, which have an inherited N-POSS order, this leads to suffixes. In the modern Flores-Lembata languages, the possessor is not always postposed but also occurs preposed to the noun. The nominal possessor is constantly pre-posed, while for the pronominal possessor there is

² It varies language by language whether the nouns themselves are classified into inalienable and alienable or whether it is rather the type of possessive relationship that is expressed by a particular construction. However, whether the alienability distinction is found on the construction level or on the word level is only of minor relevance for the topic discussed here which is concerned with the emergence of such a distinction in the Austronesian languages. One may suggest that the emergence of such an alienability distinction starts at the construction level and can lead to the formation of word classes based on alienability later.

cross-linguistic variation in position (cf. §9.2). However, possessor prefixes are not attested in the Flores-Lembata languages. This suggests that the possessor suffixes developed at a time when the pronominal possessor was still constantly postposed.

In the TAP languages, the alienability distinction can be reconstructed to Proto-TAP because it is a feature attested throughout the family without any known exception. In the Austronesian languages, the alienability distinction is an innovation. This is because there is no evidence that it is inherited from PMP as I have laid out in §8.2.1. I propose that the alienability distinction found in Western Lamaholot and Central Lamaholot is neither inherited from Proto-Flores-Lembata. This is because neither Sika and Kendang show this distinction. As I have shown in §8.2.2, the languages of Western and Central Flores, as well as of Sumba do not have a regular distinction between alienable and inalienable possession, such as attested in Lamaholot. Therefore, there is no evidence to reconstruct an alienability distinction to any level higher than Proto-Central Lamaholot (PCL) and Proto-Western Lamaholot (PWL).

The innovation of an alienability distinction in Lamaholot could be contact-induced or internally-driven. All evidence points to contact-induced change for this feature. The structural similarities of distinguishing alienable and inalienable possession in Timor-Alor-Pantar languages and in those Austronesian languages that distinguish these two types suggest structural diffusion. Alienability appears to be inherited in the TAP languages but innovated in Western Lamaholot and Central Lamaholot. Therefore, the TAP languages or languages typologically similar to TAP are a likely source of the feature that entered Western Lamaholot and Central Lamaholot. It is possible that there was contact with a now extinct branch of the TAP family that had been present in the Solor Archipelago or with a language that was typologically similar to TAP. There is considerable diversity of strategies to distinguish alienable and inalienable among Lamaholot varieties. Therefore, it is more likely that the varieties developed the pattern individually in the Central Lamaholot and in the Western Lamaholot varieties. There is also no phonological evidence that Central Lamaholot and Western Lamaholot share a common proto-language on a level below Proto-Flores-Lembata (cf. §5.3). There remains the possibility that the alienability distinction attested in the WL and CL subgroups was innovated through independent contact scenarios at the level of their respective proto-languages. However, for WL

even this hypothesis needs to be taken with precaution considering the diversity of this feature within the subgroup. It is very well possible that the alienability distinction developed after Proto-WL had split up into different varieties. For CL not enough is known about this feature in other varieties than Central Lembata to decide whether it goes back to PCL or not.

8.3 Property nouns

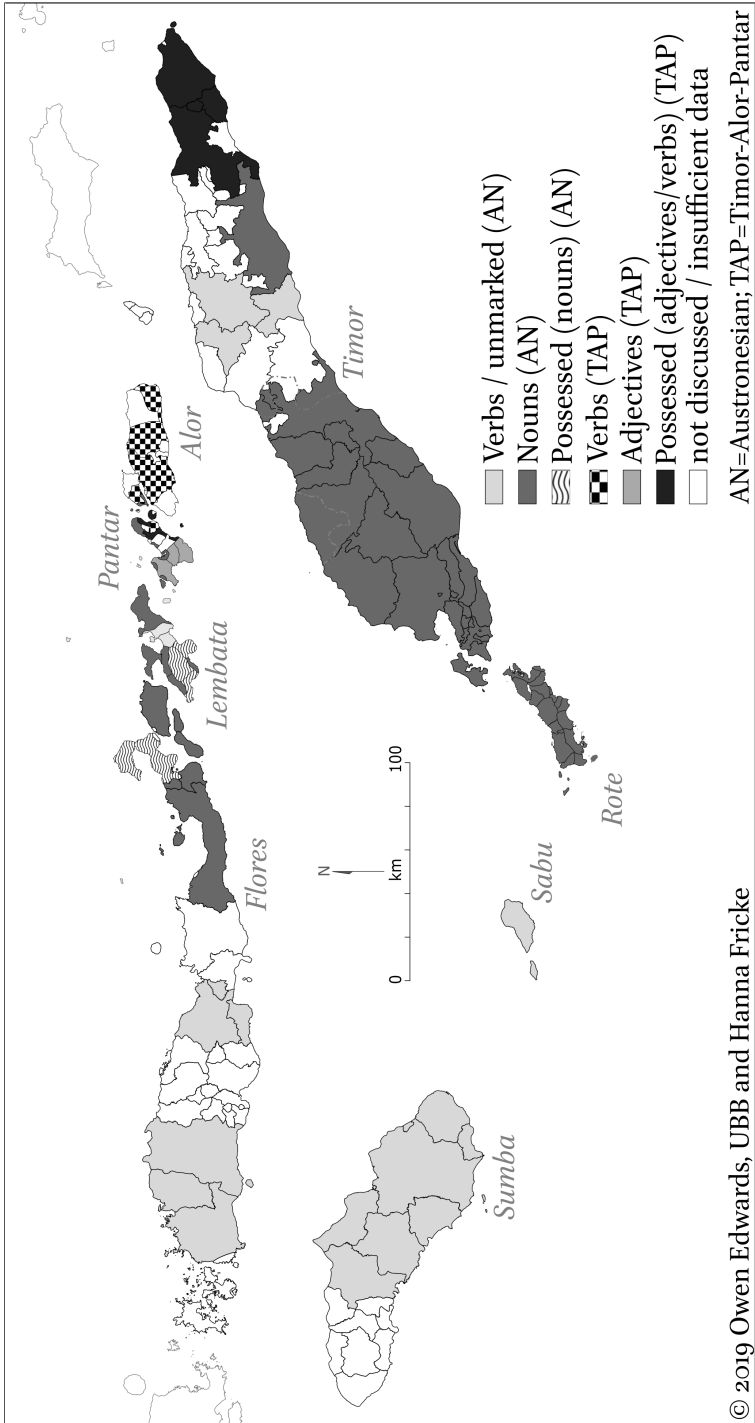
8.3.1 Overview

Cross-linguistically, properties of a nominal referent can be either expressed by adjectives, or they are expressed by words that fall into the category of nouns or of verbs (Dixon 1982). Many Austronesian languages do not have a morpho-syntactic class of adjectives. Words that express properties are usually expressed by stative verbs (Blust 2013:493). Property nouns are not considered an inherited feature in Austronesian languages. However, there are some Austronesian languages in which property words in attributive position behave morpho-syntactically in the same way as possessed nouns and these languages are spoken around New Guinea, an area where non-Austronesian languages are the majority (Blust 2013:493). Blust (2013:493) does not explicitly say that this phenomenon could be a result of contact. Nevertheless, it is striking that it is only found in the vicinity of non-Austronesian languages.

The Flores-Lembata languages are a parallel case to Blust (2013:493) for New Guinea. The Flores-Lembata languages are an Austronesian language group with properties expressed by nouns. In many cases these are clear nominalisations. Central and Western Lamaholot also use possessed nouns.

The map in Figure 8.2 shows how property words have been classified in the Austronesian and non-Austronesian languages of the area. It becomes clear that in both language families, the classification of property words as verbs and as nouns is wide-spread. Only in a few languages, property words form their own word class of adjectives. In some Austronesian, i.e. Central Lamaholot and Western Lamaholot, as well as some of the Timor-Alor-Pantar languages of Pantar and of Timor, property words can be possessed.

Figure 8.2: Classification of property words



In this section, I discuss property words, i. e. words that express adjectival concepts, and their classification as verbs or nouns. I show that property concepts are expressed by nominals in all Flores-Lembata languages, a feature which is also found in most Austronesian languages of Timor and in the TAP family but not in AN languages further west. This section is structured as follows. I have argued above that property nouns are not an inherited feature in Austronesian languages. Neither can the feature be reconstructed to a more recent ancestor, such as Proto-Bima-Lembata, as property nouns are not attested in the Austronesian languages of Sumba, Western Flores and Central Flores which are all part of this subgroup, as I demonstrate in §8.3.2. In §8.3.3 I provide evidence that in all Flores-Lembata languages property words can be classified as nouns. Based on the evidence presented before, this feature is an innovation in the Flores-Lembata languages. As it is found in all languages of the subgroup, it is likely to be an innovation of Proto-Flores-Lembata that was passed on to all modern FL languages. In §8.3.4, I show that many Austronesian languages of Timor also have property nouns. An additional feature of property nouns is found in Central Lamaholot and in WL-Lewoingu. In these languages, as I demonstrate in §8.3.5, property nouns can be possessed in attributive position and thus constitute a possessive construction with the noun that they modify semantically. Finally in §8.3.6, I examine the possibilities of language-internal development versus contact-induced innovation for the feature of property nouns in Flores-Lembata. I conclude that the nominal characteristics of property words in FL can be explained by an analogy between property attribute and possessed noun, as both occur following another noun in the NP. However, I also observe that similar structural features are attested in the TAP languages which makes (possessed) property nouns an areal feature that may have diffused either way.

8.3.2 No property nouns

In this section, I show that property words in the languages of Central Flores, Western Flores, Sumba and Sabu do not show any nominal properties, such as nominalising or possessive morphology. I also show that the structural parallel between possessed nouns and property words that is attested in the Flores-Lembata languages does not hold for the languages further west discussed in this section. This makes them structurally different from the

languages of Flores-Lembata that have property nouns. As the languages of Flores, Sumba and Sabu are the closest relatives of the Flores-Lembata family, the absence of property nouns in these languages supports the hypothesis that this feature is an innovation in the FL family.

The languages of Central Flores are largely isolating, therefore there is no morphological marking to distinguish word classes. According to Baird (2002:134), property words in Keo behave like verbs, not like nouns. Different to the Flores-Lembata languages, there is no morpho-syntactic parallel between an attributive property word and a possessed noun. In a Keo, the possessed noun, such as *iné* 'mother' (15a), precedes the other noun in the NP, while the property word, such as *petu* 'hot' (15b), follows the noun. In the Flores-Lembata languages, the possessed noun and the property word both follow the noun (cf. §8.3.6.3).

- (15) Keo
- a. *iné dako*
mother dog
'dog's mother' (Baird 2002:213)
 - b. *kopi petu*
coffee hot
'hot coffee' (Baird 2002:134)

In Manggarai, a language of Western Flores, attributive property words, such as *mese* 'big' in (16), do not show any marking.

- (16) Manggarai
- mbaru mese*
house big
'big house' (Semiun 1993:30)

Attributive property words in Manggarai as in the example above are different from possessed nouns, such as *baju* 'shirt' in (17). Possessed nouns are noun phrase initial and connected to the possessor noun with the particle *de*. Attributive property words follow the noun they modify and no particle can be inserted in between the two words. Predicative property words are also unmarked, such as *mese* 'big' in (18).

- (17) Manggarai
baju de ame
 shirt POSS father
 ‘the shirt of the father’ (Semiun 1993:41)
- (18) Manggarai
Mbaru ho'o mese.
 house this big
 ‘This house is big.’ (Semiun 1993:30)

However, Manggarai is not completely isolating as is the case for the languages of Central Flores as it has person marking enclitics (Semiun 1993:17-18), which can be attached optionally to reduplicated property words in predicative position (19). Reduplicated property words yield an intensified meaning translated with ‘very’.

- (19) Manggarai
John pintar-pintar(-n/-na).
 NAME clever-clever-3SG
 ‘John is very clever.’ (Semiun 1993:31)

The person suffixes usually attach to nouns and index the possessor (Semiun 1993:20) or to verbs and index the actor (Semiun 1993:18-19). The suffixes never attach to a property word, unless it is reduplicated, thus semantically intensified, such as in (19) above.

In Kambara on Sumba, properties are expressed by intransitive verbs (Klamer 1998:115-118). Also in Hawu on Sabu, properties are expressed by verbs (Walker 1982:25).

8.3.3 Property nouns in the FL languages

8.3.3.1 Sika

Property words in Sika do not take any productive morphological marking (Arndt 1937:12). So there is no synchronic clear morphological basis to either group them with nouns or verbs. In this section, I show that, nevertheless, Sika property words can be classified as nominals based on the derivational strategies used to generate property attributes. Part of the Sika

property nouns are generated by nominalising morphology which uses possessive suffixes.

There are three possessive suffix variants in Sika: *-n*, *-ŋ* and *-t* (cf. Lewis and Grimes 1995:607; Fricke 2014a:39-40). These suffixes mark a vowel-final noun as possessed by a 3SG possessor (20a), but also have a nominalising function (20b). Consonant-final nouns do not receive any marking (Fricke 2014a:39-40). As the same suffixes also appear on derived attributes, such as in (20c), they can be analysed as nominalisations.

- (20) a. SK-Krowe
Rapa mu'u-n
 NAME banana-POSS
 'Rapa's banana' (Lewis and Grimes 1995:608)
- b. SK-Hewa
maté-n
 dead-POSS
 'corpse' (Fricke 2014a:45)
- c. SK-Krowe
telo hena-n
 egg boil-POSS
 'a boiled egg' (Lewis and Grimes 1995:608)

Variation between the alveolar nasal *n* and the velar nasal *ng* is regional. The southeastern varieties of Sika have *n*, while the varieties further west have a velar nasal *ng* (Lewis and Grimes 1995:603). Lewis and Grimes (1995:607) state that the functional difference between the nasal suffixes and the suffix *-t* remains unclear. I observe for SK-Hewa that the choice of the possessive suffix appears to be related to the semantics of the referent. The suffix *-n* is more frequent and mainly used for non-human referents, whereas the suffix *-t* shows a tendency towards human referents and is more marginal. In addition, the variant *-t* mainly occurs on nouns in attributive position in SK-Hewa (Fricke 2014a:46). It is unknown whether these restrictions also hold for other varieties of Sika. It could be that *-t* goes back to a first person plural inclusive suffix (Fricke 2014a:47), as the nasal suffix is clearly associated with the 3SG. Lewis and Grimes (1995:607) state that there is evidence for an older system of a complete person paradigm of possessor suffixes. However, nowadays only the third person singular forms are productive.

The target category for an attribute modifying a noun, such as *hena-n* ‘boiled’ (20c), thus appears to be nominal. This explains why there are fossilised nominalising suffixes *-n* and *-t* on some Sika property words such as those in Table 8.2.

Table 8.2: Derived property nouns in Sika

Root	Derivative	Source
<i>blo</i> ‘CLASS.long’	<i>blo-n</i> ‘long’	Fricke 2014a:46
<i>du’a</i> ‘woman’	<i>du’a-t</i> ‘female’	Fricke 2014a:46
<i>da’a</i> ‘arrive’	<i>da’a-n</i> ‘raw, green’	Fricke 2014a:46
<i>mara</i> ‘thirsty’	<i>mara-k</i> ‘dried up’	Arndt 1937:12

The suffix *-k*, in contrast to *-n* and *-t*, has not been explicitly stated as a possessive marker or nominaliser for Sika. However, it is found in examples such as *mara-k* ‘dried up’. A suffix *-k* has been analysed as a nominaliser or adjectiviser for the neighbouring language WL-Lewolema (Pampus 1999:35). A suffix *-k* even appears to be more wide-spread regionally, as it is also attested as an adjectiviser in Tetun Fehan (van Klinken 1999:88-90).

The evidence presented points to a nominal character of property words. However, this is only clear when a property noun is derived from another word transparent through a nominalisation process. The majority of property words do not show any nominalisation morphology, such as SK-Hewa *gahar* ‘tall’ in (21), *geté* ‘big’ or *gu* ‘old’. These adjectives appear in attributive position without any derivation.

- (21) SK-Hewa
 tépi *gahar*
 sitting.place high
 ‘high sitting place’ (Fricke 2014a:55)

Such non-derived property nouns are not clearly nominal, but as the process to gain new property words has been shown to be nominalisation, as *hena-n* ‘boil-POSS > boiled’ above, it can be assumed that in Sika the target category to express a property is nominal.

8.3.3.2 Lamaholot

In all varieties of Lamaholot that have been described, words that express properties in attributive position can be classified as nouns. There are two morpho-syntactic properties that point to their nominal character.

First, many property words in attributive position are obviously nominalisations. This means that they are derived from a root by means of affixation, similar to what has been shown for Sika in the previous section. However, in Lamaholot these derivation processes are much more salient than in Sika and the variety of nominalisation affixes is bigger. For the varieties of Lamaholot, two subtypes of property words have been described. One of them concerns words that are inherently nominal and the other subtype contains words that are derived from an adjectival root using nominalising affixes.³

I provide details on this for two varieties of Western Lamaholot (WL) and one variety of Central Lamaholot (CL). No Eastern Lamaholot data is available on this topic.

For WL-Lewotobi, Nagaya (2011:175) differentiates adjectival nouns and adjectival verbs. Adjectival nouns, such as *wu'ũ* 'new', are nominal, based on their morpho-syntax. They can be used as NPs, as verbal objects or in predicate position, and as attributes of other nouns without any formal changes (Nagaya 2011:176). Adjectival verbs, such as *béle'* 'big', are nominalised optionally in predicate function and obligatorily in attributive function using the nominalisation suffix *-N* (nasalisation of the final vowel) or enclitic *=kẽ* (Nagaya 2011:178-179). The suffixes *-N* and *=kẽ* are otherwise used to indicate inalienable and alienable possession respectively (cf. §8.2.3.1). For property nouns, the suffix *-N* is used for "stage-level adjectives" and the suffix *=kẽ* is used for "individual-level adjectives" (Nagaya 2011:177).

For WL-Lamalera, Keraf (1978a:107-111) distinguishes adjectives that just have one form regardless in which syntactic context they occur, such as *fu* 'new' or *alus* 'good', and adjectives that have a root form and a derived form,

³ These adjectival roots are only used in particular morpho-syntactic contexts when the property concept is modified. In these contexts the nominalised forms are not allowed. This holds for all Lamaholot varieties and is described for CL-Central Lembata in §3.6.7. In addition, in Western Lamaholot varieties, the roots can also optionally be used as predicates (cf. Keraf (1978a:178-179) for WL-Lamalera and Nagaya (2011:178) for WL-Lewotobi).

such as *muulu* / *mulur* ‘straight’. There are several different derivation processes that group those adjectives that undergo derivation, such as nasalisation of the final vowel, infixation with *-n-*, or final consonant deletion or insertion. For those adjectives that have two forms, the derivatives and roots can both appear in predicate function in WL-Lamalera (Keraf 1978a:112-113), while only the derived forms are permitted in attributive function (Keraf 1978a:112). This is similar to the adjectival verbs in WL-Lewotobi, where only the nominalised adjectival verb can appear in attributive function.

In CL-Central Lembata, there are also two types of property nouns as in WL-Lamalera. Property nouns that are underived and inherently nominal, such as *weru* ‘new’ or *klemur* ‘fat’, and property nouns that became nominal through a historic derivation process, such as *knating* ‘hot’ that is derived from the root *kati* ‘hot’ by nominalisation using the infix *-n-* (cf. §3.3.6). In contrast to the Western Lamaholot varieties described above, in CL-Central Lembata, the nominalised form of property words with two forms is obligatory in predicate and attributive function (cf. §3.6.6). In addition, in Central Lamaholot varieties, property words in attributive position take possessor suffixes that otherwise only appear on possessed nouns (cf. §8.3.5).

8.3.3.3 Alorese

Although Alorese is part of the Western Lamaholot group, it has developed in its own way (cf. §1.2.4.2, §4.1.3). From a historical perspective, Alorese property words are roots, such as WL-Alorese (Baranusa) *bura* for ‘white’ (Klamer 2011:124) or fossilised nominalisation with suffix, such as in WL-Alorese (Alor Besar) *burakang* ‘white’ (Moro 2016a), which is a fossilised form of *bura-k-[a]ng* ‘white-NMLZ-3SG.POSS’. The word *bura* ‘white’ goes back to Proto-Flores-Lembata (PFL) **buda?* which again goes back to Proto-Malayo-Polynesian (PMP) **budaq* ‘white’. The presence of roots and fossilised derivatives without functional difference could mean that, at the time when Alorese split from Western Lamaholot, the nominalisation of property roots was still an active process which then was lost and the forms fossilised due to second language learning in the Alorese community, as has been proposed for all other Alorese morphology (Klamer 2011:24,26).

8.3.3.4 Kedang

In Kedang, property concepts are either expressed as verbs or as nouns. In predicate function, the property word takes a verbal suffix (22a). In attributive function, vowel-final property nouns are nominalised with suffix *-n* (22b).

(22) KD-Leuwayang

- a. *Mo' ʔaba ʔlu=né.*
 2SG.POSS necklace nice=3SG
 'Your necklace is nice.' (Samely 1991a:84)
- b. *Mo' ʔaba ʔlu-n toi'=né.*
 2SG.POSS necklace nice-NMLZ missing=3SG
 'Your nice necklace is missing.' (Samely 1991a:84)

The suffix *-n* used on property words in attributive position is a nominaliser (Samely 1991a:67). It can nominalise words from different classes, such as *déi* 'to follow' becoming *déin* 'offspring' or *maté* 'dead' becoming *matén* 'corpse'. But as in the example above, the same suffix also nominalises attributive property words.

8.3.3.5 Summary

All Flores-Lembata languages show active or fossilised nominalisation processes on all or part of their property words. In Sika and Kedang a suffix *-n* is used as a nominaliser for property words. In the varieties of Central Lamaholot a variety of nominalisation processes are attested for words expressing property concepts, which include the infix *-n-* and the suffix *-k*. In the varieties of Western Lamaholot a mix of both situations is found. There is variation in the extent to which the nominalised forms are used. In some varieties, they are only found in attributive position, such as in Kedang, whereas others also use them optionally or obligatorily in predicate position. However, nominalised forms of property words are never found in graded, intensified or negated contexts. There they always appear in the root form. In Sika, Lamaholot and Kedang, there are also property words which are not nominalised. In Kedang, these are all property words with consonant-final stems. In Lamaholot and Sika the distinction is not phonologically conditioned and seems irregular.

8.3.4 Property nouns in the AN languages of Timor

In this section, I show that there are structural similarities between property words in the Flores-Lembata languages and in some Austronesian languages of Timor.

Among the Austronesian languages of Timor, there are languages where property words have been analysed as nouns, such as in Amarasi (Edwards 2016a:233), in the Funai variety of Helong (Edwards 2018a), in the languages of Rote (Jonker 1915:457), and in Welaun (Edwards 2019). But there are also languages, such as Tetun, where property words have nominal and verbal properties (van Klinken 1999:52). In Welaun, property nouns are nominalised with affixes that are related to possessive morphology. There are also languages that have been described as having a distinct category of adjectives, such as South Mambae (Fogaça 2017:148-149) or Helong (Balle 2017) but this categorisation appears to be based on the semantics of these words and not on their formal properties.

In Amarasi there is no morpho-syntactic basis to distinguish nouns and adjectives (Edwards 2016a:233). A property word, such as *ko'u* 'big/U' in (23), is morpho-syntactically a noun.⁴

- (23) Amarasi
faut ko'u
 stone/M big/U
 '(a) big stone' (Edwards 2016a:230,243)

Superficially this construction appears similar to a possessive construction in Amarasi as the one in (24). However, a possessor noun, such as *Smara'* 'Smara'/U' in (24), is unmetathesised, marked with U in the gloss, in contrast to a modified noun, such as *faut* 'stone/M' in (23) which is marked with M in the gloss.

- (24) Amarasi
Smara' tuna-n
 Smara'/U top/U-3SG.GEN
 'top of Smara' (Edwards 2016a:261)

⁴ The abbreviations 'M' and 'U' in the glosses mean meathesised (M) and unmetathesised (U). In Amarasi, synchronic metathesis is a salient feature of multiple parts of the grammar (Edwards 2016a). The modified noun, here *faut* 'stone/M' appears in the metathesised form whereas the property noun *ko'u* 'big/U' is unmetathesised.

In Welaun, a Central Timor language, modifiers take an optional suffix *-aan* that is historically related to a third person singular genitive suffix and also functions as a nominaliser Edwards (2019). Thus modifiers, and property words as kinds of modifiers, are structurally nominals, as they are derived using nominalisation morphology. This is similar to what is described for the languages of Flores-Lembata discussed in §8.3.3.

In Tetun, property words share certain features with intransitive verbs (van Klinken 1999:52). However, there are also properties shared with nouns, such as derivational morphology (van Klinken 1999:58). Synchronically, property words in Tetun can be derived using, among others, the circumfix *k...-k* or the suffix *-k* (van Klinken 1999:87-89) but there are no records of possessive suffixes on these words. In attributive position, the property word appears in its derived form with final *k*, such as in (25). This suffix *-k* is probably ultimately related to the Lamaholot suffix *-k* used to derive property nouns (cf. §3.3.6.2).

- (25) Tetun
 tua moruk
 wine bitter
 ‘bitter wine’ (van Klinken 1999:90)

There is also a connection to possessive suffixes, in a similar way as it has been found in Sika (cf. §8.3.3.1). Tetun resultatives in attributive position are derived from verbs using the suffix *-n* which is also a genitive suffix (cf. §9.3.3.2), such as *sona-n* ‘fried’ derived from *sona* ‘fry’ (van Klinken 1999:92).

8.3.5 Possessed property nouns

In the varieties of Central Lamaholot and in the Lewoingu variety of Western Lamaholot, an additional feature of property nouns is attested. Property nouns can be morphologically possessed in attributive position. Thus, in this position they are possessed nouns. This is parallel to what Blust (2013:493) has proposed for some Austronesian languages in the New Guinea area.

In CL-Central Lembata, the possessor suffixes *-n* ‘3SG.POSS’ and *-ja/-i* ‘3PL.POSS’ are used on vowel-final inalienable nouns as well as on property nouns that are vowel-final (cf. §3.4.3), such as *weru* ‘new’ in (26a) and (26b).

(26) CL-Central Lembata

- a. *una weru-n*
house new-3SG.POSS
'a new house' (NB:34.3)
- b. *una weru-ja*
house new-3PL.POSS
'new houses' (NB:34.11)

For consonant-final stems, the third person singular possessor suffix is zero, such as on the noun *kédak* 'big' in (27a), and the third person plural suffix is *-i* '3PL.POSS', such as in (27b).

(27) CL-Central Lembata

- a. *magu kédak tu wé*
old.man big one PROX
'this big man' (S5:33)
- b. *magu kédak-i spati*
old.man big-3PL.POSS four
'the four clan heads' Lit. 'the four big men' (L3:283)

In the neighbouring variety of CL-Atadei, it appears possible to attach a third person singular suffix *-n* even to consonant-final stems (28). In this case, an epenthetic schwa <e> is inserted between the final consonant of the stem and the suffix *-n*.

(28) CL-Atadei Painara

- méja golok-en*
table tall-3SG.POSS
'a tall table' (Krauß 2016:120)

For this language, no example of a plural noun with property attribute is available. However, a plural form *-i* '3PL.POSS' is given in a list of possessor suffixes (Krauß 2016:121).

WL-Lewoingu also has suffixes on the property noun in attributive function, as shown in (29) where the property word *mite* 'black' has a possessor suffix, while the property word *béle* 'big' does not.

- (29) WL-Lewoingu
inamelaké mé'énen aho mite-nen béle
 man that dog black-3SG.POSS big
 'that man's big black dog' (Nishiyama and Kelen 2007:27)

Nishiyama and Kelen (2007:27-28) state that there is variation in the presence or absence of the possessor suffixes on property words.

Among the Austronesian languages of the area, possessed property nouns appear to be unique to some varieties of Lamaholot. As shown in §8.3.3 and §8.3.4, in the other languages of Flores-Lembata, as well as in several Austronesian languages of Timor, some property nouns are derived from other words via nominalisation using genitive suffixes. However, none of these languages has synchronically possessed property nouns as a regular pattern. In Amarasi on Timor, there are some cases of property nominals that take genitive suffixes in predicate position (Edwards 2016a:258-259) but none are attested in attributive position. This is the opposite to Central Lamaholot and WL-Lewoingu where attributive property nouns take possessor suffixes, while property nouns in predicative position cannot take suffixes. Although, Amarasi property concepts are expressed by nominals as is also the case in the languages of Flores-Lembata, they cannot be possessed in attributive position.

8.3.6 The rise of property nouns in Flores-Lembata

8.3.6.1 Summary

Table 8.3 summarises the analysis of property words in the AN languages of the area. For the Flores-Lembata languages, as well as the AN languages of Timor, nominal properties are more salient, while in the languages of Central and Western Flores and of Sumba, no nominal properties are attested.

Table 8.3: Classification of property words in AN languages

Language	Property words are ...
Flores-Lembata	
Sika	nouns
Kedang	nouns

Language	Property words are ...
Central Lamaholot	(possessed) nouns
Western Lamaholot	(possessed) nouns
Central Flores, Western Flores and Sumba	
Keo	unmarked, analysed as verbs
Manggarai	unmarked
Kambera	intransitive verbs
Hawu	verbs
Timor	
Tetun	verbs / nouns
South Mambae	unmarked
Welaun	nouns
Amarasi	nouns
Helong (Funai)	nouns

In the domain of property words, the Flores-Lembata languages do not behave like their closest Austronesian relatives of the region, the languages of Central and Western Flores and Sumba that express properties by verbs. In the Flores-Lembata languages examined here, there are still remnants of verbal qualities in some of the words that express properties. However, especially in attributive position, property words have more nominal characteristics. The main features are (i) nominalising morphology on forms that express properties and (ii) possessed nominals in attributive position.

Also most Austronesian languages of Timor, such as Amarasi or Welaun, appear to be atypical for Austronesian languages as their property concepts are not expressed by verbs, but rather by nouns. Nominalisation morphology can be found on property words, such as for example in Welaun. However, possessive marking on attributive property nouns is not found in the Austronesian languages of Timor. This makes them different from those property nouns found in some of the Lamaholot varieties but still similar to the other languages of Flores-Lembata. The Austronesian languages in Central and Western Flores and on Sumba appear to be typical Austronesian as they express properties by verbs.

From these observations, it can be concluded that the languages of Flo-

res-Lembata, as well as the Austronesian languages of Timor innovated the way they express property concepts in morpho-syntactic terms. As innovations can be contact-induced or internally-driven, both possibilities will be investigated in the following.

8.3.6.2 Evidence for contact-induced change

In this section, I provide information on the expression of properties in the Timor-Alor-Pantar (TAP) languages and discuss the question whether these languages are possible donor languages for the structural innovation of property nouns in Flores-Lembata.

Based on what is known on the languages of the Timor-Alor-Pantar family, property words either fall into the class of verbs or they form their own class. However, most TAP languages lack a distinct class of adjectives (Schapper 2014b:14). In case there is a class of adjectives, this class usually shares some features with nouns. There is a strong tendency for languages of Alor to have property words patterning with verbs, whereas on Pantar and Timor the classification of property words as verbal is less frequent and more nominal properties are attested on words expressing properties.

For the languages Abui (Kratochvíl 2007:109-110), Adang (Robinson and Haan 2014:248), Kaera (Klamer 2014b:120), Kamang (Schapper 2014a:307), Wersing (Schapper and Hendery 2014:454) and Sawila (Kratochvíl 2014:381), it has been stated that there is no distinction between adjectives and verbs. For Abui, a small set of adjectives as a separate class has been described. However, the majority of property concepts is expressed by stative verbs in this language (Kratochvíl 2007:109-110).

Except for Kaera on the island of Pantar, all the currently known languages that do not distinguish verbs and adjectives are spoken on the island of Alor. In the other languages of Pantar that have been described, namely Teiwa, Western Pantar and Blagar, there is a distinctive class of adjectives. In addition, the adjectives in Teiwa and Blagar can be possessed under certain semantic circumstances.

Western Pantar has a separate class of adjectives lacking inflectional morphology. Adjectives can appear in attributive as well as predicative position. No person prefixes are permitted on adjectives (Holton 2014:55). There is no parallel between a possessive construction and a noun-attribute construction.

In Teiwa, adjectives form their own word class (Klamer 2010:116). Adjectives can take possessor prefixes which nominalise the adjective, such as in (30), where the adjective *bunar* ‘drunk’ is nominalised by a third person singular prefix *ga-* ‘3SG’ which results in *ga-bunar* ‘the drunk one’.

- (30) Teiwa
Ga-bunar a'an!
 3SG-drunk 3SG
 ‘The drunk one [is] he!’ (Klamer 2010:119)

The possessed adjective refers to “a definite, specific entity out of a larger set” (Klamer 2010:119). To express an indefinite referent, the adjective has to be used in its bare form (31).

- (31) Teiwa
uy bunar
 person drunk
 ‘a drunk person / drunk people’ (Klamer 2010:119)

In Blagar, adjectives also form a class on their own and they are different from intransitive verbs (Steinhauer 2014:166). However, they share two features with nouns, as they can be preceded by a third person singular possessor pronoun in attributive position (32), and they take the same causative prefix as nouns (33) (Steinhauer 2014:179,193). The use of the possessor pronoun in (32) makes the referent of the noun, here *bapa* ‘crocodile’, specific in a similar way as has been described for Teiwa above.

- (32) Blagar
bapa 'e kuta
 crocodile 3SG.POSS fat
 ‘the fat crocodile (among other crocodiles)’ (Steinhauer 2014:179)

The causative prefix in (33) is a combination of a person prefix and the infix *-a-*. The infix *-a-* is only found on non-verbal roots in Blagar. Verbal roots use the infix *-i-* instead (Steinhauer 2014:195).

- (33) Blagar
Na zar 'a-bara.
 1SG.SBJ water 3SG-CAUS-hot
 ‘I made the water hot.’ (Steinhauer 2014:193)

The TAP languages of Timor optionally use a third person possessor pronoun with property words in attributive position. In Makasae and Makalero, a third person plural pronoun can precede the property word in attributive position and yield a contrastive or specific notion. This is very similar to what has been described for Teiwa and Blagar above. In (34), the Makasae noun *kareta-la* ‘car-PL’ and its attribute *supa* ‘new’ are connected by *gi* glossed as ‘LINK’ in the source. However, the linker *gi* is actually a third person possessor pronoun (Huber 2008:29) and therefore, this noun-attributive construction can be analysed as a possessive construction.

- (34) Makasae
asi kareta-la gi supa
 1SG.POSS car-PL LINK new
 ‘my new cars’ (Huber 2008:29)

The attributive proclitic *ki=* in Makalero (35) occurs with nouns, verbs and adjectives (Huber 2011:286). It is homophonous with the third person possessor prefix *ki-* (Huber 2011:226). According to Huber (2011:222), the Makasae pronoun *gi* ‘3SG’ and the Makalero possessor prefix *ki-* as well as the attributive proclitic *ki=* are cognate.

- (35) Makalero
nana ki=pere uere
 snake ATTR=big.SG 2DEM
 ‘the big snake’ (Huber 2011:285)

In both languages, Makalero and Makasae, property words also have strong verbal qualities and in Makalero they are explicitly analysed as stative verbs (Huber 2011:151). For Makasae, the classification is not entirely clear yet (Huber 2011:28).

In Fataluku, another TAP language in Eastern Timor, adjectives can be nominalised and possessed in attributive position (36). Here the attribute *tahin* ‘beautiful’ is nominalised using the suffix *-u* ‘NOM’ and it is preceded by a third person pronoun which makes the construction parallel to a possessive construction.

- (36) Fataluku
lau i tahin-u
 cloth 3SG beautiful-NOM
 ‘beautiful cloth’ (van Engelenhoven 2009:347)

In sum, a variety of patterns is found in the TAP languages. The expression of properties by means of verbs appears to be most wide-spread, as it appears on Alor, on Timor and also in one language of Pantar. On the islands of Pantar and the island of Timor, there are cases of nominalised property words and possessed property words in a similar pattern that has been described for the languages of Flores-Lembata in §8.3.3. It could be that the nominal features of property words in the TAP languages goes back to Proto-Timor-Alor-Pantar because the pattern is found on Timor and on Pantar, the two most distinct branches of the family. But it could also be that the nominal features of property words in TAP are also innovations, as they are in the Flores-Lembata languages and the Austronesian languages of Timor. There are several possible ways to explain the development in the AN and non-AN languages of the area: independent developments, borrowing from TAP languages into Austronesian languages, borrowing from Austronesian languages into TAP languages or borrowing from a third source into TAP and AN languages.

The fact that, also in the AN languages of Timor, there is a tendency for property words to have nominal features (cf. §8.3.4), makes this an areal feature that is found scattered in Austronesian as well as non-Austronesian languages all over Timor, Alor, Pantar and the Solor Archipelago but not further west. This areal feature could have diffused through these languages.

8.3.6.3 Evidence for internally-driven change

In this section, I present evidence for an internally-driven development of nominalised and possessed property words in the Flores-Lembata languages. There is a conceptual overlap between the categories of nouns and adjectives and these two categories are semantically not always distinguishable (Dixon 1982; Wierzbicka 1988; Sassoon 2013). Members of both categories can be, for example, gradable or vague (Sassoon 2013:3-4). This semantic proximity between typical nominal and adjectival concepts makes it possible that property concepts are expressed by nouns in many languages of the world.

In the Flores-Lembata languages, a crucial syntactic change in the languages, namely the shift from postposed possessor nouns to preposed possessor nouns, could have triggered the reanalysis of adjectives as nouns, and in particular as possessed nouns. As I will discuss in §9.2, the languages of

Flores-Lembata have changed the order of possessor and possessed noun, probably due to contact with non-Austronesian languages. In a typical western Austronesian language, all modifiers follow the noun they modify. These post-nominal modifiers can express possessors, wholes of a part, or properties, such as in the examples from Indonesian given in (37), where the modified noun *buah* ‘fruit’, given in bold, is followed by the possessor noun *orang* ‘person’, the holonym (whole of a part) *kelapa* ‘coconut tree’ and the property word or attribute *busuk* ‘rotten’.

- (37) Indonesian
- a. ***buah*** *orang*
fruit person
‘a person’s fruit’
 - b. ***buah*** *kelapa*
fruit coconut.tree
‘a coconut (fruit)’
 - c. ***buah*** *busuk*
fruit rotten
‘a rotten fruit’

In case of preposed possessor languages, such as the Flores-Lembata languages, nominal possessors as well as wholes of parts, which are a subtype of possessive constructions, are placed preceding the noun they modify. However, properties or attributes remain postnominal. This leads to the following structure in a language such as CL-Central Lembata in (38), where the noun *wua* / *wuak* ‘fruit’ is preceded by the possessor noun *ata* ‘person’, the holonym *tapu* ‘coconut tree’ but followed by the attributive property *knewak* ‘rotten’. The modified noun is given in bold to highlight the inconsistency in position compared to the Indonesian example in (37) above.

(38) CL-Central Lembata

- a. *ata wuak*
person fruit
'a person's fruit'
- b. *tapu wuak*
coconut.tree fruit
'a coconut (fruit)'
- c. *wua knewak*
fruit rotten
'a rotten fruit'

After this syntactic change of preposing the possessor and the holonym in a noun phrase but not property words, property words have the same syntactic position as possessed nouns, and not the same as possessors, as would be in the case of a typical Austronesian language. This realignment could have caused the reanalysis of a property words as a possessed noun or part of whole, which are typically nominals. In Central Lamaholot, the construction has been made into a full possessive construction by allowing possessor suffixes on the property noun under certain conditions (cf. 3.3.4.2).

8.3.6.4 Conclusions

In the two preceding sections, evidence for a contact-induced as well as an independent development of the nominal character of property words in the Flores-Lembata languages was presented.

I argue that nominalised and possessed property words in the languages of Flores-Lembata are a result of the preceding contact-induced word order change in the adnominal possessive construction (cf. §9.2). The preposing of the possessor led to the reanalysis of the property word as a possessed noun in attributive position. Only in Central and Western Lamaholot, these nouns are morphologically marked for possession but in all Flores-Lembata languages, they are nominal. This internally-driven process of reanalysis could have been reinforced by contact with languages that had or have nominal property words, such as the Timor-Alor-Pantar languages, as shown in §8.3.6.2. Property nouns could have been a feature of Proto-Timor-Alor-Pantar because it is attested in the TAP languages of Timor as well as in the TAP languages of Pantar. These languages belong to two very distant

branches of the TAP family. Therefore, it could be that the feature of property nouns is inherited from Proto-TAP, a proto-language with which the Flores-Lembata languages or rather their ancestor PFL could have been in contact.

Also in the Austronesian languages of Timor the attributive property aligns with the possessed noun and the part of whole in a noun phrase and is nominal in most languages. Thus, in the Austronesian languages of Timor, the same reanalysis of property words as (possessed) nouns can be proposed. These languages also have undergone a syntactic change from postposed to preposed nominal possessors.

I conclude, that property nouns are an areal feature which possibly has diffused between Austronesian and non-Austronesian languages but it is not always possible to determine the direction of diffusion. However, there is some evidence that the feature is an old feature in TAP (going back to PTAP) rather than an innovation.

8.4 Plural marking on nouns

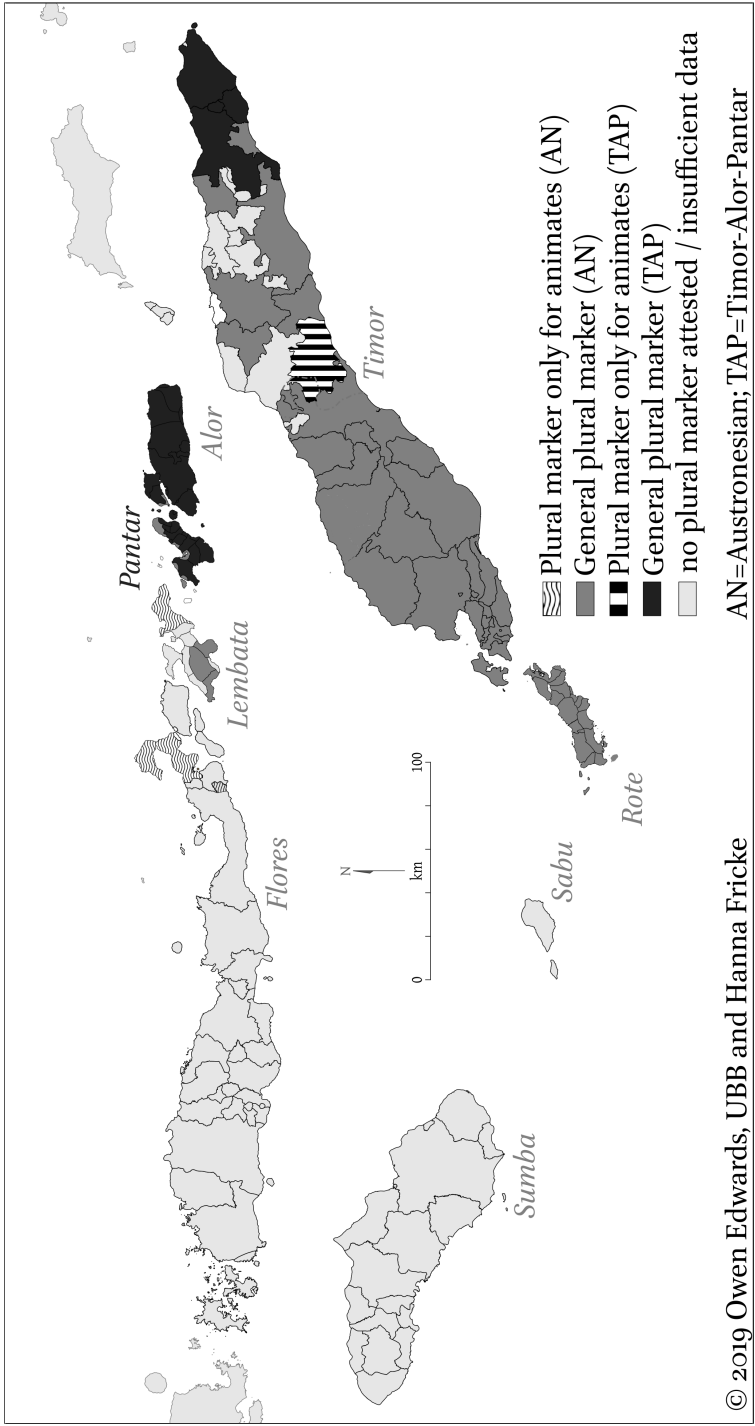
8.4.1 Overview

In East Nusa Tenggara and Timor-Leste, most Austronesian languages do not mark number on nouns in a formal way. However, there are several languages in the eastern part of the area that have developed plural words or plural morphemes to mark nouns, as indicated on the map in Figure 8.3. For the languages indicated as having plural marking, this marking can be either by plural words or plural suffixes. In the following section, the strategies are explained in more detail per language.

In Austronesian languages in general, plurality marking on nouns is not a wide-spread phenomenon. Nevertheless, a putative plural infix **-ar-* has been reconstructed for PAN (Blust 2013:389). However, the cases of plural marking described for the area studied here cannot be explained as reflexes of this affix. They must be considered innovations.

In this section, morpho-syntactic strategies for plural marking on nouns are discussed. Plural marking by means of suffixes or plural words is a feature of all TAP languages and some Austronesian languages of Timor and some of the Flores-Lembata group.

Figure 8.3: The spread of plural marking



I have argued above that, although plural marking of nouns does occur in several Austronesian languages of all branches, the phenomenon is not an inherited feature for the Austronesian languages of the Flores-Lembata family as well as those on Timor where it also occurs in many languages.

In §8.4.2, I show that plural marking on nouns is rare in the Flores-Lembata languages in general but there are clear cases of it. I introduce the emergence of a plural suffix in CL-Central Lembata which has developed parallel to a specificity suffix which is only used in singular contexts. Then I discuss the cases in Western Lamaholot, the plural word *hiré* in WL-Alorese and the associative plural suffix *-wé* in WL-Lewoingu. Finally, there are non-obligatory plural words in SK-Hewa and in Kedang for human or animate referents. In §8.4.3, I demonstrate that in the Austronesian languages of Timor several instances of nominal plural marking are attested. In §8.4.4, I show that plural marking on nouns appears to be a feature of all Timor-Alor-Pantar languages. In §8.4.5, I suggest that the concept of plural marking on nouns was borrowed from TAP languages or from languages typologically similar to TAP into some Austronesian languages of the area. But instead of also borrowing actual forms, universal strategies were applied, such as the grammaticalisation of third person plural pronoun.

8.4.2 Plural in the Flores-Lembata languages

8.4.2.1 Central Lembata

Central Lembata, a variety of Central Lamaholot, has innovated a pair of number marking suffixes on alienable nouns: *-u* for singular specific and *-ja* for plural. The two suffixes behave morpho-syntactically in the same way and they are mutually exclusive. At the end of §3.3.2, I discussed the question of whether these two suffixes belong to the category of number and concluded that, synchronically, there is no evidence for that. It appears that *-u* ‘SPEC’ encodes specificity while *-ja* ‘PL’ encodes plural number. Whether these two suffixes have been historically one category is difficult to decide. However, I suggest that they emerged within a common development for the following reasons: they attach to the same subset of nouns, namely alienable nouns, they cannot be combined with each other, and they both yield morpho-phonological processes when attaching to the nominal stem.

The plural suffix *-ja* and the pronoun *da* ‘3PL’ trace back to Proto-Central

Lamaholot (PCL) *da ‘3PL’. When PCL *da ‘3PL’ attached to vowel-final nouns to mark them for plurality, the regular sound change of PCL *d > j [dʒ] in intervocalic position applied (cf. §5.2.2) and the suffix became -ja, such as in *ao-ja* ‘dog-PL’ from vowel-final *au* ‘dog’ (< PMP *asu ‘dog’). This process caused the preceding vowel to be lowered from high vowel to mid vowel, thus here from *u* to *o*. When PCL *da ‘3PL’ attached to a consonant-final stem two developments are attested. First, the final consonant was deleted and the same sound change of PCL *d > j in intervocalic position applied, such as in *angin* ‘wind’ (< PMP *haŋin ‘wind’) which becomes *angi-ja* ‘wind-PL’. Or second, the final consonant of the stem is retained and the initial consonant of the suffix is lost, such as in *wetem* ‘millet’ (< PMP *bətəm ‘millet sp.’) which becomes *wetem-a* ‘millet-PL’.

The specificity and the plural suffix do not appear on inalienable nouns, for two reasons. First, the suffix slot is already occupied by a possessor suffix and second, number as a category is less relevant for inalienable nouns (cf. §3.3.1 and 3.3.2).

Number marking with suffixes as described here for Central Lembata appears to spread over all varieties of Central Lamaholot but not over Western and Eastern Lamaholot. Evidence for this is given in §3.3.3 about coda alternation because these two phenomena are connected in Central Lamaholot.

8.4.2.2 Western Lamaholot

Two unrelated instances of plural marking are attested in WL-Alorese and in WL-Lewoingu.

The WL language Alorese has a plural word *hiré* which is placed after a noun to mark it as plural, such as in *mato anang hiré* ‘frog small PL’ = ‘(the) small frogs’ (Moro 2018:178). The word *hiré* is a grammaticalisation of the third person plural pronoun inherited from PWL *hira ‘3PL’ (< PMP *si ida ‘they’). This grammaticalisation process has been shown to be a result of contact with Alor-Pantar languages (Moro 2018).

WL-Lewoingu has a plural suffix -wé which attaches to personal names and nouns (Nishiyama and Kelen 2007:43). The origin of this suffix is unknown. However, this suffix does not appear very frequent as it is only found on two examples in the grammar of Nishiyama and Kelen (2007:43): *Lado-wé* ‘Lado and the folks’ and *guru-wé* ‘teachers’. Also in both examples

the referents are human, thus this is probably rather a case of associative plural, meaning 'X and its people', than a general plural.

8.4.2.3 Sika and Kedang

Plural words are attested in the Sika variety of Hewa and in Kedang. However, these plural words are not obligatory and their use is restricted to human or animate referents. SK-Hewa has a plural word *'ahan* only used for human referents, such as in *dédi' anak 'ahan* 'child small PL' = 'the small children' (Fricke 2014a:14). The word *'ahan* originally means 'all; entire' (Pareira and Lewis 1998:2). In Kedang, there is the word *ata* which can follow animate nouns to mark them as plural, such as in *au ata oyo* 'dog PL sideways' = 'those dogs' (Samely and Barnes 2013:73). The use of these plural words is very restricted and thus different from plural words attested in WL-Alorese or from the plural suffix attested in CL-Central Lembata.

8.4.3 Plural in the AN languages of Timor

In the Austronesian languages of Timor, several cases of plurality and specificity marking are attested but it is not a feature that is found in all Austronesian languages of Timor. There are cases of plural words, as well as plural suffixes or enclitics.

In Tetun, spoken in the eastern part of Timor, the plural word *sia* '3PL' is attested, in addition to a singular specific article *ida* 'one' which is homophonous with the numeral 'one' (van Klinken 1999:123-124). The sources for the Tetun plural word and the specificity marker exactly match with the sources for the CL-Central Lembata plural and specificity suffixes: the plural marker is grammaticalised from the third person plural pronoun and the specificity marker comes from the numeral 'one'. Further cases of plural words are attested in Welaun, Mambae, Naueti and Waima'a. Welaun has *hira* 'PL' derived from PMP *pija 'how many' (Owen Edwards, pers. comm.). South Mambae (Fogaça 2017:159), Naueti (Veloso 2016) and Waima'a (Himmelmann et al. 2006) all use the language-internal 3PL pronoun as a plural word. Also the Malay variety used in Kupang, the biggest city in western Timor, has calqued the structure of the local languages and developed a plural word *dong* from its third person plural pronoun.

The language Amarasi in western Timor has a plural enclitic =*n* and a set of determiner enclitics that mark specificity (Edwards 2016a:214,237,239). Helong has a plural suffix -*s* (Edwards 2018a). The Rote languages have plural enclitics that are related to the the 3PL pronoun *ala/ara* (Jonker 1908:260; Jonker 1915:297).

8.4.4 Plural in the Timor-Alor-Pantar languages

All known Timor-Alor-Pantar (TAP) languages mark plurality on nouns, except for Bunaq on Timor where nouns are unmarked for number (Schapper 2009:197). The languages of the Alor-Pantar branch of the TAP family have plural words following nouns. Several of them go back to Proto-Alor-Pantar (PAP) *non 'PL', which functions as a plural word (Klamer et al. 2017:376-377).

In the Timor branch of the family, plural suffixes or enclitics are attested. Makasae has a nominal plural suffix -*la* (Huber 2008:14), such as in *asukai-la* 'man-PL'. In Makalero, two optional plural suffixes are attested for nouns: -*raa* for kinship terms and -*laa* for others (Huber 2011:236-237). Fataluku has a plurality enclitic =*éré* to mark nouns, such as in *olo=éré* 'bird=PL' meaning 'birds' (Heston 2015:21). Although in Bunaq nouns are generally unmarked for number, there is a plural marker for human referents which derives from the third person plural pronoun *hala'i*. It is mainly used in its full form but can also be reduced to =*i* 'HUM.PL' (Schapper 2009:197-199).

8.4.5 The emergence of plural marking

Plural marking on nouns is attested in scattered subset of the languages of Timor, in the Solor Archipelago and the Alor archipelago but not further west (cf. map in Figure 8.3). All TAP languages have plural marking, most AN languages of Timor also mark plurality on nouns. Among the Flores-Lembata languages, plural marking is less wide-spread.

Plural marking is more common in the non-AN Timor-Alor-Pantar (TAP) languages of the area than in the Austronesian languages. However, plural marking in the TAP languages goes back to two unrelated developments. The Alor-Pantar branch developed plural words deriving from PAP *non 'PL', while the Timor branch developed plural morphemes that attach to nouns as suffixes or enclitics and are unrelated to PAP *non 'PL'.

In the Austronesian languages that have plural marking, this sometimes goes hand in hand with specificity marking which could be interpreted as the singular counterpart of the plural marker. However, synchronically it is not always clear if these two types of marking belong to the same category (cf. §3.3.2). Plural marking is an innovation in the Austronesian languages of the region. The different cases of plural marking cannot be reconstructed to a common ancestor.

The development of a third person plural pronoun into a plural marker is a common process in Austronesian languages as well as world-wide (Holm 2000:215-217; Michaelis 2008:205; Wu 2017:61). The CL-Central Lembata suffix *-ja* goes back to the third person plural pronoun PCL **da*. On Timor, there are several cases of 3PL pronouns as sources for plural markers (cf. §8.4.3). In Tetun, for example, the plural word *sia* is still identical in form with the synchronic 3PL pronoun of the language but has broadened its function.

The plural morphemes in the TAP languages Makalero *-raa/-laa*, Makasae *-la* and Fataluku *=éré* are of unknown origin. However, taking the very common pathway of 3PL pronoun to plural marker into account, a possible relation to Austronesian forms of surrounding languages, such as Lakalei, Idate or Tetun Dili, emerges, as these languages all have *sira* (< PMP **sida* '3PL') as their third person plural pronoun. The second syllable *ra* could have developed into the plural marker of the non-Austronesian languages of the area. The language-internal 3PL pronoun, such as *kiloora* in Makalero, is not a possible source. It is derived from the 3SG pronoun *kiloo* by adding the plural suffix *-raa* (Huber 2011:221).

Two main conclusions can be drawn. Firstly, all TAP languages have plural marking on nouns, while in the Austronesian languages, especially in the eastern part of the area, there is a tendency towards plural marking. Secondly, plural marking emerges independently in most Austronesian languages and the plural words or morphemes cannot be reconstructed to higher levels. In contrast, in the AP languages, a plural word can be reconstructed to Proto-Alor-Pantar. The more ancient plural marking in the AP languages and the more recent developments in the Austronesian languages suggest that the Austronesian languages innovated plural marking. It cannot be proven that this happened due to contact with AP languages as the forms, reflexes of PAP **non* 'PL' for AP languages and reflexes of PMP **sida* '3PL' for Austronesian languages of the area, are different. However, there is the possibility

that only the concept was borrowed but not the form. Moro (2018) describes the case of Alorese where the concept of marking plurality, by means of a plural word, was borrowed from neighbouring AP languages, but not the form. Alorese grammaticalised a 3PL pronoun going back to Proto-Western Lamaholot (PWL) *hira ‘3PL’ into a plural word. The same scenario could be imagined for Central Lembata and several languages of Timor. However, in the cases of Central Lembata and Timor, the potential donor language which uses a plural marker is not spoken in the direct vicinity any more but could have been there in the past.

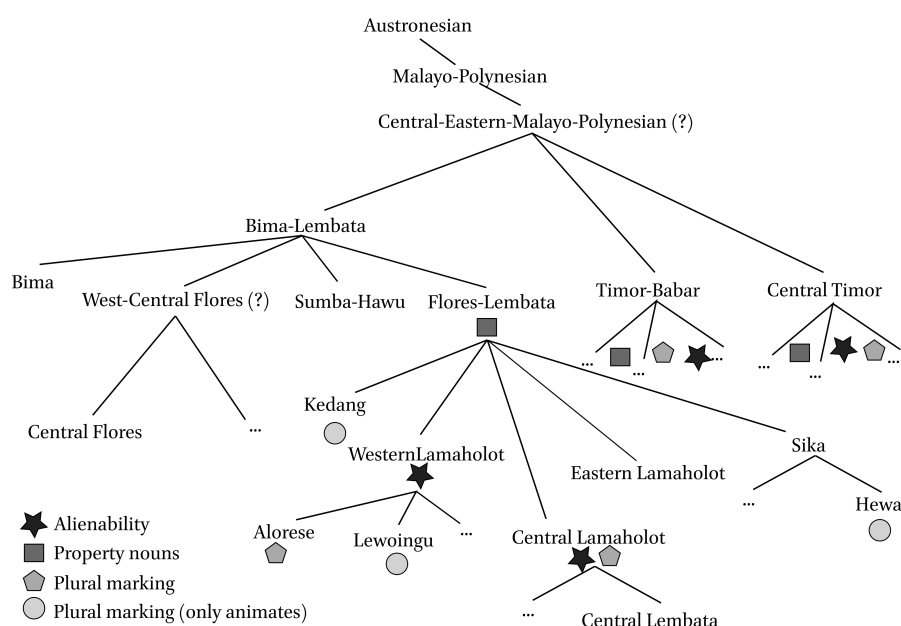
8.5 Conclusions

In this chapter, I discussed three morpho-syntactic innovations in the nominal domain that are attested in the Flores-Lembata family and also among the Austronesian languages of Timor. One of these innovations, property nouns, can be reconstructed to Proto-Flores-Lembata (PFL). The other two features, an alienability distinction and plural marking on nouns, cannot be reconstructed to PFL but they are rather innovated in some of the FL subgroups.

In Figure 8.4, I illustrate in which languages or language groups the innovations occur. For the AN languages of Timor, the symbols are placed below the subgroup because the features are not found in all languages of these groups and most likely do not go back to the proto-languages but rather developed on lower levels. As the scope of this dissertation does not include Timor languages in detail, in this figure, I do not provide further details in which languages of Timor the features occur. The tree structure is based on the current knowledge on the languages investigated (cf. §1.3.1).

Each of these three features in the Flores-Lembata languages appears to have a slightly different history. A formal distinction between alienable and inalienable possession is most likely a non-Austronesian feature that diffused into Western Lamaholot and Central Lamaholot (§8.2). Similarly, the concept of plural marking on nouns possibly diffused from non-Austronesian languages into several Flores-Lembata languages (cf. §8.4). WL-Alorese and Central Lamaholot have general plural marker. In SK-Hewa, Kedang and WL-Lewoingu, the plural markers attested are restricted to animates or humans.

Figure 8.4: Innovations in the nominal domain



An alienability distinction and plural marking are both also attested in the Alor-Pantar (AP) languages which are located in geographic proximity to Flores-Lembata. The features are found in all AP languages and can be reconstructed as features of Proto-Alor-Pantar (PAP). The fact that alienability and plural marking are old features in AP but new features in some of the FL languages suggests possible structural borrowing from AP languages into some of the FL languages. However, the donor languages could also have been languages that are typologically similar to the TAP languages. As no lexical material is borrowed, there is no clear proof that TAP languages were involved in the development of these morpho-syntactic features in the Flores-Lembata languages.

Also in some of the AN languages of Timor, the same two features of plural marking and an alienability distinction are attested. These feature may very well also have emerged due to contact with non-Austronesian languages but in a process distinct and independent from the developments in Flores-Lembata.

The PFL feature of property nouns is different, as it is less clear whether

it is an old AP features (cf. §8.3.6.4). Property nouns have been innovated in PFL as a consequence of the preposing of the possessor in the noun phrase that happened earlier in the language's history (cf. §8.3). As the preposing of the nominal possessor happened due to non-Austronesian influence (cf. §9.2), nominalised and possessed property words can be seen as an indirect consequence of contact. Although this cannot be attributed directly to contact, it must be noted that property nouns appear to be an areal feature only found in Flores-Lembata, partly on Timor and in TAP, but not on Sumba and Flores, which corresponds exactly to the distribution of the other two nominal features: alienability distinction and plural marking.

CHAPTER 9

Innovations in the noun phrase

9.1 Introduction

In this chapter, I discuss three syntactic features of the noun phrase that have been innovated in all Flores-Lembata languages, as well as in some closely related Austronesian languages. In the unrelated Timor-Alor-Pantar (TAP) languages the same features are attested. These areal features are (i) pre-nominal possessor (POSS-N), (ii) post-nominal locative nouns (N-LOC) and (iii) post-nominal numerals (N-NUM). The three features are not inherited from an Austronesian ancestor but are innovated in the Austronesian languages due to contact with non-Austronesian languages.

Table 9.1 provides information on their distribution among the languages of Nusa Tenggara Timur and Timor-Leste (cf. map in §7.1). A minus (-) means that the feature is not attested. A plus (+) means that the feature is attested in all or nearly all languages of this group. In this table, Sumba also includes the island of Sabu, and Timor also includes the island of Rote. All three features are attested throughout the languages of Flores-Lembata, the Austronesian languages of Timor and the non-Austronesian languages of the Timor-Alor-Pantar family. Only post-nominal numerals (N-NUM) also spread further west to Central Flores.

Table 9.1: The spread of innovations in the noun phrase

Feature	AUSTRONESIAN					Non-AN
	Sumba	W Flores	C Flores	Flores-Lemb.	Timor	TAP
POSS-N	-	-	-	+	+	+
N-LOC	-	-	-	+	+	+
N-NUM	-	-	+	+	+	+

W=West; C=Central

In this chapter, I investigate each of these features in detail following the methodology described in §7.2. In §9.2, I discuss the development of POSS-N in the Flores-Lembata languages. In §9.3, I show that post-nominal locatives are a consequence of the previous change in the possessive construction. In §9.4, I discuss the emergence of N-NUM in Flores-Lembata. In §9.4.4, I show traces of non-decimal counting systems in the numerals of Lamaholot and Kedang which may also be attributed to non-AN influence. In §9.5, I summarise the findings of this chapter and conclude that all three syntactic features of the noun phrase in Flores-Lembata, and also in the Austronesian languages of Timor and Central Flores, are not inherited from an Austronesian ancestor but rather innovated through contact-induced change. The syntactic changes most likely occurred in Proto-Flores-Lembata, in the proto-languages of the languages of Timor (cf. §1.3.1), and in Proto-Central Flores, due to separate events of contact with non-Austronesian languages.

Glosses and transcriptions from other sources are adapted to the conventions of this dissertation (cf. §7.3). A list of adapted glosses and re-transcribed sounds with their original representations is found in Appendix C.

9.2 Pre-nominal possessor

9.2.1 Overview

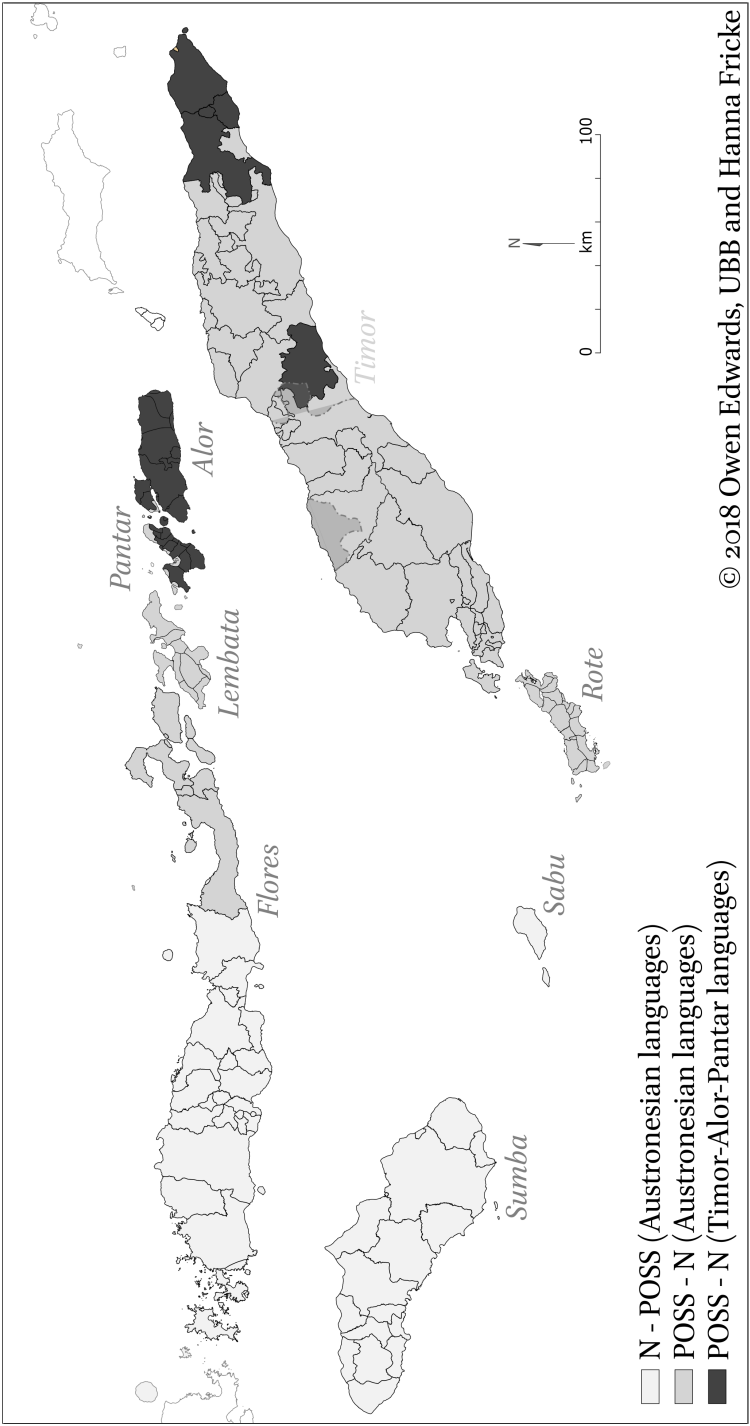
This section discusses the order of a nominal possessor (POSS) and its possessed noun (N) in a possessive construction. The languages of Nusa Tenggara Timur and Timor-Leste include languages with N-POSS order inherited from Proto-Malayo-Polynesian, as well as languages with POSS-N order which

has been argued to have emerged in the Austronesian languages due to contact with non-Austronesian languages. The POSS-N order in adnominal possessive constructions is one of the most discussed typological features of Austronesian languages in eastern Indonesia. This feature was initially mentioned by Brandes (1884:20-27) and became known as the ‘reversed genitive’ contrasting with the N-POSS order found in Austronesian languages further west. Himmelmann (2005:112-113,175) proposes two major types of Austronesian languages of Asia and Madagascar based on typological criteria: symmetrical voice languages and preposed possessor languages. Blust (2013:455) notes that Himmelmann’s pre-posed possessor languages correspond to Brandes’ languages showing the ‘reversed genitive’. Blust also suggests that POSS-N may be the result of substrate influence. Klammer et al. (2008:123) lists the POSS-N order as one of the areal features in East Nusantara shared among Austronesian and non-Austronesian languages.

The map in Figure 9.1 shows that the Austronesian languages of Flores-Lembata and Timor, as well as the Timor-Alor-Pantar (TAP) languages have POSS-N, while the Austronesian languages of Central Flores, Western Flores, Sabu and Sumba have N-POSS word order. Although not every single language is discussed in this section, on the map all languages are classified because this word order pattern appears to be distributed without known exceptions.

In §9.2.2, I describe adnominal possessive constructions with N-POSS order in the western languages of the area of study, i. e. Sumba, Western Flores and Central Flores, to show that these languages follow the typical Austronesian word order. As these languages are the closest relatives of the Flores-Lembata family, their N-POSS order strengthens the argument that POSS-N is an innovation in Flores-Lembata. In §9.2.3, I show that all Flores-Lembata languages have POSS-N and that the same order is also found in the Austronesian languages of Timor as well as in the non-Austronesian TAP languages. In §9.2.4, I discuss the development of the non-Austronesian word order in the possessive construction in the Austronesian languages of Flores-Lembata and Timor. I argue that the Flores-Lembata languages, and also the AN languages of Timor, innovated POSS-N due to separate contact scenarios with non-Austronesian languages.

Figure 9.1: Order of Nominal Possessor and Noun



9.2.2 N-Poss

In this section, I show that the Austronesian languages of Central Flores, Western Flores, Sabu and Sumba have N-Poss word order in their possessive constructions. This word order is an inherited Austronesian feature as explained in §9.2.1.

As shown in example (1) from Keo, in Central Flores, the possessed noun *'udu* 'head' is directly followed by the nominal possessor *wawi* 'pig'. There is no morphological marking.

- (1) Keo
'udu wawi
 head pig
 'pig's head' (Baird 2002:210)

Keo also has a possessive construction which uses the possessive particle *ko* or *ko'o* placed in between N and Poss (cf. Baird 2002:204, 217). The word order always remains N-Poss.

In Manggarai, spoken in Western Flores, N and Poss are connected by the preposition *dé*, such as in (2). The word order is equally N-Poss.

- (2) Manggarai
baju dé amé
 shirt POSS father
 'the father's shirt' (Semiun 1993:41)

Kambera and Hawu, on the islands Sumba and Sawu, both also have N-Poss order, as shown in (3) and (4).

- (3) Kambera
ana-na=i Ndilu
 child-3SG.POSS=ART NAME
 'Ndilu's child' (Klamer 1998:51)
- (4) Hawu
emu duaé
 house king
 'king's house' (Walker 1982:48)

9.2.3 POSS-N

9.2.3.1 The Flores-Lembata languages

In all Flores-Lembata languages, the POSS-N order is attested. Therefore, a preposed nominal possessor can be reconstructed to Proto-Flores-Lembata (PFL) with high confidence. There is evidence from pronominal possessors in the Flores-Lembata languages that the original order was N-POSS at some point in the past and POSS-N is an innovation of PFL. Free possessor pronouns occur pre- or post-nominally, with a tendency from postposition in the west to preposition in the east. Bound possessor pronouns are all suffixing. Possessive prefixes never occur in the Flores-Lembata languages.

In case alienable and inalienable possession are expressed by different constructions, the nominal possessor is pre-posed in both constructions. In most cases, the possessed noun can be marked with a bound possessor pronoun, while the possessor noun is never marked morphologically. In this section, examples of adnominal possessive constructions are given for all Flores-Lembata languages, with the exception of Eastern Lamaholot varieties due to a lack of data on possessive constructions in this language.

The westernmost Flores-Lembata language Sika (SK), juxtaposes nominal possessor and possessed noun, while marking the possessed noun with a suffix, as in (5).

- (5) SK-Krowe
Rapa mu'u-n
 NAME banana-POSS
 'Rapa's bananas' (Lewis and Grimes 1995:608)

The suffix *-n* 'POSS' marking the possessed noun is only found on vowel-final nouns, whereas consonant-final nouns have zero marking (Fricke 2014a:39-40). For more details on possessive suffixes in Sika see §8.3.3.1.

In varieties of Western Lamaholot, possessor and possessed noun are also directly juxtaposed and a suffix or enclitic can be added to the possessed noun, as can be seen in the examples (6a), (6b) and (6c).

- (6) a. WL-Lewotobi
Hugo lango'-kě
 NAME house=NMLZ
 'Hugo's house' (Nagaya 2011:24)

- b. WL-Lewoingu
guru oto-ka
 teacher car-3PL
 ‘teachers’ (pl) car’ (Nishiyama and Kelen 2007:26)
- c. WL-Solor
lango wolar=ẽ
 house rooftop=3SG.POSS
 ‘rooftop of a house’ (Kroon 2016:69)

In the WL variety of Lewotobi, only a third person singular marker is left on the possessed nouns and this marker also functions as nominaliser (Nagaya 2011:111). This is parallel to the situation in the neighboring SK variety of Hewa (Fricke 2014a:45). In contrast, the WL varieties of Lewoingu and Solor still retain a full person paradigm of possessor suffixes (Nishiyama and Kelen 2007:23; Kroon 2016:67).

In WL-Lewoingu, there is also a pattern of N-Poss (7). Here, the possessor is expressed twice, by a noun and by a pronoun.

- (7) WL-Lewoingu
oto Lado na'én
 car NAME 3SG.POSS
 ‘Lado’s car’ (Nishiyama and Kelen 2007:25)

As this N-Poss order in WL-Lewoingu only appears with an additional pronoun, this construction is different from the constructions discussed so far. The construction in (7) might be literally translated as ‘the car, Lado’s possession’ where the second part *Lado na'én* ‘Lado’s possession’ again shows Poss-N order. The first part is then analysed as a preposed adjunct.

Alorese, belonging to the WL subgroup, has developed an analytic possessive construction but maintains the order of Poss-N (8), where the third person singular pronoun *no* and the possessive particle *ning* are in between the nominal possessor *ruha* ‘deer’ and the possessed noun *kotong* ‘head’. Historically, there is a fossilised possessive nasal suffix on the possessed noun *kotong* ‘head’ (Klamer 2011:26).

- (8) WL-Alorese
ruha no ning kotong
 deer 3SG POSS head
 ‘the deer’s head’ (Klamer 2011:55)

Central Lamaholot (CL) varieties have possessive constructions where a pronoun occurs between POSS and N and constructions where POSS and N are juxtaposed.¹ In the example from CL-Atadei Painara in (9a), POSS and N are separated by the third person singular proclitic *ne*. The possessed noun is marked by a possessor suffix of which a full paradigm exists in this variety (Krauß 2016:121). In the example (9b) from CL-Central Lembata, no free pronoun is found between POSS and N. Similar to CL-Atadei Painara, a possessor suffix *-n* '3SG.POSS' is attached to the possessed noun *bapa* 'father'. CL-Central Lembata also has a full paradigm of possessor suffixes (cf. §3.4.3).

- (9) a. CL-Atadei Painara
Daniel ne una-n
 NAME 3SG house-3SG.POSS
 'Daniel's house' (Krauß 2016:121)
- b. CL-Central Lembata
kopong bapa-n
 child father-3SG.POSS
 'the child's father' (P:14)

The language Kedang also uses a possessive construction with a pronoun (10), where the possessor *sétang* 'devil' is followed by a third person singular possessor pronoun *né'* which is followed by the possessed noun *ning* 'nose'.

- (10) KD-Leuwayang
sétang né' ning
 satan 3SG.POSS nose
 'the devil's nose' (Samely 1991a:157)

9.2.3.2 The AN languages of Timor

The Austronesian languages of Timor pattern typologically with the Flores-Lembata languages in having pre-nominal possessors, as seen in the following example (11) from Amarasi with POSS-N order. A third person singular

¹ In CL-Central Lembata the distinction between connecting POSS and N with a pronoun or by juxtaposition is used to mark alienability (cf. §3.3). In CL-Central Lembata, possessor suffixes are used to mark inalienable possession only. In CL-Atadei Painara, the choice of juxtaposition or free pronoun construction rather appears to differentiate possessive relations from part-whole or attributive relations (Krauß 2016:121).

pronoun *in* is inserted between the possessor *faafj=é* ‘pig=DIST’ and the possessed noun *éku-n* ‘neck-3SG.POSS’ which is marked with a possessor suffix.

- (11) Amarasi
faafj=é in éku-n
 pig=3DET 3SG neck-3SG.POSS
 ‘the pig’s neck’ (Edwards 2016a:258)

This analytical construction with a pronoun *in* between the two nouns is similar to what has been described for Alorese, Central Lamaholot and Kedang in §9.2.3.1. However, in Amarasi, the insertion of the third person possessive pronoun is optional (Edwards 2016a:256). This is different to the Central Lamaholot varieties, where the insertion of the 3SG pronoun is functional. In CL-Central Lembata, the construction with a pronoun *in* between POSS and N is reserved for alienable possessive relations, and in CL-Atadei Painara, the construction with a pronoun is only found in attributive constructions, but not in part-whole relations.

Tetun allows preposed as well as postposed possessors. However, the preposed possessors are with 80% by far the most frequent word order found (van Klinken 1999:142). Juxtaposition of POSS and N (12a), as well as a possessive construction with an inserted third person pronoun (12b) are found.

- (12) Tetun
 a. *fahi na’in*
 pig noble
 ‘the pig’s owner’ (van Klinken 1999:143)
 b. *Bui Hirik ni naran*
 NAME NAME 3SG name
 ‘Bui Hirik’s name’ (van Klinken 1999:143)

In addition to stylistic, dialectal and structural factors, alienability plays a role in the choice of the possessive construction in Tetun (cf. §8.2).

9.2.3.3 The Timor-Alor-Pantar languages

The non-Austronesian Timor-Alor-Pantar languages all have a basic word order of POSS-N (Schapper 2014b:14). Juxtaposition of POSS and N and constructions with a third person pronoun *in* between POSS and N are both attested in TAP.

Western Pantar, for example, has a construction with a linking pronoun (13a), where the possessor *wénang* ‘man’ is followed by a third person pronoun *gai* and the possessed noun *bla* ‘house’, but also a construction, such as in (13b) where the possessor *yattu* ‘tree’ is directly followed by the possessed noun *ga-’ung* ‘3SG-head’ that is marked with a possessor prefix.

(13) Western Pantar

- a. *wénang gai bla*
 man 3SG.POSS house
 ‘the man’s house’ (Holton 2014:61)
- b. *yattu ga-’ung*
 tree 3SG-head
 ‘the head (top) of the tree’ (Holton 2014:63)

The use of these two different constructions is determined by the alienability of the possessive relationship expressed. In an alienable possessive construction, the linking pronoun is obligatory, while in inalienable possessive constructions, it is optional but the possessor prefix is obligatory (Holton 2014:60,62).

9.2.4 The emergence of Poss-N

I propose that the Poss-N word order in the Austronesian languages of Nusa Tenggara Timur and Timor-Leste is an innovation which emerged due to contact with non-Austronesian languages. In §9.2.1, I have proposed already that Poss-N in the adnominal possessive constructions is not inherited from Proto-Malayo-Polynesian. In addition, the retention of possessor suffixes in many of the Austronesian languages that have preposed nominal possessors, support the presence an earlier postposed possessor order (Klamer et al. 2008:128). According to Himmelmann (2005:113), not taking into account Oceanic languages, Austronesian languages with a preposed possessor are found on Timor, in the Moluccas and in West Papua. These are all areas where non-Austronesian languages are spoken alongside Austronesian languages until today and contact across language families is likely.²

² However, at least one exception is known. Himmelmann (2005:114) states that the Formosan language Pazeh spoken on Taiwan has pre-nominal possessors. It is unknown whether this construction is a retention or innovation in this language.

Nusa Tenggara Timur and Timor-Leste is such an area as well. The Timor-Alor-Pantar (TAP) languages in this region all have POSS-N order which thus can be considered an old feature that is reconstructible to Proto-TAP. I have shown in the §9.2.3.1 and §9.2.3.2 that in this area, the languages geographically closer to TAP languages, i. e. the Flores-Lembata languages and the AN languages of Timor, have innovated POSS-N, while those AN languages further west, i. e. the languages of Central Flores, Western Flores and Sumba, have retained N-POSS. Therefore, I conclude that contact with non-Austronesian languages has either introduced or reinforced the pattern of preposed possessors in the Austronesian languages with which they had contact. This is in line with Reesink and Dunn (2018:947) who suggest that the preposed possessor “gained popularity in communalects that came into contact with speakers of Papuan languages.”

Although the Flores-Lembata languages and the AN languages of Timor have both innovated POSS-N, no shared development can be proposed. First, this word order feature is certainly typological and not subgroup defining (Himmelman 2005:114). This feature cannot be used to propose that the Austronesian languages of Timor and the Flores-Lembata languages form a subgroup. In addition, phonological evidence has been presented in §5.5 that the Flores-Lembata languages subgroup with their western neighbours on Flores and Sumba, and not directly with the AN languages of Timor. Thus, the innovation of POSS-N most likely occurred in Proto-Flores-Lembata and not earlier because the closest relatives of Flores-Lembata retain N-POSS, as I have shown in §9.2.2.

9.3 Post-nominal locatives

9.3.1 Overview

This section is about the noun phrase word order of the noun (N) and locative expressions that encode topological information (LOC). Topological information is encoded in words that express a spatial relation between an object or a person and a point of reference. In English, these are prepositions, such as ‘in’, ‘near’, ‘on’, ‘behind’ and the like (Levinson 2003:71-72). The point of reference in the expressed spatial relation, such as a house or a tree, is usually expressed by a noun, also referred to as ‘relatum’ or ‘ground’

in spatial relation theory (Levinson 2003:7,35).

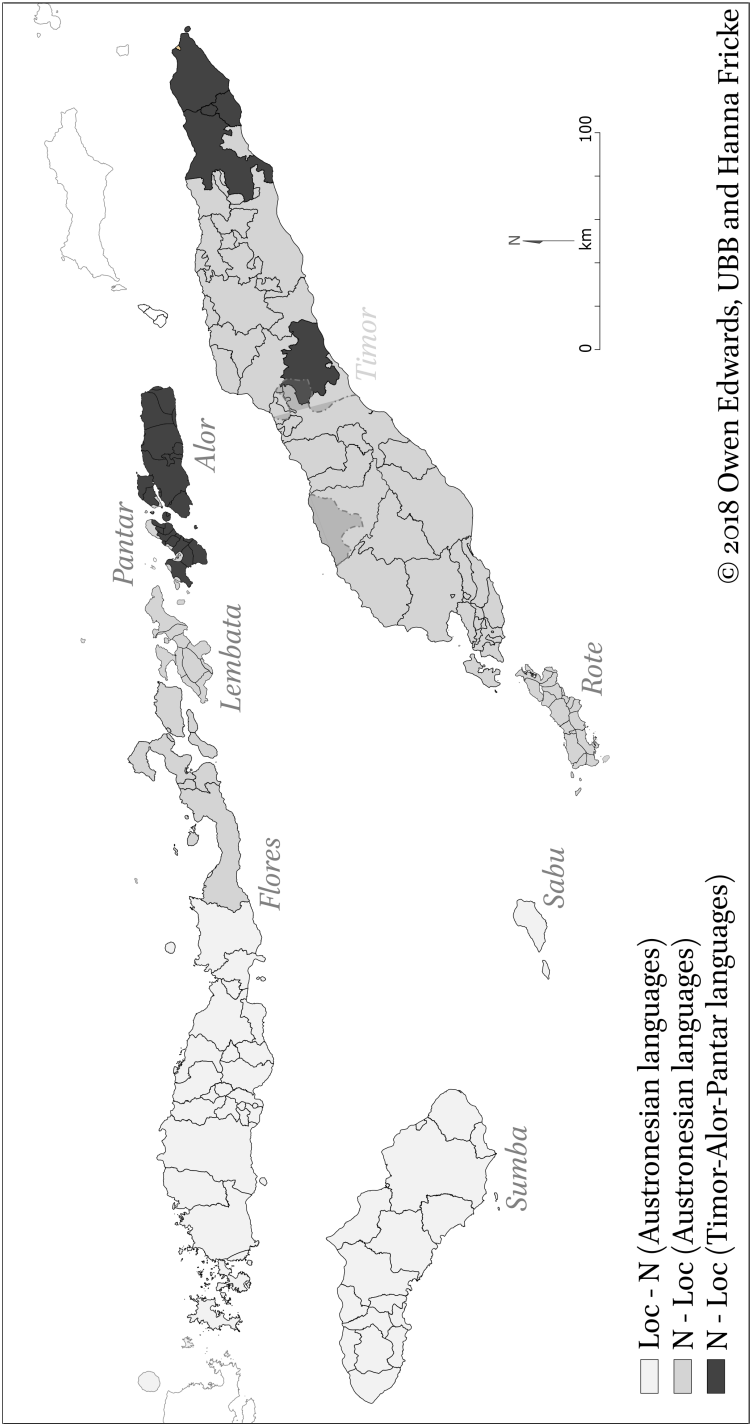
According to Blust (2013:309), locative expressions in Austronesian languages are generally composed of a generic locative preposition followed by a locative noun that expresses a topological relation. These two elements, preposition and locative noun, form a complex preposition together and precede the ground noun. This inherited Austronesian order of LOC-N is reversed in the Flores-Lembata languages as well as in the AN languages of Timor. In this section, I argue that the innovation of N-LOC in these languages is a consequence of the innovation of preposed possessors as discussed in §9.2, and thus also a result of contact with non-Austronesian languages.

On the map in Figure 9.2, languages with LOC-N and those with N-LOC are displayed. The word order of LOC-N is attested in the Austronesian languages of Central and Western Flores, as well as on Sumba and Sabu. However further to the east, in the AN languages of Timor, the Austronesian Flores-Lembata languages and in the non-Austronesian Timor-Alor-Pantar languages, the word order N-LOC is found. As for the order of nominal possessor and possessed noun in §9.2, not every single language in the area of study is discussed in this section but as no exceptions are known, on the map all languages are either indicated as N-LOC or LOC-N.

Typical locational constructions in the languages of this area often contain two elements that express different semantic aspects of the location. The first element can be a default locational preposition or a locational that carries information about distance or direction to the deictic centre. These can be words meaning ‘seawards’ or ‘upwards’. The second element, often a nominal such as ‘inside’, ‘top’ or ‘lower part’, encodes topological information.³ In this section, I am concerned with the second element, henceforward “topological nouns” or “locative nouns” (LOC), and its position relative to the ground noun (N).

³ There are also languages which only have one locational element per locational phrase. This could be a single adposition, such as in Keo in Central Flores which has prepositions or in some Timor-Alor-Pantar languages which have postpositions. These adpositions either express spatial information on distance and direction or topological information.

Figure 9.2: Order of Locative and Noun



This section is structured as follows. §9.3.2 and §9.3.3 concern the position of locative nouns in the noun phrase by showing a pattern of LOC-N word order towards the west and N-LOC word order further east. The languages of Sumba, Western Flores and Central Flores, discussed in §9.3.2, have retained the typical Austronesian LOC-N word order and thus provide evidence for the hypothesis that the N-LOC order in the Flores-Lembata languages (§9.3.3.1), as well as in the AN languages of Timor (§9.3.3.2), is an innnovation. In §9.3.3.3 I show that the non-Austronesian TAP languages also have LOC-N, which is not an innovation in these languages but most likely a retention from their ancestor Proto-TAP. This fact supports structural diffusion from TAP languages into Flores-Lembata and the AN languages of Timor. In §9.3.4, I discuss the etymological origins of locative nouns in Flores-Lembata. In §9.3.5, I compare and analyse the fossilised possessive suffixes found on locative nouns in a subset of the Flores-Lembata languages. In §9.3.6, I argue that the word order change from inherited Austronesian LOC-N to N-LOC arose due to contact with non-Austronesian languages of the area. In addition, the change was facilitated by the previous change of N-Poss to Poss-N because the locative noun and the ground noun structurally represent a possessive construction, where a locative noun, such as ‘inside’ is the possessed noun and the ground noun, such as ‘house’, is the possessor.

9.3.2 LOC-N

In the Austronesian languages of Central Flores, Western Flores, Sabu and Sumba, all locational information precedes the ground noun. This information can contain up to two elements, a general locational and an element that specifies the topological relation with the ground noun.

In the language Keo in Central Flores, a prepositional structure is attested (14), where the preposition *'oné* precedes the ground NP *kéka go'o* ‘small house’.

- (14) Keo
 'oné kéka go'o
 inside hut small
 ‘inside a small hut’

(Baird 2002:344)

There is also a general locational preposition *éna* 'LOC' (Baird 2002:340) but it is not combined with other locational preposition such as *'oné* 'inside'. Thus, in Keo only one element can precede the ground noun.

Also in Manggarai, a set of locative prepositions is attested. Besides directional, elevational or distance information, such as *éta* 'a higher place' or *wa* 'lower place', some of these prepositions also convey topological information, such as *oné* 'inside' or *musi* 'behind' (Semiun 1993:34-35).

In the languages Kambera on Sumba (15), as well as in Hawu on Sabu (16), a general locative preposition is combined with a topological expression preceding the ground noun.

- (15) Kambera
la lumbu topu
 LOC under mat
 'under a mat' (Klamer 1998:124)
- (16) Hawu
pa dida kéraja behi nadé
 LOC top cage iron DEM.2SG
 'on top of this iron cage' (Walker 1982:26)

In sum, the topological information precedes the noun in these languages. The words conveying the topological information have been analysed as prepositions or as nouns by different researchers. In Kambera, for example, they are named "prepositional nouns" (Klamer 1998:123) and some of them can still function as full nouns but most of them are grammaticalised and have to be combined with the general locational *la* 'LOC' as in (15).

The topological constructions can also be analysed as possessive constructions, with the topological noun (LOC) being the possessed noun (N_P) and the ground noun (N_G) being the possessor (POSS). When adopting this analysis, the LOC- N_G order is explained by the order of the possessive construction which is N_P -POSS (cf. §9.2). The analysis of the topological nouns as prepositions, especially in the languages of Flores, can be regarded as a further stage in grammaticalisation. In Kambera and Hawu a default locational preposition is still needed, which suggests that the topological nouns have not yet fully grammaticalised into prepositions.

9.3.3 N-Loc

9.3.3.1 The Flores-Lembata languages

In this section, I show that the word order N-LOC is consistent throughout all Flores-Lembata languages, using examples with three different topological relations, namely ‘in’, ‘on’ and ‘under’. As for the POSS-N order, also the innovation of N-LOC can be reconstructed to Proto-Flores-Lembata. The etymology of the locative nouns to express these relations is discussed in §9.3.4.

In Sika, topological information is expressed by a postposed locative noun. However, in some cases, the information is encoded in the combination of the semantics of the preposition and the postposed locative noun. In SK-Hewa, locative nouns are clearly marked as possessed by a possessive suffix *-n*. Thus locative constructions in Sika can be analysed as possessive constructions.

For SK-Krowe, there is little descriptive information available. However, from example sentences in the dictionary by Pareira and Lewis (1998) it appears that final nasals are fossilised on some locative nouns, such as *uneng* ‘inside’. In example (17), the ‘in’ relation is expressed by the posposed locative noun *uneng* ‘inside’ and the prepositional element *wawa* ‘downwards’ adds directional information.

- (17) SK-Krowe
wawa tana uneng
 downwards earth inside
 ‘in the earth’ (Pareira and Lewis 1998:204)

The topological relations expressing ‘under’ and ‘on’ are expressed by the combination of two elements, a preposition and a locative noun. The SK-Hewa example in (18) shows the use of the possessed locative noun *lewu* ‘space underneath’ to express ‘below’ or ‘under’. The reading of ‘under’ is enforced by the preposition *wawa* ‘downwards’ which points to a lower direction than the location of the deictic centre.

- (18) SK-Hewa
Ia méong wawa méja lewu-n.
 exist cat downwards table space.underneath-POSS
 ‘There is a cat under the table.’ (HC_SR:31)

To express the ‘on’ relation, a locational *réta* ‘hillwards’ or *lala* ‘upwards’ pointing to a higher location and the noun *wutu* meaning ‘end’ are combined. In example (19) from the Krowe variety both, the locational *réta* ‘hillwards’ indicating the upward direction and the locational noun *wutu* ‘end’ follow the ground noun *méja* ‘table’, while the prepositional slot is filled with the default locational *é’i* LOC.

- (19) SK-Krowe
Lopa deri é’i méja réta wutu.
 PROH 2SG.sit LOC table hillwards end
 ‘Don’t sit on the table.’ (Pareira and Lewis 1998:218)

In the sources on Sika, there is no descriptive analysis of a topological ground phrase in SK-Krowe. However, the example in (19) from the Sika dictionary by Pareira and Lewis (1998) suggests that the topological information is combined in the two elements *réta* ‘hillwards’ and *wutu* ‘end’ which are both placed after the ground noun *méja* ‘table’. Possibly, these are two locational phrases, *é’i méja* ‘at the table’ and *réta wutu* ‘at the end in hillward direction’.

In the Hewa construction in (20), similar elements are found but they appear in a different order and the locative noun has a possessive suffix. The locational *lala* ‘upwards’, which does not exist in SK-Krowe according to the dictionary of Pareira and Lewis (1998), precedes the ground noun ‘*ai*’ ‘tree’ and the locative noun *wutu-n* ‘end-POSS’ follows the ground noun.

- (20) SK-Hewa
lala ‘ai wutu-n
 upwards tree end-POSS
 ‘on the tree’ (HC_SR:43)

The locationals *lala* ‘upwards’ in SK-Hewa and *réta* ‘hillwards’ in SK-Krowe both point to an upward direction from the deictic center. The postposed locative noun *wutu* ‘end’ indicates that the location is on the (upper) end or surface.

Western Lamaholot (WL) varieties use postposed locative nouns for ‘inside’ that are cognate with the Sika noun *uneng* ‘inside’. The locative nouns show remnants of possessive suffixes that vary across varieties, as can be seen in the examples in (21a) and (21b).

- (21) a. WL-Lewotobi
 ia boti one'
 LOC bottle inside
 'inside the bottle' (Nagaya 2011:259)
- b. WL-Alorese
 oro sepatu unung
 LOC shoe inside
 'inside the shoe' (Klamer 2011:77)

To express 'on', WL varieties use either *wutu* 'end; surface' (22a), as also found in Sika, or *lolo* 'top; surface' (22b) which is also found in Central and Eastern Lamaholot varieties, as well as in Kedang. The distribution of these two options seems to occur in free variation, as in most varieties, both options are attested. Also these locative nouns show fossilised possessive suffixes. In WL-Lewotobi, the possessive suffix is realised by nasalisation of the final vowel.

- (22) WL-Lewotobi
- a. *ia uri' wutĩ*
 LOC bed top
 'on the bed' (Nagaya 2011:612)
- b. *pé méja lolõ*
 DIST table top
 'on the table' (Nagaya 2011:286)

To express the 'under' relation a postposed locative noun is also used, such as in (23), where *wewelen* 'under' follows the ground noun *méja* 'table'. The final syllable *-en* is a fossilised possessive suffix.

- (23) WL-Lewoingu
 Busan to'u pé méja wewelen.
 cat one at table under
 'There is a cat under the table' (Nishiyama and Kelen 2007:90)

Central Lamaholot (CL) as well as Eastern Lamaholot (EL) varieties also use postposed locative nouns, however these nouns do not show any trace of

a possessive suffix.⁴ One could analyse them as grammaticalised postpositions.

For Central Lamaholot, the forms meaning ‘inside’ are *or* or *ora* depending on dialectal variation, as can be seen in the Central Lembata example (24a) and the Atadei Painara example (24b) respectively.

- (24) a. CL-Central Lembata
 jéné una or
 upwards house inside
 ‘inside the house’ (FH2:10)
- b. CL-Atadei Painara
 wel una ora
 sideways house inside
 ‘inside the house’ (Krauß 2016:126)

In Central Lamaholot, the locative noun *lolo* is used for ‘top’, as in (25), while *leng* is attested for ‘below’, such as in (26).

- (25) CL-Central Lembata
 jéné ulik lolo
 upwards bed top
 ‘on the bed’ (Fl:10)
- (26) CL-Central Lembata
 wo méja leng
 DIST table space.underneath
 ‘under the table’ (FH1:4)

For Eastern Lamaholot, the locative noun *bolo* ‘inside’ (27a), *lol’o* ‘top’ (27b) and *langu* ‘below’ (27c) are attested.

⁴ For Eastern Lamaholot almost no data is available. However, I elicited the expressions in (27a) to (27c) from a speaker of the Eastern Lamaholot variety of Lewoeleng. As these sentences were not recorded but written down by the speaker himself, the representation of phonemes might not correspond to the exact realisation of the words but the structure could be captured well. The EL-Lewoeleng speaker uses an apostrophe in several words. Probably, the apostrophe symbolises a glottal stop as this is the way people in the region usually represent the glottal stop in writing. However, the apostrophe appears in between consonants and vowels which is an unusual place for a glottal stop in the languages of the region. It is unknown whether this is the actual place where a glottal stop occurs in this language or whether the glottal stop rather occurs after the vowel, thus at the end of the word.

(27) EL-Lewoeleng

- a. *do ul'i bolo*
 LOC house inside
 'inside the house' (NB:82)
- b. *do lub'a lol'o*
 LOC cooking.pot top
 'on the cooking pot' (NB:82)
- c. *do méjaa langu*
 LOC table under
 'under the table' (NB:82)

Kedang (KD) also exhibits post-nominal locative nouns, such as in example (28), where *laleng* 'inside' is placed after the ground noun *huna* 'house'.

(28) KD-Leuwayang

- bè huna laléng*
 LOC house inside
 'in the house' (Samely and Barnes 2013:330)

It is not entirely clear whether there are fossilised possessive suffixes on the locative noun in Kedang or not. Synchronically, Kedang does not have a suffix *-ng* (Samely 1991a:62). On the one hand, the final nasal of the noun *laléng* could be a fossilised suffix cognate with those found in Western Lamaholot and Sika. But as *laléng* goes back to PMP **daləm* 'in', the final nasal might also come from **m*. Given that **m* > *ng* is attested word finally for Kedang, it is more likely that the final nasal is part of the stem and that there is no suffix on the noun.

For the locative noun *lolo'* expression 'top', such as in example (29), no clear suffix is found. There is no data of a full topological construction expressing 'under' in Kedang.

(29) KD-Leuwayang

- bè éné lolo'*
 LOC sand top
 'on the sand' (Samely 1991a:158)

9.3.3.2 The AN languages of Timor

The languages Amarasi (30) and Helong (31), both spoken in the Western part of Timor island, have a postposed noun meaning ‘inside’. In Amarasi, the locative noun *nana-* ‘inside’ has a third person singular genitive suffix *-n* that goes back to the same PAN source **ni* as the fossilised suffixes found in most Flores-Lembata languages (cf. §9.3.5). However, in Helong, no suffix is found on the locative noun; the ground noun *uma* ‘house’ and the locative noun *dalé* ‘inside’ are simply juxtaposed.

- (30) Amarasi
 ét ooj=é nana-n
 IPFV.LOC water=3DET inside-3SG.POSS
 ‘in the water’ (Edwards 2016a:400)

- (31) Helong
 né uma dalé
 at house inside
 ‘inside the house’ (Balle 2017:96)

This kind of juxtaposition is also found in South Mambae (32a), an Austro-nesian language spoken in the eastern part of Timor Island. In South Mambae the locative noun *lala* ‘inside’ is placed after the ground noun, here *ai* ‘tree’, and the two nouns form a compound. The same structure is also found with other South Mambae locative nouns as in (32b).

- (32) South Mambae
 a. *ni ai lala*
 LOC tree inside
 ‘in the jungle’ (Edwards 2016a:60)
 b. *ээр lau*
 water top
 ‘on the water’ (Grimes et al. 2014:28)

In the constructions from South Mambae above, there is no evidence for genitive suffixes. However, there is evidence that locative constructions can be analysed as possessive constructions with a lost genitive suffix *-n* on the

possessum. In other varieties of Mambae many body parts occur in the citation for with a final *-n*. Secondly, in South Mambae nouns ending in CV generally metathesise their final syllable. However, locative nouns are unmetathesised. The historic presence of a final consonant *n* in these words could have blocked metathesis at the time when final syllable metathesis was a productive process (Edwards 2016a:64).

Tetun again has the same structure and possibly a fossilised nasal suffix at the locative noun *laran* ‘inside’ (33). van Klinken (1999:161) gives a list of locative nouns and all of them except for one *kotuk* ‘behind; back’ end in *n*. van Klinken (1999:150) analyses the final *n* on locative nouns as “fossilised genitive marking”.

- (33) Tetun
iha ké'an laran
 LOC house inside
 ‘inside the house’ (van Klinken 1999:161)

The word order of ground noun followed by a locative noun appears consistent across the Timor area and is parallel to the word order found in Flores-Lembata languages. Among Timorese languages surveyed here, possessive marking on the locative noun is only found in Amarasi and Tetun. Taking into account that the possessive marker goes back to a PAN form, it could be assumed that it was lost in South Mambae and Helong whereas Amarasi and Tetun can be considered as more conservative in this respect. The highly isolating nature of South Mambae with only one prefix (Grimes et al. 2014:8) supports the loss of other affixes.

9.3.3.3 The Timor-Alor-Pantar languages

The non-Austronesian languages of the Timor-Alor-Pantar family have topological verbs or postpositions, such as *ta* ‘on’ in Kaera (34).

- (34) Kaera
gang a-bat ta ga-dag
 3SG 3SG.POSS-leg on 3SG-leave
 ‘She puts him on her lap.’ (Klamer 2014b:119)

But it is also possible to encode topological information with nominals such as shown in the following examples from Kaera (35a) and Western Pantar (35b). The noun denoting ‘inside’ is marked with a possessive person prefix, *g-* ‘3SG.POSS’ in Kaera (35a) and *ga-* ‘3SG’ in Western Pantar (35b). In Western Pantar, the same person prefixes, such as *ga-* ‘3SG’, are used to mark inalienable possessors on nouns as well as core arguments on verbs (Holton 2014:70).

- (35) a. Kaera
 abang g-om
 village 3SG.POSS-inside
 ‘in the village’ (Klamer 2014b:110)
- b. Western Pantar
 bla ga-umé
 house 3SG-inside
 ‘in the house’ (Holton 2014:48)

Not all topological nouns in Timor-Alor-Pantar languages are marked with a possessive prefix. They can also be simply postposed to the ground noun without any morphological marking as in the examples from Western Pantar (36a) and Makasae (36b).

- (36) a. Western Pantar
 méa gégung
 table under
 ‘under the table’ (Holton 2014:48)
- b. Makasae
 kaisa lé'ééré mutu é'é
 box PROX inside PROX
 ‘inside this box’ (Huber 2008:36)

In the examples above, it can be seen that the Timor-Alor-Pantar languages consistently express topological relations with an element that appears after the ground noun, the word order is thus N-LOC. There are cases where the topological element has been analysed as a postposition, such as in (34).

However, as stated for Western Pantar, the postposition is “similar to obligatorily possessed nouns” (Holton 2014:47) as it carries an obligatory person prefix. Holton (2014:48) also gives an alternative analysis of the locative postpositions as they “share many properties with verbs”. Thus, the topological element can be either analysed as a noun or a verb that in some cases grammaticalised into a postposition. This situation is also described for Teiwa, another Alor-Pantar language. Klammer (2010:208-210) discusses possessed locative nouns in Teiwa and concludes that some of them are possibly “multifunctional items that can be used as either nouns or verbs”. As possessor prefixes for nouns and verbal prefixes are often homophonous or very similar in the languages of Alor-Pantar, it might be difficult to differentiate nouns and verbs in this case.

In sum, TAP languages express topological information either by verbs or nouns but these topological words are always postposed to the ground noun, thus following the order of N-LOC which is also attested in Flores-Lembata and on Timor.

9.3.4 Etymology of the Flores-Lembata locative nouns

In this section, I compare the lexemes that are used to express topological information among Flores-Lembata languages and provide suggestions on their origin and development. Table 9.2 shows four cognate sets that can be established when comparing the Flores-Lembata locative nouns for ‘inside’ from the example sentences in §9.3.3.1. In the lower part of the table, a selection of cognates found among other languages of Flores and Timor are listed. The data in this section is taken from the sources of the example sentences in the previous sections. Additional lexical data in this section is taken from the same sources as used in Part II of this dissertation (cf. §4.2). For cognates which underwent semantic shift and have a different meaning than ‘inside’, this meaning is indicated next to the form.

Two of the cognate sets are of PMP origin and spread over wider areas. Cognates going back to PMP *qunəj ‘soft core; pith’ are found in Western Lamaholot (WL), Sika (SK) and in Central and Western Flores. Cognates going back to PMP *daləm ‘in, area within, inner part of something; between; below, under; deep; mind, feelings’ are found in Kedang (KD) as well as in the Austronesian languages of Timor and in the Western Flores language Komodo.

Table 9.2: Locative nouns for ‘inside’ and their origins

PMP	*qunəj	*daləm	-	-
PCL	-	-	*ora	-
Reflexes in Flores-Lembata languages				
SK-Hewa	<i>uné n</i>	-	-	-
SK-Krowe	<i>une ng</i>	-	-	-
WL-Lewotobi	<i>one '</i>	-	-	-
WL-Lewoingu	<i>ono 'on</i>	-	-	-
WL-Solor	<i>one ě</i>	-	-	-
WL-Alorese (Baranusa)	<i>unu ng</i>	-	-	-
CL-Central Lembata	<i>una</i> ‘house’	-	<i>or</i>	-
CL-Atadei Painara	<i>una</i> ‘house’	-	<i>ora</i>	-
EL-Lewoeleng	-	-	-	<i>bolo</i>
KD-Leuwayang	<i>huna</i> ‘house’	<i>laléng</i>	-	-
KD-Leubatang	-	<i>laléng</i>	-	-
Reflexes in languages of Central Flores and Western Flores				
Keo	<i>'oné</i>	-	-	-
Ende	<i>oně</i>	-	-	-
Manggarai	<i>oné</i>	-	-	-
Komodo	-	<i>lalé</i>	-	-
Reflexes in Austronesian languages of Timor				
Amarasi	-	<i>nana-n</i>	-	-
Lole (Rote)	-	<i>dalé-k</i>	-	-
Helong-Funai	-	<i>dalé n</i>	-	-
Tetun	-	<i>lara n</i>	-	-
Mambae	-	<i>lala</i>	-	-

PMP=proto-Malayo-Polynesian; PCL=Proto-Central Lamaholot

In the languages of Flores-Lembata initial PMP *d regularly changes into *l* if there is another liquid in the word, which explains the initial *l* in the Kedang form *laléng* ‘inside’ going back to PMP *daləm (cf. §5.2.2). This means that, although the Flores-Lembata languages and the AN languages of Timor have changed their word order from LOC-N to N-LOC, these languages kept

Austronesian lexemes as locative nouns. In contrast to most languages discussed here, Central Lamaholot and Eastern Lamaholot did not retain Austronesian forms for ‘inside’. Central Lamaholot has innovated the form PCL *ora ‘inside’ which probably also meant ‘liver’ as this is an alternate meaning in CL-Central Lembata, while Eastern Lamaholot has innovated PEL *bolo ‘inside’. No clear cognates of these items could be found in other languages.⁵

Locative nouns in Flores-Lembata languages that denote ‘top’ can be grouped into two cognate sets given in Table 9.3. None of those two sets goes back to a PMP form, such as for example PMP *babaw ‘upper surface’. Both cognate sets appear to be innovations at the level of Proto-Flores-Lembata (PFL). No clear cognates have been found in other languages. Reflexes in square brackets are attested in the language for which they are listed but cannot be used to express the topological relation ‘on top of’. Reflexes marked with a question mark are not clearly cognate.

Table 9.3: Locative nouns for ‘top’ and their origins

PFL	*lolo ‘top’	*wutu ‘end; tip’
SK-Hewa	? [<i>lala</i> ‘upwards’]	<i>wutu</i> <i>n</i> ‘top; end’
SK-Krowe	-	<i>wutu</i> ‘top; end’
WL-Lewotobi	<i>lolō</i>	<i>wutū</i> ‘top’
WL-Lewoingu	<i>lolo</i> <i>n</i>	<i>wutu</i> <i>n</i> ‘top; last’
WL-Solor	<i>lolō</i>	<i>wutū</i> ‘top; end, tip, edge’
WL-Adonara	<i>lolō</i>	-
WL-Alorese (Baranusa)	<i>lulu</i> <i>ng</i>	-
WL-Alorese (Alor Kecil)	<i>lolo</i> <i>ng</i>	-
WL-Alorese (Alor Besar)	<i>lolo</i> <i>ng</i>	[<i>futu</i> <i>ng</i> ‘end’]
CL-Central Lembata	<i>lolo</i>	[<i>wutu</i> ‘end’]
CL-Lewokukung	<i>lolo</i>	-
EL-Lewoeleng	<i>lol’o</i>	-
KD-Leuwayang	<i>lolo’</i>	[<i>wutu</i> ‘tip’]
KD-Leubatang	-	? ‘ <i>oté</i> ‘top’

⁵ Several languages of western Timor have similar forms meaning ‘hole’ which could be cognate with Lewoeleng *bolo* ‘inside’. Examples include Termanu (Rote) *bolo-k* ‘hole, cavity, hollow, pit’ and Helong *bolo* ‘hole’. However the semantic relation of ‘hole’ and ‘inside’ is not straightforward.

PFL *lolo ‘top’ could be related to #blolo/golo ‘tall’ which has reflexes in Lamaholot and Kedang and also cognates in the Alor-Pantar languages, such as Blagar *blolu* ‘tall’ and Reta *balolu* ‘tall’. The putative cognate *lala* ‘upwards’ in SK-Hewa is not clearly related. The vowel change appears to be irregular and also the function and semantics are slightly different. The Hewa word *lala* ‘upwards’ is a preposition that points to a place that is located at a higher place and it does not function as a locative noun following the ground noun.

PFL *wutu ‘end; tip’ has reflexes throughout all Flores-Lembata languages but it can only be used as a locative noun in Sika and some Western Lamaholot varieties. In CL-Central Lembata for instance, the semantics of *wutu* are restricted to an actual end of something as in (37) and cannot be extended to the meaning ‘surface’ or ‘top’. The same restriction holds for WL-Alorese and KD-Leuwayang. The use of *wutu* ‘end’ as a locative noun appears to be innovated in Sika and some WL varieties.

- (37) CL-Central Lembata
 wo [*kaju wutu*]
 DIST wood end
 ‘at the end of the piece of wood’ (S2:24)

For the expression of the topological relation ‘under; below’ in form of a locative noun, three cognate sets can be established for the Sika and Lamaholot varieties. These sets are given in Table 9.4. For Kedang, there is no information available on how this relation is expressed.

Table 9.4: Locative nouns for ‘under’ and their origins

PMP	-	*burit ‘back’	-	-
LH	-	-	-	#langu
PWL	-	-	*kəwəl	-
SK-Hewa	<i>lewu n</i>	-	-	-
WL-Lewotobi	-	<i>wuĩ</i>	-	-
WL-Lewoingu	-	-	<i>wewel en</i>	-
WL-Solor	-	-	<i>kenewel ẽ</i>	<i>langũ</i>
WL-Alorese	-	-	-	<i>lang</i>
CL-Central Lembata	-	-	-	<i>léng</i>
EL-Lewoeleng	-	-	-	<i>langu</i>

Only one of the cognates sets goes possibly back to a PMP form, namely PMP *burit ‘hind part, rear, back’. The initial *w* is a regular reflex of PMP *b, and PMP *r appears to have merged with PMP *R which becomes glottal stop or is deleted in WL. Final consonants are frequently dropped in Western Lamaholot (cf. §5.2). The other sets must be regarded as innovations at lower levels, as at the current stage no clear cognates in other languages could be found. The forms most formally similar to #langu are Timor-Alor-Pantar forms, such as *alolang* ‘below’ in Blagar-Pura and *galolang* ‘below’ in Reta. The WL-Solor form *kenewel* could be a derivation of *kewel* using the nominalising infix *-n-*. WL-Lewoingu *wewel* could be a partial reduplication of *wel*, a short form of *kewel*. However, *kewel* or *wel* are known to be found as independent forms in Lewoingu. Therefore this analysis has to be kept tentative.

9.3.5 Suffixes on Flores-Lembata locative nouns

In the previous sections, I showed that in Sika, Western Lamaholot, but not in Central Lamaholot, Eastern Lamaholot and possibly Kedang, a fossilised nasal suffix is found on the topological nouns. In the case of Kedang, it remains unclear whether there are fossilised suffixes on locative nouns or not because there is only little data on locative nouns in this language. Table 9.5 provides an overview of the nasal suffixes and the varieties that have them.

Table 9.5: Suffixes on locative nouns

PAN	*ni
PFL	*-n
SK-Krowe	-ng
SK-Hewa	-n
WL-Lewotobi	-' / - $\tilde{V}$
WL-Lewoingu	-'on / -n / -en
WL-Solor	-' $\tilde{e}$ / - $\tilde{e}$ / - $\tilde{V}$
WL-Alorese	-ng

Sika and WL-Alorese display much more regularity in the shape of the suffix than Western Lamaholot. The variation between SK-Krowe and SK-Hewa is

a dialectal difference between the southwestern tana 'Ai varieties of Sika that merge final *n* and *ng* to *n* and those further west that do not (Lewis and Grimes 1995:603). In Western Lamaholot, the realisation of the suffix is mainly nasal, in some cases, a glottal stop <'> is combined with the nasal or found on its own. Nasals are either alveolar, velar or reduced to a nasality feature of the final vowel.

In many Austronesian languages a nasal suffix is associated with genitive or possessive marking. PAN **ni* has been reconstructed as a 'marker of possession and part-whole relationships'. In several languages this has developed into a suffix (Blust and Trussel 2010). As Sika and Western Lamaholot show traces of a suffix going back to PAN **ni*, leads to the conclusion that a reflex has been present in Proto-Flores-Lembata and that Central Lamaholot, Eastern Lamaholot and Kedang have lost this suffix on locative nouns. In possessive constructions, a reflex is still found in Central Lamaholot, such as CL-Central Lembata *-n(u)* '3SG.POSS' (cf. §9.2.3.1). For Kedang, a reflex *-n* is still found on adjectives in attributive position (cf. §8.3). Also taking the Kedang reflex *-n* into account, makes it likely to reconstruct PFL **-n* for the possessive suffix found on locative nouns. In the western varieties of Sika and in WL-Alorese, PFL final **-n* appears to have changed into *ŋ*.

9.3.6 The emergence of N-LOC

The distribution pattern of LOC-N and N-LOC over the languages of Nusa Tenggara Timur and Timor-Leste shows exactly the same pattern as the distribution of N-POSS and POSS-N discussed in §9.2. This is no coincidence, as in most languages with N-LOC, the locational constructions are structurally possessive constructions. The locative noun is formally a possessed noun and is marked as such by a possessive affix. Thus, the non-Austronesian word order of POSS-N is also found in the locational construction. This has been stated explicitly for Lamaholot and Alorese (Klamer 2012b:82-83; Klamer to appear).

The fact that locative constructions are expressed as possessive constructions alone is not a non-Austronesian feature. This feature is also found in Austronesian languages in western Indonesia. Also cross-linguistically, it is very common that possessive constructions with a nominal possessor are employed to cover a wide range of meanings, which are not all strictly possessive. The most common semantics expressed by possessive constructions

(after ownership) are kinship relations, part-whole relations and other kinds of abstract relations (Shopen 2007:178). As described in this section, the languages of Flores-Lembata and of Timor also use (inalienable) possessive constructions to express locational constructions. The Flores-Lembata languages extend the possessive construction even further, namely to attributive property constructions (cf. §8.3).

Nevertheless, there are more parallels than the word order within the LOC-N languages of the area on the one hand and the N-LOC languages on the other hand. The locative nouns of the LOC-N languages are not marked morphologically for possession. In contrast, locative nouns in the languages with N-LOC — both Austronesian and non-Austronesian — usually are possessed nouns that are marked morphologically for possession. Although possessed locative nouns are not a non-Austronesian feature in general, they might indeed be a non-Austronesian feature of this region. Thus, the re-analysis of the locative nouns as possessed nouns in the Austronesian languages of this area could have been reinforced due to contact.

Possibly the re-analysis of locative nouns as possessed nouns preceded the word order change in the possessive construction. Considered to be possessed nouns, locative nouns would automatically be affected by any change in the possessive construction. Therefore, assuming this order of changes, it becomes evident that the change of N-POSS to POSS-N automatically also caused the change of LOC-N to N-LOC. The fact that POSS-N and N-LOC is attested in exactly the same sample of Austronesian languages supports the hypothesis that the two developments are connected. For Flores-Lembata, both features were probably innovated at the level of Proto-Flores-Lembata because they are found throughout all FL languages but not in their closest relatives to the west, i. e. the languages of Central Flores, Western Flores and Sumba. This innovation in PFL is a structural innovation of word order but there is no shared innovation in forms. Some of the locative nouns used in the Flores-Lembata languages are retentions from PMP, such as words going back to PMP *qunəj ‘soft core; pith’ used to express ‘inside’, but others are innovations in the subgroups, such as for example PCL *ora ‘inside’ or the word *lolo* ‘top’ which is found in WL, EL, CL and KD but not in Sika. Without a Sika reflex, it remains unclear whether this item can be reconstructed to PFL or it spread through the languages after they had already split up into subgroups.

9.4 Post-nominal numerals

9.4.1 Overview

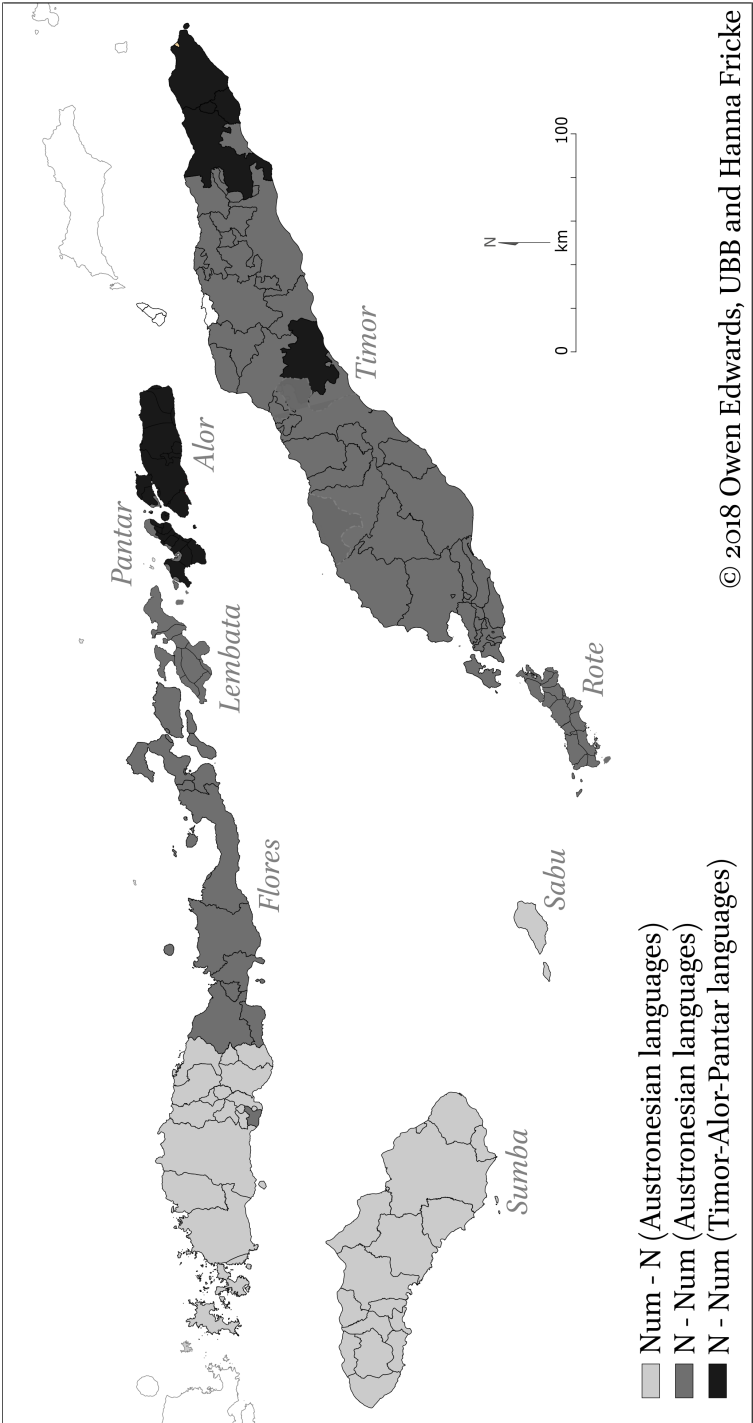
The modern-day Austronesian languages show three different patterns: (i) Noun-Numeral (N-NUM) as the dominant order, (ii) Numeral-Noun (NUM-N) as the dominant order or (iii) no dominant order (Dryer and Haspelmath 2013). However, the distribution of these patterns shows a clear tendency for NUM-N in western Indonesia, Malaysia, the Philippines, Taiwan and N-NUM in eastern Indonesia and Papua-New-Guinea. Thus, a dominant N-NUM word order is more frequent in areas where non-Austronesian languages are also spoken. It has been suggested that a consistent pattern of N-NUM in Austronesian languages could have been caused by contact with non-Austronesian languages (Reesink and Dunn 2018:934). N-NUM is the most frequent word order found in non-Austronesian languages of the area of eastern Indonesia and Papua-New-Guinea (Reesink and Dunn 2018:930).

Reesink and Dunn (2018:934) have reconstructed both orders of noun and attributive numeral as possible ancestral states for Proto-Austronesian (PAN) and Proto-Malayo-Polynesian (PMP). The fact that both ancestral states can be reconstructed to PAN and PMP suggests that a N-NUM pattern could have been inherited from an Austronesian ancestor as a minor pattern. Due to contact with non-Austronesian languages this pattern became the exclusive order in languages that were in contact with non-Austronesian speakers.⁶

Looking closely at the languages of Nusa Tenggara Timur and Timor-Leste, precisely again the division of west and east relating to NUM-N and N-NUM becomes visible, as displayed on the map in Figure 9.3. Western Flores, Sumba and Sabu have NUM-N word order following the inherited Austronesian pattern. In contrast, the Austronesian languages of Timor, the Flores-Lembata languages, as well as the Central Flores languages have N-NUM following the typical non-Austronesian pattern, also found in the Timor-Alor-Pantar languages.

⁶ It has to be noted that there are also a few examples of Austronesian languages with a dominant N-NUM order that are located outside of the contact area of eastern Indonesia. Dryer and Haspelmath (2013) lists the language Ma'anyan on Borneo and the language Paiwan on Taiwan as having N-NUM as their dominant order. There is no evidence to explain these cases of word order by contact.

Figure 9.3: Order of Numeral and Noun



This section is structured as follows. In §9.4.2, I describe the NUM-N order found on Sumba and in Western Flores to show that these languages, which are among the closest relatives of the Flores-Lembata languages, retained the typical Austronesian order in the numeral phrase. In §9.4.3, I show that differently to the languages in the western part of the area, the languages of Central Flores, Flores-Lembata, the AN languages of Timor have N-NUM order and that the same order is also found in the non-Austronesian Timor-Alor-Pantar languages. §9.4.4 is an overview of the etymology of numerals in the Flores-Lembata languages. In contrast to the non-Austronesian characteristic found in the order of the numeral phrase, the numeral lexemes themselves are overwhelmingly of Austronesian origin. Nevertheless, traces of non-decimal counting systems are attested which is different to the inherited Austronesian decimal system (Blust 2013:278). In §9.4.5, I discuss the emergence of the non-Austronesian N-NUM order in the Austronesian languages of the region, with a focus on the Flores-Lembata languages.

9.4.2 NUM-N

In the languages of Sumba, Sabu and Western Flores, numerals are pre-nominal, thus follow the typical Austronesian pattern. Frequently, numerals are followed by a classifier. This classifier can be obligatory or optional depending on the language. In some languages, such as Hawu and Manggarai, the combination of classifier and numeral can also be placed post-nominally. However, this is a minor pattern, possibly showing topicalisation of nouns.

In Kambera, numerals are pre-nominal. In a noun phrase, a numeral is always combined with a classifier that follows the numeral. Kambera has five different classifiers that are used for different semantic classes of nouns, such as humans, animals, long objects or flat objects (Klamer 1998:139). An example is given in (38), where the classifier *ngiu* is placed in between the numeral *dua* ‘two’ and the noun *kamambi* ‘goat’.

- (38) Kambera
dua ngiu kamambi
 two CLF goat
 ‘two goats’ (Klamer 1998:139)

In the Western Flores language Manggarai, numerals are also pre-nominal, as shown in (39a) and (39b). In Manggarai, there are two main classifiers

that can optionally be used between the numeral and the noun. These are *mongko* for things and animals, and *ngata/tau* for humans (Semiun 1993:13-14). The classifier *ngata* is used for one person, the classifier *tau* is for more than one as in (39b).

- (39) Manggarai
- a. *telu gelas*
three glass
'three glasses' (Semiun 1993:6)
 - b. *sua-tau mensia*
two-CLF human
'two people' (Original translation: 'two persons') (Semiun 1993:43)

In Hawu on Sabu (40), numerals are also pre-nominal. Hawu has several different classifiers used according to the shape and other semantic properties of the nominal referent (Walker 1982:18-20).

- (40) Hawu
- héo bela bafu*
nine CLF blouse
'nine blouses' (Walker 1982:19)

In Manggarai (41a) and in Hawu (41b), a minor pattern of postposed numerals with classifier following the numeral is attested.

- (41) a. Manggarai
- tanah a-lingko*
land one-CLF
'an area of land' (Semiun 1993:43)
 - b. Hawu
 - tudi due bengu*
knife two CLF
'two knives' (Walker 1982:19)

The word order of N-NUM-CLF in the examples (41a) and (41b) indicates that the classifiers in languages like Hawu and Manggarai are best analysed as

nouns and the other nouns, here 'land' and 'knife', are pre-posed or topicalised. Under this analysis, the word order of NUM-N remains, as the numeral is preceding the nominal classifier.

In sum, the languages of Western Flores, Sumba and Sabu all have NUM-N word order. Some of these languages have the additional option of pre-posing the noun, while the classifier remains after the numeral, which leads to a N-NUM-CLF order.

9.4.3 N-NUM

9.4.3.1 The Central Flores languages

In the Central Flores language Keo (42), numerals follow the noun and an obligatory classifier is placed in between the noun and the numeral. There is a range of different classifiers available for count nouns (Baird 2002:145-147).

- (42) Keo
sapa di'é tedu
 canoe CLF three
 'three canoes' (Baird 2002:145)

This CLF-NUM combination following the noun (N-CLF-NUM) is different to the topicalised noun followed by a numeral and a classifier (N-NUM-CLF) in languages such as Hawu and Manggarai, discussed in §9.4.2 above. In the languages of Central Flores, the classifier is in between the noun and the following numeral. Analysing classifiers as nouns makes the numerals in these languages entirely post-nominal, thus following the N-NUM pattern. In the languages Hawu and Manggarai, the noun is preposed and the order of NUM-N is retained by the numeral and the classifier which is originally a noun.

9.4.3.2 The Flores-Lembata languages

Throughout all Flores-Lembata languages, numerals are post-nominal. If a classifier is employed, it is placed between the noun and the numeral following the template N-CLF-NUM. With the numeral 'one', in most cases, no classifiers are used. For higher numbers, there are classifiers. However, there

is a lot of variation found in the number of different classifiers and in the degree of optionality of classifiers among the languages of Flores-Lembata.

In Sika, there are a couple of different obligatory classifiers for referents of different shapes and sizes (Fricke et al. 2019). In example (43) below, the classifier *teman* is used in between a compound noun and the numeral *hutu* ‘four’. The classifier *teman* is widely used for any round or bulky object of middle to big size.

- (43) SK-Hewa
Ia pau wua-n teman hutu.
 DIST mango fruit-POSS CLF four
 ‘There are four mangoes.’ (Fricke 2014a:40)

Western Lamaholot possesses several numeral classifiers that are used for measurements (Kroon 2016:152). They are mostly used for fruits, parts of plants, basically to make uncountable nouns countable, such as *lépa* ‘CLF.box’ in (44) which probably means something like ‘container’ and is used to make the uncountable noun *tahan* ‘rice’ countable.

- (44) WL-Lewoingu
tahan lépa to’u
 rice CLF.box one
 ‘a box/container or rice’ (Nishiyama and Kelen 2007:29)

However, for many WL nouns, no classifier is needed, such as in example (45) from WL-Lewoingu, where the numeral *telo* ‘three’ follows directly the noun *hépe* ‘knife’. Also in WL-Alorese, we find a similar pattern. Nouns and numerals are merely juxtaposed, such as in (46).

- (45) WL-Lewoingu
hépe telo
 knife three
 ‘three knives’ (Nishiyama and Kelen 2007:44)
- (46) WL-Alorese
ni ning aho rua
 3SG POSS dog two
 ‘his two dogs’ (Klamer 2011:45)

In CL-Central Lembata, there is only one general classifier *ua* 'CLF' which is obligatorily used with any numeral above one. An example is in (47) where the noun *lamé* 'man' is followed by the general classifier *ua* and the numeral *jua* 'two'.

- (47) CL-Central Lembata
lamé ua jua
 man CLF two
 'two men' (S2:1)

In Kedang, there are two optional classifiers that can accompany the numeral. The general classifier *urén* (48), and the more specific classifier *ulu* 'seed, kernel, stone' for round objects (Samely 1991a:96).

- (48) KD-Leuwayang
wéta' urén sué
 house CLF two
 'two houses' (Samely and Barnes 2013:662)

In sum, all Flores-Lembata languages all have post-nominal numerals. In Sika and Central Lamaholot, numerals higher than the numeral for 'one' are obligatorily accompanied by a classifier. In Kedang, there is the optional general classifier *urén*. Finally, in Western Lamaholot and Alorese, no general classifier exists but a couple of specific classifiers that are mainly used to make otherwise uncountable nouns countable.

9.4.3.3 The AN languages of Timor

The Austronesian languages of Timor also have post-nominal numerals and optional classifiers, such as in Tetun (49a) and South Mambae (49b), or no classifiers, such as in Helong (49c) and Amarasi (49d).

- (49) a. Tetun
ata nia-k rua né'é
 slave 3SG-POSS two this
 'these two slaves of her' (van Klinken 1999:133)
- b. South Mambae
haru nora ruu
 shirt CLF two
 'two shirts' (Fogaça 2017:154)

c. Helong

Kaim sos asu tilu.

1PL.EXCL buy cow three

'We bought three cows.'

(Balle 2017:15)

d. Amarasi

n-fee naan toon teun

3-give DIST year three

'That one has been given three years.'

(Edwards 2016a:265)

9.4.3.4 Non-AN Timor-Alor-Pantar languages

In the non-AN Timor-Alor-Pantar (TAP) languages numerals are also post-nominal. Classifiers between noun and numeral are common in most of the languages of Alor and Pantar but cannot be reconstructed to Proto-Alor-Pantar (Klamer 2014a:150-152; Klamer 2014c). In the Western Pantar noun phrase in (50) the optional classifier *haila* 'CLF:area' is placed in between the noun *bla* 'house' and the numeral *ye* 'one'.

(50) Western Pantar

bla (haila) yé

house CLF:area one

'one house'

(Holton 2014:56)

As in this example, the use of classifiers in the Alor-Pantar languages is not necessarily obligatory and there may be specific restrictions, such as in Kaera, where classifiers are not used with animals (Klamer 2014b:123) or in Adang, where numerals higher than one are mainly used with classifiers (Robinson and Haan 2014:249). In Wersing, classifiers are even not found at all (Schapper and Hendery 2014:466). Klamer (2014a) suggested that the classifiers in the Alor-Pantar languages developed through contact with Austronesian languages of the region (Klamer 2014a:159; Klamer 2014c).

In the TAP languages of Timor, classifiers are present but used less extensively. In Makasae for example, classifiers are mainly used for human or animal referents (51) and there is a tendency for classifiers to be optional with the numeral 'one'.

- (51) Makasae
Gi bai boku u suri.
 3SG pig CLF one shoot
 ‘He shot a pig.’ (Huber 2008:23)

In Makalero, classifiers are mainly associated with counting and there is no suitable classifier for all nouns (Huber 2011:120-121).

9.4.3.5 Summary

The Austronesian languages of Central Flores, Flores-Lembata and Timor consistently postpose the numeral in relation to the noun that is counted. Classifiers are obligatory, optional or largely absent depending on the individual language. There is a tendency for the AN languages surveyed here that the further east the less important the use of classifiers appears to be. In the Central Flores language Keo, classifiers are obligatory, in the Flores-Lembata languages a mixed picture emerges, while in the languages of Timor classifiers are optional or lost.

The non-Austronesian Timor-Alor-Pantar languages have the same word order in the numeral phrase as the AN languages discussed in this section, thus postposed numerals in relation to the noun (N-NUM). Classifiers in between noun and numeral are a common but non-inherited feature in the languages of Alor-Pantar. In the TAP languages of Timor classifiers are also attested but used less.

9.4.4 Etymology of numerals in Flores-Lembata

Flores-Lembata basic numerals are etymologically mainly of Austronesian origin and follow a decimal counting system inherited from Proto-Austronesian (Blust 2013:278). However, a few exceptions are found in the numerals for ‘one’, ‘four’, ‘eight’ and ‘ten’. Table 9.6 gives an overview of Flores-Lembata basic numerals from one to ten in comparison to PMP forms. There is no data on Eastern Lamaholot numerals. Numerals that are not or not straightforwardly inherited from PMP are marked in bold and are discussed below.

Table 9.6: Basic numerals in the Flores-Lembata languages

	PMP	PFL	Sika	WL (AL)	WL (LWI)	CL (KK)	Kedang
‘one’	*əsa	*ha	<i>ha</i>	<i>tou</i>	<i>to’u</i>	<i>tu(né)</i>	<i>‘udé’</i>
‘two’	*duha	*dʒua	<i>rua</i>	<i>rua</i>	<i>rua</i>	<i>jua</i>	<i>suwé</i>
‘three’	*təlu	*təlu	<i>telu</i>	<i>talo</i>	<i>telo</i>	<i>telu</i>	<i>télu</i>
‘four’	*əpat	*əpat	<i>hutu</i>	<i>paa</i>	<i>paa</i>	<i>paat</i>	<i>apa’</i>
‘five’	*lima	*lima	<i>lima</i>	<i>léma</i>	<i>léma</i>	<i>léém</i>	<i>lémé</i>
‘six’	*ənəm	*ənəm	<i>ena</i>	<i>namu</i>	<i>nemung</i>	<i>enem</i>	<i>énéj</i>
‘seven’	*pitu	*pitu	<i>pitu</i>	<i>pito</i>	<i>pito</i>	<i>pito</i>	<i>pitu</i>
‘eight’	*walu	*valu	<i>walu</i>	<i>buto</i>	<i>buto</i>	<i>buto</i>	<i>butu rai</i>
‘nine’	*siwa	*siva	<i>hiwa</i>	<i>hiwa</i>	<i>hiwa</i>	<i>siwa</i>	<i>lémé apa’</i>
‘ten’	*sa-ŋa-puluq	*spulu	<i>pulu</i>	<i>kertou</i>	<i>pulo</i>	<i>spulo</i>	<i>pulu</i>

AL=Alorese, LWI=Lewoingu, KK=Kalikasa (Central Lembata)

For the numeral ‘one’, Sika *ha* is a clear reflex of PMP *əsa ‘one’ as PMP *s > *h* is regular in Sika. All other Flores-Lembata languages have innovated the form for the numeral ‘one’. The Lamaholot forms *tou*, *to’u* and *tuné* are all cognate. The second syllable of the CL word *tuné* is an insertion. The Kedang form *‘udé’* ‘one’ is probably not cognate with the other forms for ‘one’. The origins of these numerals for ‘one’ in Lamaholot and Kedang remain unknown. There is a possible connection with the Timor-Alor-Pantar languages, as the vowel *u* also appears in the TAP numerals for ‘one’. PAP *nuk ‘one’ (Schapper and Klamer 2017:311), Makalero *u(n)* ‘one’ and Makalero *tu* ‘first’ may point to borrowing from TAP languages. However, as the words are very short, the evidence appears to be too thin and difficult to prove.

The Kedang numeral *suwé* ‘two’ is not straightforwardly cognate with the other forms for ‘two’ which are reflexes of PMP *duha ‘two’, as a change of PMP *d- > *s* is not regularly attested. However, it is likely that for this form an irregular change of PMP *d- > PFL *dʒ already occurred at the level of PFL, with subsequent *dʒ > *s* in Kedang. Based on the Sika and Lamaholot forms for ‘two’, PFL *dʒua ‘two’ has to be reconstructed. The reconstruction of PFL *dua would not account for the present-day Lamaholot forms with initial *r* and *j* (cf. §5.2.2.6, §5.2.2.1). A change of final *a > *e* also occurs in other Kedang words, such as PMP *ina > Kedang *iné* ‘mother’. The *w* in *suwé* ‘two’ could be an insertion in between the two vowels. Thus, Kedang *suwé*

‘two’ is most likely cognate with the other forms.

The Sika numeral *hutu* ‘four’ is most likely a borrowing from neighbouring Lio *sutu* ‘four’, going back to Proto-Central-Flores (PCF) *wutu ‘four’ (Elias 2018:120). The change of *s > h is regular in Sika (cf. §5.2.3.2).

The Kedang numeral *lémé apa* ‘nine’, composed out of *lémé* ‘five’ and *apa* ‘four’, shows a quinary system, in exactly the same way as found in the languages of Alor-Pantar which also combine numerals for ‘five’ and ‘four’ to create a numeral for ‘nine’ (Schapper and Klamer 2017:288). Thus, Kedang probably borrowed the quinary system for its numeral ‘nine’ from Alor-Pantar languages. Schapper and Klamer (2017:307) say that the borrowing of the quinary system could very well be related to cultural contact between the speakers of Kedang and Alor-Pantar languages. However, as discussed further below, the Kedang numeral for ‘nine’ could also be a remnant of an older non-decimal system.

The first part *ker* ‘ten’ of the Alorese numeral *kertou* ‘ten’ is clearly a borrowing from Alor-Pantar languages. For PAP *qar ‘ten’ has been reconstructed (Schapper and Klamer 2017:311). The second part *tou* ‘one’ of the numeral ‘ten’ is a retention from Proto-Western Lamaholot (PWL) *toʔu ‘one’. Since Alorese speakers settled on Pantar and Alor until today, there has been intense contact with speakers of Alor-Pantar languages (Moro 2018), therefore borrowing is very likely to occur.

The Kedang form *butu rai* ‘eight’ probably points to a now obsolete non-decimal counting system, as had been suggested by Schapper and Hammarström (2013:428) and Schapper and Klamer (2017:306). The numeral is composed of two words, *butu* and *rai* ‘many’. The form *butu* standing alone does not have any meaning in the modern language, apart from being used as an abbreviation for *butu rai* ‘eight’ (Samely and Barnes 2013:135,557).⁷ As *rai* means ‘many’, *butu* could originally have meant ‘four’ in Kedang leading to the literal gloss ‘four many’ *butu rai* for the numeral ‘eight’. Although nowadays, the putative original *butu ‘four’ has been replaced by *apa?* ‘four’, a numeral of Austronesian origin. Kedang’s former numeral *butu* ‘four’ probably ultimately comes from the same source as Proto-Central Flores (PCF) *wutu ‘four’. The change of *b > w is regular in Central Flores (Elias 2018:90).

⁷ Schapper and Hammarström (2013:428) argue that *rai* means ‘two’, similar to the numerals in the Central Flores languages where the numeral ‘eight’ is composed of ‘two’ and ‘four’. However, as Kedang has *suwé* ‘two’ and *rai* ‘many’ this interpretation is unlikely.

The source could be PCEMP **butu* ‘group, crowd, flock, school, bunch, cluster’ which has been reconstructed to Proto-Central-Eastern-Malayo-Polynesian (PCEMP) by Blust and Trussel (2010). In Kedang, this original meaning is still found in the fixed expression *butu bongan* ‘sit and visit’ (Samely and Barnes 2013:135). Here the word *butu* most likely retains a meaning more closely to the reconstructed PCEMP **butu* ‘crowd, group, bunch’. The second word *bongan* in the expression *butu bongan* ‘sit and visit’ also means ‘visit, make a social call, sit and visit’ (Samely and Barnes 2013:120). From this evidence, I propose that Proto-Kedang (PK) **butu* meant ‘group; bunch’ and also ‘four’. Also Proto-Alor-Pantar (PAP) **buta* ‘four’ could be connected with PCEMP **butu* ‘group, crowd, flock, school, bunch, cluster’.

The Lamaholot numeral *buto* ‘eight’ also remains unexplained when assuming a purely decimal system because it is etymologically related to numerals meaning ‘four’ in other languages, such as the Central Flores languages. If assuming that the Lamaholot counting system was also a mix of decimal and non-decimal, as I suggested for Kedang, and *buto* indeed originally meant ‘four’, then *buto* ‘eight’ in Lamaholot could be explained by originally having a second element similar to Kedang *butu rai* ‘eight’. The putative second part of the Lamaholot numeral ‘eight’, such as *rai* ‘many’ in Kedang, became unnecessary because the numeral ‘four’ was replaced by *paat* (< PMP **əpat* ‘four’) and the mixed system was replaced by a decimal system which led to the interpretation of *buto* alone as ‘eight’. As mentioned earlier, also in Kedang the numeral *butu rai* ‘eight’ can be abbreviated to *butu* ‘eight’.

In sum, there was possibly a non-decimal system of ‘one, two, three’ and the word ‘bunch, group’ for any higher numeral in Kedang, Lamaholot and maybe also in the Alor-Pantar languages. Later the word for ‘bunch, group’ gained a more specific meaning, namely ‘four’, and by multiplication, this led to the numeral ‘eight’. It remains unknown whether there were more non-decimal numerals in this system because no traces are left. Nowadays in Kedang and Lamaholot, a decimal system with mainly Austronesian lexemes is the norm. But it is possible that there was a time where both systems were used. Such a situation is still found today in the AN language Mambae spoken on Timor. There are varieties of Mambae that use a decimal counting system, while others use a non-decimal system. The wordlists in Fogaça (2017:243-245) show that Central Mambae has a purely decimal system, while Northwest Mambae and South Mambae have a non-decimal sys-

tem for the numerals six to ten.

For Central Flores, it has been argued that the non-decimal system in the Austronesian languages of that area is a feature of non-Austronesian origin (Schapper and Klamer 2017:315). The remnants of non-decimal systems in the numerals ‘eight’ in Lamaholot and Kedang and also ‘nine’ for Kedang, could also be attributed to non-Austronesian contact. Or more precisely, the numeral ‘eight’ points to a pre-Austronesian non-decimal system that was probably present in these languages in addition to the new decimal system introduced from an Austronesian ancestor. Possibly both systems were competing at a certain point in time and finally the decimal system won, while the non-decimal system was lost, being kept in Kedang a bit longer than in Lamaholot. Meanwhile, the numeral ‘ten’ in Alorese points to more recent borrowing from Alor-Pantar languages. The numeral ‘nine’ in Kedang could also be a recent calque from AP languages, or it could be a retention of an older non-decimal system which was reinforced in Kedang (but not other FL languages) by contact with AP.

9.4.5 The rise of post-nominal numerals

In the previous sections, I have shown that the languages of Central Flores, of Flores-Lembata and the Austronesian languages of Timor underwent a change in word order from NUM-N to N-NUM. The N-NUM order is considered an innovation because Austronesian languages further west, such as in western Flores and on Sumba, as well as Austronesian languages in general have an inherited NUM-N word order, as has been shown in §9.4. At the same time, the counting systems of several languages that underwent this syntactic change show traces of non-decimal counting systems, as I have shown in §9.4.4. This again is an innovation in these Austronesian languages as a decimal counting system can be reconstructed to Proto-Austronesian. These two innovations, in word order and in counting system, are not necessarily connected directly but they both could have been caused by contact situations with non-Austronesian languages.

The change in word order that is attested in Central Flores, Flores-Lembata and in the AN languages on Timor could have either happened once when an ancestor of those languages that display the N-NUM order today was in contact with speakers of non-Austronesian languages, or the feature could have diffused through the various Austronesian subgroups concerned.

The scenario of a single event of change in an undefined ancestor language becomes unlikely when considering the genealogical relations. The closest relatives of Central Flores are the western Flores languages (Fernandez 1996). These did not undergo the change from NUM-N to N-NUM, neither do they show traces of non-decimal counting systems. In §5.5, I provide evidence for subgrouping the Central and Western Flores languages with the languages of Flores-Lembata and of Sumba. Therefore, any scenario of shared innovation for NUM-N would have to explain why the languages of Western Flores and Sumba do not show this change. Therefore, I suggest that at least three separate contact scenarios led to the pattern that is attested in Nusa Tenggara Timur and Timor-Leste today. Proto-Central Flores, Proto-Flores-Lembata and the proto-languages of Timor all underwent the change from NUM-N to N-NUM independently.

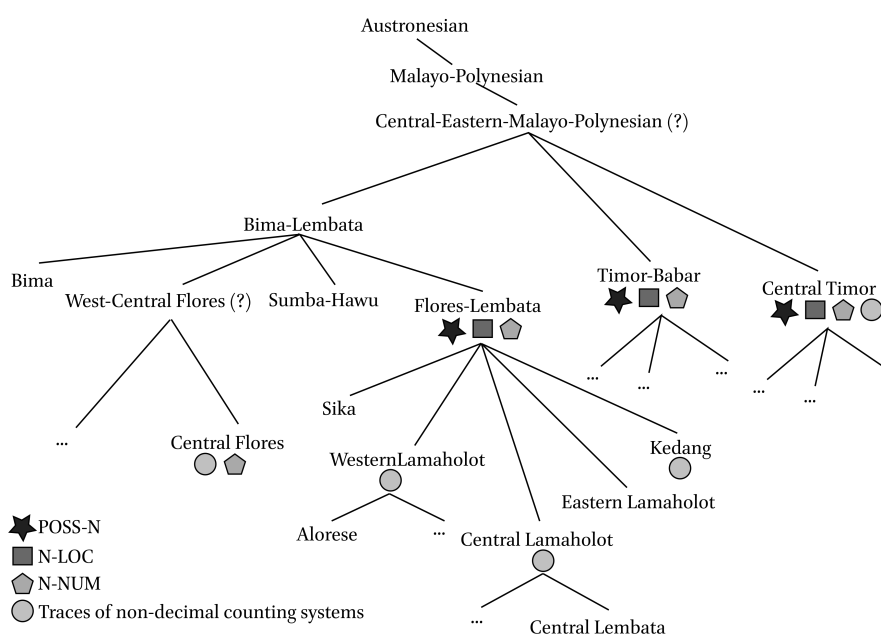
9.5 Conclusions

In this chapter, three non-Austronesian syntactic features of the noun phrase and the traces of non-decimal counting systems have been discussed. The syntactic features are found in a subset of the Austronesian and in all non-Austronesian languages of Nusa Tenggara Timur and Timor-Leste. In the Austronesian languages of this area these features are innovations, while in the non-Austronesian Timor-Alor-Pantar languages they are inherited. POSS-N, N-LOC and N-NUM are attested in the Flores-Lembata languages and in the AN languages of Timor. The feature of N-NUM spreads further including the Central Flores languages. I propose that these contact-induced word order changes took place in several proto-languages independently.

In Figure 9.4, the innovation of features is indicated with different symbols. The symbol is placed below the name of a subgroup in which all languages show the respective feature. Consequently, the feature can be reconstructed to the proto-language of the subgroup. The tree structure is based on the current knowledge on the languages investigated (cf. §1.3.1).

Proto-Flores-Lembata and the proto-languages of the AN languages of Timor each underwent independently the same two syntactic changes leading to N-NUM and POSS-N. In addition, as locative nouns were reanalysed as possessed nouns, they were moved to a post-nominal position, leading to N-LOC.

Figure 9.4: Innovations in the NP



Proto-Central Flores only switched the order of numeral and noun and gained N-NUM. I propose independent developments because, as long as there is no lexical or phonological evidence that could prove that Flores-Lembata and Timor form an innovation defined subgroup within the Austronesian languages of the area, there is no reason to argue for a single contact event. In contrast, there is evidence for a Bima-Lembata subgroup including PFL, Central Flores and the Austronesian languages further west on Flores and Sumba (cf. §5.5). Assuming word order changes on a higher level, such as Bima-Lembata or even higher, would demand an explanation why there are languages that did not inherit the new pattern.

Traces of non-decimal counting systems are an additional innovation found in the languages that underwent the word order change in the numeral phrase. These innovations are possibly also attributed to contact with non-Austronesian languages but there is not necessarily a causal connection with the change in word order. For Central Flores, the non-decimal counting system is found in all Central Flores languages. Therefore, it can be suggested that it emerged in Proto-Central Flores. In the Flores-Lembata

languages this is not the case. Only in Kedang, and possibly in Lamaholot, traces of non-decimal systems are found. Reconstruction of a non-decimal system to PFL is not straight-forward, though it is possible that decimal and non-decimal systems were in competition at this stage, leading to the mixed (though mostly decimal) systems seen in the present day languages.

CHAPTER 10

Innovations in the clause

10.1 Introduction

In this chapter, I discuss two clausal structural features of the Flores-Lembata languages that are considered atypical for Austronesian languages. These features are (i) clause-final deictic motion verbs (DMV) and (ii) clause-final negations (NEG). In Table 10.1, the Flores-Lembata languages and other languages of Nusa Tenggara Timur and Timor-Leste are evaluated for these features.

For each language, the features found in this language are marked with a plus (+) and the features not found in this language are marked with a minus (-). When a feature is only found incidentally or in some varieties of a language group, it is marked with minus/plus (-/+). In this table, Sumba also includes the island of Sabu, Flores refers to western and central Flores, and Timor also includes the island of Rote. The languages Sika (SK), Kedang (KD), Western Lamaholot (WL), Central Lamaholot (CL) and Eastern Lamaholot (EL) are subgroups of the Flores-Lembata language family (cf. §5.3). For Eastern Lamaholot (EL) no data is available on deictic motion verbs, hence it is marked with a question mark. The only non-Austronesian (non-AN) languages in the area are the Timor-Alor-Pantar languages (TAP).

Table 10.1: The spread of clause-final deictic motion verbs and negators

Feature	Sumba	Flores	AUSTRONESIAN					Timor	Non-AN TAP
			SK	KD	WL	CL	EL		
Fin. DMV	-	-	+	+	+	+	?	-/+	+
Fin. NEG	-	-	-/+	-	+	+	+	-/+	+

DMV=deictic motion verbs, NEG=negation

In the Austronesian languages, the two features are only attested in languages of the eastern part of the area studied in this dissertation, namely in the Flores-Lembata languages and in the Austronesian languages of Timor. In these Austronesian languages, final NEG and DMV arose out of contact with non-Austronesian languages, possibly with TAP languages as they also have these features and are the only non-Austronesian languages attested in immediate proximity.

Clause-final DMV are found in all Flores-Lembata languages, but not at all, or only occasionally, in their closest Austronesian neighbours on Timor and Flores. This suggests a shared innovation of clause-final DMV at the level of Proto-Flores-Lembata (PFL). In contrast, clause-final negation is found in all varieties of Lamaholot but not in Kedang (KD) and only incidentally in Sika (SK), namely in the variety of Hewa. This suggests innovation of clause-final NEG at the levels of the Lamaholot subgroups.

This chapter follows the methodology laid out in §7.2 to demonstrate that these two features are innovations that is likely to have arisen due to contact with non-Austronesian languages. In §10.2, I discuss the semantics and the development of clause-final deictic motion verbs. In §10.3, I discuss the development of clause-final negation in the Lamaholot varieties. In Section 10.4, I summarise the findings and draw conclusions.

Glosses and transcriptions from other sources are adapted to the conventions of this dissertation (cf. §7.3). A list of adapted glosses and re-transcribed sounds with their original representations is found in Appendix C.

10.2 Clause-final deictic motion verbs

10.2.1 Overview

Deictic motion verbs (DMV) are verbs denoting a movement into a direction, either away from the deictic centre ('go') or towards the deictic centre ('come'). The deictic center is most of the time the speaker. In most of the Flores-Lembata languages, DMV do not only encode deictic direction but also elevation, i. e. upward or downward direction. In §10.2, I discuss the clausal position of DMV, their semantics and their etymology in the languages of Flores-Lembata and the languages surrounding this language family.

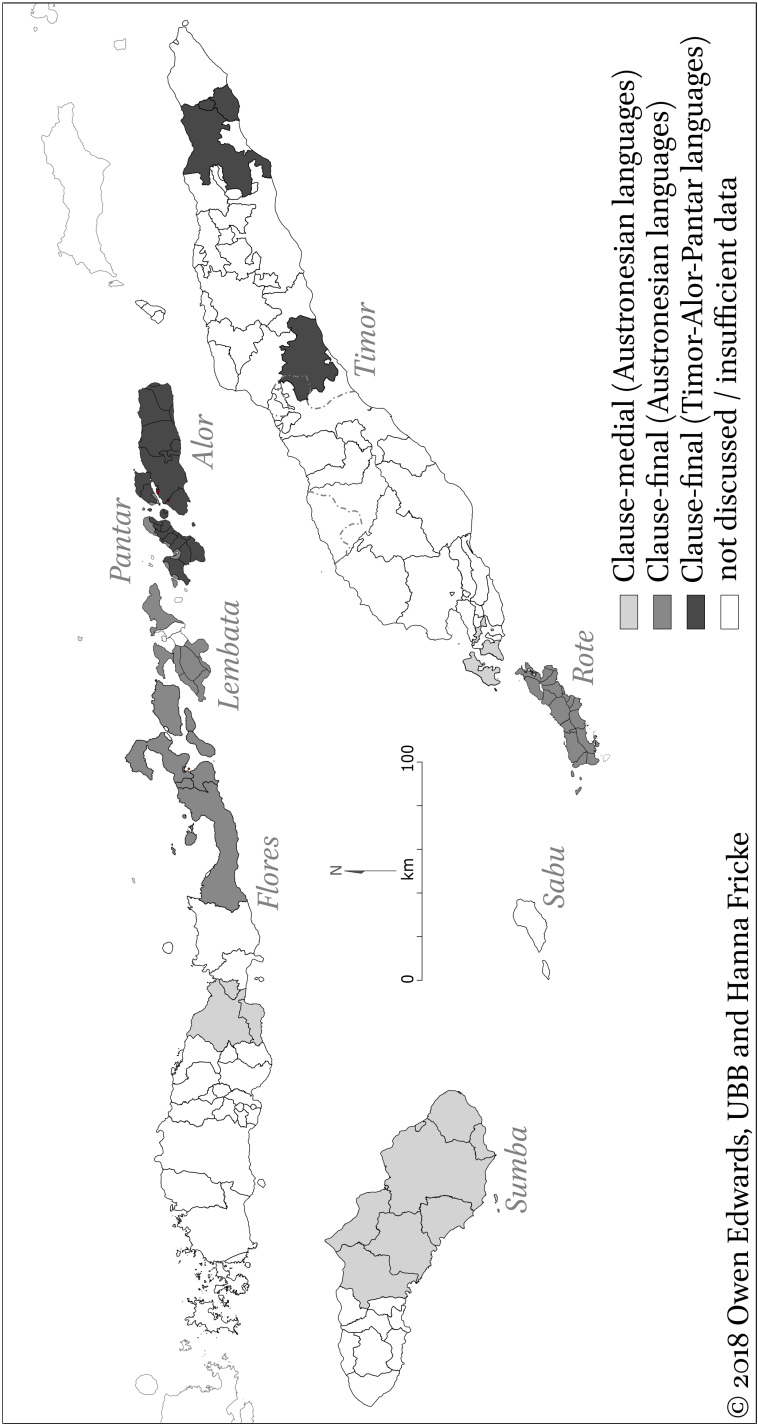
In Austronesian languages, the basic word order is predominantly verb-medial and also frequently verb-initial (Blust 2013:461). However, in contact areas of Austronesian and Papuan languages, there are cases of Austronesian languages with a verb-final word order (Blust 2013:470). These typological observations are supported by phylogenetic evidence from Reesink and Dunn (2018:950) who reconstructed verb-initial order as a very likely ancestral states for Proto-Malayo-Polynesian. Verb-final order could not be reconstructed for any node in the Austronesian language family (Reesink and Dunn 2018:950).

Generally, the Austronesian languages of Nusa Tenggara Timur and Timor-Leste show verb-medial word order (Klamer 2002:374) and thus are typical Austronesian in respect to this feature. However, there are Austronesian languages of the area where clauses with deictic motion verbs (DMV) have verb-final word order. I propose that these verbs changed their position due to influence from non-Austronesian languages.¹

The map in Figure 10.1 shows the distribution of medial and final DMV in the AN and non-AN languages of Nusa Tenggara Timur and Timor-Leste. Clause-final DMV are found in all Flores-Lembata languages and incidentally in the AN languages of Timor, namely in the Rote languages. Clause-final DMV are attested throughout all Timor-Alor-Pantar (TAP) languages. TAP languages generally have verb-final order (Klamer 2017:15). Therefore, in the TAP languages final DMV are not an exception.

¹ For WL-Lewotobi, Nagaya (2011:46) has suggested that deictic motion verbs appearing in SOV order are possibly a Papuan feature.

Figure 10.1: Position of deictic motion verbs



The following sections are structured as follows. In §10.2.2, I show that in the closest Austronesian relatives of Flores-Lembata, the AN languages of Flores and of Timor, DMV are clause-medial as any other verbs in these languages. These languages do not show any change in this feature in respect to the typical Austronesian word order pattern. The languages of Rote are the only exception, as clause-final DMV are attested here. In §10.2.3, I show that different to their closest relatives in the west, the Flores-Lembata languages all have clause-final DMV, while other verbs in these languages are in medial position of the clause. In this section, I also illustrate the patterns of clause-final DMV in the languages of Flores-Lembata. In §10.2.4, I present DMV in the Alor-Pantar languages, the non-Austronesian neighbours of Flores-Lembata to the east. I stress the similarities in the DMV system of Alor-Pantar and Flores-Lembata. In §10.2.5, I discuss the possible etymologies of the Flores-Lembata DMV. In §10.2.6, I propose contact-induced change for the emergence of clause-final DMV in Flores-Lembata due to contact with AP languages.

10.2.2 Medial DMV on Flores, Sumba and Timor

In the Austronesian languages of Central and Western Flores, deictic motion verbs (DMV) expressing ‘come’ and ‘go’ are found clause-medially, in the same position as other verbs in those languages. There are no clause-final verbs in these languages. A clause-medial DMV in a locational clause is exemplified below in (1) for Keo, a Central Flores language. The deictic verb *kai* ‘go’ may be cognate with the deictic verbs for ‘go’ found in Flores-Lembata languages reconstructed to LH-SK #-ai ‘go’ (cf. §10.2.3). However, in contrast to the majority pattern in Flores-Lembata languages, in Keo, this deictic verb meaning ‘go’ is clause-medial.

- (1) Keo
'Imu kai mbana pasa rédé So'a.
 3SG go walk market east NAME
 ‘She went to the market in So’a.’ (Baird 2002:301)

In the Austronesian languages on Sumba, such as Kambara, DMV are also clause-medial. An example is the verb *laku* ‘go’ in (2).

(2) Kambera

... *na-laku la Umalulu.*

3SG-go LOC Melolo

‘... he went to Melolo.’

(Klamer 1998:345)

In locational clauses from the Austronesian language Helong (3), DMV are also clause-medial, as here the verb *maa* ‘come’ (< PMP *ma(R)i ‘come’).

(3) Helong

Un pait maa lui la lo.

3SG return come boat DIST NEG

‘He did not return to the boat.’

(Jacob and Grimes 2011:355)

An exception to the languages just discussed are the Austronesian languages on the island of Rote. In these languages, clause-final DMV are found, as in example (4) from Lole.

(4) Lole

Boé ma Yohanis kalua némé uma dalé mai.

then NAME exit from house inside come

‘Then John came out of the house.’ (Jacob and Grimes 2011:359)

The structure of the Lole locational clause with a clause-final DMV is very similar to the structures found in locational clauses in the Flores-Lembata languages which are discussed in §10.2.3 below. However, as the languages of Rote are the only Austronesian languages of Timor known to have clause-final DMV and they are in other respects more closely related to the languages of Timor than to the Flores-Lembata languages, this is the result of an independent innovation in those languages.

In sum, virtually all the Austronesian languages of Sumba, Flores and Timor have medial DMV, and thus retain the typical Austronesian word order for their DMV.

10.2.3 Final DMV in the Flores-Lembata languages

10.2.3.1 Overview

In the Flores-Lembata languages, DMV in locational clauses appear in clause-medial or clause-final position, with clause-final position being the most

frequent choice. Other verbs always appear in clause-medial position. The clause-final DMV in the Flores-Lembata languages does not occur in the typical Austronesian position but appear to be moved to the end of the clause, the typical position for verbs in the non-Austronesian Timor-Alor-Pantar (TAP) languages.

Table 10.2 provides an overview of clause-final DMV in the Flores-Lembata languages. The use of these forms is illustrated in the subsequent sections 10.2.3.2, 10.2.3.3 and 10.2.3.4. A more detailed discussion on the etymology of the Flores-Lembata DMV is §10.2.5.

DMV do not only encode deictic direction but also elevation. There are three categories: *LEVEL* denoting a movement unspecified for elevation, *HIGH* denoting a movement towards or away from a higher place than the location of the deictic centre, and *LOW* denoting a movement towards or away from a lower place than the location of the deictic centre. The words in the category *LEVEL* often denote a movement from or towards a place with the same height as the position of the deictic centre, thus no change in elevation. Therefore, the category is labelled *LEVEL*. However, when combined with locationals meaning ‘upwards’ or ‘downwards’ in a clause, an upward or downward movement can be expressed. For this reason, I gloss the *LEVEL* DMV with plain ‘come’ and ‘go’, while DMV in the categories *HIGH* and *LOW* are glosses as ‘go.up’ / ‘come.up’ and ‘go.down’ / ‘come.down’ respectively.

For each elevation category, *LEVEL*, *HIGH* and *LOW*, there are usually two motion verbs, one which expresses the ‘go’ direction and one that expresses the ‘come’ direction. Stem-final consonants in Central Lamaholot, given in brackets in the table, are only realised when a suffix is added. Variants are separated by a slash. The words given for one concept are not always all cognates. The forms (most likely) not relating to the reconstructions provided above are given in square brackets. The shared Flores-Lembata lexemes LH-SK #-ai ‘go’ and LH-KD #gedi ‘go.up’ are not reconstructible to PFL at the current stage of knowledge because for #-ai no cognate is found in Kedang and for #gedi no cognate is found in Sika. For the concept ‘come.down’ no PFL reconstruction is possible with the data available.

Table 10.2: Deictic motion verbs in the Flores-Lembata languages

	LEVEL		HIGH		LOW	
	'go'	'come'	'go.up'	'come.up'	'go.down'	'come.down'
PMP	-	*ma(R)i	-	*sakay	-	-
PFL	#-ai	*mai	#gedi	*hakay	*lodon	?
Clause-final position						
SK	-a	mai	-	-	-	-
KD	-	ma	kéu/[dau]	ḁ'	do'/[bunu']	nè
CL	-ai	méné(k)	géji/géwi	aka(j)	lodo	nau(n)
WL (SL)	-ai	pai	géré/[dopa]	haka/[dai]	lodo	hau
WL (LTB)	-ai	dai	-	haka	-	hau
WL (AL)	-éi	méné	géré	-	lodo	-
Cognates in clause-medial position						
SK	-	-	-	-	lodong	-
WL (LTB)	-	-	géré	-	lodo	-
CL	-	-	-	-	[bunu]	-

LTB=Lewotobi, SL=Solor, AL=Alorese

In Sika and in WL-Lewotobi, the concept of 'return' is also expressed by clause-final verbs, namely Sika *balér* and WL-Lewotobi *gwali*. The concept of 'return' is a specific kind of 'come', as it means to come towards the deictic centre but with the additional component that the person who is coming had been at that place before. In addition CL-Central Lembata has a small additional set of non-motion verbs that can optionally appear clause-finally in combination with locational phrases that precede the verb (cf. §3.7.2.2).

The verbs listed in Table 10.2 can all appear in clause-final position. This does not exclude their use in medial position. For WL-Lewotobi, Nagaya (2011:336) states that deictic motion verbs can appear clause-medially, but in this position they do not carry their deictic meaning but are pure motion verbs. How far this semantic difference related to the position of the verb also holds for the other varieties of Flores-Lembata is not clear from the sources available.

For the clause-final verbs in Kedang listed as 'go' verbs in Table 10.2, it ap-

pears that the semantics do not include a clear deictic component. They can be used for a movement towards the speakers ('come'), as well as away from the speaker ('go') (Samely 1991a:130), thus I refer to them as "non-deictic" (cf. §10.2.3.3). However, in the table, these Kedang verbs are listed with the 'go' paradigm for the sake of symmetry, as Kedang also has a separate set of 'come' verbs (cf. §10.2.3.3). An additional reason to group them with the 'go' verbs is that some of the non-deictic Kedang verbs are cognate with 'go' verbs in other Flores-Lembata varieties.

In the following, I first describe the system in the Lamaholot varieties in §10.2.3.2 as it appears to be most complete and elaborate. Clause-final DMV are attested with all three levels of elevation and their semantics all convey a clear deictic component. In Kedang (§10.2.3.3) also all levels of elevation are attested in the set of clause-final verbs, however not all clause-final verbs appear to differentiate between 'come' and 'go'. In Sika only LEVEL DMV verbs are attested (§10.2.3.4).

10.2.3.2 The Lamaholot varieties

Throughout all known varieties of the three Lamaholot subgroups, clause-final DMV are found that encode elevation and direction in relation to the position of the deictic centre. A complete set of DMV in Lamaholot contains three verbs for 'go' and three verbs for 'come'. These sets of three verbs each have one verb that is unspecified for elevation (LEVEL), one verb that encodes a higher location (HIGH) and one verb that encodes a lower location relative to the deictic centre (LOW).

The Central Lamaholot variety of Central Lembata has a complete set of clause-final DMV with three 'go' verbs: *-ai* 'go', *géji/géwi* 'go.up' and *lodo* 'go.down', and three 'come' verbs: *méné(k-)* 'come', *aka(j-)* 'come.up' and *nau(n-)* 'come.down'. Their use is discussed in detail in the Central Lembata grammar sketch in §3.7.2.2.

In the Western Lamaholot variety Solor, as documented by Kroon (2016), the set of DMV is also complete. There are 'go' verbs for all three elevation levels: *-ai* 'go', *dopa/géré* 'go.up' and *lodo* 'go.down', illustrated in (5), and also 'come' verbs for all three elevation levels *pai* 'come', *haka/dai* 'come.up' and *hau* 'come.down'. However, Kroon (2016) only provides an example of *dai* 'come.down', given in (6).²

² Glosses in the examples from Kroon (2016:Appendix 5) are added, as the source does not

(5) WL-Solor: 'go' verbs

- a. *Ra'é lua wata r=ai.*
 3PL descend beach 3PL-go
 'They went (descended) down to the beach.' (Kroon 2016:217)
- b. *Tite t=eté wa'a ilé dopa.*
 1PL.INCL 1PL.INCL=bring burden mountain go.up
 'We bring these goods up to the mountain.' (Kroon 2016:217)
- c. *... ula wé gelora téti wato lolo-n lodo.*
 snake DIST glide.down upwards stone top-3SG.POSS go.down
 '... a snake gliding down from above the stone.' (Kroon 2016:Appendix5)

(6) WL-Solor: 'come'-verbs

- Na'é lau watã dai.*
 3SG seawards beach come.up
 'He is coming from the beach.' (Kroon 2016:Appendix5)

In some WL varieties, the set of clause-final DMV is not entirely complete. WL-Lewotobi possesses one clause-final 'go' verb *-ai* 'go', but three 'come' verbs: *dai* 'come', *hau* 'come.down' and *haka* 'come.up' (Nagaya 2011:330-331). There are the WL-Lewotobi verbs *géré* 'go.up' or *lodo* 'go.down' but they are not clause-final.

In Alorese, an offspring of WL, the following clause-final DMV are attested in Klammer (2011:64-65): *méné* 'come', *-éi* 'go', *lodo* 'go.down' and *géré* 'go.up'. It remains to be investigated if Alorese also has clause-final DMV meaning 'come.down' and 'come.up'.

For WL-Lewoingu, a list of DMV is given but their use is not sufficiently described to deduce which of them are clause-final (Nishiyama and Kelen 2007:91,116-117). Nevertheless from the examples in Nishiyama and Kelen (2007:116-117), it becomes clear that at least *-a'i*-1 go' can be clause-final.

To sum up, all Lamaholot varieties have clause-final DMV. The exact number of DMV varies across varieties. A complete set of clause-final DMV, with three 'go' verbs and three 'come' verbs, is documented for CL-Central Lembata and for Alorese. In WL-Lewotobi an incomplete set, three clause-final 'come' verbs and only one 'go' verb, is attested. For other WL varieties,

provide glossing.

such as Solor or Lewoingu, the documentation of the DMV in use is not sufficient to decide whether the full set of DMV is can appear in clause-final position.

10.2.3.3 Kedang

Kedang has two sets of clause-final motion verbs, as documented by Samely (1991a:129-131). Table 10.3 shows that one set is non-deictic, glossed with ‘move’, only encoding elevation. The second set is deictic, encoding elevation and the direction towards the speaker, glossed with ‘come’. A set that solely encodes direction away from the speaker ‘go’ does not exist. To express this meaning, a non-deictic motion verb is used and the deictic direction away from the speaker is interpreted from the context.

Table 10.3: Clause-final motion verbs in Kedang

Non-deictic	
‘move.up’	<i>dau</i>
‘move.up’ (vertically)	<i>kéu</i>
‘move.down’	<i>bunu’</i>
‘move.down’ (vertically)	<i>do’</i>
Deictic	
‘come.sideways’	<i>ma</i>
‘come.seawards’	<i>nè</i>
‘come.hillwards’	<i>ḡ’</i>

The non-deictic set is based on the up-down axis, whereas the deictic set is based on the sea-land axis. A non-deictic motion verb such as *dau* ‘move.up’ can be preceded by the prepositions *ḡté* ‘upwards’, as in (7a), or *ḡlé* ‘downwards’. A deictic motion verb such as *ḡ’* ‘come.hillwards’, which is based on the sea-land axis, can be preceded by the prepositions *ḡjo* ‘sideways’, *ḡli* ‘hillwards’ and *ḡwé* ‘seawards’, as in (7b).

(7) KD-Leuwayang

- a. *Nuo ɔté wéla dau.*
 3SG upwards inland move.upwards
 'He goes up into the interior.' (Samely 1991a:130)
- b. *Wèngpitu wati' ko' a'é ʔana' ʔabé ʔwé ʔ'.*
 CLF seven again 1SG.POSS older.sibling child male seawards
 come.hillwards
 'In seven days, my older brother comes here.' (Samely 1991a:131)

With the non-deictic motion verbs, nevertheless, it is possible to express deixis, thus a 'come' or 'go' movement. The combination of the preposition *ɔté* 'upwards' and the non-deictic motion verb *dau* 'move.upwards' (7a) yields a reading of 'go'. The use of the opposite preposition *ɔlé* 'downwards' with the same deictic motion verb leads to a reading of 'come', such as in the phrase *ɔlé dau* 'to come upwards from a lower position' (Samely 1991a:131). Thus, when using non-deictic motion verbs, 'go' and 'come' can be differentiated by the choice of the preceding preposition.

In contrast, for the deictic motion verbs meaning 'come', the prepositions on the sea-land axis have to be used in semantic accordance to the DMV, so *ʔwé* 'seawards' can only be combined with *ʔ'* 'come.hillwards'. In (7b), the preposition *ʔwé* which encode the origin of the movement does not add any semantic value to the phrase, as the verb *ʔ'* 'come.hillwards' already encodes the origin of the movement as being a location at sea level or overseas.

To sum up, Kedang possesses a set of clause-final 'come' verbs and a set of clause-final non-deictic motion verbs, 'move' verbs, which, in combination with prepositions, can express the deictic component of 'go' or 'come'.

10.2.3.4 Sika

Sika has only two main clause-final deictic motion verbs: *-a* 'go' and *mai* 'come'. In addition, the verb *balér* 'return' is also clause-final. In Sika, elevation is not encoded in deictic motion verbs. The verb *-a* 'go' is mainly found in expressions such as 'Where are we going to?' or 'She walks there', as shown in examples (8a) and (8b). An example of the verb *mai* 'come' is given in (8c).

(8) a. SK-Hewokloang

Ita épai t-a-t?

1PL.INCL where 1PL.INCL-go-1PL.INCL

'Where are we going to?'

(Rosen 1986:62)

b. SK-Nita

Nimu gawi lau n-a?

3SG walk to 3SG-go

'She walks there.'

(Rosen 1986:59)

c. SK-Hewa

A'u 'ia Weri mai.

1SG LOC NAME come

'I'm coming from Weri.'

(Fricke 2014a:61)

Both verbs *-a* 'go' and *mai* 'come' can be inflected by pronominal affixes. In many varieties, these affixes have been eroded completely or generalised to a default nasal /n/ or /ŋ/. The prefixes are retained with more stability than the suffixes. The most conservative variety in this respect is spoken in the Hewokloang area located in the inland towards the east of the regency capital Maumere (Rosen 1986:62). In this variety, a full paradigm of pronominal prefixes and suffixes is retained. The erosion of the inflection suggests that the verb *-a* 'go' is losing its verbal qualities and is on the way to grammaticalise into a clause-final directional particle (Rosen 1986:59-62; Fricke 2014a:35).

10.2.4 Final DMV in the Alor-Pantar languages

In Table 10.4, the forms of deictic motion verbs (DMV) in four present-day Alor-Pantar languages and in Proto-Alor-Pantar (PAP) are reproduced from Schapper (2017:274). The four languages are chosen to represent a wide geographic as well as genealogical distance. They belong to different subbranches of the Alor-Pantar (AP) languages and Western Pantar is the westernmost AP languages, while Wersing is one of the easternmost AP languages of Alor. The DMV are grouped according to the same three levels of elevation

and the deictic categories of ‘go’ and ‘come’ as in the Flores-Lembata languages (§10.2.3.1). The Wersing verb *a* is cognate with the set of ‘come.down’ but underwent a semantic change and now means ‘go.down’.

Table 10.4: Cognate sets of deictic motion verbs in the AP languages

	LEVEL		HIGH		LOW	
	‘go’	‘come’	‘go’	‘come’	‘go’	‘come’
PAP	*wai	*mai	*mid(a)	*medai(ŋ)	*pia	*ya(ŋ)
Western Pantar	<i>wa</i>	<i>ma</i>	<i>mia</i>	<i>middaŋ</i>	<i>pia</i>	<i>yaŋ</i>
Teiwa	<i>wa</i>	<i>ma</i>	<i>mir</i>	<i>daa</i>	-	<i>yaa</i>
Blagar	<i>va</i>	<i>ma</i>	<i>mida</i>	<i>da</i>	<i>ʔipa</i>	<i>ya</i>
Wersing	<i>wai</i>	<i>mai</i>	<i>mid</i>	<i>dai</i>	-	<i>a</i> ‘go’

All Timor-Alor-Pantar languages are verb-final (Klamer 2017:15), and motion verbs occur in clause-final position, as in the Teiwa example in (9). Here the DMV *mir* appears clause-finally after the subject pronoun and the location.

- (9) Teiwa
Iman ta hafan u ga-mir-an, maan?
 they TOP village DIST 3SG-go.up-REAL NEG
 ‘They didn’t go up to that village, did they?’ (Klamer 2010:160)

The data from the Alor-Pantar languages show two similarities with the Flores-Lembata languages. First, the deictic motion verbs occur in ‘go’ and ‘come’ set with each three levels of elevation. Second, DMV appear clause-finally in both language groups. This word order is the norm in AP, while it is an exception to the general word order in Flores-Lembata, where other verbs occur in clause-medial position.

10.2.5 Etymology of the Flores-Lembata DMV

10.2.5.1 The Flores-Lembata ‘go’ verbs

The lexical items used in clause-final position as deictic motion verbs (DMV) in the Flores-Lembata languages are a mix of innovations and inherited

forms. The ‘go’ verbs are innovated, whereas some of the ‘come’ verbs appear to be inherited. This holds for those DMV that are unmarked for elevation and for those that are encoding elevation.

In Table 10.5, I list the forms of elevated and unelevated DMV meaning ‘go’ in the languages of Flores-Lembata including their reconstructions. No PMP sources are attested for these cognate sets. Only the verb ‘go.down’ can be reconstructed to Proto-Flores-Lembata (PFL) *lodoŋ ‘go.down’ with certainty. The other two forms #-ai ‘go’ and #gedi ‘go.up’ do not have cognates in all Flores-Lembata languages which makes the reconstruction to PFL not possible at the current stage.

Table 10.5: ‘Go’ verbs and their proto-forms

	‘go’	‘go.up’	‘go.down’
PFL	#-ai	#gedi	*lodoŋ
Sika	-a	-	<i>loding</i>
Kedang	-	<i>kéu</i>	<i>do’</i>
Central Lamaholot	-ai	<i>géji</i> (/ <i>géwi</i>)	<i>lodo</i>
Western Lamaholot	-ai	<i>géré</i>	<i>lodo</i>

For the DMV #-ai ‘go’, cognates are found in all varieties of Sika and Lamaholot. Kedang is missing a cognate and also any other unelevated clause-final motion verb to express a movement away from the speaker. In Sika, the *i* of #-ai is dropped and pronominal suffixes can be added.

LH-SK #-ai ‘go’ cannot be linked to any known Proto-Malayo-Polynesian (PMP) form with the meaning ‘go’ and cognates are largely absent in the related Austronesian languages of the area. The languages of Central Flores, such as Keo, have a form *kai* ‘go’ which is similar to the Flores-Lembata forms (Baird 2002:301). However, the apparent loss of initial *k* in Flores-Lembata cannot be explained as usually *k* is retained in these languages (cf. §5.2.1). Similarly, the language Welaun on Timor has *kai* ‘go’ (Edwards 2019:19). However, it is unclear how or if these forms are related as they are found in geographically disparate languages.

Another possible source for the innovation of #-ai ‘go’ in Sika and Lamaholot are the neighbouring non-Austronesian Alor-Pantar languages. The

innovated form #-ai ‘go’ could be related to Proto-Alor-Pantar (PAP) *wai ‘go’ (Schapper 2017:274). However, it remains unclear if PFL borrowed from PAP why the initial *w* has been lost in PFL as these languages usually preserve *w* (cf. §5.2.6).

For the LH-KD #gedi ‘go.up’, no cognate has been found in Sika. In Kedang, *g > k and *d > y/Ø are regular changes (cf. §5.2.2). However, the vowel change of *i > u in final position appears irregular. The reflexes of CL *géji* and WL *géré* are regular, as PFL *-d- > Western Lamaholot -r- and Central Lamaholot -j-. Final vowel lowering, as *i > e, is frequently found in Western Lamaholot. The additional form *géwi* in Central Lamaholot remains unexplained.

For PFL *lodoŋ ‘go.down’, there are cognates attested in all FL languages. The cognate in Sika is only found in clause-medial position (Pareira and Lewis 1998:122). The final consonant in the reconstructed form is probably ŋ, reduced to glottal stop in Kedang and zero in Lamaholot.

10.2.5.2 The Flores-Lembata ‘come’ verbs

In Table 10.6, I list the ‘come’ forms in the Flores-Lembata languages and their reconstructions. Most forms in the set of ‘come’ and ‘come.up’ go back to Proto-Malayo-Polynesian (PMP) forms, whereas for the ‘come.down’ set no PMP form is attested.

Table 10.6: ‘Come’ verbs and their proto-forms

	‘come’	‘come.up’	‘come.down’
PMP	*ma(R)i	*sakay	
PFL	*mai	*hakay	
Sika	<i>mai</i>	-	-
Kedang	<i>ma</i>	<i>ḁʼ</i>	<i>nè</i>
Central Lamaholot	<i>méné(k-)</i>	<i>aka(j-)</i>	<i>nau(n-)</i>
WL-Solor	<i>pai</i>	<i>haka/dai</i>	<i>hau</i>
WL-Lewotobi	<i>dai</i>	<i>haka</i>	<i>hau</i>
WL-Alorese	<i>méné</i>	-	-

Sika *mai* ‘come’ and Kedang *ma* ‘come’ are cognate, and go back to PFL *mai

and ultimately to PMP *ma(R)i ‘come’. Cognates of this set are also found in the Austronesian languages of Flores and Timor (Blust and Trussel 2010). Strikingly, exactly the same proto-form *mai ‘come’ has also been reconstructed for PAP (Schapper 2017:274). This raises the question whether PAP has borrowed from its Austronesian neighbours or whether this similarity in form is due to coincidence. However, as no other DMV are borrowed from Austronesian languages into PAP, this appears rather to be a coincidence. This issue is also discussed by Klammer (2010:325–326).

Reflexes of PMP *sakay ‘climb; ascend’ are found in Kedang, Central Lamaholot and Western Lamaholot varieties. For Proto-Flores-Lembata, I reconstructed PFL *hakay based on the regular sound changes of PMP *s > PFL *h > PCL / PK *Ø, the sound change PMP *y > PCL *dʒ, and finally, PMP *k > PK *ʔ (cf. §5.2).

Central Lamaholot and Alorese appear to have innovated a new form *mene ‘come’ of unknown origin. The first syllable of *mene ‘come’ could be connected to PFL *mai ‘come’ but the evidence for this remains unclear. Reflexes *mene are attested in Central Lamaholot and Alorese. As *mene is only attested in these two languages, it cannot be reconstructed to PFL.

Western Lamaholot varieties have innovated *dai* or *pai* ‘come’ and *dai* ‘come.up’. The verb *dai* appears related to Proto-Alor-Pantar (PAP) *medai(ŋ) ‘come.up’ (Schapper 2017:270). Also in some Alor-Pantar languages, the reflexes of this form show a loss of the initial syllable, such as Teiwa *daa* or Wersing *dai* (Schapper 2017:274). Possibly, Western Lamaholot varieties borrowed *dai from AP languages. In WL-Lewotobi, a semantic change from ‘come.up’ to general ‘come’ occurred. Whether WL-Solor *pai* ‘come’ is connected to *dai ‘come.up’ or *mai ‘come’ remains unclear.

The etymology of the forms for ‘come.down’ is more difficult to reconstruct. It is not clear if all the Flores-Lembata forms for ‘come.down’ are cognate. As further evidence is missing, they have to be regarded as independent forms of unknown origin.

10.2.6 The rise of clause-final deictic motion verbs

I argue that the Flores-Lembata languages innovated clause-final verbs in locational clauses due to contact with non-Austronesian languages that had a canonical verb-final constituent order. DMV in clause-final position as in the Flores-Lembata languages are not found in their closest Austronesian

relatives, as shown in §10.2.2, nor are they inherited from Proto-Malayo-Polynesian (cf. §10.2.1). Therefore, clause-final DMV must have been innovated at the level of Proto-Flores-Lembata. This innovation has probably started with the unelevated forms of LH-SK #-ai ‘go’ and PFL *mai ‘come’, as clause-final reflexes of these verbs are attested throughout virtually all Flores-Lembata languages, with the exception of Kedang, where no reflex of LH-SK #-ai ‘go’ is attested. The elevated clause-final forms might have developed later, as they are only attested in Kedang and Lamaholot and not in Sika.

For the innovation of clause-final DMV, a language-internal development is unlikely. Only a small subset of verbs, namely deictic motion verbs, became clause-final in the Flores-Lembata languages. All other verbs remain clause-medial. Therefore, it is unlikely to assume a universal principle, such as a tendency to converge towards a certain word order pattern, as the cause for the change.

I put forward three arguments to strengthen the hypothesis that the innovation of clause-final verbs in locational clauses arose due to contact with non-Austronesian languages. First, there is evidence from other studies that clause-final verbs in several Austronesian languages have evolved in contact zones with non-Austronesian languages. According to Blust (Blust 2013:461), New Guinea has many verb-final Austronesian languages and the Solomon Islands have a few. The hundreds of non-Austronesian languages found especially in New Guinea have verb-final word order (Blust 2013:470) which suggests that the verb-final order in many of the Austronesian languages of New Guinea developed due to contact with the non-Austronesian languages of that area (Blust 2013:471).

Second, the Flores-Lembata languages are far from being entirely verb-final. Only a certain set of verbs, mainly deictic motion verbs, is used clause-finally when appearing with a locational phrase. The reason for the syntactic change only affecting DMV may lay in their semantic properties of encoding elevation which is a feature the Flores-Lembata DMV share with the non-AN Alor-Pantar languages. Therefore, I suggest that both the syntactic position and the semantics of these verbs have been calqued from non-Austronesian languages. Most likely from a language related or typologically similar to the neighbouring Alor-Pantar languages, as these languages have clear parallels to the semantic features of Flores-Lembata DMV discussed in §10.2.3 and 10.2.4. In the languages of Flores-Lembata, as well as the Alor-Pantar languages, DMV encode elevation on a HIGH-LOW axis

in addition to the pure deictic component of ‘come’ and ‘go’. Finally, there is strong evidence that the form WL *dai* ‘come.up’ has been borrowed from AP languages and weaker evidence that LH-SK #-ai ‘go’ has also been borrowed (cf. §10.2.5). It could be that the semantic system of the DMV was calqued first and subsequently, this semantic conformance caused the change in position based on the clause-final position of the same types of verbs in the non-Austronesian contact language.

10.3 Clause-final negation

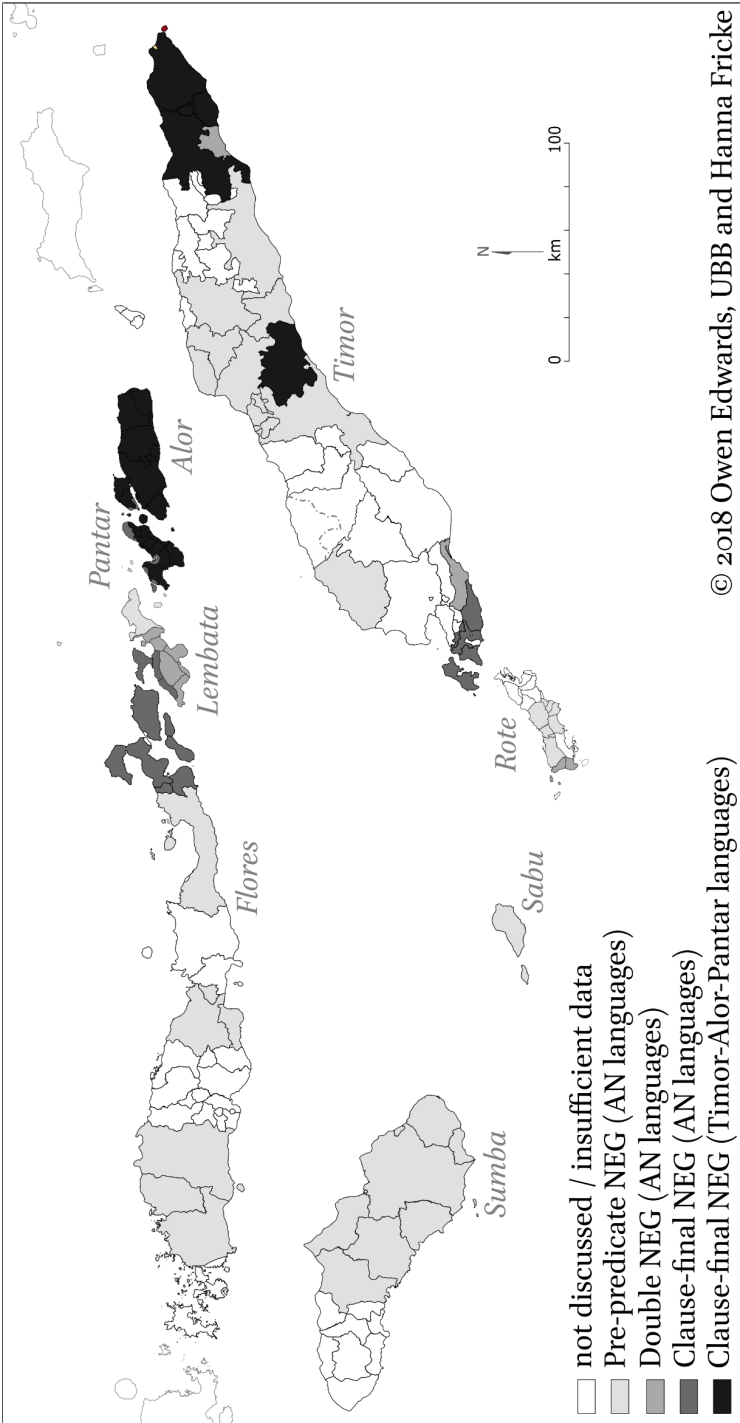
10.3.1 Overview

Clause-final negation is defined as a negation pattern with a clause-final negator.³ This also includes cases of double negation where one negator is in pre-predicate position and the second negator is clause-final. In Austronesian languages, clause-final negators are innovations. As with 70 % of the world’s languages (Vossen 2016:4), Austronesian languages typically show pre-predicate negation marking. However, there is a considerable number of instances of clause-final negation among Austronesian languages located in areas where Austronesian and Papuan languages are close to each other (Klamer 2002:375; Vossen 2016:119-121,202; Reesink and Dunn 2018:936). This pattern of clause-final negation marking in Austronesian languages has been argued to be of Papuan origin (Reesink 2002:246). Clause-final negation has been identified as a Papuan feature of Lamaholot (Klamer 2012a:76).

The map in Figure 10.2 shows the distribution of pre-predicate negation, double negation and clause-final negation in the Austronesian and non-Austronesian languages of Nusa Tenggara Timur and Timor-Leste. While in the west, only pre-predicate negation is attested, some of the Austronesian languages towards the east have developed clause-final negators in single or double negation patterns. Among the Flores-Lembata languages, only the Lamaholot varieties and SK-Hewa have clause-final negators. On Timor a mixed picture of pre-predicate, double and clause-final single negation emerges. In the non-Austronesian Timor-Alor-Pantar (TAP) languages, only clause-final negation is attested.

³ An earlier version of §10.3 has been published as Fricke (2017).

Figure 10.2: Negation patterns



§10.3 discusses patterns of declarative sentence negation in the Flores-Lembata languages and is structured as follows. §10.3.2 provides the theoretical background on the model of a Jespersen Cycle to explain the diachronic development of negation in Flores-Lembata languages. §10.3.3 shows that in the closest Austronesian relatives of the Flores-Lembata languages, on Flores and Sumba, only pre-predicate negation is attested. These languages thus retain the typical Austronesian pattern for this feature. §10.3.4 presents different negation patterns in the languages of Flores-Lembata. It is shown that also some Flores-Lembata languages, i. e. Sika and Kedang, retain pre-predicate negation, but that in the Lamaholot varieties a new pattern of clause-final negation is innovated. §10.3.5 gives a brief overview of negation patterns found in the Austronesian languages of Timor. In these languages, similar to the Flores-Lembata group, all negation patterns are attested. §10.3.6 discusses etymologies of the negators used in the Flores-Lembata languages. §10.3.7 discusses the proposal of contact-induced change giving rise to clause-final negation.

10.3.2 Theoretical background: Jespersen Cycle

A Jespersen Cycle is a diachronic change from pre-predicate single negation, to double negation, to post-predicate single negation (van der Auwera and Du Mon 2015:412).⁴ A well known example is the Romance language French which underwent all three stages of a Jespersen Cycle. Initially, French had a single pre-predicate negator *ne*, nowadays standard French uses double negation *ne ... pas* and colloquial varieties of French already reached the final stage of the Jespersen Cycle by only using the single post-predicate negator *pas*.

In many cases, a Jespersen Cycle starts with the need for emphatic negation (van der Auwera and Du Mon 2015:412). An element is added to the negated sentence to emphasise its negative meaning, over time this emphatic meaning is bleached and the new strategy becomes the neutral negation pattern (van der Auwera 2009:41). Then, in the following stage, the original negator is lost and the new element becomes the only negator in the clause.

⁴ It has been shown that, cross-linguistically, there is not only one kind of Jespersen Cycle but several patterns that can be referred to as Jespersen Cycle. For the purpose of this chapter, I will not go into details of this cross-linguistic diversity. For further information see van der Auwera (2009) and Vossen (2016).

New negators can come from different sources. Cross-linguistically common sources are indefinites, like a word for ‘nothing’, negative main verbs, like ‘refuse’ or ‘lack’, (van Gelderen 2008:196) or negative existentials (Croft 1991:6) that grammaticalise into a negator. Other sources are partitives, like ‘a little’, nominalisers, possessives or a copy of the original negator (Vossen 2016:36-37).

Jespersen Cycles have been attested in many languages all over the world (van Gelderen 2008; Vossen 2016). According to Vossen (2016:202), Jespersen Cycles are common in the Austronesian language family, especially in those parts of the family which are in contact with Papuan languages. Papuan languages typically have clause-final negation in line with their verb-final word order. This observation suggests that the start of a Jespersen Cycle can be triggered by language contact. An example is provided by Vossen (2016:123). She proposes a Jespersen Cycle in the Austronesian Markham Valley languages in Papua New Guinea. This group has languages in all three stages of a Jespersen Cycle. Some of the Markham Valley languages have pre-predicate single negation, most have double negation and two of them have reached the final stage of post-predicate single negation. The fact that the same pre-predicate negator is shared by languages that have single pre-predicate negation and languages that have double negation is seen as a strong evidence for a Jespersen Cycle. The same holds for the shared post-predicate negator in varieties that have double and single negation (Vossen 2016:125).

To prove the existence of a Jespersen Cycle, one would, ideally, show the change using historical data of the language under study. However, this can only be done for languages with a long written tradition. Nevertheless, the case study of the Markham Valley languages mentioned above and other recent studies (van der Auwera and Vossen 2016; Vossen and van der Auwera 2014) have shown that it is equally possible to propose a Jespersen Cycle on the ground of synchronic data from several related languages. Here, I show this for the languages of Flores-Lembata.

10.3.3 Pre-predicate negation on Flores and Sumba

This section provides examples of pre-predicate negation attested throughout the Austronesian languages of Flores, Sumba and Sawu. These languages have retained the inherited Austronesian pattern for this feature. In the lan-

guages of Western and Central Flores, negators are in pre-predicate position, as Keo *mona* in (10) and Manggarai *toe* in (11).

- (10) Keo
Imu mona nai nio.
 3SG NEG climb coconut
 'He didn't climb the coconut tree.' (Baird 2002:333)
- (11) Manggarai
Joni toe mo le sekola-e.
 NAME NEG go LOC school-3SG
 'John does not go to school.' (Semiun 1993:92)

Also in Kambera (12), on Sumba, pre-predicate negation is attested. In Hawu on the island of Sawu, the negation can be pre-verbal (13a) or post-verbal (13b) but not clause-final.

- (12) Kambera
Nda ku-ngangu-a iyang.
 NEG 1SG.NOM-eat-MOD fish
 'I don't eat fish.' (Klamer 1998:108)
- (13) Hawu
 a. *Do peda Ø faa, haku do fega lodo dě.*
 STAT be.sick ABS 1SG result NEG work day PROX
 'I'm sick, so I'm not working today.' (Walker 1982:47)
 b. *Pidé do ri Ubu Naba Ø né naléhu puné.*
 pick.up NEG ERG NAME NAME ABS ART handkerchief PROX
 'Ubu Naba did not pick up the handkerchief.' (Walker 1982:47)

10.3.4 Negation in the Flores-Lembata languages

10.3.4.1 Overview of negation patterns

In the languages of Flores-Lembata all three stages of a Jespersen Cycle are attested: 1) pre-predicate single negation, 2) double negation and 3) clause-final single negation. These patterns and the varieties in which they occur are laid out in Table 10.7.

Table 10.7: Negation patterns in Flores-Lembata languages

Negation Pattern				Occurrence
Pre-predicate:	NEG	V		Kedang Sika
Double:	NEG	V	NEG	SK-Hewa EL-Lewoeleng CL-Central Lembata WL-Lamalera
Clause-final:		V	NEG	WL-Lewoingu WL-Lewotobi WL-Solor WL-Alorese

Pre-predicate single negation is found in the peripheral varieties of Flores-Lembata, in Sika and Kedang, whereas clause-final negation is clustered in the centre of the area, in the Western Lamaholot varieties. In between, there are varieties that show the intermediate state of double negation. These are SK-Hewa, in between other Sika varieties with pre-predicate negation and Western Lamaholot varieties with clause-final negation, and Central Lamaholot, Eastern Lamaholot and WL-Lamalera in between other Western Lamaholot varieties and Kedang.

10.3.4.2 Pre-predicate single negation

Pre-predicate single negation is the main negation pattern in Kedang (14) and in Sika (15).

- (14) KD-Leuwayang
Wèi oha' in=u.
 water NEG drink.1SG=1SG
 'I don't drink water.' (Samely 1991a:74)

- (15) Sika
A'u éné ra'intang.
 1SG NEG know
 'I don't know.' (Arndt 1931:42)

This pattern presents the first stage of a Jespersen Cycle. The second stage, double negation, might be a minor pattern in Sika. Arndt (1931) gives one example of double negation (16), however he does not specify exactly in which context the sentence appears.

- (16) Sika
Nimu éné léta ata natar péhang e'ong.
 3SG NEG invite person village other NEG
 'He did **not** invite the people from the other village.' (Arndt 1931:42)

In the original translation of example (16), Arndt presents the negator *not* in bold but does not explain why he does so. The bold font may suggest that the double negation functions as an emphatic negation. If this interpretation is correct, this can be taken as evidence for the emphatic phase of double negation when this pattern is not yet semantically bleached to become the general pattern.

10.3.4.3 Double negation

Double negation uses a pre-predicate negator in combination with a clause-final negator. Double negation is the main pattern in SK-Hewa (17), CL-Central Lembata (18), EL-Lewoeleng (19) and WL-Lamalera (20). All these Lamaholot varieties are found on the island of Lembata.

- (17) SK-Hewa
Dedi' anak e'on puas iwa.
 child little NEG satisfied NEG
 'The little child is not satisfied.' (Fricke 2014a:9)
- (18) CL-Central Lembata
Ta na=mojip si.
 NEG 3SG=live NEG
 'It does not live.' (Fricke 2019)
- (19) EL-Lewoeleng
Go'e ta hab'u wa.
 1SG NEG bathe NEG
 'I don't take a shower.' (NB:81)

- (20) WL-Lamalera
T-ai fulã pe tana di taku tegel hala.
 1PL.INCL-go market DIST land also NEG see NEG
 ‘When we went to the market, we didn’t see the island.’ (Keraf 1978a:232)

Double negation in SK-Hewa and CL-Central Lembata is not obligatory, either of the negators can be left out. However, in my corpus data of these varieties, double negation is the most frequent negation pattern used. The data available in Keraf (1978a) for WL-Lamalera show consistent double negation as well. For EL-Lewoeleng only a few elicited sentences are available, they all show double negation.

10.3.4.4 Clause-final single negation

Clause-final single negation is the main negation pattern in all Lamaholot varieties documented outside of the island of Lembata. These varieties are in WL-Alorese (21), WL-Lewotobi (22), WL-Lewoingu (23) and WL-Solor (24).

- (21) WL-Alorese
No pana ha néi tahi lahé.
 3SG walk this 3SG.go.to sea NEG
 ‘He did not go to the sea.’ (Klamer 2011:87)
- (22) WL-Lewotobi
Go kě ikě hua hela’.
 1SG 1SG.eat fish fish.sp NEG
 ‘I don’t eat *hua* fish.’ (Nagaya 2011:392)
- (23) WL-Lewoingu
Go berin na hala’.
 1SG hit 3SG NEG
 ‘I did not hit him.’ (Nishiyama and Kelen 2007:69)
- (24) WL-Solor
Ema déna wata la.
 mother cook rice NEG
 ‘Mom is not cooking rice.’ (Kroon 2016:158)

In the WL varieties outside of Lembata, there are remnants of a minor pattern of double negation. In example (25), the pre-predicate negator *gara* is combined with the clause-final negator *hala*. Example (25) is taken from a grammar by Arndt (1937) on Lamaholot varieties outside of Lembata.

- (25) Lamaholot (outside Lembata)
Goe gara taka kan' hala.
 1SG NEG steal eat NEG
 'I did not steal and eat it' (Arndt 1937:99)

Arndt (1937) notes that this double negation pattern is infrequent in Lamaholot. Taking the Jespersen Cycle into account, this infrequent pattern can be seen as evidence that these varieties previously went through a stage of double negation.

10.3.5 Negation in the AN languages of Timor

This section presents a brief overview of negation patterns found in the AN languages of Timor. Similar to the Flores-Lembata languages, all stages of a Jespersen Cycle can be observed in these languages. The pre-predicate negator can be found as the only negator, as in (26), or it is combined with the clause-final negator to a double negation pattern, as in (27) and (28). In addition, cases of only clause-final negation without pre-predicate negator are attested, as illustrated in (29) and (30).

In Tetun of Eastern Timor, there is a pre-predicate negator *la* and a post-predicate negator *ha'i* (van Klinken 1999:228) which can be combined or used separately, as in the example below where only *la* is used. According to van Klinken (1999:228), pre-predicate negation only as in (26) occurs in 75% of the cases in Tetun.

- (26) Tetun
... nia la karian.
 3SG NEG work
 '..., he doesn't work' (van Klinken 1999:228)

Also in most languages of Rote pre-predicate negation is the only option, such as in Termanu, Lole and Dengka (Jonker 1915:587).

However, in Dela on Rote, and in Amarasi in western Timor, double negation is the norm, as in the examples below. In the Kotos variety of Amarasi,

the clause-final negator =*fa* is occasionally left out in free speech. However, speakers of Kotos Amarasi consider the double negation with both negators as the only correct form (Edwards 2016a:133).

- (27) Dela
Ne-nori' naa nda matetu' sa.
 NMLZ-teaching that NEG straight NEG
 'That teaching is not good.' (Tamelan 2007:8)
- (28) Amarasi-Kotos
 ... *ka=na-sai=fa.*
 NEG=3-flow=NEG
 '... it didn't flow.' (Edwards 2016a:132)

Clause-final negation without pre-predicate negator is attested in the Ro'is variety of Amarasi (29) and in Helong (30).

- (29) Amarasi-Ro'is
Au k-oka maé'.
 1SG 1SG-with NEG
 'I won't come along.' (Owen Edwards, pers. comm.)
- (30) Helong
Kaim laok lo.
 1PL.EXCL walk NEG
 'We are not going.' (Balle 2017:15)

From this very brief overview, it can be suggested that also in the Austro-nesian languages of Timor a Jespersen Cycle appears to be starting. Pre-predicate as well as double negation are found. Clause-final negation may occur optionally and in some languages, clause-final negation is the main pattern. According to current knowledge, clause-final negators, in double or single negation, are only found in languages of the Timor-Babar subgroup. Central Timor languages have pre-predicate negation. See §1.3.1 for more information on the Timor subgroups.

Despite the similarities of the Timor case and the Lamaholot case, I assume that both developments are independent because the rise of clause-final negation cannot be reconstructed to any proto-language that would encompass Lamaholot and languages of Timor with clause-final negators, but

exclude those languages that remain with pre-predicate negation only, such as Kedang, most Sika varieties, and some languages of Rote and Timor. In addition, the Flores-Lembata languages are more closely related to the AN languages of Flores and Sumba than to the AN languages of Timor (cf. §5.5) and the AN languages of Flores and Sumba do not show any case of clause-final negation. Therefore it is very unlikely that the Flores-Lembata languages have inherited clause-final negation from a common ancestor with the Timor languages.

10.3.6 Etymology of the Flores-Lembata negators

10.3.6.1 Overview of negators

§10.3.6 examines the etymology of pre-predicate and clause-final negators in the Flores-Lembata languages. Table 10.8 gives an overview of the negators grouped according to their negation pattern. Only negators appearing in major patterns are discussed.

Table 10.8: Negators in the Flores-Lembata languages

Variety	Pre-predicate		Clause-final
KD-Leuwayang	<i>oha'</i>	...	
SK	<i>éné</i>	...	
SK-Hewa	<i>é'o(n)</i>	...	<i>iwa</i>
EL-Lewoeleng	<i>ta</i>	...	<i>wa</i>
CL-Central Lembata	<i>ta / tak</i>	...	<i>si(né)</i>
WL-Lamalera	<i>taku</i>	...	<i>hala</i>
WL-Lewotobi		...	<i>hela'</i>
WL-Lewoingu		...	<i>hala'</i>
WL-Solor		...	<i>la</i>
WL-Alorese		...	<i>lahé</i>

^a The form *tak* is used for negation of nominals or other non-verbal elements.

10.3.6.2 Pre-predicate negators

For the pre-predicate negators, Kedang *oha'*, Sika *éné* and SK-Hewa *é'o(n)*, no Austronesian proto-form can be found.

For Sika *éné* several possibly related forms are found in languages of the area. Possibly related to the Sika negator *éné* are Kedang *anung* ‘reject, refuse, decline’ (Samely 1991a:162) or the post-verbal negator *néné* in the Timor-Alor-Pantar language Adang (Robinson and Haan 2014). For the Kedang pre-predicate negator *oha*’ no convincing cognates or related forms are found.

For the SK-Hewa negator *é'o(n)*, the homophonous negative existential is a language-internal source. Negative existentials have been shown to be a cross-linguistically common source of verbal negators by Croft (1991:6).

Among the pre-predicate negators in Table 10.8, only the negators *tak/ta* in CL-Central Lembata, *ta* in EL-Lewoeleng, and *taku* in WL-Lamalera can be clearly traced back to an Austronesian form. Proto-Malayo-Polynesian (PMP) *ta ‘no, not’ has been reconstructed as a negative marker (Blust and Trussel 2010). Reflexes of PMP *ta ‘no, not’ are found all over Austronesian languages and even, as a result of diffusion, in several non-Austronesian languages (Reesink 2002:246; Vossen 2016:161). The origin of the second part of the pre-predicate negators, *k* in CL-Central Lembata and *ku* in WL-Lamalera, remains unclear. This element is also present in the negative existential *take* which could be the direct source for the verbal negator in CL-Central Lembata and WL-Lamalera.

10.3.6.3 Clause-final negators

Assuming that the ancestral pattern of the Flores-Lembata languages was to have a pre-predicate negator (cf. §10.3.7), it is the origin of the clause-final negator that we are most interested in. There are four possibilities to explain the origin of the clause-final negator, (i) a loan, (ii) a spontaneous innovation of unknown origin, (iii) a pre-predicate negator that is copied and used in clause-final position, or (iv) an inherited word that became a negator due to grammaticalisation. In this section, I show that the clause-final negators in the Central and Western Lamaholot varieties are inherited words that grammaticalised, whereas the clause-final negators in SK-Hewa and EL-Lewoeleng are most likely innovations.

The SK-Hewa negator *iwa* and the EL-Lewoeleng negator *wa* are most likely cognate. This is important because these languages are not very closely related and their closest relatives do not have a cognate of this negator. When looking for possible sources of SK-Hewa *iwa* and EL-Lewoeleng *wa* several

possibilities come up. These negators could be related to #ba/#βa/#(u)wa, a cognate set of negators which has been argued to be of Papuan origin and is attested all over Austronesian and Papuan languages (Reesink 2002:246; Vossen 2016:158).⁵ Possibly also related to these could be the pre-predicate negators *mbiwa* (Antonius and Ruskhan 1997:25) in Rongga and *iwa* in Ende (Aoki and Nakagawa 1993), both Austronesian languages of central Flores. The Lamaholot prohibitive *nawa* found in WL-Lewotobi and WL-Solor may also be connected to this cognate set, as well as finally, the post-predicate particle *fā* ‘a little’ in some varieties of Rote (Jonker 1908:117). The Rote particle *fā* is also found in negations. The final negator *fā* is also found in several varieties of Meto (Edwards 2016b:56).

The CL-Central Lembata post-predicate negator *si(né)* is homophonous with the CL-Central Lembata word *si(né)* ‘a little’. Thus, I proposed that *si(né)* ‘a little’ grammaticalised into a negator in CL-Central Lembata. The same grammaticalisation path from ‘a little’ to a negator can be proposed for the Rote-Meto particle *fā* above.

All WL varieties have clause-final negators that are cognates of *hala* and go back to PMP *salaq ‘wrong, mistake’ (Blust and Trussel 2010). In the languages concerned, *s > h is a regular sound change (cf. §5.2.3). Apart from this regular sound change, in some varieties, *a* is weakened to schwa and final *q became glottal stop or is lost. In WL-Solor the syllable *ha* is lost leading to the negator *la*, and in WL-Alorese, syllable metathesis and vowel rising of *a* to *e* word-finally leads to *lahé*. Most of the varieties discussed still have a cognate of this word with the original meaning ‘mistake’.

In sum, CL-Central Lembata grammaticalised a word for ‘a little’ into a clause-final negator, whereas all WL varieties grammaticalised a word for ‘wrong, mistake’ into a clause-final negator. In contrast to grammaticalising language-internal material, the SK-Hewa and the EL-Lewoeleng clause-final negator have to be regarded as an innovation of unknown origin.

10.3.7 The rise of clause-final negation

In the languages of Flores-Lembata, a Jespersen Cycle has been documented. The varieties of Lamaholot added a second negator to their pre-predicate

⁵ The geographically closest non-Austronesian language that has a negator possibly related to this set is the Timor-Alor-Pantar language Western Pantar with *kauwa* ‘NEG’ (Holton 2014:51).

negation, gained double negation, and most Western Lamaholot varieties eventually lost the pre-predicate negator and were left with clause-final single negation. Section 10.3.6 has shown which Flores-Lembata negators are inherited Austronesian words and which are innovations of unknown origin.

The strongest evidence for contact induced grammaticalisation is provided by Vossen's (2016) typological study on Jespersen Cycles in Austronesian languages. Vossen (2016:88,120) only finds evidence of double negation and clause-final single negation in areas which are located eastwards to the islands of Borneo and Java.⁶ Reesink and Dunn (2018) equally note that clause-final negation is only found in eastern Indonesia and in Oceania. Further west, no instances of double or clause-final negation are found. This strongly indicates a connection between the existence of clause-final negation and the location of Papuan languages.

The ancestor of all Flores-Lembata languages had pre-predicate single negation, because pre-predicate single negation is the most common negation pattern in Austronesian languages (Vossen and van der Auwera 2014:61) and it is also the most common pattern universally (van der Auwera and Du Mon 2015:411). In addition, the Lamaholot varieties retain reflexes of the Austronesian pre-predicate negator PMP *ta in pre-predicate position. As there are still Flores-Lembata varieties, namely Sika and Kedang, for which pre-predicate single negation is the only negation pattern (cf. §10.3.4.2), it is very unlikely that the Jespersen Cycle already started in an ancestor language including all the languages of Flores-Lembata. It must have started on a lower level, namely independently in SK-Hewa and the three Lamaholot subgroups.

The origin of the clause-final negators in Lamaholot can be explained by contact-induced grammaticalisation (Heine and Kuteva 2003; Heine and Kuteva 2005). Language internal material is used to generate a new grammatical word, a clause-final negator. An additional argument for independent developments in each Lamaholot subgroup is that each of the subgroups uses a different clause-final negator. All WL varieties have a cognate of *hala* going back to PMP *salaq 'wrong'. CL-Central Lembata has *si(né)* as the clause-final part of its double negation. And EL-Lewoeleng has *wa* which is cognate

⁶ There are two exceptions to this. Chamic and Moken languages, which are Austronesian minority languages in Vietnam, Cambodia and Thailand, show double and post-predicate negation (van der Auwera and Vossen 2015; Vossen 2016:92,117,118).

with the clause-final negator *iwa* attested in SK-Hewa.

In contrast to the separate developments of clause-final negators in the three Lamaholot subgroups, all traces of pre-predicate negators found in the Lamaholot subgroups are cognates and reflexes of the PMP negator *ta ‘no, not’. Thus, I propose that Proto-Flores-Lembata had a pre-predicate negator PFL *ta(kV) ‘no, not’ which is lost in most of the WL varieties by now but it survives in EL, CL and the WL varieties on Lembata.

Potential contact languages for Lamaholot are the languages of the Alor-Pantar (AP) family. In the AP languages, negators are clause-final, such as in the Western Pantar example in (31).

- (31) Western Pantar
Anaga was arugga kauwa.
 today sun hot NEG
 ‘It’s not hot today.’ (Holton 2014:51)

Table 10.9 provides a selection of negators in languages of the AP family taken from the grammar sketches in Schapper (2014b). These negators, except for possibly Western Pantar *kauwa* are not cognate with any of the Flores-Lembata negators. However, the clause-final position of the AP negators is parallel to the position of several negators in Lamaholot and one of the negators in SK-Hewa. Thus, it remains unknown whether the contact language was an Alor-Pantar language, as no actual language material was borrowed. The contact language could also have been another unknown language that became extinct.

Table 10.9: Negators in the Alor-Pantar languages

Language	Negator
Western Pantar	<i>kauwa</i>
Kaera	<i>bino</i>
Blagar	<i>nian</i>
Adang	<i>nane</i>
Kamang	<i>naa</i>
Sawila	<i>naanuu</i>
Wersing	<i>nanu</i>

I have shown that there is evidence to posit contact-induced grammaticalisation as the cause for clause-final negation in the Lamaholot varieties and in SK-Hewa. The absence of Jespersen Cycles in Austronesian languages of Borneo, Java and further west supports this hypothesis. In addition, the presence of several other potential Papuan features in Lamaholot as discussed in this dissertation strengthens the proposal that contact-induced change led to double and clause-final negation.

An open question is why the Hewa, Sika and Kedang pre-predicate negators are so diverse and why do they show innovations and no reflexes of PMP *ta ‘no, not’ as attested in Lamaholot. While we do not know the answer to these questions, this fact at least suggests that Hewa, Sika and Kedang have distinct histories from the Lamaholot varieties.

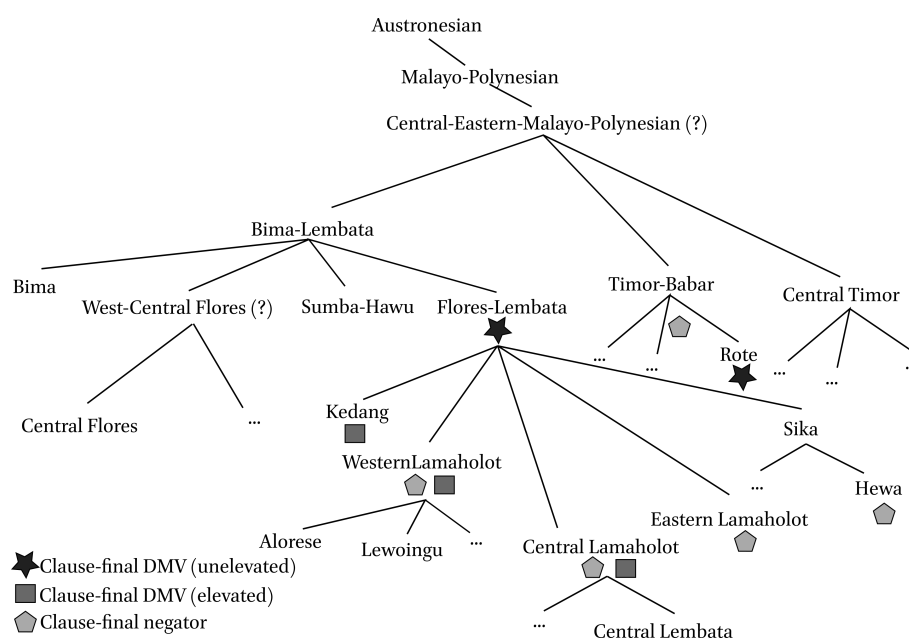
10.4 Conclusions

In this chapter, I discussed two clause-final structural elements in the languages of Flores-Lembata: deictic motion verbs (§10.2) and negation (§10.3). Both features are innovations as the inherited position for verbs is clause-medial following the subject, and for negation, the inherited position is pre-predicate. These two clausal innovations are very likely attributed to contact with non-Austronesian languages.

In Figure 10.3, the innovation of each feature is indicated with a different symbol. The symbol is placed below the name of a subgroup in which all languages show the respective feature. Consequently, the feature can be reconstructed to the proto-language of the subgroup. As the scope of this dissertation does not include Timor languages in detail, in this figure, I do not provide further details in which languages of Timor the features occur. The respective sections of this chapter on the AN languages of Timor provide more details. The tree structure is based on the current knowledge on the languages investigated (cf. §1.3.1).

Clause-final deictic motion verbs were innovated in Proto-Flores-Lembata (PFL) due to contact with a language possibly related or typologically similar to the Alor-Pantar languages. At the level of PFL, only the unelevated deictic motion verbs ‘go’ and ‘come’ were introduced. These verbs have cognates in all Flores-Lembata languages today, with the exception of Kedang missing a cognate for the verb #-ai ‘go’.

Figure 10.3: The innovation of clause-final DMV and negation



Further, in Lamaholot and Kedang, elevated forms encoding higher and lower places were added to the system. This further development could also have been reinforced by continuous contact with an Alor-Pantar language, as these languages also encode elevation in their deictic motion verbs, or at least with a language that has a typologically similar profile to the AP languages.

The reasons for proposing AP languages or typologically similar languages as donors for the innovations above are the following. (i) AP languages also have clause-final verbs. (ii) AP languages have a set of DMV encoding elevation (cf. §10.2.4) similar to what has been described for the Flores-Lembata languages (§10.2.3). (iii) Both features are found in all AP languages, thus are likely to be inherited features from Proto-Alor-Pantar. (iv) There are two potential lexical borrowings from AP languages, Western Lamaholot *dai* 'come; come.up' from PAP *medain 'come.up', and LH-SK #-ai 'go' from PAP *wai 'go' (cf. §10.2.5). However, especially the second potential borrowing remains controversial as in case of borrowing from PAP or one of its descendants, the loss of the initial glide would require an additional step which remains unexplained.

Clause-final negation is exclusively found in the Lamaholot varieties

and in SK-Hewa (cf. §10.3.4). To gain a second, clause-final negator, WL varieties grammaticalised a word meaning ‘wrong; mistake’, CL varieties grammaticalised a word meaning ‘a little’ and EL varieties and SK-Hewa innovated a word of unknown origin. There is some evidence that also this word in EL and SK-Hewa goes back to a word meaning ‘a little’. Assuming that EL *wa* and SK-Hewa *iwa* are cognate with *fa* ‘a little; NEG’ in the Rote languages and Meto varieties, this word originally could have also meant ‘a little’ (cf. §10.3.6.3). In §10.3.7, it has been shown that contact with non-Austronesian languages is a very likely cause for clause-final negation in the Lamaholot varieties and in SK-Hewa. Whether these non-Austronesian languages were related to the Alor-Pantar languages is difficult to prove for this case. It is not implied that every single speech variety was in contact with the donor languages directly but the feature could also have diffused among the Lamaholot varieties and SK-Hewa.

CHAPTER 11

Discussion and conclusions

11.1 Overview

The Flores-Lembata languages are an innovation-defined subgroup and the Flores-Lembata phonology traces back to Proto-Malayo-Polynesian (PMP) with regular sound correspondences in consonants and vowels (cf. Chapter 5). The closest relatives of the Flores-Lembata languages are the Austronesian languages of central and western Flores, Sumba, Sawu and Bima. These languages form together the innovation-defined Bima-Lembata subgroup, as I have shown in §5.5.

While in lexicon and phonology they group together, morpho-syntactically, the Flores-Lembata languages diverge considerably from their western relatives, as I have shown in Part III of this dissertation. A number of their grammatical features are shared with both the Austronesian languages of Timor and with the non-Austronesian Timor-Alor-Pantar languages that are spoken towards the east of the Flores-Lembata family. These features are innovations in the Flores-Lembata languages and they are atypical for Austronesian languages in general. I argue that these aberrant features in the Flores-Lembata languages are the result of contact with non-Austronesian languages in the area. Together with the non-Austronesian vocabulary pre-

sented in Chapter 6, these features are part of a non-Austronesian substrate in the Flores-Lembata languages. However, not all Flores-Lembata languages are influenced to the same extent by this substrate. The Flores-Lembata languages in the geographical centre of the family, the Lamaholot subgroups, have more substrate features than those at the edges.

In §11.2, I summarise and analyse the lexical and structural substrate features in more detail and relate them to Proto-Flores-Lembata or individual subgroups of Flores-Lembata. §11.3, I propose a scenario that could have led to the contact-induced innovations attested. In §11.4, I draw final conclusions and in §11.5, I provide general implications and suggestions for further research.

11.2 Non-Austronesian substrate features

Table 11.1 provides an overview of the contact-induced innovations in the Flores-Lembata languages that were discussed in this dissertation. The table is arranged by the level at which the innovation occurred. Some of the innovations are reconstructible to Proto-Flores-Lembata (PFL) but most are not, thus they must have occurred in lower-level subgroups. Some innovations can only be traced to the level of individual languages.

I divide the features into four types concerning their level of occurrence. First, there are features that trace back to PFL. These features are attested in all five subgroups of Flores-Lembata and within these subgroups, all documented varieties have these features. Second, there are features that are found in more than one subgroup of Flores-Lembata but not in all. These features are attested in all documented varieties of these subgroups. For the purpose of this comparison, I have grouped the three Lamaholot subgroups together into one category. Lamaholot here means that the feature is attested in all three Lamaholot subgroups. However, for the shared lexical items, the occurrence in at least two of the three subgroups was regarded as sufficient. For the feature of alienability, no data on Eastern Lamaholot is available. Thus, here the categorisation as Lamaholot is only based on data from Western Lamaholot (WL) and Central Lamaholot (CL).

Table 11.1: Contact-induced innovations in the Flores-Lembata languages

Level	Feature	Domain	Section
Reconstructible innovations for PFL			
PFL (or higher?)	39 PFL forms	Lexicon	§6.2.3
PFL	POSS-N	Syntax	§9.2
	N-LOC	Syntax	§9.3
	N-NUM	Syntax	§9.4
	Property nouns	Morpho-Syntax	§8.3
	Clause-final DMV ‘come’ and ‘go’	Syntax; lexicon	§10.2
Innovations shared among subgroups			
Lamaholot-Sika	42 shared lexical innovations	Lexicon	§6.3.3
Lamaholot-Kedang	75 shared lexical innovations	Lexicon	§6.3.2
	Clause-final DMV ‘come.up/down’ and ‘go.up/down’	Syntax; lexicon	§10.2
	Traces of non-decimal counting systems	Semantics	§9.4.4
Lamaholot	74 shared lexical innovations	Lexicon	§6.3.4
	Negation with clause-final negator	Syntax; lexicon	§10.3
	Alienability distinction	Morpho-syntax; semantics	§8.2
Reconstructible innovations for individual subgroups			
Kedang	Plural word for animates	Semantics	§8.4.2.3
Central Lamaholot	Plural suffix	Morpho-syntax; semantics	§8.4.2.1
Innovations in individual languages			
SK-Hewa	Plural word for animates	Semantics	§8.4.2.3
	Double negation with clause-final negator	Syntax; lexicon	§10.3
CL-Central Lembata	Additional clause-final verbs	Syntax	§3.7.2.2
WL-Lewoingu	Plural suffix for animates	Morpho-syntax; semantics	§8.4.2.2
WL-Alorese	Plural word	Semantics	§8.4.2.2

Third, there are features that are only attested in one subgroup and reconstructible to its proto-language. These features are attested in all known varieties of the subgroup. Fourth, there are features that are only found in one language or variety within a subgroup. In addition to the level of occurrence, I indicate the linguistic domain for each feature and the section of this dissertation where the feature has been discussed.

The non-Austronesian substrate of PFL mainly affects the syntax of the noun phrase and the numeral phrase, which is part of the noun phrase in a wider sense. I have shown that the changes in the noun phrase, except for the change in the numeral phrase, all go back to the word order change of the possessive construction (§8.3.6.4; §9.5). Another syntactic change took place in PFL. Deictic motion verbs were moved to the end of the clause, while the basic word order of PFL was otherwise verb-medial. Based on current knowledge, the non-Austronesian component of the PFL lexicon is not particularly large with only about 20% of reconstructions not having a clear AN origin. In addition, one third of this non-AN vocabulary is found in other AN languages and thus possibly reconstructs to a higher level, meaning that it was not innovated in PFL.

After the split of PFL into five subgroups, these subgroups must have also been in contact with several substrate languages. The new features at the level of the subgroups not only concern the lexicon and the morpho-syntax of the languages, but also the semantics, as new semantic categories, such as plural number and alienability, were grammaticalised.¹ The non-Austronesian features accumulate towards the centre of the Flores-Lembata language family with Central Lamaholot having most substrate features and also the highest amount of non-Austronesian vocabulary, followed closely by Western Lamaholot (cf. §6.5). By subgroup, the following features were added after the split of PFL.

¹ At the current stage, it remains unclear if the extension of clause-final deictic motion verbs to elevated forms, such as 'go.up' and 'go.down', also includes a semantic innovation or not. Verbs expressing upward or downward motion are attested in virtually all languages of the region. However, the differentiation of a 'come' or 'go' movement, thus a deictic element, may not always be part of the semantics of these verbs. A verb meaning 'ascend', thus expressing an upward movement, could mean 'go.up' as well as 'come.up'. This is still the case for some of the motion verbs in Kedang (cf. §10.2.3.3). As long as the exact semantics of the movement verbs are not known, it cannot be decided if the deictic component in the semantics of the elevated motion verbs is an innovation in Kedang and Lamaholot or has been inherited.

Sika did not gain any additional structural features. However, new lexical items (42 in my database) entered the Sika subgroup. The same lexical items are also attested in at least one Lamaholot subgroup and do not trace back to PFL.

Kedang developed a plural word for animates, non-decimal numerals for 'eight' and 'nine', and extended the clause-final position to other deictic motion verbs that encode elevation. Kedang also gained new lexical items (75 in my database) that are shared with at least one Lamaholot variety.

The Lamaholot subgroups have gained most new features, compared with Sika and Kedang. First, the Lamaholot subgroups show a large number of new lexical items that are partly shared with Sika (42 in my database), partly with Kedang (75 in my database) and many of them only attested among the Lamaholot subgroups themselves (74 in my database). New structural features in the Lamaholot subgroups are an alienability distinction in the possessive construction, a non-decimal numeral for 'eight', clause-final deictic motion verbs that encode elevation and clause-final negators that are used exclusively in WL and in combination with a pre-predicate negator in CL and EL. In addition, Central Lamaholot has developed a plural suffix.

In sum, about 34% of the vocabulary of WL and CL has been added since PFL times. About 24% of the Kedang vocabulary and about 19% of the Sika vocabulary has been added since that time. WL and CL have each gained two new syntactic features (clause-final negation and elevated deictic motion verbs) and three new semantic features (alienability, plural marking and a non-decimal numeral), Kedang has gained one syntactic feature (clause-final elevated deictic motion verbs) and two semantic feature (plural marking for animates, and non-decimal numerals), while Sika has not gained any new structural features since PFL.²

Following from the summaries above, three main types of substrate effects that affect more than one subgroup can be differentiated. (i) The main substrate effect in the west is the addition of lexical items, and it affects Sika and Lamaholot. (ii) In the east, the substrate effect also had a large lexical component which is attested in Kedang and Lamaholot. Additionally, the extension of the clause-final deictic motion verbs to elevated motion verbs,

² I am not taking into account Eastern Lamaholot here because too little data is available for this subgroup to make the numbers comparable.

the development of the non-decimal numerals ‘eight’ and ‘nine’, and the development of a plural word in Kedang can be attributed to contact. (iii) In the centre of the Flores-Lembata area, the largest amounts of substrate effects are attested. In addition to an equally large amount of lexical items, the alienability distinction, clause-final negation in all Lamaholot subgroups and the plural marking in Central Lamaholot are substrate effects.

On the level of individual languages within the subgroups, further innovations are attested. However, it is not clear whether innovations in individual varieties can be attributed to contact with non-Austronesian substrate languages. SK-Hewa has gained a clause-final negator and a plural word for animates. WL-Lewoingu has developed a plural suffix for animates. CL-Central Lembata has added more intransitive clause-final verbs with locative phrases and WL-Alorese has developed a plural word. The history of the WL-Alorese plural word is well attested. The Alorese speakers developed this plural word due to contact with speakers of Alor-Pantar language during the past 500-600 years since they moved to the islands of Alor and Pantar (Moro 2018). The additional clause-final verbs in CL-Central Lembata may have been developed by analogy with the clause-final DMV that already existed. However, the developments of the clause-final negator in SK-Hewa and the plural markers in Sika-Hewa and WL-Lewoingu remain obscure.

11.3 Contact scenarios

Depending on the circumstances, contact-induced language change can affect any feature of a language (Thomason and Kaufman 1988:14). The social scenario in which the contact takes place plays an important role in determining constraints on contact-induced change for a particular contact situation (Muysken 2010). Analysing the outcome of languages contact, such as the innovated features of the Flores-Lembata languages discussed in this dissertation, a possible contact scenario can be reconstructed.

The contact outcomes in case of the Flores-Lembata languages are new vocabulary, morpho-syntactic changes and grammaticalisation of new semantic categories. As shown in §11.2, the subgroups of Flores-Lembata have been affected to different degrees by these innovations. Most affected is Central Lamaholot, followed by Western Lamaholot and then Kedang. Least affected is Sika. Contact-induced innovations that can be reconstructed to

PFL are morpho-syntactic changes and new vocabulary but not grammaticalisation of new semantic categories.

Convergence in word order and new morpho-syntactic categories based on semantic distinctions have been described as a result of prolonged bilingualism over several generations, involving all age groups in the society (Muysken 2010:272). The changes in the syntax of the noun phrase in PFL, as well as the position of deictic motion verbs and the new negators in some of the Flores-Lembata languages are results of this kind of convergence in word order. The development of the new grammatical categories plural number and alienability in possessive constructions are examples of the emergence of new morpho-syntactic categories. The phenomena observed in the Flores-Lembata languages fall under bilingual copying. According to Ross (2013:23), there are three stages of bilingual copying: (i) lexical calquing (loan translation), (ii) grammatical calquing and (iii) syntactic restructuring (=metatypy). As the source languages are not spoken any more, it is difficult to prove calquing in the Flores-Lembata languages. However, some of the new features could indeed be the result of calquing. The grammaticalisation of words meaning 'a little' in Central Lamaholot and 'wrong' in Western Lamaholot to negators could be a case of lexical calquing. It is possible that the source language also used a word meaning 'a little' or 'wrong' to express negation. Additionally, the Sika expression *alan ?rouy* for 'hair' literally meaning 'leaf of the head' is very likely an example of lexical calquing (cf. Donohue and Grimes 2008:147-148). Only bilingual speakers who are fluent in both languages are able to recognise such a connection and copy it.

The development of different constructions for alienable and inalienable possession could be the result of grammatical calquing which translates a construction morpheme by morpheme (Ross 2013:23). The grammaticalisation of the 3PL pronoun into a plural marker in Central Lamaholot could also be a case of lexical calquing that led to grammatical calquing. However, this grammaticalisation pathway is also universally attested. Therefore, the argument for calquing is less strong here. All other changes are cases of syntactic restructuring: The changes in the noun phrase, as well as the clause-final position of the deictic motion verbs. These changes do not reach the stage of metatypy where the whole syntax of the language is restructured. However, the noun phrase is affected considerably and some verbs are moved. Also the new negator in the Lamaholot varieties is placed clause-finally.

The proposed prolonged bilingual situation led to additional grammatical features in some of the Flores-Lembata languages. In PFL only syntactic changes are attested but no additional features. The same holds for Sika. In Kedang and the Lamaholot varieties, features were added and this means an increase in complexity (Ross 2013:32).

PFL and its descendants have all added new vocabulary. However, the increase of new lexical items in PFL and Sika is lower than in Kedang and the Lamaholot varieties. The large amount of new vocabulary could be a remnant of code-switching by highly proficient bilinguals. The new vocabulary is basic as well as special vocabulary. No specific semantic domain is favoured. A social situation that can lead to such an unsystematic mixing of vocabulary is a community where all speakers are bilinguals and where code-switching is the most common form of communication. This concern in particular congruent lexicalisation, a form of code-switching by fluent bilinguals where lexical items from two or more sources are randomly inserted into a common frame (Muysken 2008:364). The "fossilisation" of such type of code-switching can lead to a so-called bilingual mixed language (Thomason 2001:198,215).

The new structural features, as well as the additional vocabulary, point to bilingual communities with more than one contact scenario of a similar kind. PFL is most likely the result of bilingual mixing, as are Kedang and the Lamaholot subgroups.

For Sika, this is less clear. Possibly, Sika has a contact history which is different to the Lamaholot varieties and Kedang. There is some new vocabulary in Sika but no significant new structural features. In addition, the morphology of Sika appears to be much more simplified than in Kedang and Lamaholot. Only traces of person marking on verbs are left in Sika (Rosen 1986), while Lamaholot and Kedang have a much more elaborate system (cf. Samely 1991a; §3.4; §3.5). Therefore, the case of Sika points to simplification rather than complexification over time. This could be a sign of rapid language shift (Ross 2013:30,37). Only a short period of bilingualism with more adult learners than children may have preceded language shift. This situation did not allow for the addition of new features because additional grammatical features are usually the result of prolonged bilingualism involving children and adolescents as stated above.

A similar contact scenario to the one proposed for Sika could have affected Proto-Central Flores, as in the Central Flores languages virtually all

morphology is lost. In addition, these languages have a non-decimal counting system which has been attributed to non-Austronesian contact (cf. Elias 2018:119; §9.4.4). Also similar contact effects, namely morphological simplification, have been attested in Alorese (cf. Klammer 2012b). For this language, a large number of second language learners has been attested which most likely has caused the simplification (Moro 2019).

Nowadays no non-Austronesian languages are spoken anymore in the whole area of Flores-Lembata. Therefore, also the contact scenarios of Proto-Kedang, and the proto-languages of Lamaholot must have reached the stage of language shift towards the Austronesian languages at some point. When finally all speakers switched, the languages had already been heavily influenced by the non-Austronesian languages due to a long and intensive period of bilingualism. It may even be possible that, as the whole society became bilingual, speakers did not differentiate the languages any more but the mixed code became their new language. Nevertheless, the Kedang and Lamaholot languages remain overall more Austronesian than non-Austronesian in lexicon and grammar. Therefore, assuming a mixed code does not mean an equal mix that leads to doubts on the genealogical affiliation of these languages. However, the non-Austronesian component in lexicon and grammar is considerably large, going beyond some instances of borrowing. This amount of non-AN features suggest a language mixing based on long-term bilingualism with code-switching practices, at least up to a certain degree.

From the comparative study in Part III, it appears that the non-Austronesian substrate languages that influenced the Flores-Lembata languages must have been typologically very similar to the Timor-Alor-Pantar languages. This typological similarity alone is not enough evidence to conclude that these substrate languages belonged to the TAP family. There are a few clear cases of related lexical items between the non-Austronesian vocabulary of Flores-Lembata and the vocabulary of the TAP family, especially the Alor-Pantar (AP) branch, as I have indicated in Chapter 6. But some of them may also be borrowings from Flores-Lembata languages, especially from Alorese, into AP languages. In addition, I suggested in §10.2.5 that the deictic motion verbs #-ai 'go' in the Lamaholot subgroups and Sika and *dai* 'come up' in Western Lamaholot could have come from an AP language. However, the current lexical evidence is too weak to conclude that the non-Austronesian substrate languages were related to the AP languages. It is also possible that they were part of a different non-Austronesian family that was also in con-

tact with the AP languages.

Finally, I have argued that PFL as well as its subgroups were in contact with non-Austronesian languages that caused structural and lexical change. The time and place of these contact scenarios is a question that linguistics alone is probably not able to answer entirely. Nevertheless, there are a few indications about timing. It is known that the first speakers of Austronesian languages arrived in the area of East Nusa Tenggara and Timor-Leste about 3,500 years ago (Pawley 2005:17). As PFL descends from a proto-language also covering several Austronesian languages towards the west of PFL on the islands of Flores and Sumba, it might have been this proto-language, Proto-Bima-Lembata (PBL), whose speakers arrived around 3,500 years ago. The point in time when PFL was spoken must thus have been later than that, allowing enough time for PBL to diversify into several subgroups, one of them being PFL. PFL was then acquired by speakers of non-Austronesian languages with different syntactic patterns. It is unknown if these non-Austronesian languages were related to the ones which later influenced the descendants of PFL. After PFL broke up into five subgroups, these either remained in contact with the same non-Austronesian languages and/or underwent contact with other non-Austronesian languages. However, the intensity of contact differed. Given that Proto-Sika did not gain any new structural features after it split off from PFL, this indicates that the contact Sika experienced was less intense and most importantly did not involve much long-term and pre-adolescent bilingualism. Given that Proto-Kedang, PWL and PCL all gained more new features also after their split off from PFL, this indicates that these subgroups experienced more contact involving long-term bilingualism and children acquiring more than one language at the same time.

Concerning the question where the Flores-Lembata family originates, the “centre of gravity principle” can be taken as an indicator (Sapir 1916:79; Robbeets and Savelyev 2017:6-8). According to this principle, the place where the highest diversity of a language family is attested is likely to be the homeland of the language family, the point from where the proto-language dispersed. For the Flores-Lembata family, the diversity hotspot is clearly the island of Lembata, where four out of five subgroups are attested. Only Sika is not spoken on Lembata. As the island of Lembata appears to be the homeland of Proto-Flores-Lembata, as well as the present-day home of the subgroups that have most substrate features, this all points to the historic pres-

ence of non-Austronesian languages on the island of Lembata.

PFL speakers came to this island, and were in contact with speakers of non-Austronesian languages. These languages co-existed further until after the break-up of PFL into subgroups. All the subgroups, except for Sika, gained more non-Austronesian features after the split. This suggests that Proto-Sika speakers left Lembata first as is also corroborated by the current location of the Sika varieties, which are located furthest away from Lembata on the island of Flores. There might have been contact between Sika and other non-Austronesian languages but if there was, it was much less intense than on the island of Lembata.

Proto-Kedang speakers moved to the eastern edge of Lembata island and Proto-Western Lamaholot speakers moved out towards the west as the Proto-Sika speakers did before them. But before that, they had gained more non-Austronesian features. It remains unknown until how recently the non-Austronesian languages on Lembata might have been spoken. However, it must have been at least several hundred years ago that they died out, as there are no known records of them in oral or written sources.

11.4 Summary of main conclusions

PFL and its subgroups have been in contact with non-Austronesian languages. This contact scenario was characterised by long-term bilingual communities with fluent bilinguals of all age groups. As a result of this language contact, PFL and its descendants gained new grammatical features and lexical items. These speakers either finally shifted to the Flores-Lembata languages or their bilingual code became the “new” language. Slightly different scenarios are assumed for PFL and the proto-languages of the lower-level subgroups. Possibly, Proto-Sika is the result of more rapid language shift, while subgroups with more non-Austronesian features, such as Central Lamaholot, may have had a longer period of bilingualism with the result of being a bilingual mixed language. The non-Austronesian contact languages were typologically similar to the neighbouring non-Austronesian Timor-Alor-Pantar languages that are still spoken today. But it remains unclear if the Flores-Lembata substrate languages were closely related to the TAP languages as there is no sufficient lexical evidence for such a connection.

11.5 Directions for further research

The internal structure of the Bima-Lembata group was not within the scope of this dissertation but is an important step in further revealing the history of the Austronesian languages of East Nusa Tenggara. Also the wider spread of the non-AN vocabulary in PFL needs to be addressed in more detail.

From a broader perspective, an interesting point for further research are the Austronesian languages of Timor which show a strikingly similar typological picture to the Flores-Lembata languages. In addition, the vocabulary of the AN Timor languages also appears to be mixed (Edwards in prep) and two layers of regular sound correspondences have been attested (Edwards 2016b; Edwards 2018b). As the results of this dissertation show, the Flores-Lembata languages and the Austronesian languages of Timor do not have a shared history. However, they appear to have undergone a similar history of contact with non-Austronesian languages.

For the Flores-Lembata languages in particular, more lexical and most importantly also structural data of Eastern Lamaholot varieties will allow a more fine-grained picture of the three Lamaholot subgroups. Further research should also investigate in more detail shared lexical items between the individual Lamaholot subgroups with Kedang and Sika. All possible combinations among the five Flores-Lembata subgroups need to be considered. For the lexical analysis of shared lexical innovations in Chapter 6, the Lamaholot subgroups had been considered as a unit. This was done because they are located in the centre of the family and they are lexically more similar to each other than to Sika and Kedang. Also the limited data on Eastern Lamaholot would have made it impossible to consider the three Lamaholot subgroups in a comparable way.

In general, this dissertation has shown that it is important to take all major parts of languages into account, the lexicon, the phonology and the morpho-syntax, to reveal their history. In particular, it has been proven to be of high relevance to investigate not only inherited features but also those that entered the languages at various stages of their history. For this, the genealogical affiliation of the languages needs to be known. Without knowing to which subgroup a language belongs, it is impossible to decide whether a feature is inherited or innovated. Once such an investigation is done, it enables us to reconstruct parts of the social histories of language contact in an-

cient times on the basis of languages as they are spoken today. In particular the Austronesian languages of eastern Indonesia, with virtually no records of their past, have proven to be fruitful ground for such investigations.

APPENDICES

APPENDIX A

Central Lembata phoneme minimal pairs

A.1 Overview

In this appendix, the phonemic status of the consonants and vowels in Central Lembata, introduced in §3.2.1, is demonstrated by giving minimal pairs of phonetically close oppositions. Stress is not marked in the phonemic transcriptions because stress in Central Lembata is not always stable (cf. §3.2.5).

For the consonants, I start with the discussion of plosives in §A.2 and nasals in §A.3. Subsequently, the fricatives and the only affricate are discussed together in §A.4, as they are articulatorily close and form a natural class. Finally, the trill and the lateral are discussed together under the common label ‘liquid’ in §A.5. For the vowels, I first discuss the high vowels /i/ and /u/ in Section A.6, then mid vowels /e/, /o/ and /ə/ in §A.7 and finally the low vowel /a/ in §A.8.

In the tables with minimal pairs in the following sections, I mark a gap with n.a. (=not applicable) when no minimal pair can be found because a sound cannot occur in the respective position. An example are minimal pairs of /p/ and /b/ in final position. As voiced /b/ cannot occur in final position, no final minimal pairs with /p/ and /b/ is possible. A gap due to missing data is marked with [...].

A.2 The plosives /p, t, k/ and /b, d, g/

Table A.1 shows the plosives in all permitted positions and provides minimal pairs to prove their phonemic contrast. For all three voicing contrasts, no minimal pair in final position can be provided because voiced plosives cannot appear in final position (cf. §3.2.2).

Table A.1: Voicing contrast between plosives

Phonemes	Position	Minimal pair	Gloss
/p/ - /b/	initial	/pu.uk/	‘stem’
		/bu.uk/	‘night’
	intervocalic	/ə.pak/	‘palm (of hand/foot)’
		/ə.bak/	‘small river’
	final	/o.kap/ n.a.	‘chase’
/t/ - /d/	initial	/ta.ni/	‘weave’
		/da.ni/	‘hit (a drum)’
	intervocalic	/pa.ta/	‘boil’
		/pa.da/	‘fast, forbid’
	final	/mə.rit/ n.a.	‘knife’
/k/ - /g/	initial	/ka.i/	‘say’
		/ga.i/	‘painful’
	intervocalic	/so.ka/	‘dance’
		/so.ga/	‘lift’
	final	/ma.nuk/ n.a.	‘chicken’

n.a. = not applicable

A.3 The nasals /m, n, ŋ/

Table A.2 provides data that show the phonemic contrast between the bilabial /m/ and the alveolar /n/ in all positions, while Table A.3 provides evid-

ence for the phonemic contrast of /n/ and /ŋ/ in intervocalic and final position. The velar nasal /ŋ/ cannot occur in initial position.

Table A.2: Contrast between the nasals /m/ and /n/

Phonemes	Position	Minimal pair	Gloss
/m/ - /n/	initial	/ma.i/	‘thirsty’
		/na.i/	‘breathe’
	intervocalic	/ta.ma/	‘enter’
		/ta.na/	‘soil’
	final	/pa.dʒam/	‘papaya’
		/pa.dʒan/	‘accompany’

Table A.3: Contrast between the nasals /n/ and /ŋ/

Phonemes	Position	Minimal pair	Gloss
/n/ - /ŋ/	initial	/nə.bu/	‘not thirsty’
		n.a.	
	intervocalic	/na.ne/	‘3SG’
		/na.ŋe/	‘swim’
	final	/ə.van/	‘accuse’
		/ə.vaŋ/	‘animal’

n.a. = not applicable

A.4 The fricatives /v, s/ and the affricate /dʒ/

Table A.4 shows the phonemic status of the coronal fricative /s/ by contrasting it to the coronal plosive /t/. The voiceless plosive /t/ is chosen for contrast as it only differs from /s/ in manner of articulation. In final position, a minimal pair contrasting final /s/ and /t/ could not be found because final /s/, with less than 30 occurrences in my corpus, is rather rare. However, a minimal pair contrasting /s/ with /k/ could be found, namely /pu.us/ ‘wash’ and /pu.uk/ ‘stem’.

Table A.4: Contrast between the alveolars /s/ and /t/

Phonemes	Position	Minimal pair	Gloss
/s/ - /t/	initial	/si.va/	‘nine’
		/ti.va/	‘throw away’
	intervocalic	/tu.su/	‘breast’
		/tu.tu/	‘tell’
	final	/pu.us/ [...]	‘wash’

[...] = gap due to insufficient data

Table A.5 shows the phonemic status of the labiodental /v/ by contrasting it with the bilabial stop /b/. The voiced stop /b/ is chosen to demonstrate the contrast because it only differs from /v/ in manner of articulation and slightly in place of articulation. As the voiced stop /b/ does not occur in final position, minimal pairs are only given for initial and intervocalic position.

Table A.5: Contrast between the labiodental /v/ and bilabial /b/

Phonemes	Position	Minimal pair	Gloss
/v/ - /b/	initial	/vi.ti/	‘goat’
		/bi.ti/	‘execute a kind of weaving process’
	intervocalic	/na.vak/	‘body’
		/na.bak/	‘big coconut bowl’
	final	/ma.luv/ n.a.	‘be hungry’

n.a. = not applicable

Table A.6 shows the phonemic contrast between the affricate /dʒ/ and the plosive /d/. The voiced plosive /d/ is chosen to provide evidence for the phoneme status of /dʒ/ as these two phonemes only differ in manner of articulation and slightly in place of articulation.

Table A.6: Contrast between postalveolar /dʒ/ and alveolar /d/

Phonemes	Position	Minimal pair	Gloss
/dʒ/ - /d/	initial	/dʒu/	‘downwards’
		/du.u/	‘see’
	intervocalic	/pu.dʒu/	‘tie’
		/pu.du/	‘wind’ (v.)
	final	/so.badʒ/ n.a.	‘swell’

n.a. = not applicable

Initial /dʒ/ is marginal. Therefore, only a near minimal pair could be found for this position.

A.5 The liquids /l, r/

Table A.7 provides minimal pairs contrasting the liquids /l/ and /r/ with each other.

Table A.7: Contrast between the liquids /l/ and /r/

Phonemes	Position	Minimal pair	Gloss
/l/ - /r/	initial	/le.do/	‘be tired’
		/re.do/	‘shake’
	intervocalic	/a.la/	‘make a sound’
		/a.ra/	‘but’
	final	/ma.val/	‘kind of basket’
		/ma.var/	‘share (a field)’

A.6 The high vowels /i, u/

The following minimal pairs show the phonemic status of the high vowels /i/ and /u/ by contrasting them with their phonetically close counterparts

of /e/ and /o/ in Table A.8 and Table A.9 respectively, then further with each other in Table A.10.

Table A.8: Contrast between the front vowels /i/ and /e/.

Phonemes	Position	Minimal pair	Gloss
/i/ - /e/	initial	/i.kan/	‘fish’
		/e.kan/	‘garden; condition’
	medial	/ki.a/	‘first; now’
		/ke.a/	‘turtle’
	final	/kpa.i/	‘remember’
		/kpa.e/	‘be silent’

Table A.9: Contrast between the back vowels /u/ and /o/

Phonemes	Position	Minimal pair	Gloss
/u/ - /o/	initial	/u.la/	‘snake’
		/o.la/	‘cut’
	medial	/snu.kar/	‘difficult’
		/sno.kar/	‘digging stick’
	final	/lo.du/	‘wash’
		/lo.do/	‘descend’

Table A.10: Contrast between the high vowels /i/ and /u/

Phonemes	Position	Minimal pair	Gloss
/i/ - /u/	initial	/i.da/	‘ladder’
		/u.da/	‘reach, touch’
	medial	/ni.dʒu/	‘nose’
		/nu.dʒu/	‘spoon’
	final	/di.i/	‘also’
		/du.u/	‘see’

A.7 The mid vowels /e, ə, o/

Having contrasted the mid vowels /e/ and /o/ with their high vowel counterparts in the previous section, I contrast them here with each other in Table A.11. In Table A.12 and Table A.13, I contrast the central vowel /ə/ with /e/ and /o/ respectively. In Table A.12, there is a data gap as /ə/ and /e/ both occur word initially but no minimal pair could be found. This might be due to the fact that vowel initial words are less frequent than consonant initial words in general. The vowel schwa /ə/ does not appear in final position, therefore no minimal pairs with schwa can be provided for final position.

Table A.11: Contrast between the mid vowels /e/ and /o/

Phonemes	Position	Minimal pair	Gloss
/e/ - /o/	initial	/e.do/	‘shake’
		/o.do/	‘give’
	medial	/te.na/	‘boat’
		/to.na/	‘star’
	final	/tu.ne/	‘one’
		/tu.no/	‘grill’

Table A.12: Contrast between the mid vowels /ə/ and /e/

Phonemes	Position	Minimal pair	Gloss
/ə/ - /e/	initial	[...]	
	medial	/rə.kət/	‘sharp’
		/rə.ket/	‘bite into pieces’
	final	n.a.	

[...] = gap due to insufficient data, n.a. = not applicable

Table A.13: Contrast between the mid vowels /ə/ and /o/

Phonemes	Position	Minimal pair	Gloss
/ə/ - /o/	initial	/ə.bak/	‘river’
		/o.bak/	‘spill’
	medial	/sə.ba/	‘wet’
		/so.ba/	‘wear’
	final	n.a.	

n.a. = not applicable

A.8 The low vowel /a/

In this section, I contrast the central low vowel /a/ with the central mid vowel /ə/ in Table A.14. Further, I contrast the low vowel /a/ with the front mid vowel /e/ in Table A.15 and with the back mid vowel /o/ in Table A.16. There are no minimal pairs with schwa /ə/ in final position as, phonemically, this vowel never occurs in final position.

Table A.14: Contrast between the central vowels /a/ and /ə/

Phonemes	Position	Minimal pair	Gloss
/a/ - /ə/	initial	/a.nəm/	‘thing’
		/ə.nəm/	‘six’
	medial	/pa.ta/	‘boil’
		/pə.ta/	‘split’
	final	n.a.	

n.a. = not applicable

Table A.15: Contrast between the low vowel /a/ and the mid vowel /e/

Phonemes	Position	Minimal pair	Gloss
/a/ - /e/	initial	/a.an/	‘what’
		/e.en/	‘only’
	medial	/ba.ra/	‘carry’
		/be.ra/	‘fast’
	final	/to.ba/	‘fall’
		/to.be/	‘sit’

Table A.16: Contrast between the low vowel /a/ and the mid vowel /o/

Phonemes	Position	Minimal pair	Gloss
/a/ - /o/	initial	/a.pi	‘fire’
		/o.pi	‘buy’
	medial	/ba.ti	‘hunt’
		/bo.ti	‘assemble’
	final	/la.ka/	‘keep; not give away’
		/la.ko/	‘civet cat’

APPENDIX B

Flores-Lembata lexeme sets

This appendix contains the lexeme sets that are used in Part II of this dissertation. The sets are extracted from the LexiRumah database (Kaiping et al. 2019) with occasional additions from other sources. For a full list of the Flores-Lembata varieties used here, their sources and abbreviations see §4.2. The forms in this appendix are not retranscribed and are given in IPA as they are transcribed in the source. The sets are sorted alphabetically by the concept that is assigned to the lexemes in LexiRumah, here given in bold. As this appendix is sorted by semantic concept and not by formal lexeme set, the same set can appear under more than one concept.

The Proto-Flores-Lembata (PFL) forms (marked by an asterisk) and the potential reconstructions that are at the current stage of research not reconstructible (marked by a hashtag) are taken from Part II of this dissertation. Other proto-forms, such as Proto-Malayo-Polynesian (PMP) and other Austronesian subgroups, are not indicated in this appendix. For those lexeme sets which are discussed in Part II, these proto-forms are given there. This appendix also contains several lexeme sets without reconstruction but which are potentially reconstructible to PFL or other (higher-level) proto-languages. These sets have not been included into the main body of this dissertation, therefore no reconstruction has been made. As they were part of the database, they are nevertheless included in this appendix.

1pl excl:

(1) PFL ***kami**, CL-IL *kame*, CL-KK *kam*, CL-KK *kame*, CL-KK 'kame, CL-LK *kam*, CL-LP *kame*, CL-LR *kam*, CL-LR *kame*, CL-LWK *kame*, CL-LWT *kame*, CL-MR *kame*, CL-PN *kam*, EL-LE *ome*, EL-LMT *ame*, KD-LB *ke:*, KD-LW *ε*, KD-LW *kε*, SK-HEW *ami*, SK-HEW ?*ami*, SK-MM ?*ami*, SK-TA ?*ami*, WL-AB *kame*, WL-AD 'kame, WL-BL *kame*, WL-BM *kame*, WL-BN *kame*, WL-BP *kame*, WL-BT *kame*, WL-DL *kame*, WL-HRW *kame*, WL-IA *kame*, WL-KO *kame*, WL-LG *kame*, WL-LH *kame*, WL-LMK *kame*, WL-LML *kame*, WL-LTB *kame*, WL-LWI *ka'mε*, WL-LWL *kame*, WL-MD *kame*, WL-ML *kame*, WL-MS *kame*, WL-PD *kame*, WL-PK *kame*, WL-RE *kame*, WL-TJ *kame*, WL-WB *kame*, WL-WK *kame*, WL-WT *kame*, WL-WW *kame*.

1pl incl:

(1) PFL ***kita**, CL-IL *tite*, CL-KK *tite*, CL-KK 'tite, CL-LK *tite*, CL-LP *tite*, CL-LR 'tite, CL-LWK *tite*, CL-LWT *tite*, CL-MR *tite*, CL-PN *tite*, EL-LE *gite*, EL-LMT *gite*, KD-LB *te:*, KD-LW *tε*, SK-HEW *ita*, SK-HEW ?*ita*, SK-MM ?*ita*, SK-TA ?*ita*, WL-AB *ite*, WL-AD 'tite, WL-BL *tite*, WL-BM *tite*, WL-BN ,?i'tε, WL-BP *tite*, WL-BT *tite*, WL-DL *tite*, WL-HRW *tite*, WL-IA *tite*, WL-KO *tite*, WL-LG *tite*, WL-LH *tite*, WL-LMK *tite*, WL-LML *tite*, WL-LTB *tite*, WL-LWI 'tite, WL-LWL *tite*, WL-MD *tite*, WL-ML *tite*, WL-MS *ite*, WL-MS *tite*, WL-PD *tite*, WL-PK *tite*, WL-RE *tite*, WL-TJ *tite*, WL-WB *tite*, WL-WK *tite*, WL-WT *tite*, WL-WW *tite*.

1sg:

(1) PFL ***aku**, CL-IL *go*, CL-KK *go*, CL-KK *gone*, CL-KK *gone*, CL-LK *go*, CL-LP *go*, CL-LR *gɔn*, CL-LWK *gone*, CL-LWT *go*, CL-MR *go*, CL-PN *gon*, EL-LE *goe*, EL-LMT *go?e*, KD-LB *ε?i*, KD-LW *εi*, KD-LW *ε?i*, SK-HEW *au*, SK-HEW *a?u*, SK-MM *a?u*, SK-TA *a?u*, WL-AB *go*, WL-AD 'go?ε, WL-BL *goe*, WL-BM *go*, WL-BN *go:*, WL-BP *go?e*, WL-BT *go?e*, WL-DL *go*, WL-HRW *go?e*, WL-IA *go*, WL-KO *go?e*, WL-LG *go*, WL-LH *go?e*, WL-LMK *go?e*, WL-LML *goe*, WL-LTB *go*, WL-LWI *go:*, WL-LWL *go?e*, WL-MD *go?e*, WL-ML *go?e*, WL-MS *go*, WL-PD *go*, WL-PK *go*, WL-RE *go?e*, WL-TJ *go*, WL-WB *go?e*, WL-WK *go?e*, WL-WT *go?e*, WL-WW *goe*.

2pl:

(1) PFL ***miu**, CL-KK 'mio, CL-LR 'mio vo'lām, KD-LB *me:*, KD-LW *mε*, SK-HEW *miu*, SK-MM *miu*, SK-TA *miu*, WL-AB *mi*, WL-AD 'mio, WL-BN ,mi sa'kali, WL-LWI 'mio, WL-MS *mi*, WL-PD *mi*.

2sg informal:

(1) PFL ***mo**, CL-IL *mo*, CL-KK *mo*, CL-KK *mone*, CL-LK *mo*, CL-LP *mo*, CL-LR *mɔn*, CL-LWK *mone*, CL-LWT *mo*, CL-MR *mo*, CL-PN *mon*, EL-LE *moʔe*, EL-LMT *moʔe*, KD-LB ɔ, KD-LW ɔ, WL-AB *mo*, WL-AD 'moʔe, WL-BL *moe*, WL-BM *mo*, WL-BN *mɔ*, WL-BP *moʔe*, WL-BT *moʔe*, WL-DL *moʔe*, WL-HRW *mo*, WL-IA *mo*, WL-KO *moʔe*, WL-LG *mo*, WL-LH *moʔe*, WL-LMK *moʔe*, WL-LML *moe*, WL-LTB *mo*, WL-LWI *mo*, WL-LWL *moʔe*, WL-MD *moʔe*, WL-ML *moe*, WL-MS *mo*, WL-PD *mo*, WL-PK *mo*, WL-RE *moʔe*, WL-TJ *mo*, WL-WB *moʔe*, WL-WK *moʔe*, WL-WT *mo*, WL-WW *moe*.

(2) PFL ***miu**, CL-IL *mio*, CL-LK *mio*, CL-LP *mio*, CL-LWT *mio*, CL-PN *mio*, EL-LE *mio*, SK-HEW *miu*, WL-BM *mio*, WL-BP *mio*, WL-BT *mio*, WL-DL *mio*, WL-HRW *mio*, WL-KO *mio*, WL-LG *mio*, WL-LMK *mio*, WL-LML *mio*, WL-LTB *mio*, WL-LWL *mio*, WL-ML *mio*, WL-PK *mio*, WL-RE *mio*, WL-TJ *mio*, WL-WB *mio*, WL-WK *miʔo*, WL-WT *mio*, WL-WW *mio*.

(3) PFL ***kau**, SK-HEW ʔau, SK-TA ʔau.

3pl:

(1) PFL ***hida**, CL-IL *daro*, CL-KK *da*, CL-KK *dane*, CL-LK *dan*, CL-LP *daro*, CL-LR 'dane, CL-LWK *jane*, CL-LWT *da*, CL-MR *daro*, CL-PN *dane*, EL-LE *roʔe*, EL-LMT *roʔe*, KD-LW *suo*, SK-HEW *rimu*, SK-MM *rimu*, SK-TA *rimu*, WL-AD 'raʔe, WL-BL *rae*, WL-BM *ra*, WL-BP *raʔe*, WL-BT *raʔe*, WL-DL *raʔe*, WL-HRW *raʔe*, WL-IA *ra*, WL-KO *raʔe*, WL-LG *ra*, WL-LH *raʔe*, WL-LMK *raʔe*, WL-LML *rae*, WL-LTB *ra*, WL-LWI 'ra, WL-LWL *raʔe*, WL-MD *raʔe*, WL-ML *rae*, WL-PK *ra*, WL-RE *raʔe*, WL-TJ *ra*, WL-WB *raʔe*, WL-WK *raʔe*, WL-WT *raʔe*, WL-WW *rae*.

(2) PAL ***we**, WL-AB *fe*, WL-BN *fe: sa'kali*, WL-MS *we*, WL-PD *we*.

3sg:

(1) PFL ***na**, CL-IL *na*, CL-KK *na*, CL-KK *nane*, CL-LP *na*, CL-LWK *nane*, CL-LWT *na*, CL-MR *na*, CL-PN *nane*, EL-LE *noʔe*, EL-LMT *noʔe*, KD-LB *nuc*, KD-LW *ni*, KD-LW *nuc*, SK-HEW *nimu*, SK-TA *nimu*, WL-AB *no*, WL-AD 'naʔe, WL-BL *nae*, WL-BM *na*, WL-BN *nɔ*, WL-BP *naʔe*, WL-BT *naʔe*, WL-DL *na*, WL-HRW *naʔe*, WL-IA *na*, WL-KO *naʔe*, WL-LG *na*, WL-LH *nae*, WL-LMK *naʔe*, WL-LML *nae*, WL-LTB *na*, WL-LWI *na*, WL-LWL *naʔe*, WL-MD *naʔe*, WL-ML *nae*, WL-MS *no*, WL-PK *na*, WL-RE *naʔe*, WL-TJ *naʔe*, WL-WB *naʔe*, WL-WK *naʔe*, WL-WT *naʔe*, WL-WW *naʔe*.

(2) WL-MS *ro*, WL-PD *ro*.

above:

(1) LH #blolo/golo ‘tall’, CL-KK *'lolo*, CL-KK *'golo*, WL-AB *ɔɔ 'lɔɔŋ*, WL-AB *iti 'lɔɔŋ*, WL-AD *'bloʔlõ*, WL-AD *lolõ*, WL-BN *lɔɔŋ*, WL-LWI *teti 'lɔɔŋ*, WL-MS *lɔɔŋ*, WL-PD *lɔɔŋ*.

(2) LH-SK #dʒe ‘upwards; above’, CL-KK *dʒe*, CL-KK *dʒene*, CL-LR *rə'dʒe*.

(3) SK-HEW *lala nan*, SK-HEW *lala vutun*.

alive:

(1) PFL *modip, CL-IL *modʒipa*, CL-KK *modʒip*, CL-KK *mo'dʒip*, CL-LK *modʒip*, CL-LP *modʒip*, CL-LR *modʒi'pən*, CL-LWK *modʒip*, CL-LWT *modʒip*, CL-MR *modʒip*, CL-PN *modʒipən*, EL-LE *mori*, EL-LMT *mori*, SK-HEW *moret*, SK-MM *moret*, SK-TA *morə-t*, WL-AB *məri*, WL-AD *'morit*, WL-BL *morip*, WL-BM *mori*, WL-BN *məri*, WL-BP *mori*, WL-BT *mori*, WL-DL *morʔi*, WL-HRW *morit*, WL-IA *moriw*, WL-KO *morip*, WL-LG *mori*, WL-LH *morip*, WL-LMK *morʔi*, WL-LML *mori*, WL-LTB *mori*, WL-LWI *'məri*, WL-LWL *mori*, WL-MD *mori*, WL-ML *mori*, WL-MS *mərik*, WL-PD *mərik*, WL-PK *mori*, WL-RE *mori*, WL-TJ *mori*, WL-WB *mori*, WL-WK *morʔi*, WL-WT *morʔi*, WL-WW *mori*.

(2) KD-LB *bita*, KD-LW *bita*, KD-LW *bita matε*.

all:

(1) WL-BN *ʔata sa'kali*, WL-MS *s'kaliʔ*, WL-MS *sə'kali*, WL-PD *skali*.

(2) KD-LB *rɛiʔ*, KD-LW *rɛiʔ*.

(3) WL-AD *'wahã 'kaeʔ*, WL-BM *waokae*, WL-BP *wakajkae*, WL-BT *wahajkae*, WL-DL *wahajkae*, WL-HRW *wahajkae*, WL-IA *wəkae*, WL-KO *wahajkae*, WL-LG *wəkãe*, WL-LMK *wahok*, WL-LML *fakahae*, WL-LWI *wau 'kaʔeŋ*, WL-LWL *wakajkae*, WL-MD *wahakae*, WL-ML *fəka*, WL-RE *wokae*, WL-TJ *wahak*, WL-WB *wakajkae*, WL-WK *wakae*, WL-WT *wahajkae*, WL-WW *wahajkae*.

(4) CL-IL *pua*, CL-LP *pua*, CL-LWT *pua*, CL-MR *puari*, WL-BL *pua*.

(5) LH #svaol, CL-KK *sva'ɔl*, CL-LK *səwaol*, CL-LR *va'ɔl*, CL-PN *waol*, EL-LE *rowao*, EL-LMT *rowao*, WL-LH *waol*, WL-LTB *waokãe*.

(6) SK-HEW *leuvaiha*, SK-HEW *leʔu vaʔi ha*, SK-TA *haha*.

and:

(1) CL-IL *na*, CL-KK *no*, CL-LP *no*, CL-LR *no na'ne*, CL-LWK *no*, CL-PN *no*, EL-LE *nu*, EL-LMT *ʔo*, KD-LB *no're*, SK-HEW *nora*, SK-MM *nora*, WL-AB *nɔŋ*, WL-AD *nɔʔ*, WL-BL *no*, WL-BM *no*, WL-BN *nɔ̃*, WL-BP *nõ*, WL-BT *noʔõ*, WL-DL *noʔo*, WL-HRW *noʔõ*, WL-IA *no*, WL-KO *noʔõ*, WL-LG *noʔõ*, WL-LMK

noʔõ, WL-LML *nã*, WL-LTB *noʔõ*, WL-LWI *'nɔʔɔŋ*, WL-LWL *noʔõ*, WL-ML *nɔ*, WL-MS *naŋ*, WL-MS *nəŋ*, WL-PD *naŋ*, WL-PK *noʔõ*, WL-RE *noʔõ*, WL-TJ *noʔo*, WL-WB *nã*, WL-WK *no*, WL-WT *noʔõ*, WL-WW *noʔõ*.

(2) CL-LK *tal*, WL-AD *'taliʔ*.

(3) CL-KK *ko*, WL-LH *kã*, WL-MD *ko*.

(4) CL-KK *mo*, EL-LMT *mo*, SK-TA *mole*, WL-LH *mã*, WL-MD *mo*.

(5) CL-LWT *fɛ*, CL-MR *we*.

animal:

(1) CL-LWK *binataŋ*, SK-HEW *binataŋ*, WL-AB *bi'nataŋ*, WL-BP *binataŋ*, WL-LG *binataŋ*, WL-LTB *binataŋ*, WL-LWI *bi'nataŋ*, WL-MS *bi'nata*, WL-PD *bi'nata*, WL-PK *binataŋ*. From Indonesian/Malay *binatang* 'animal'.

(2) LH-KD **əvaŋ*, CL-IL *əfaŋ*, CL-KK *ə'va*, CL-KK *ə'vaŋ*, CL-KK *əwaŋ*, CL-LK *əwaŋ*, CL-LP *əfaŋ*, CL-LR *ə'vaŋ* *'lakut*, CL-LWT *əfaŋŋ*, CL-MR *əfau*, CL-PN *əwaŋ* *ləku*, EL-LE *əwã*, EL-LMT *əwã*, KD-LB *ʔewaŋ*, WL-AD *ʔɜ'wã*, WL-BL *əwan nawu*, WL-BM *əwã*, WL-BT *əwã*, WL-DL *əwã*, WL-HRW *əwã*, WL-IA *əwaŋnawun*, WL-KO *əwã*, WL-LH *əwã*, WL-LMK *əwã*, WL-LML *əfã*, WL-MD *əwã*, WL-ML *əwə*, WL-RE *əwã*, WL-TJ *əwã*, WL-WB *əwã*, WL-WK *əwã*, WL-WT *əwã*, WL-WW *əwã*.

ant:

(1) WL-AB *ka'rɛ*, WL-AD *kɜ'mɜrɜk*, WL-MS *kə'rɛk*, WL-PD *ka'rɛ*.

(2) PFL **təmiɜsi*, CL-KK *təmi'sɛr*, CL-KK *təmi'si*, CL-LR *temi'her*, KD-LB *mɛhi* *'wileŋ*, SK-HEW *mihe*, WL-LWI *tə'mihe*.

arm:

(1) PFL **lima* 'hand; arm', CL-KK *li'ma*, CL-LR *li'ma*, KD-LB *liŋ*, SK-HEW *lima*, WL-AB *limaŋ*, WL-AD *'lima'kɜt*, WL-BN *limaŋ*, WL-LWI *'limaŋ*, WL-MS *limaŋ*, WL-MS *limaʔ*, WL-PD *limaŋ*.

arrow:

LH-SK *#supɛl*, CL-KK *'supɛl*, CL-LR *'hupɛl*, SK-HEW *huper*, WL-AB *hupɜ*, WL-MS *hupɜl*, WL-PD *hupɜl*.

(2) CL-LR *əmat*, KD-LB *na'mɛʔ*, WL-AD *ʔamɜt*, WL-LWI *'amɜ*.

ash:

(1) PFL **ʔavu*, CL-KK *ka'vɔk*, CL-LR *'kavok*, SK-HEW *avu*, SK-MM *avu*,

WL-AB *afɔ*, WL-AD *kɜʔawuk*, WL-LWI *kəʔawu*, WL-MS *awo 'kaha*, WL-PD *awɔ'kaha*, WL-PD *awo'kaha*.

axe:

- (1) KD-LB *ba'liʔ*, WL-AB *bali*, WL-BN *bali*, WL-MS *bali*, WL-PD *bali*.
- (2) CL-KK *'kapa*, CL-KK *'kapak*, SK-HEW *kapak*, WL-AD *soru 'kapa*. From Indonesian/Malay *kapak* 'axe'.

back:

- (1) EL-LMT *punuʔu*, WL-AB *punuŋ*, WL-MS *punuŋ*, WL-MS *punuʔ*, WL-PD *punuŋ*.
- (2) KD-LB *ʔobi*, KD-LW *ɔbi*.
- (3) CL-LK *wurəsa*, KD-LW *ɣhur*, WL-BP *hurut*.
- (4) CL-LR *vu'huk*, CL-PN *wuhuk*, EL-LE *woho*, WL-AD *'wohʔo'kət*, WL-BL *usuk*, WL-BM *uhuk*, WL-DL *uhuk*, WL-IA *uhuk*, WL-KO *woho*, WL-LG *uhuk*, WL-LMK *woho*, WL-MD *uhukət*, WL-ML *uhuk*, WL-RE *uhuk*, WL-WK *uhuk*, WL-WT *wohɔ̃*, WL-WW *uhuk*.
- (5) CL-IL *kətubəŋəs*, CL-LP *kətubəŋ*, CL-LWT *kətubəŋha*, CL-MR *kətubəsa*, WL-LML *kətubəŋ*.
- (6) LH *#knaru*, CL-KK *kəna'ru*, CL-KK *kənarun*, CL-LWK *kənaruga*, WL-LH *kənaruk*.
- (7) WL-LWI *'kɔlaŋ*, WL-LWL *kolaʔ*, WL-TJ *kolaʔ*, WL-WB *kolaʔ*.
- (8) SK-HEW *toʔe*, SK-HEW *tuʔe*, SK-TA *toʔə-ŋ*.
- (9) SK-MM *ləpar*, SK-MM *leʔar*, SK-TA *leʔar*.

banana:

- (1) PFL **muku*, CL-KK *'muku*, CL-KK *'mukɔr*, CL-LR *'mukor*, KD-LB *mu'ʔü*, SK-HEW *muʔu*, WL-AB *mukɔ*, WL-AD *'muko*, WL-LWI *'mukɔ*, WL-MS *mukɔ*, WL-PD *mukɔ*.

bark of tree:

- (1) CL-KK *'kad͡ʒu ka'mak*, CL-LR *kaj 'kamək*, KD-LB *ʔai 'ʔamaʔ*, KD-LW *ai amaʔ*, WL-AB *kad͡ʒɔ 'kamaŋ*, WL-AD *kajo 'kaʔmã*, WL-BN *kad͡ʒu 'kamaŋ*, WL-MS *kad͡ʒɔ 'kamaŋ*, WL-PD *kad͡ʒɔ 'kamaŋ*.
- (2) SK-HEW *ʔai klokok*, WL-LWI *kə'loka*.

bat:

(1) KD-LB *mi'ʔi*, SK-HEW *niʔi*, WL-AD *pə'nike*, WL-LWI *pəni'ke*, WL-MS *mike*, WL-PD *mike*, WL-PD *pə'nike*.

(2) CL-KK *krɛ'be*, CL-KK *krɛ'beŋ*, CL-LR *kəre'beŋ*.

bean:

(1) WL-AB *kepa*, WL-MS *kepa*, WL-PD *kepa*.

(2) PFL **uta* 'bean; vegetable', CL-KK *u'ta*, CL-KK *u'tan*, CL-LR *'utan*, KD-LB *u'tan*, WL-AD *'utã*, WL-LWI *utaŋ*.

because:

(1) WL-BN *karena*, WL-PD *karna*. From Indonesian/Malay *karena* 'because'.

(2) WL-BN *se'bab*, WL-KO *səbab*. From Indonesian/Malay *sebab* 'reason'.

(3) EL-LE *ʔele*, KD-LB *ʔele*, KD-LW *ele*, SK-HEW *ale*.

(4) CL-IL *puka*, CL-KK *pukən*, CL-KK *'pukən*, CL-LK *pukən*, CL-LP *pukã*, CL-LWK *puk*, CL-LWT *puk*, CL-MR *pukɔ*, CL-PN *pukən*, EL-LMT *puʔẽ*, SK-HEW *puʔan poi*, WL-AD *'pukɜ*, WL-BL *puk*, WL-BM *pukã*, WL-BP *pukã*, WL-BT *pukã*, WL-DL *pukã*, WL-HRW *pukã*, WL-IA *pukən*, WL-LG *pukən*, WL-LH *pukã*, WL-LMK *pukã*, WL-LML *pukãŋ*, WL-LTB *pukã*, WL-LWI *'pukəŋ*, WL-LWL *pukẽ*, WL-MD *pukã*, WL-ML *pukə*, WL-PK *pukã*, WL-RE *pukã*, WL-TJ *pukã*, WL-WB *pukã*, WL-WK *puka*, WL-WT *pukã*, WL-WW *pukã*.

bedbug:

(1) PFL **təmayuŋ*, CL-KK *təma'dʒuŋ*, CL-KK *təma'dʒu*, CL-LR *təma'dʒuŋ*, KD-LB *maeŋ*, SK-HEW *ma:in*, WL-AB *ta'madʒɔŋ*, WL-AD *tə'majũ*, WL-LWI *tə'majuŋ*, WL-MS *madʒuŋ*, WL-PD *madʒuŋ*, WL-PD *tə'madʒuŋ*.

betel nut areca:

(1) PFL **vua* 'betelnut; fruit', KD-LB *uwe*, SK-HEW *vua*, WL-AB *ua*, WL-AD *'wuaʔ*, WL-BN *ʔufa*, WL-LWI *'wua*, WL-MS *ua*, WL-PD *ua*.

(2) CL-KK *kle'ruk*, CL-LR *'kleruk*.

betel vine:

(1) LH-KD #*malu*, CL-KK *ma'lu*, CL-KK *ma'lr*, CL-LR *'malor*, KD-LB *mal*, WL-AB *malu*, WL-AD *'maluʔ*, WL-BN *malu*, WL-MS *malu*, WL-PD *malu*.

big:

- (1) WL-AB *beij*, WL-BN *bĩ*, WL-MS *beaʔ*, WL-MS *peij*, WL-PD *bea*.
 (2) PFL **raya*, CL-IL *radžan*, CL-LR *raʔjan*, KD-LB *riʔa*, KD-LW *ria baraʔ*.
 (3) CL-KK *kedak*, CL-KK *keʔdak*, CL-LK *keda*, CL-LR *keʔdak*, CL-LWK *ke-dak*, CL-LWT *keda*, CL-MR *kedak*, CL-PN *kedakən*, KD-LW *kədəŋ kenaʔ*.
 (4) EL-LE *bəle*, EL-LMT *bəlaʔā*, WL-AD *ʔbele*, WL-BL *bel*, WL-BM *belə*, WL-BP *beleʔ*, WL-BT *belʔā*, WL-DL *belā*, WL-HRW *belʔā*, WL-IA *belən*, WL-KO *belʔā*, WL-LG *belā*, WL-LH *belā*, WL-LMK *belʔā*, WL-LML *belā*, WL-LTB *belā*, WL-LWI *ʔbelə*, WL-LWL *belē*, WL-MD *belā*, WL-ML *belɤ*, WL-PK *bəlʔā*, WL-RE *belā*, WL-TJ *belə*, WL-WB *belā*, WL-WK *belʔē*, WL-WT *belʔā*, WL-WW *bele*.
 (5) SK-HEW *gəte*, SK-HEW *gəteʔ*, SK-MM *gəte*, SK-TA *gəte*.

bird:

- (1) CL-IL *koloŋ*, CL-LP *koloŋ*, CL-LWK *koloŋ*, CL-LWT *koloŋu*, EL-LMT *olō*, SK-HEW *olon*, SK-MM *olon*, WL-AB *koloŋ*, WL-BL *kolo*, WL-BM *kolō*, WL-BN *koloŋ*, WL-BP *kolō*, WL-BT *kolō*, WL-KO *kolō*, WL-LG *kolō*, WL-LH *kolo*, WL-LMK *kolō*, WL-LML *kolo*, WL-LWI *ʔkoloŋ*, WL-LWL *kolō*, WL-MD *kolo*, WL-ML *kolo*, WL-MS *koloŋ*, WL-PD *koloŋ*, WL-PK *kolō*, WL-RE *kolō*, WL-TJ *kolō*, WL-WB *kolō*, WL-WK *kolo*, WL-WW *kolō*.
 (2) KD-LB *buʔruŋ*, WL-DL *burō*.
 (3) EL-LE *manuʔotō*, WL-HRW *manuk utā*, WL-WT *manuk utā*.
 (4) CL-LR *kuʔkak*, WL-IA *kukak*.
 (5) CL-KK *kəlieŋ*, CL-KK *ʔklie*, CL-KK *kliʔeŋ*, CL-LK *kəlieŋ*, CL-PN *kəlieŋ*.
 (6) SK-HEW *kəna horon*, SK-HEW *kənahoron*, SK-TA *horo-ŋ*.

bitter:

- (1) PFL **m-paʔit*, CL-KK *pəʔnait*, CL-LR *ʔpeta ʔpeta*, CL-LR *pnetəʔnən*, KD-LB *peiʔ*, SK-HEW *baʔit*, WL-AB *pei*, WL-AD *ʔpait*, WL-LWI *ʔpaik*, WL-MS *paiʔ*, WL-PD *pai*.

black:

- (1) PFL **mitəm*, CL-IL *mitəm*, CL-KK *mitəm*, CL-KK *miʔtəm*, CL-LK *mitəm*, CL-LP *mitəm*, CL-LR *miʔtəm*, CL-LR *mitəʔman*, CL-LWK *mitəm*, CL-LWT *mitəm*, CL-MR *mitəmə*, CL-PN *mitəməən*, EL-LE *mitā*, EL-LMT *mitəwā*, KD-LB *mitəŋ*, KD-LW *mitəŋ kajəʔ*, SK-HEW *mitan*, SK-MM *miṭak*, SK-MM *miṭan*, SK-TA *miṭa-ŋ*, SK-TA *mitak*, WL-AB *mitəŋ*, WL-AD *ʔmitə*, WL-BL *mitəm*, WL-BM *mitə*, WL-

BN *mit:ey*, WL-BP *mitən*, WL-BT *mitã*, WL-DL *mitã*, WL-HRW *mitã*, WL-IA *mitəm*, WL-KO *mitɤ*, WL-LG *mitã*, WL-LH *mitəm*, WL-LMK *mitã*, WL-LML *mitãŋ*, WL-LTB *mitã*, WL-LWI *'mitəŋ*, WL-LWL *mitẽ*, WL-MD *mitəm*, WL-ML *miten*, WL-MS *mitəŋ*, WL-PD *mitəŋ*, WL-PK *mitã*, WL-RE *mitõ*, WL-TJ *mitə*, WL-WB *mitõ*, WL-WK *mitəŋ*, WL-WT *mitã*, WL-WW *mitẽ*.

blood:

- (1) PFL **ra:*, WL-AB *ra*, WL-BN *ra:*, WL-MS *ra*, WL-MS *raŋ*, WL-PD *ra*.
- (2) KD-LB *wəi?*, KD-LW *vəi?*.
- (3) LH-SK *(*sə*)*mei*, CL-IL *səmeisa*, CL-KK *səmei*, CL-KK *'smein*, CL-LK *smei*, CL-LP *səmei*, CL-LR *hə'mei*, CL-LWK *səmei*, CL-LWT *həmeiha*, CL-MR *səmei*, CL-PN *həmein*, EL-LE *me?i*, EL-LMT *me?i*, SK-HEW *mei*, SK-MM *mei*, SK-TA *mei*, SK-TA *mei-ŋ*, WL-AD *mei*, WL-BL *mei*, WL-BM *mei*, WL-BP *mei*, WL-BT *mei*, WL-DL *mei*, WL-HRW *mei*, WL-IA *mei*, WL-KO *mei*, WL-LG *mei?*, WL-LH *mei*, WL-LMK *mei*, WL-LML *mei*, WL-LTB *mei*, WL-LWI *'mei*, WL-LWL *mei*, WL-MD *mei*, WL-ML *mehi*, WL-PK *mei*, WL-RE *mei*, WL-TJ *mei*, WL-WB *mei*, WL-WK *səmei*, WL-WT *mei*, WL-WW *mei*.

body:

- (1) WL-MS *badak*, WL-MS *badəŋ*.
- (2) PFL **vəki*, CL-KK *və'ki na'vak*, KD-LB *wə?*, SK-HEW *vi?in*, WL-AB *fakiŋ*, WL-AD *'wəki'kət*, WL-LWI *'wəkiŋ*, WL-PD *wə'k:iŋ*.
- (3) LH **navak*, CL-KK *və'ki na'vak*, CL-LR *na'vak*, LH-Lewolema (Pampus) *nawak*.

body hair :

- (1) PFL **vulu-k* 'body hair; feather', KD-LB *wurun*, SK-HEW *vulu?*, WL-AB *fu'lukuŋ*, WL-MS *wuluk*, WL-PD *wuluk*.
- (2) LH #*ravuk* 'body hair; feather', CL-KK *ra'vøk*, CL-LR *ri'au*, WL-AD *'rawu'kət*, WL-LWI *rawu'kəŋ*.

bone:

- (1) PFL **luri*, CL-IL *riukəs*, CL-KK *riuk*, CL-KK *ri'uk*, CL-LK *riuksa*, CL-LP *riuk*, CL-LR *'riuk*, CL-LWK *riuk*, CL-LWT *riuk*, CL-MR *riuk*, CL-PN *riuk*, EL-LE *ri?ũ*, EL-LMT *ri?ũ*, KD-LB *lu'rin*, KD-LW *hara? lurin*, SK-HEW *lurin*, SK-MM *lurin*, SK-TA *luri-ŋ*, WL-AB *ruĩ*, WL-AD *'ri?ũ*, WL-BL *riuk*, WL-BM *ri?u*, WL-BN *ru'iy*, WL-BP *ri?uk*, WL-BT *ri?ũ*, WL-DL *ri?ũ*, WL-HRW *ri?uk*, WL-IA *riuk*,

WL-KO *riʔũ*, WL-LG *riʔu*, WL-LH *riʔu*, WL-LMK *rũ*, WL-LML *riuk*, WL-LTB *riʔu*, WL-LWI *'riʔu*, WL-LWL *riʔuk*, WL-MD *riuk*, WL-ML *riuk*, WL-MS *rue*, WL-MS *ruij*, WL-PD *ruij*, WL-PK *riʔuk*, WL-RE *riʔuk*, WL-TJ *riʔuk*, WL-WB *riʔuk*, WL-WK *riʔũ*, WL-WT *riʔũ*, WL-WW *riuk*.

bow:

(1) PFL **vuhur*, CL-KK *vu'vu*, CL-KK *vu'vɔr*, CL-LR *'uvor*, KD-LB *wur*, SK-HEW *vuhur*, WL-AB *fihu*, WL-AD *wuhuʔ*, WL-LWI *'wuhu*, WL-MS *wuhu*, WL-PD *wuhu*.

branch:

- (1) KD-LB *li'maʔ*, WL-AB *kadʒɔ 'liman*, WL-MS *kadʒɔ 'liman*, WL-PD *kadʒɔ 'liman*.
 (2) WL-MS *kadʒɔ k'raŋak*, WL-PD *kadʒɔ kə'raŋak*.
 (3) CL-KK *vi'kil*, CL-LR *'vigit*, WL-AD *'wikʔi*.

breast:

(1) PFL *(t)*usu*, CL-KK *tu'su*, CL-LR *tu'hor*, KD-LB *tuʔu*, SK-HEW *uhu*, SK-TA *uhu-ŋ*, WL-AB *tuhũ*, WL-AD *'tuho*, WL-BN *tu'hɔ:*, WL-LWI *'tuho*, WL-MS *tuho*, WL-PD *tuhɔ*.

bride price:

- (1) PFL **veli* 'price; bride price; expensive; buy', CL-KK *ve'li ə'lan*, CL-LR *ve'lin*, SK-HEW *belis*, WL-AB *feliŋ*, WL-AD *weli*, WL-LWI *'weliŋ*, WL-MS *k'wae 'weliŋ*, WL-PD *weli*, WL-PD *weliŋ*.
 (2) KD-LB *pae*, WL-AB *palan*.

broom:

- (1) CL-LR *bə'namur*, WL-AB *namɔ*, WL-AD *'namu*, WL-MS *namɔ*, WL-PD *namo*.
 (2) CL-KK *kno'rɔt*, WL-LWI *kə'nuru*.

burden stick:

- (1) WL-AB *kə'lapan*, WL-MS *ləp'lepaʔ*.
 (2) PFL **leba*, CL-KK *bəle'ba*, CL-KK *bəle'bar*, CL-LR *bə'lebar*, SK-HEW *leba*, WL-AD *bə'lebaʔ*, WL-LWI *bə'lebaʔ*.

butterfly:

(1) PFL **(kə)bukal*, CL-KK *kəbu'kal*, CL-LR *kəbu'kal*, KD-LB *ʔebal*, SK-HEW *mata bulak*, WL-AD *kə'bukare*, WL-LWI *buka'laka*.

canoe:

(1) PFL **tena*, CL-KK *te'nar*, CL-KK *tena*, CL-LR *'tenar*, KD-LB *tenɛ*, SK-HEW *tena*, WL-AB *tena*, WL-AD *tena*, WL-BN *tena*, WL-LWI *'tena*, WL-MS *tena*, WL-PD *tena*.

cassava:

(1) WL-AB *kura 'dʒafa*, WL-BN *kur'a*, WL-MS *kura 'dʒawa*, WL-PD *kura*.
(2) CL-KK *ska'dʒo u'sər*, CL-LR *'hure 'kajor*, KD-LB *ʔa'leuʔ uwaʔ*, WL-AD *uwe 'kaju*, WL-LWI *uwe kajo*.

cheek:

(1) PFL **pipi/*klipi*, CL-KK *pi'pi*, CL-LR *kli'pi*, KD-LB *pipin*, SK-HEW *pipi*, WL-AB *pipij*, WL-AD *'pipi'kət*, WL-LWI *kə'lipi*, WL-MS *pipij*, WL-MS *pipiʔ*, WL-PD *pipij*.

chest:

(1) LH-KD *#korok*, CL-KK *kə'rək*, CL-LR *ko'rok*, KD-LB *ə'rəkʔ*, WL-LWI *'kərək*, WL-MS *kələŋ*, WL-MS *kələʔ*, WL-PD *kələŋ*.
(2) WL-AB *farɛ*, WL-PD *war:ɛŋ*.

chicken:

(1) PFL **manuk* 'chicken; bird', CL-KK *ma'nuk*, CL-LR *ma'nuk*, KD-LB *ma'nuʔ*, SK-HEW *manuʔ*, WL-AB *manu*, WL-AD *'manuk*, WL-LWI *'manu*, WL-MS *manuʔ*, WL-PD *manu*.

child:

(1) PFL **anak*, CL-IL *ana*, CL-LP *ana*, CL-LR *ana'kən*, CL-LWT *anakdʒa*, CL-MR *anudʒa*, CL-PN *anaken*, EL-LE *ana*, EL-LMT *ana ranã*, KD-LB *anaʔ*, KD-LW *anaʔ*, SK-HEW *dedi anak*, WL-AB *anaŋ*, WL-AD *'anaʔ*, WL-BL *ana*, WL-BM *anaʔ*, WL-BN *'anaŋ*, WL-BP *ana*, WL-BT *ana*, WL-DL *anaʔ*, WL-HRW *anaʔ*, WL-IA *ana*, WL-KO *anaʔ*, WL-LG *anaʔ*, WL-LH *ana*, WL-LMK *ana*, WL-LML *ana*, WL-LTB *anaʔ*, WL-LWI *'anaʔ*, WL-LWL *ana*, WL-MD *ana*, WL-ML *ana*, WL-MS *ana*, WL-MS *anaŋ*, WL-MS *anaʔ*, WL-RE *ana*, WL-TJ *ana*, WL-WB

ana, WL-WK *ana*, WL-WT *ana*, WL-WW *ana*.

(2) CL-KK *kɔ'pɔ*, CL-KK *kɔ'pɔŋ*, CL-KK *kopoŋ*, CL-LK *kopoŋ*.

(3) SK-HEW *dediʔ*, SK-MM *dedi*.

(4) SK-HEW *me:*, SK-MM *me:*, SK-TA *me:*.

civet cat:

(1) PFL **laku*, CL-KK *la'kɔr*, CL-KK *la'kɔ*, CL-LR *la'kɔr*, KD-LB *la'ʔu*, SK-HEW *laʔu*, WL-AB *lako*, WL-AD *'lako*, WL-LWI *'lakɔ*, WL-MS *lakɔ*, WL-PD *lako*.

clean:

(1) CL-KK *sla'ek*, WL-AB *lahiŋ*, WL-AD *'laɛʔ*, WL-LWI *'laɛŋ*, WL-MS *lahɛk*, WL-MS *lahɛŋ*, WL-PD *lahɛŋ*.

clothing:

(1) KD-LB *pa'kejan*, SK-HEW *pakian*, SK-HEW *pakaian*, WL-AB *pa'kiãŋ*, WL-LTB *pakaian*, WL-MS *pə'kiaŋ*.

(2) WL-BN *kondʒo*, WL-PD *kɔndɔ*.

(3) CL-IL *alolo*, CL-KK *'ale 'lɔɔ*, CL-KK *'ale 'lolo*, CL-KK *alelolo*, CL-LK *alelolo*, CL-LP *alelol*, CL-LR *'ale 'lolo*, CL-LWT *alelolodʒa*, CL-MR *alelolo*, CL-PN *alelolo*, EL-LMT *alalolo*, WL-AD *ale'lolo*, WL-BL *alelolo*, WL-BT *alelolo*, WL-DL *alelolo*, WL-HRW *alelolo*, WL-IA *alelolo*, WL-KO *alelolo*, WL-LH *alelolo*, WL-LMK *alelolo*, WL-LML *alelolo*, WL-MD *alelolo*, WL-ML *alelolo*, WL-RE *alelolö*, WL-TJ *alelolo*, WL-WK *alelolo*, WL-WT *alelolo*, WL-WW *alelolo*.

(4) WL-BM *agopake*, WL-BP *agopake*, WL-LG *akəpake*, WL-LWL *agopake*, WL-PK *akopake*, WL-WB *agopake*.

(5) CL-KK *mə'kaŋ-a*, WL-LWI *ə'kaŋ*.

cloud:

(1) LH-KD #*kova*, CL-IL *kofar*, CL-KK *ko'va*, CL-KK *ko'var*, CL-KK *kowa*, CL-LK *kowar*, CL-LP *kofaru*, CL-LR *'kovar*, CL-LWK *kowaru*, CL-LWT *kofaru*, CL-MR *kowa*, CL-PN *kowar*, EL-LE *ʔowa*, EL-LMT *ʔowa*, KD-LB *ʔɔ'wa*, SK-HEW *kova*, SK-HEW *kova*, WL-AB *kɔfā*, WL-AD *'kɔ:wa*, WL-BL *kowa*, WL-BM *kowa*, WL-BP *kowa*, WL-BT *kowa*, WL-DL *kowa*, WL-IA *kowa*, WL-KO *kowa*, WL-LG *kowa*, WL-LH *kowa*, WL-LMK *kowa*, WL-LML *kɔfā*, WL-LTB *kowa*, WL-LWI *kɔwa*, WL-LWL *kowa*, WL-MD *kowa*, WL-PK *kowa*, WL-RE *kowa*, WL-TJ *kowa*, WL-WB *kowa*, WL-WK *kowa*, WL-WT *kowa*.

(2) WL-HRW *abak*, WL-MS *nabaʔ*, WL-WW *abak*.

coconut:

- (1) LH-KD #*tapu*, CL-KK *tapu*, CL-KK *tapɔr*, CL-LR *'tapor*, KD-LB *taʔ*, WL-AB *tapo*, WL-AD *'tapo*, WL-LWI *'tapo*, WL-MS *tapɔ*, WL-PD *tapo*.
 (2) PFL **niur*, SK-HEW *niur*.

cold:

- (1) WL-AB *kaluaj*, WL-BN *ka'luaj*, WL-MS *k'luaj*, WL-PD *k'luaj*.
 (2) KD-LB *ʔemi*, KD-LW *æmi pana*, KD-LW *æmi ruku*, KD-LW *mi ælæʔ*.
 (3) CL-IL *lātəs*, CL-KK *bəlātəs*, CL-KK *lə'təs*, CL-LK *lātəs*, CL-LR *lātə'hən*, CL-LWK *lātəs*, CL-LWT *lātəh*, CL-MR *lātəs*, CL-PN *lātəhən*, EL-LE *lato*, EL-LMT *lətha*, SK-HEW *bəlatan*, SK-HEW *blatan*, SK-MM *blaʔan*, SK-TA *blata-ŋ* *blaʔur*, SK-TA *blata-ŋ* *blira-ŋ*, SK-TA *blata-ŋ*, WL-AD *glətɔ*, WL-BL *lātəs*, WL-BM *gəlātā*, WL-BP *gəlātā*, WL-BT *gəlātā*, WL-DL *gəlātəŋ*, WL-HRW *gəlātɕ*, WL-IA *lātən*, WL-KO *gəlātɕ*, WL-LG *gəlātə*, WL-LH *lata*, WL-LMK *gəlātā*, WL-LML *gəlātā*, WL-LWI *'gəlātə*, WL-LWL *gəlātē*, WL-MD *lātā*, WL-ML *leta*, WL-PK *gātā*, WL-RE *gəlātā*, WL-TJ *gəlātā*, WL-WB *gəlātə*, WL-WK *lātəs*, WL-WT *gəlātā*, WL-WW *gəlātā*.
 (4) CL-LP *ləkə*, WL-LTB *loʔi*.

comb:

- (1) LH #*kiri*, CL-KK *ki'ri*, CL-KK *ki'rer*, CL-LR *'kirer*, WL-AB *kiri*, WL-AD *ki'riʔ*, WL-BN *ki'ri*, WL-LWI *'kiri*, WL-MS *kiri*, WL-PD *kiri*.

corn, maize:

- (1) PFL **vatar*, KD-LB *watar*, SK-HEW *vatar*, WL-AD *'wataʔ*, WL-LWI *'wata*, WL-MS *wata*, WL-PD *wata*.
 (2) CL-KK *kva'rɔr*, CL-KK *'kvaru*, CL-LR *'kəvaror*.

correct:

- (1) WL-BN *mallɔŋ*, WL-PD *mɔlɔŋ*.
 (2) LH-KD #*mudəŋ*, CL-KK *mu'dʒək*, KD-LB *muar*, WL-AB *muro*, WL-AD *muʔ'rɕ*, WL-BL *mur*, WL-BM *murə*, WL-BP *mu*, WL-BT *murā*, WL-DL *murʔā*, WL-HRW *murʔā*, WL-KO *murʔā*, WL-LH *nā murā*, WL-LMK *murʔā*, WL-LML *murā*, WL-LWI *'murəŋ*, WL-MD *murā*, WL-ML *murɕ*, WL-MS *murɔk*, WL-PD *murok*, WL-RE *nə murā*, WL-TJ *murə*, WL-WB *nə murə*, WL-WK *nā mure*, WL-WT *murʔā*, WL-WW *murō*.
 (3) WL-IA *apadikən*, WL-LG *naʔandikā*, WL-LTB *nādikā*, WL-LWL *apadiken*.

(4) CL-IL *na moi*, CL-KK *nər moi*, CL-LK *nər moi*, CL-LP *nara moi*, CL-LR *ne'moi*, CL-LWK *na moi*, CL-LWT *mui*, CL-MR *namui*, CL-PN *nəmoi*.

(5) EL-LE *na toʔō*, EL-LMT *toʔa*, WL-PK *tokeʔē*.

crocodile:

(1) LH-KD #*bəpəp*, CL-KK *bəpapa* 'monitor lizard', KD-LB *bapa'ju*, WL-AB *bapa*, WL-BN *bapa*, WL-MS *bapa*, WL-PD *bapa*.

(2) LH-SK #*kəbu*, CL-KK *'kobu*, CL-LR *kəb*, SK-HEW *kobu*, WL-AD *'kobu*, WL-LWI *'kobu*.

dark:

(1) WL-AB *kui*, WL-BN *kuiŋ*, WL-MS *kuiŋ*, WL-PD *kuiŋ*.

(2) CL-KK *'mitəm*, CL-LR *eka mitə'man*, KD-LB *mitəŋ*, WL-LWI *'mitəŋ*.

(3) CL-KK *rə'mak*, SK-HEW *rumaŋ*.

day:

(1) PFL **lədav* 'sun; day', CL-IL *lədʒo*, CL-KK *le'dʒo*, CL-KK *lədʒon*, CL-LK *lədʒon*, CL-LP *lədʒonən*, CL-LR *la'jonan*, CL-LWK *ləjon*, CL-LWT *lədʒon*, CL-MR *lədʒo*, CL-PN *ləjonən*, EL-LE *ləro*, EL-LMT *ləro*, KD-LB *lə'jo*, KD-LW *ləjo*, SK-HEW *ləron*, SK-MM *ləron*, SK-TA *ləro-ŋ*, WL-AB *la'ra*, WL-AD *rə' rō*, WL-BL *ləro*, WL-BM *lərō*, WL-BN *la'ra*, WL-BP *rərō*, WL-BT *rərō*, WL-DL *rərō*, WL-HRW *rərō*, WL-IA *ləron*, WL-KO *rərō*, WL-LG *lərō*, WL-LH *lərō*, WL-LMK *rərō*, WL-LML *ləro*, WL-LTB *lərō*, WL-LWI *lə'rəŋ*, WL-LWL *rərō*, WL-MD *ləro*, WL-ML *ləro*, WL-MS *la'ra*, WL-PD *la'ra*, WL-PK *lərō*, WL-RE *lərō*, WL-TJ *lərō*, WL-WB *lərō*, WL-WK *ləro*, WL-WT *rərō*, WL-WW *rərō*.

deer:

(1) FL #*rusa*, CL-KK *ru'sar*, CL-KK *ru'sa*, CL-LR *ru'har*, KD-LB *ru'ha*, SK-HEW *ruha*, WL-AB *ruha*, WL-AD *'ruha*, WL-LWI *'ruha*, WL-MS *ruha*, WL-PD *ruha*.

dew:

(1) CL-IL *apu*, CL-KK *apun*, CL-LK *apun*, CL-LP *apu*, CL-LWK *apunū*, CL-LWT *apun*, CL-MR *apu*, CL-PN *apun*, EL-LE *apũ*, EL-LMT *apũ*, SK-HEW *apun*, WL-BL *apu*, WL-BM *apu*, WL-BP *apu*, WL-BT *apũ*, WL-DL *apũ*, WL-HRW *apũ*, WL-KO *apũ*, WL-LG *apũ*, WL-LH *apũ*, WL-LMK *apu*, WL-LML *apu*, WL-LTB *apũ*, WL-LWL *apũ*, WL-MD *apũ*, WL-ML *apu*, WL-PK *apũ*, WL-RE *apu*, WL-TJ

apu, WL-WB *apũ*, WL-WK *apũ*, WL-WT *apũ*, WL-WW *apũ*.

dibble stick:

- (1) CL-KK *u'a*, CL-KK *u'aŋ*, CL-LR *'uan*.

dirty:

- (1) WL-AB *ka'lita*, WL-BN *ka'lit:a*, WL-MS *k'lita?*, WL-PD *k'lita*.

- (2) KD-LB *mitɛŋ*, KD-LW *mitɛŋ kajo?*.

(3) CL-IL *milana*, CL-KK *milan*, CL-KK *mi'lan*, CL-LK *mila*, CL-LP *milan*, CL-LR *bəmi'lan*, CL-LWK *milan*, CL-LWT *mila*, CL-MR *mila*, CL-PN *milan*, SK-HEW *milan*, WL-AD *mi'lã*, WL-BM *milã*, WL-BP *milã*, WL-BT *milã*, WL-DL *milã*, WL-HRW *milã*, WL-IA *milan*, WL-KO *milã*, WL-LG *milã*, WL-LH *milã*, WL-LMK *milã*, WL-LML *milã*, WL-LTB *milã*, WL-LWI *'milaŋ*, WL-LWL *milã*, WL-MD *milã*, WL-ML *milɤ*, WL-PK *milã*, WL-RE *milã*, WL-TJ *milã*, WL-WB *milã*, WL-WK *mila*, WL-WT *milã*, WL-WW *milã*.

dog:

(1) PFL **ahu*, CL-IL *aor*, CL-KK *a'or*, CL-KK *au*, CL-LK *aor*, CL-LP *aoru*, CL-LR *a'or*, CL-LWK *aoru*, CL-LWT *aoru*, CL-MR *aoru*, CL-PN *aho*, EL-LE *aho*, EL-LMT *aho*, KD-LB *au*, KD-LW *au vavi*, SK-HEW *ahu*, SK-HEW *ahu*, SK-MM *ahu*, SK-TA *ahu*, WL-AB *aho*, WL-AD *'aho*, WL-BL *ao*, WL-BM *aho*, WL-BN *a'ho*, WL-BP *aho*, WL-BT *aho*, WL-DL *aho*, WL-HRW *aho*, WL-IA *aho*, WL-KO *aho*, WL-LG *aho*, WL-LH *aho*, WL-LMK *aho*, WL-LML *ao*, WL-LTB *aho*, WL-LWI *'aho?*, WL-LWL *aho*, WL-MD *aho*, WL-ML *aho*, WL-MS *aho*, WL-MS *hao*, WL-PD *aho*, WL-PK *aho*, WL-RE *aho*, WL-TJ *aho*, WL-WB *aho*, WL-WK *aho*, WL-WT *aho*, WL-WW *aho*.

door:

(1) WL-AB *pitu*, WL-AD *'pita*, WL-MS *pite 'mataŋ*, WL-MS *pitu 'mataŋ*, WL-PD *pite*.

(2) LH **knavi*, CL-KK *'knaver*, CL-KK *kna'vɛr*, CL-KK *'knavi*, CL-LR *kə'naver*, CL-LR *'knave*, WL-LWI *kə'nawe*.

drizzle:

(1) LH-KD *#ritik*, CL-KK *u'dʒan grə'tik*, CL-LR *u'dʒa krenik*, KD-LB *riti?*, WL-AD *gri'ʔɛna*, WL-LWI *kite'riteŋ*.

dry:

(1) PFL **mada* ‘dry; thirsty’, CL-IL *madzakən*, CL-KK *madzak*, CL-KK *ma'dzak*, CL-LK *madz*, CL-LP *madza*, CL-LR *maja'kən*, CL-LWK *madzak*, CL-LWT *madzakən*, CL-MR *madzak*, CL-PN *majakən*, EL-LE *məw*, EL-LMT *marʔă*, KD-LB *ma'jaʔ*, KD-LW *majaʔ midε*, SK-HEW *marā*, SK-HEW *maran*, SK-MM *maran*, SK-TA *marā*, WL-AB *marā*, WL-AD *'maʔră*, WL-BL *marā*, WL-BM *marā*, WL-BN *marā*, WL-BP *marā*, WL-BT *marʔă*, WL-DL *marā*, WL-HRW *marā*, WL-IA *marā*, WL-KO *marʔă*, WL-LG *marā*, WL-LH *marā*, WL-LML *mară*, WL-LTB *mară*, WL-LWI *'marā*, WL-LWL *mară*, WL-MD *marʔă*, WL-ML *marā*, WL-MS *marak*, WL-PD *marak*, WL-PK *marăʔă*, WL-RE *mară*, WL-TJ *marā*, WL-WB *marā*, WL-WK *marā*, WL-WT *marā*, WL-WW *marā*.

(2) PFL **m-tuʔur*, KD-LW *turi rəka*, SK-HEW *duʔur*, SK-MM *duʔur*, SK-TA *duʔur*, WL-LMK *tuʔuk*.

dry in sun:

(1) PFL **pa-vari*, CL-KK *pa'ri*, CL-LR *pa'riy*, KD-LB *pa'ri:*, SK-HEW *vorī*, WL-AB *peij 'pakiaŋ*, WL-AD *'pahe*, WL-AD *'paʔi*, WL-BN *paē:*, WL-LWI *pa'ʔiŋ*, WL-MS *paiŋ*, WL-PD *paiŋ apa'hire*.

dust:

(1) PFL **ka-ʔavu* ‘ash; dust’, CL-IL *kafok*, CL-KK *kawok*, CL-LK *kawok*, CL-LP *kafoku*, CL-LWK *kawok*, CL-LWT *kafoku*, CL-MR *kawok*, CL-PN *kawok*, EL-LE *awo*, EL-LE *ʔawə*, SK-HEW *avu*, SK-HEW *avu*, WL-AB *afə*, WL-BL *kawuk*, WL-BM *kəawuk*, WL-BM *kʔawuk*, WL-BN *a'fo:*, WL-BP *kəawuk*, WL-BT *kəawuk*, WL-DL *kawuk*, WL-HRW *kəʔawuk*, WL-IA *awo*, WL-KO *kəʔawuk*, WL-LG *keawuk*, WL-LG *kəʔawuk*, WL-LH *kawu*, WL-LMK *kaʔwuk*, WL-LML *kafu*, WL-LTB *əwu*, WL-LTB *kəwu*, WL-LWI *kə'ʔawu*, WL-LWL *kəawuk*, WL-MD *kawu*, WL-ML *kawu*, WL-PK *kəwuk*, WL-RE *awo*, WL-RE *kawuk*, WL-TJ *kəawuk*, WL-WB *kəawuk*, WL-WK *kawok*, WL-WT *kəʔawuk*, WL-WW *kəʔawuk*.

(2) SK-TA *ʔreβu*, WL-MS *krawuʔ*, WL-PD *kə'rawu*, (3) CL-IL *əmut*, CL-KK *əmut*, CL-KK *ə'mət*, CL-LK *əmut*, CL-LP *əmut*, CL-LR *'əmut*, CL-LWK *amutu*, CL-LWT *əmutu*, CL-MR *əmutu*, CL-PN *əmut*, EL-LMT *əmu*, WL-AD *ʔə'mut*, WL-BL *əmut*, WL-BT *əmut*, WL-DL *əmut*, WL-HRW *əmut*, WL-IA *əmut*, WL-KO *əmut*, WL-LH *əmut*, WL-LMK *əmu*, WL-LML *əmut*, WL-MD *əmut*, WL-ML *əmu*, WL-WK *əmut*, WL-WT *əmut*, WL-WW *əmut*.

ear:

(1) EL-LE *tilũ*, EL-LMT *tilu*, KD-LB *til*, KD-LW *til*, SK-HEW *tilu*, SK-MM *tilun*, SK-TA *tilu*, SK-TA *tilu-ŋ*, WL-AB *tiluŋ*, WL-AD *'tilu'kət*, WL-BL *tiluk*, WL-BM *tilu*, WL-BN *tiluŋ*, WL-BP *tilu*, WL-BT *tilũ*, WL-DL *tilũ*, WL-HRW *tiluk*, WL-IA *tiluk*, WL-KO *tiluk*, WL-LG *tilu*, WL-LH *tilukə*, WL-LMK *tilũ*, WL-LML *tilu*, WL-LTB *tilu*, WL-LWI *'tiluŋ*, WL-LWL *tilu*, WL-MD *tiluk*, WL-ML *tilu*, WL-MS *tiluŋ*, WL-MS *tilu?*, WL-PD *tiluŋ*, WL-PK *tilu*, WL-RE *tilũ*, WL-TJ *tilu*, WL-WB *tilu*, WL-WK *tilu*, WL-WT *tilũ*, WL-WW *tilu*.

(2) CL-IL *kiŋasa*, CL-KK *ki'ŋa*, CL-KK *kiŋan*, CL-LK *kiŋa*, CL-LP *kiŋa*, CL-LR *ki'ŋa*, CL-LWK *kiŋa*, CL-LWT *kinaha*, CL-MR *kiŋasa*, CL-PN *kiŋag*.

Earth:

(1) CL-KK *'tana 'ekan*, CL-LR *'tana 'ekan*, EL-LE *tana*, EL-LMT *tana*, SK-HEW *nian tana*, SK-HEW *tana*, SK-MM *ṭana*, SK-TA *tana*, WL-AD *tana 'ʔekā*, WL-BL *tana*, WL-BM *tanʔa*, WL-BN *tana*, WL-BP *tana*, WL-BT *tana*, WL-DL *tana*, WL-HRW *tana*, WL-IA *tana*, WL-KO *tana*, WL-LG *tana*, WL-LH *tana*, WL-LMK *tana*, WL-LML *tana*, WL-LTB *tana*, WL-LWI *tana 'ʔekəŋ*, WL-LWL *tana*, WL-MD *tana*, WL-ML *tana*, WL-MS *tana*, WL-PK *tana*, WL-RE *tana*, WL-TJ *tana*, WL-WB *tana*, WL-WK *tana*, WL-WT *tana*, WL-WW *tana*.

(2) WL-AB *bumi*, WL-PD *bumi*.

(3) KD-LB *ʔau?*, KD-LW *au?*.

(4) CL-IL *ənaj*, CL-KK *ənad̥ʒ*, CL-LK *ənad̥ʒ*, CL-LP *ənad̥ʒ*, CL-LWK *ənad̥ʒ*, CL-LWT *ənad̥ʒ*, CL-MR *ənad̥ʒ*, CL-PN *ənaj*.

earthworm:

(1) WL-AB *fali 'fali*, WL-BN *ʔula 'fale*, WL-MS *ula 'wali*, WL-PD *ula 'wali*.

(2) CL-IL *salafalak*, CL-KK *səlvə'lak*, CL-LK *səlvad̥ʒa*, CL-LR *helva'lak*, CL-LWK *səwalaku*, CL-LWT *salafalaku*, CL-MR *salafalak*, EL-LE *səwalo*, EL-LMT *səwalo*, KD-LW *hala mukəl*, WL-LML *salafalā*, WL-WK *salawalak*.

(3) WL-AD *kala'wala*, WL-IA *keləwalo*, WL-LWI *ula'wala*.

(4) WL-BM *kəbeŋə*, WL-BP *kəbeŋə*, WL-KO *kəbeŋā*.

(5) CL-KK *ekod̥ʒ*, CL-LP *ekod̥ʒ*, CL-PN *egoj*, WL-BL *eko*, WL-WB *gego*.

(6) WL-BT *kəbeŋə*, WL-DL *kəbeŋə*, WL-HRW *kəbeŋə*, WL-LG *kəbeŋā*, WL-LMK *kəbeŋer*, WL-LTB *kəbeŋā*, WL-LWL *kəbeŋer*, WL-RE *kəbeŋā*, WL-TJ *kəbeŋə*, WL-WT *kəbeŋər*, WL-WW *kəbeŋe*.

(7) WL-LH *kodəl*, WL-MD *kodəl*, WL-ML *kodɤ*.

(8) SK-HEW *doho*, SK-TA *doho*.

egg:

(1) PFL **təlu*, CL-IL *təluk*, CL-KK *təluk*, CL-KK *tə'lok*, CL-LK *təluk*, CL-LP *təluk*, CL-LR *tə'luk*, CL-LWK *təluk*, CL-LWT *təluk*, CL-MR *təluk*, CL-PN *təlukən*, EL-LE *təlü*, EL-LMT *təluʔü*, KD-LB *tə'lor*, KD-LW *manuʔ təlor*, SK-HEW *təlon*, SK-MM *təlo*, SK-TA *manu təlo*, SK-TA *təlo*, WL-AB *ta'lu*, WL-AD *təluk*, WL-BL *təlu*, WL-BM *təluk*, WL-BN *taluk*, WL-BP *təluk*, WL-BT *təlü*, WL-DL *təʔü*, WL-HRW *təluk*, WL-IA *təluk*, WL-KO *təlu*, WL-LG *təluk*, WL-LH *təʔu*, WL-LMK *təluk*, WL-LML *təlu*, WL-LTB *təlü*, WL-LWI *təlu*, WL-LWL *təluk*, WL-MD *təʔü*, WL-ML *təlu*, WL-MS *tə'luk*, WL-PD *taluk*, WL-PK *təʔü*, WL-RE *təluk*, WL-TJ *təluk*, WL-WB *təluk*, WL-WK *təlu*, WL-WT *təʔü*, WL-WW *təluk*.

eight:

(1) LH-KD #*butu* 'eight; bunch; group', CL-KK *'buto*, CL-LR *'buto*, KD-LB *butu 'rai*, WL-AB *buto*, WL-AD *butoʔ*, WL-BN *butʔ*, WL-LWI *'butʔ*, WL-MS *buto*, WL-PD *buto*.

(2) PFL **valu*, SK-HEW *valu*.

excrement:

(1) PFL **taʔi*, CL-KK *ta'i*, CL-KK *ta'er*, CL-LR *taer*, SK-HEW *taʔi*, WL-AB *tae*, WL-AD *tae*, WL-LWI *'tae*, WL-MS *sapi 'taiŋ*, WL-MS *witi 'taiŋ*, WL-MS *taiŋ*, WL-PD *sapi 'taiŋ*, WL-PD *witi 'taiŋ*.

expensive:

(1) PFL **veli* 'price; bride price; expensive; buy', CL-KK *ve'lin*, CL-LR *'veli veli*, KD-LB *welin*, SK-HEW *velin*, WL-AB *feliŋ*, WL-AD *'welĩ*, WL-LWI *'weliŋ təgəra*, WL-MS *weliŋ b'ələʔ*, WL-PD *weliŋ*.

eye:

(1) PFL **mata*, CL-KK *matan*, CL-KK *ma'ta*, CL-LK *mata*, CL-LR *ma'ta*, CL-LWK *mata*, CL-PN *matag*, KD-LB *mato*, KD-LW *matʔ*, SK-HEW *mata*, SK-MM *maʔan*, SK-TA *mata*, SK-TA *mata-ŋ*, WL-AB *mataŋ*, WL-AD *'matak*, WL-AD *matakʔt*, WL-BL *matak*, WL-BM *mata*, WL-BN *mataŋ*, WL-BP *mata*, WL-BT *mata*, WL-DL *mata*, WL-HRW *matak*, WL-IA *matak*, WL-KO *mata*, WL-LG *mata*, WL-LH *matakə*, WL-LMK *matã*, WL-LML *matã*, WL-LTB *mata*, WL-LWI *'mata*, WL-LWL *mata*, WL-MD *matakə*, WL-ML *matɤ*, WL-MS *mataŋ*, WL-MS *mataʔ*, WL-PD *mataŋ*, WL-PK *mata*, WL-RE *mata*, WL-TJ *mata*, WL-WB *mata*,

WL-WK *mata*, WL-WT *mata*, WL-WW *mata*.

(2) CL-IL *kulukəs*, CL-LP *kuluka*, CL-LWT *kuluk*, CL-MR *kuluk*, EL-LE *ulu*, EL-LMT *ulʔũ*.

far:

(1) PFL **doa* 'far; long', CL-IL *doa*, CL-KK *doak*, CL-KK *do'ak*, CL-LK *doak*, CL-LP *doa*, CL-LR *doa'kən*, CL-LWK *doa*, CL-LWT *doa*, CL-MR *doek*, CL-PN *doakən*, EL-LE *doa*, EL-LMT *doə*, KD-LB *doa*, KD-LW *də tɛvɛl*, WL-AB *d̥ʒoan̩*, WL-AD *d̥oã*, WL-BL *doa*, WL-BM *doã*, WL-BN *d̥ʒuan̩*, WL-BP *doã*, WL-BT *doʔẽ*, WL-DL *doã*, WL-HRW *doã*, WL-IA *doa*, WL-KO *bədoʔẽ*, WL-LG *doã*, WL-LH *doa*, WL-LMK *doã*, WL-LML *doe*, WL-LTB *d̥oẽ*, WL-LWI *'doan̩*, WL-LWL *doã*, WL-MD *doe*, WL-ML *doa*, WL-MS *doan̩*, WL-PD *doan̩*, WL-PK *doã*, WL-RE *doʔẽ*, WL-TJ *doa*, WL-WB *doã*, WL-WK *doʔe*, WL-WT *doã*, WL-WW *doã*.

(2) LH-SK #*blavir* 'long; far', SK-HEW *bəlavir*, SK-HEW *blavir*, SK-MM *blavir*, SK-TA *blaβir*.

fat:

(1) PFL **voda-k*, CL-IL *fod̥ʒakən*, CL-KK *vo'd̥ʒak*, CL-KK *wod̥ʒak*, CL-LK *wod̥ʒak*, CL-LP *fod̥ʒak*, CL-LR *vo'jak*, CL-LWK *wod̥ʒak*, CL-LWT *fod̥ʒak*, CL-MR *fod̥ʒak*, CL-PN *wod̥ʒak*, EL-LE *wor̥*, EL-LMT *woraʔã*, KD-LB *w̥jaʔ*, KD-LW *v̥jaʔ ɸval*, SK-HEW *vura*, SK-HEW *vura*, SK-MM *vuran*, WL-AB *f̥o'rakan̩*, WL-BL *wora*, WL-BM *wora*, WL-BP *wora*, WL-BT *worʔã*, WL-IA *worak*, WL-KO *worʔã*, WL-LG *worʔã*, WL-LH *worʔã*, WL-LMK *worʔã*, WL-LML *forã*, WL-LTB *worã*, WL-LWI *'w̥ran̩*, WL-LWL *worã*, WL-MD *worã*, WL-ML *wor̥*, WL-MS *w̥rak*, WL-PD *w̥rak*, WL-PK *woraʔã*, WL-RE *worã*, WL-TJ *worã*, WL-WB *worã*, WL-WK *worʔa*, WL-WT *worʔã*.

(2) WL-AD *'kahare*, WL-DL *kənaha*, WL-HRW *kənaharə*, WL-WW *kənaha*.

father:

(1) CL-IL *ama*, CL-KK *'ama*, CL-KK *ama*, CL-LP *ma*, CL-LR *ama*, CL-LWK *ama*, CL-LWT *ama*, CL-MR *ama*, CL-PN *amaj*, EL-LE *ama*, EL-LMT *ama*, KD-LB *ʔame*, KD-LW *am̥*, KD-LW *ame*, SK-HEW *ama*, SK-MM *ama*, SK-TA *ama*, WL-AB *amã*, WL-AD *'ama*, WL-BL *ama*, WL-BN *ʔama*, WL-DL *ama*, WL-HRW *ama*, WL-IA *ama*, WL-KO *ama*, WL-LH *amakə*, WL-MD *ama*, WL-MS *ama*, WL-MS *amaj*, WL-MS *mama*, WL-PD *mama*, WL-WK *ama*, WL-WW *ama*.

(2) CL-LK *bapa*, WL-BM *bapa*, WL-BP *bapaʔ*, WL-BT *bapa*, WL-LG *bapa*,

WL-LMK *bapa*, WL-LML *bapa*, WL-LTB *bapa*, WL-LWI *'bapa?*, WL-LWL *baba*, WL-ML *bapa*, WL-PK *ba?*, WL-RE *bapa*, WL-TJ *bapa*, WL-WB *bapa*, WL-WT *bapa*. From Indonesian/Malay *bapak* 'father'.

feather:

(1) KD-LB *u'run*, SK-HEW *vulu*, SK-HEW *vulu*, SK-TA *βulu-ŋ*, WL-AB *fu'lukuy*, WL-BN *wu'luk*, WL-MS *wuluk*, WL-PD *wuluk*.

(2) CL-IL *rafukəj*, CL-KK *'klie ra'vok*, CL-KK *rawuk*, CL-LK *rawuk*, CL-LP *rafuk*, CL-LWK *rawuk*, CL-LWT *rafuk*, CL-MR *rafuk*, EL-LE *rawũ*, EL-LMT *rawu?ũ*, WL-AD *'rawuke*, WL-BL *rawuk*, WL-BM *rawuk*, WL-BP *rawuk*, WL-BT *rawũ*, WL-DL *rawuk*, WL-HRW *rawuk*, WL-IA *rawuk*, WL-KO *rawuk*, WL-LG *rawuk*, WL-LH *rawukə*, WL-LMK *raʔwũ*, WL-LML *rafuk*, WL-LTB *rawu*, WL-LWI *rawu'kəŋ*, WL-LWL *rawuk*, WL-MD *rawuk*, WL-ML *rawuk*, WL-PK *rawu?ũ*, WL-RE *rawuk*, WL-TJ *rawuk*, WL-WB *rawuk*, WL-WK *rawũ*, WL-WT *rawuk*, WL-WW *rawuk*.

(3) CL-LR *ku'kak 'riav*, CL-PN *riaw*.

fence:

(1) CL-LR *'nihar*, KD-LB *ni'he*, SK-HEW *niha*, WL-AB *ni'ha*, WL-AD *niha*, WL-LWI *'niha*, WL-PD *niha*.

finger:

(1) CL-KK *ki'kil*, CL-LR *ki'kil*, SK-HEW *kikir*.

(2) WL-MS *numak*, WL-PD *liməŋ numak*.

(3) WL-AD *'limakət 'ʔa:nʔa*, WL-LWI *liməŋ 'anaŋ*.

finished:

(1) LH *#vahak*, CL-KK *vaka*, CL-KK *vak*, WL-AD *'wahak*, WL-LWI *'waha*.

(2) WL-MS *hapu?*, WL-PD *ka'p:o*.

fire:

(1) PFL **api*, CL-IL *aper*, CL-KK *ape*, CL-KK *a'per*, CL-KK *a'pi*, CL-LK *aper*, CL-LP *aperu*, CL-LR *'aper*, CL-LWK *aperu*, CL-LWT *aperu*, CL-MR *aperu*, CL-PN *aper*, EL-LE *ape*, EL-LMT *ape*, KD-LB *a'pi*, KD-LW *api padu*, SK-HEW *api*, SK-HEW *api*, SK-MM *api*, SK-TA *api*, WL-AB *ape*, WL-AD *'ʔape*, WL-BL *ape*, WL-BM *ape*, WL-BN *a'pe:*, WL-BP *ape*, WL-BT *ape*, WL-DL *ape*, WL-HRW *ape?*, WL-IA *ape*, WL-KO *ape*, WL-LG *ape*, WL-LH *ape*, WL-LMK *ape*, WL-

LML *ape*, WL-LTB *ape*, WL-LWI *'ape?*, WL-LWL *ape*, WL-MD *ape*, WL-ML *ape*, WL-MS *ape*, WL-PD *ape*, WL-PK *ape*, WL-RE *ape*, WL-TJ *ape*, WL-WB *ape*, WL-WK *ape*, WL-WT *ape*, WL-WW *ape*.

firewood:

(1) CL-LR *kaj 'api*, KD-LB *?ai 'japi*, WL-AB *kad͡ʒo 'ape*, WL-AD *kajo 'ape*, WL-LWI *kajo 'ape*, WL-MS *kad͡ʒo 'ape*, WL-PD *kad͡ʒo 'ape*.

fish:

(1) PFL **ikan*, CL-IL *ikan*, CL-KK *ika*, CL-KK *i'ka*, CL-KK *i'kan*, CL-LK *ikan*, CL-LP *ikan*, CL-LR *i'kan*, CL-LWK *ikan*, CL-LWT *ikanu*, CL-MR *ikau*, CL-PN *ikaj*, EL-LE *?i?ã*, EL-LMT *i?ã*, KD-LB *i?a*, KD-LW *i?a ɛlan*, SK-HEW *i?an*, SK-HEW *i?an*, SK-MM *i?an*, SK-TA *i?a-ŋ*, WL-AB *ikan*, WL-AD *'ikã*, WL-BL *ika*, WL-BM *ikã*, WL-BN *ikan*, WL-BP *ikã*, WL-BT *ikã*, WL-DL *ikã*, WL-HRW *ikã*, WL-IA *ikan*, WL-KO *ikã*, WL-LG *ikã*, WL-LH *ikã*, WL-LMK *ikan*, WL-LML *ikã*, WL-LTB *ikã*, WL-LWI *'ikan*, WL-LWL *ikã*, WL-MD *ikã*, WL-ML *ikɔ*, WL-MS *ikan*, WL-PD *ikan*, WL-PK *ikã*, WL-RE *ikã*, WL-TJ *ikã*, WL-WB *ikã*, WL-WK *ikã*, WL-WT *ikã*, WL-WW *ikã*.

fishing hook:

(1) PFL **kavil*, CL-KK *ka'vil*, CL-LR *'kavil*, KD-LB *?awil*, SK-HEW *kavir*, WL-AB *kafi*, WL-AD *kawi*, WL-LWI *'kawi*, WL-MS *kawil*, WL-PD *kawil*.

fishnet:

(1) PFL **pukət*, CL-KK *'pukət*, CL-LR *'pukat*, KD-LB *pue?*, WL-AB *pukɔ*, WL-AD *'pukət*, WL-MS *pukɔ?*, WL-PD *pukɔ*.

(2) SK-HEW *a:la*, WL-AD *ala*, WL-LWI *'ala?*.

fist:

(1) FL #*sadok*, CL-IL *tad*, CL-KK *sadok*, CL-LK *sadok*, CL-LP *sadok*, CL-LWT *tado*, CL-MR *tadu*, CL-PN *hadok*, EL-LE *sado*, EL-LMT *sado*, SK-HEW *sadok*, WL-BM *sadok*, WL-LG *sado?*, WL-LML *tado*, WL-LWL *sadok*, WL-MD *sadok*, WL-ML *sado*, WL-PK *sadok*, WL-TJ *sadok*, WL-WK *tado*.

(2) CL-LWK *tubuk*, WL-BL *tubuk*, WL-BT *tubuk*, WL-DL *tubuk*, WL-HRW *tubuk*, WL-IA *tubuk*, WL-KO *tubuk*, WL-LH *tubuk*, WL-LMK *tubuk*, WL-LTB *tubuk*, WL-RE *tubuk*, WL-WB *tubuk*, WL-WT *tubuk*, WL-WW *tubuk*.

five:

(1) PFL *lima ‘arm; hand; five’, CL-KK *lem*, CL-LR *'lema*, KD-LB *lemε*, KD-LW *lemε*, SK-HEW *lima*, SK-TA *lima*, WL-AB *lema*, WL-AD *'lema*, WL-BN *lem:a*, WL-LWI *'lema*, WL-MS *lema*, WL-PD *lema*.

flat:

- (1) CL-PN *nədəkən*, WL-MS *nadak*, WL-PD *nadak*.
 (2) PFL *napu-k ‘flat; stream; river’, KD-LB *na'po?*, WL-AB *na'pukuŋ*, WL-IA *nəpa*, WL-PD *napuk*.
 (3) WL-BM *əbʔã*, WL-BP *əbʔã*, WL-BT *əbʔã*, WL-DL *əbʔã*, WL-HRW *əbʔã*, WL-KO *əbʔã*, WL-LG *əbʔã*, WL-LWI *'əbaʔaŋ*, WL-LWL *əbʔã*, WL-RE *əbã*, WL-TJ *əbʔã*, WL-WB *əbaʔ*, WL-WW *əbʔã*.
 (4) LH #lerek, CL-IL *lerekən*, CL-KK *lerek*, CL-KK *le'rek*, CL-LK *lerek*, CL-LP *lereken*, CL-LR *lere'kən*, CL-LWK *lerek*, CL-LWT *lerekən*, CL-MR *lerek*, EL-LMT *lerʔẽ*, WL-BL *lere*, WL-LH *lerʔe*, WL-LML *lere*, WL-MD *lere*, WL-ML *lere*, WL-WK *lere*.
 (5) WL-BN *hama*, WL-LMK *hama*.
 (6) SK-HEW *dətu*, SK-TA *dətu-ŋ*.

flea:

- (1) PFL *təməla, CL-KK *təmə'lar*, CL-KK *təmə'la*, CL-LR *təmə'lar*, KD-LB *mεle*, SK-HEW *məla*, WL-AB *aho ta'mela*, WL-AD *tə'məla*, WL-MS *mə'r:ε*.
 (2) WL-MS *bəgal*, WL-PD *begal*.

flower:

- (1) PFL *buŋa/*puŋa, CL-IL *puŋan*, CL-KK *'buŋa*, CL-KK *puŋan*, CL-LK *puŋa*, CL-LP *puŋa*, CL-LR *'buŋa*, CL-LWK *buŋaru*, CL-LWT *puŋa*, CL-MR *puŋa*, CL-PN *puŋan*, WL-BL *puŋa*, WL-BN *buŋa*, WL-LTB *buŋa*, WL-PD *buŋa*.
 (2) PFL *puhun, EL-LE *puhũ*, EL-LMT *puhũ*, KD-LB *pu'hun*, KD-LW *puhun*, SK-HEW *puhun*, SK-MM *puhun*, SK-TA *puhu-ŋ*, WL-AB *puhuŋ*, WL-AD *'puhũ*, WL-BM *puhũ*, WL-BP *puhũ*, WL-BT *puhũ*, WL-DL *puhũ*, WL-HRW *puhũ*, WL-IA *puhun*, WL-KO *puhũ*, WL-LG *puhũ*, WL-LH *puhũ*, WL-LMK *puhũ*, WL-LML *pu*, WL-LWI *'puhuŋ*, WL-LWL *puhũ*, WL-MD *puhũ*, WL-ML *puhu*, WL-MS *puhuŋ*, WL-PK *puhũ*, WL-RE *puhũ*, WL-TJ *puhũ*, WL-WB *puhũ*, WL-WK *puhu*, WL-WT *puhũ*, WL-WW *puhũ*.

fly:

(1) PFL ***kə-namuk**, CL-KK *kəma'muk*, CL-LR *kəma'muk*, WL-AB *ka'namuŋ*, WL-LWI *kə'namu*, WL-MS *namuʔ*, WL-PD *k'namu*.

flying fox:

(1) WL-MS *mi'hua*, WL-PD *mihua*.

(2) CL-KK *kəmi'ker*, CL-KK *kəmi'ki*, CL-LR *mi'ker*.

foam:

(1) PFL ***vuda**, CL-KK *pə'vudʒa*, CL-LR *'vudʒan*, KD-LB *wuran*, SK-HEW *vuran*, WL-AB *furaŋ*, WL-AD *wurhã*, WL-LWI *wu'rahaŋ*, WL-MS *buran*, WL-PD *wuraŋ*.

foot:

(1) LH-KD ***lei**, CL-IL *linasa*, CL-KK *lir*, CL-KK *liga*, CL-LK *linaga*, CL-LP *leiha*, CL-LR *li*, CL-LWK *liga*, CL-LWT *linaha*, CL-MR *lidʒesa*, CL-PN *lig*, EL-LE *lei*, EL-LMT *lei*, KD-LB *lei*, KD-LW *læi*, WL-AB *leiŋ*, WL-AD *'leikʒt tʒ'nəpʒã*, WL-BL *leik*, WL-BM *lei*, WL-BN *leiŋ ka'lumak*, WL-BP *lei*, WL-BT *lei*, WL-DL *leĩ*, WL-HRW *leik*, WL-IA *leik*, WL-KO *leik*, WL-LG *lei*, WL-LH *leik*, WL-LMK *leik*, WL-LML *leite*, WL-LTB *lei*, WL-LWI *kəan 'leiŋ*, WL-LWL *lei*, WL-MD *leik*, WL-ML *lei*, WL-MS *leiŋ*, WL-MS *leiʔ*, WL-PD *leiŋ*, WL-PK *lei*, WL-RE *lei*, WL-TJ *lei*, WL-WB *lei*, WL-WK *leik*, WL-WT *lei*, WL-WW *leĩ*.

(2) PFL ***vaʔi**, SK-HEW *vaʔi*, SK-HEW *vaʔi*, SK-MM *vaʔin*, SK-TA *βaʔi*, SK-TA *βaʔi-ŋ*.

footprint:

(1) LH-SK ***əpak** 'palm of hand; footprint', CL-KK *liga əpak*, CL-LK *liga əpak*, CL-LP *leiha əpak*, CL-LWK *liga əpak*, CL-LWT *linaha əpak*, CL-PN *lig əpak*, EL-LE *lei pã*, EL-LMT *lei əpaʔãŋ*, SK-HEW *əpak*, WL-BL *əpa*, WL-BP *lei əpã*, WL-BT *lei tənəpã*, WL-DL *leĩ tənəpã*, WL-HRW *leik əbʔã*, WL-IA *əpakən*, WL-KO *leik əbã*, WL-LH *əpaŋ*, WL-LMK *leik əpã*, WL-LWL *lei əpʔã*, WL-MD *leik əpã*, WL-PK *əbʔã*, WL-RE *lei əbã*, WL-WB *lei əpã*, WL-WK *leik əpã*, WL-WW *leĩ əbʔã*.

(2) WL-BM *makat*, WL-LG *makat*, WL-LTB *maka*, WL-TJ *makat*.

(3) CL-IL *linasa or*, CL-MR *lidʒesa or*, WL-ML *lei onã*.

forehead:

- (1) WL-AB *kala'raŋ*, WL-MS *k'larak*, WL-PD *kə'larak*.
 (2) LH **kənito*, CL-KK *kəni'tɔ*, CL-LR *kəni'to*, WL-AD *kə'ni:to'kət*, WL-LWI *kəni'tɔʔ*.

forest:

- (1) KD-LW *uŋaŋ*, WL-BN *u'taŋ*, WL-LMK *utan*, WL-LWI *'hutaŋ*, WL-PD *utaŋ*.
 (2) WL-MS *dəhək*, WL-PD *də'hək*.
 (3) PFL **tuŋan*, CL-KK *'duan*, KD-LB *tu'wen*, SK-HEW *tuan bəleŋ*, SK-HEW *tuŋan*, SK-MM *tuŋan*, WL-KO *dua*, WL-MS *duaŋ*, WL-TJ *duã*.
 (4) KD-LW *ɛʔa kain*, KD-LW *kain laraŋ*.
 (5) WL-AD *rukʔã*, WL-BT *rukʔã*, WL-DL *rukʔã*.
 (6) WL-BM *woʔonã*, WL-BP *nəwa onã*, WL-LG *nəwoʔona*, WL-LWL *nəwa onã*.
 (7) CL-IL *ekamaŋan*, CL-KK *maŋak*, CL-KK *ma'ŋak*, CL-LK *maŋak*, CL-LP *ekamaŋan*, CL-LWK *maŋak*, CL-LWT *eka maŋan*, WL-BL *eka maŋan*.
 (8) WL-HRW *məgã*, WL-IA *məgən*, WL-ML *məgə*, WL-WK *əka məge*.
 (9) WL-LH *bənəwa*, WL-MD *bənəwa*.
 (10) CL-MR *ekapənət*, WL-LML *eka pənət*, WL-RE *pənɤ*.
 (11) CL-LR *'knovot*, CL-PN *kowoten*.
 (12) WL-PK *ruka onã*, WL-WT *rukʔã*, WL-WW *rukã*.

four:

- (1) PFL **əpat*, CL-KK *pa:t*, CL-LR *'patə*, KD-LB *a'paʔ*, KD-LW *apaʔ*, WL-AB *pa*, WL-AD *'pa:t*, WL-BN *pa:*, WL-LWI *pa:*, WL-MS *paʔ*, WL-PD *pa*.
 (2) SK-HEW *hutu*, SK-TA *hutu*.

friend:

- (1) WL-AB *ba'r:ə ba'r:ə*, WL-PD *ba'r:ə*.
 (2) LH-KD *(*kəne*) *breuŋ*, CL-KK *kə'neu*, CL-LR *kəne'breuŋ*, SK-HEW *deuŋ*, WL-AD *'reu*, WL-LWI *'breuŋ*.

frog:

- (1) CL-KK *kəmə'tɔr*, WL-AB *matɔ*, WL-AD *'mɜto*, WL-LWI *'mɜtɔ*, WL-MS *mə'tɔ*, WL-PD *ma'tɔ*.
 (2) CL-KK *'tulɔr*, CL-KK *tulu*, KD-LB *tepatu'le*.

(3) WL-AB *i'hikiŋ*, WL-MS *i'hik*, WL-PD *ihik*.

fruit:

(1) PFL **vua*, CL-IL *fuak*, CL-KK *vu'ak*, CL-KK *wuak*, CL-LK *wuak*, CL-LP *fuak*, CL-LR *'vuak*, CL-LWK *wuak*, CL-LWT *uak*, CL-MR *fuak*, CL-PN *wuakən*, EL-LE *wuʔɔ*, EL-LMT *wuʔã*, KD-LB *uan*, KD-LW *yan*, SK-HEW *vuan*, SK-HEW *vuan*, SK-MM *vuan*, SK-TA *βua-ŋ*, WL-AD *kajo 'wuã*, WL-BL *wua*, WL-BM *wuã*, WL-BP *wuã*, WL-BT *wuã*, WL-DL *wuã*, WL-HRW *wuã*, WL-IA *wuan*, WL-KO *wuã*, WL-LG *wuʔã*, WL-LH *wuã*, WL-LMK *wuã*, WL-LML *fuã*, WL-LTB *wuã*, WL-LWI *'wuaŋ*, WL-LWL *wuã*, WL-MD *wuã*, WL-ML *wuɣ*, WL-PK *wuʔã*, WL-RE *wuã*, WL-TJ *wua*, WL-WB *wuã*, WL-WK *wuã*, WL-WT *wuã*, WL-WW *wuã*.

full:

(1) PFL **m-pənu*, CL-KK *mə'nuk*, CL-LR *mənu'kən*, KD-LB *penu*, KD-LW *pənu mənu*, SK-HEW *bənu*, SK-TA *bənu*, WL-AB *pan:o*, WL-AD *'pɜnɔ*, WL-BN *pan:ɔŋ*, WL-LWI *'mənuŋ*, WL-MS *pə'nɔŋ*, WL-PD *pa'n:ɔŋ*.

(2) KD-LW *ɪhi*, KD-LW *ɪhiʔ*.

garden:

(1) CL-KK *eka*, CL-KK *ekan*, WL-AB *ekaŋ*, WL-MS *ekaŋ 'butaʔ*, WL-MS *ekaŋ*, WL-PD *ekaŋ*.

(2) PFL **uma*, KD-LB *lumar*, SK-HEW *uma*, WL-AD *mã:*, WL-LWI *ma:ŋ*.

good:

(1) LH-KD #*dikə-n* 'good; person', EL-LE *diʔe*, KD-LB *diʔən*, KD-LW *diʔen herun*, WL-AB *dike*, WL-BN *dik:ɛ*, WL-MS *dikek*, WL-PD *dikek*.

(2) CL-LP *məla*, CL-LWT *məla*, CL-MR *məla*, WL-AD *'mɜʔlaʔ*, WL-BL *məla*, WL-BT *məlʔã*, WL-DL *məlaʔ*, WL-HRW *məlʔã*, WL-KO *məlʔã*, WL-WK *məlä*, WL-WW *məlä*.

(3) CL-IL *sənaɾəkən*, WL-BM *əre*, WL-IA *sare*, WL-LG *əreʔ*, WL-LMK *ərʔě*, WL-LML *sənaɾeŋ*, WL-LTB *sarě*, WL-LWI *ə'reŋ*, WL-RE *sənaɾe*, WL-WT *əre*.

(4) LH-KD #*alus*, CL-KK *alus*, CL-KK *a'lus*, CL-LK *alus*, CL-LR *alu'hən*, CL-LWK *alus*, CL-PN *aluhən*, WL-LH *alus*, WL-MD *alusã*, WL-ML *alusɣ*.

(5) WL-BP *mae*, WL-LWL *măe*, WL-TJ *mae*, WL-WB *mae*.

(6) SK-HEW *əpan*, SK-MM *ʔəpan*, SK-TA *ʔepa-ŋ*.

grass:

- (1) WL-AB *la'duɲ*, WL-BN *la'duɲ*, WL-MS *lə'duɲ*, WL-PD *laduɲ*.
- (2) PFL **udu*, KD-LB *wuru*, KD-LW *ruʔ vadɛʔ*, KD-LW *ɣru*, SK-TA *uru-ŋ*.
- (3) WL-AD *kə'dəku*, WL-KO *kəduku*, WL-WT *kədəku*.
- (4) WL-BP *kəramət*, WL-BT *kəramət*, WL-HRW *kəramət*, WL-IA *kəramət*, WL-LG *kəramə*, WL-LH *kəramət*, WL-LML *kəramet*, WL-LWI *kəramə*, WL-LWL *kəramət*, WL-MD *kəramət*, WL-ML *kəramə*, WL-RE *kəramə*, WL-WW *kəramət*.
- (5) WL-BM *luʔo*, WL-DL *luʔo*, WL-LMK *luʔo*, WL-LTB *luʔo*, WL-PK *luʔo*, WL-WB *luʔo*.
- (6) CL-IL *kərakasa*, CL-KK *kərakak*, CL-KK *krə'kak*, CL-LK *kərakak*, CL-LP *kərakə*, CL-LWT *kərakə*, WL-BL *kərakak*.
- (7) EL-LE *gəpe*, EL-LMT *gəpe*.
- (8) CL-LR *'mevau*, CL-LWK *mewak*, CL-PN *mewaw*.
- (9) SK-HEW *vaʔan*, SK-MM *vaʔan*, SK-TA *βaʔa-ŋ*.
- (10) SK-HEW *riʔi*, SK-TA *riʔi*.

grasshopper:

- (1) LH #*madu*, CL-KK *ma'dʒu*, CL-KK *ma'dʒɔr*, CL-LR *ma'yor*, WL-AB *ka'-marɔ*, WL-AD *kə'marɔʔ*, WL-MS *marɔʔ*, WL-PD *k'marɔ*, WL-PD *marɔ*.

guilty:

- (1) PFL **sala* 'wrong; guilty', CL-KK *sna'lak*, CL-LR *hənalə'kən*, SK-HEW *dəna'hala*, WL-AB *hala*, WL-AD *'nalã*, WL-LWI *'nalay*, WL-MS *mariŋ'go'hala*, WL-PD *hala*.

guts:

- (1) CL-KK *tai ka'luɲ*, KD-LB *tein*, SK-HEW *tain vaten*, SK-TA *taʔi-ŋ*, WL-AB *teiŋ ka'rikiŋ*, WL-BN *teiŋ 'ɔnɔŋ*, WL-LWI *'taiŋ 'onəʔəŋ*, WL-MS *taiŋ 'beiŋ*, WL-MS *taiŋ kə'rik*, WL-PD *taiŋ 'kar:ik*.
- (2) CL-KK *kəbɔ'ti*, CL-LR *kəbotin*, WL-AD *kə'botĩ*.

hair:

- (1) LH #*latar*, CL-IL *latarəs*, CL-KK *latar*, CL-KK *la'tar*, CL-LK *latar*, CL-LP *latar*, CL-LR *la'tar*, CL-LWK *rata*, CL-LWT *latarha*, CL-MR *lata*, CL-PN *latar*, EL-LE *əte ratã*, EL-LMT *ratã*, WL-AB *ratã*, WL-AD *'rataʔ*, WL-BL *ratak*, WL-BM *rata*, WL-BN *rata*, WL-BP *rata*, WL-BT *rata*, WL-DL *'rataʔ*, WL-HRW *rata*, WL-IA *rata*, WL-KO *'rataʔ*, WL-LG *'rataʔ*, WL-LH *rata*, WL-LMK *rata*, WL-LML

rata, WL-LTB *rata*, WL-LWI 'rata?', WL-LWL *rata*, WL-MD *rata*, WL-ML *rata*, WL-MS *rata*, WL-PD *rata*, WL-PK *ratak*, WL-RE *rata*, WL-TJ *rata*, WL-WB *rata*, WL-WK *rata*, WL-WT *rata*, WL-WW *ratak*.

(2) PFL **vuhak*, KD-LB *ʔu'ha*, KD-LW *uha*.

(3) SK-HEW *loʔe*, SK-HEW *loʔe*, SK-HEW *loʔe roun*.

(4) SK-MM *alan roun*, SK-TA *ala*, SK-TA *ala-ŋ*, SK-TA *ala-ŋ ʔrou-ŋ*.

hand:

(1) PFL **lima* 'hand; arm; five', CL-IL *limasa*, CL-KK *li'ma*, CL-KK *limaga*, CL-LK *limaga*, CL-LP *limaha*, CL-LR *li'ma*, CL-LWK *limaga*, CL-LWT *limaha*, CL-MR *limasa*, CL-PN *limag*, EL-LE *lima*, EL-LMT *lima*, KD-LB *liŋ*, KD-LW *liŋ*, SK-HEW *lima*, SK-MM *liman*, SK-TA *lima-ŋ*, WL-AB *limanŋ*, WL-BL *limak*, WL-BM *lima*, WL-BN *limanŋ ka'lumak*, WL-BP *lima*, WL-BT *lima*, WL-DL *lima*, WL-HRW *limak*, WL-IA *limak*, WL-KO *limak*, WL-LG *lima*, WL-LH *limak*, WL-LMK *limak*, WL-LML *limate*, WL-LTB *lima*, WL-LWI 'limanŋ', WL-LWL *lima*, WL-MD *limak*, WL-ML *lima*, WL-MS *limanŋ*, WL-MS *limaʔ*, WL-PD *limanŋ*, WL-PK *lima*, WL-RE *lima*, WL-TJ *lima*, WL-WB *lima*, WL-WK *limak*, WL-WT *limã*, WL-WW *limak*.

head:

(1) LH #*kote*, EL-LE *ate*, EL-LMT *ate*, WL-AB *kɔtɔŋ*, WL-AD 'kotɔk, WL-AD *kotɔm*, WL-BL *kotək*, WL-BM *kote*, WL-BN *kɔtɔŋ*, WL-BP *kotət*, WL-BT *kotək*, WL-DL *kotəŋ*, WL-HRW *kotək*, WL-IA *kotɔ*, WL-KO *kotək*, WL-LG *kotã*, WL-LH *kotəkə*, WL-LMK *kotʔã*, WL-LML *kotãte*, WL-LTB *kotə*, WL-LWI 'kotəʔ', WL-LWL *kotə*, WL-MD *kotəkət*, WL-ML *kotə*, WL-MS *kɔtɔŋ*, WL-MS *kɔtɔʔ*, WL-PD *kɔtɔŋ*, WL-PK *kotã*, WL-RE *kotã*, WL-TJ *kotək*, WL-WB *kotə*, WL-WK *kotẽ*, WL-WT *kotʔã*, WL-WW *kotã*.

(2) KD-LB *tu'bar*, KD-LW *tubar*.

(3) CL-IL *ulusa*, CL-KK *u'lu*, CL-KK *ulun*, CL-LK *ulu*, CL-LP *ulu*, CL-LR *u'lu*, CL-LWK *uluga*, CL-LWT *uluha*, CL-MR *ulu*, CL-PN *ulug*.

(4) SK-HEW *loʔe*, SK-HEW *loʔen*.

(5) SK-MM *alan*, SK-TA *ala-ŋ*.

heart:

(1) WL-AB *kubanŋ*, WL-MS *tap'kubanŋ*, WL-PD *tapɔ 'kubanŋ*.

(2) KD-LB *taʔ'mutiʔ*, KD-LW *mutiʔ*.

(3) PFL **puhum*, CL-IL *puosa*, CL-KK *puo*, CL-KK *pu'o*, CL-LK *puo*, CL-LP

puoha, CL-LR *puon*, CL-LWK *puo*, CL-LWT *puoha*, CL-MR *puo*, CL-PN *puon*, EL-LE *pũhũ*, EL-LMT *puho*, SK-TA *puhu-ŋ βua-ŋ*, WL-AD *'puho*, WL-BL *puo*, WL-BM *puhũ*, WL-BT *puho*, WL-DL *puho*, WL-HRW *puho*, WL-IA *puho*, WL-KO *puho*, WL-LG *puhũ*, WL-LH *puhokə*, WL-LMK *puho*, WL-LML *puo*, WL-LTB *puhũ*, WL-MD *puho*, WL-ML *puhu*, WL-RE *puho*, WL-TJ *puhũ*, WL-WB *puhũ*, WL-WK *puho*, WL-WT *puho*, WL-WW *puho*.

(4) SK-HEW *vuon*, WL-BP *pəwuã*, WL-LWL *wuak*.

heavy:

(1) PFL **bərat*, CL-IL *bəra*, CL-KK *bənə'rat*, CL-KK *bərat*, CL-LK *bərat*, CL-LP *bəra*, CL-LR *bənəra'tən*, CL-LWK *bərat*, CL-LWT *bera*, CL-MR *bəra*, CL-PN *bənəratən*, EL-LE *bəra*, EL-LMT *bəra*, KD-LB *ba'ra?*, KD-LW *bara? betɛ*, SK-HEW *bərat*, SK-MM *bərat*, SK-TA *bərat*, WL-AB *ba:*, WL-AD *'ba?at*, WL-BL *bāt*, WL-BM *ba?at*, WL-BN *ba:*, WL-BP *ba?at*, WL-BT *ba?at*, WL-DL *ba?at*, WL-HRW *ba?at*, WL-IA *baat*, WL-KO *ba?at*, WL-LG *ba?a*, WL-LH *ba?at*, WL-LMK *ba?a*, WL-LML *batã*, WL-LTB *ba?a*, WL-LWI *'ba?a*, WL-LWL *ba?at*, WL-MD *bəna?at*, WL-ML *baa*, WL-MS *bak*, WL-MS *ba?*, WL-PD *ba:*, WL-PK *ba?at*, WL-RE *ba?a*, WL-TJ *ba?at*, WL-WB *ba?at*, WL-WK *ba?at*, WL-WT *ba?at*, WL-WW *bat*.

heel:

(1) WL-AB *leiŋ 'ubɔŋ*.

(2) LH *#nudəp*, CL-KK *nu'dəp*, CL-LR *'udəm*, WL-AD *'nudəɾɪt*, WL-LWI *'ude*, WL-MS *lei 'udek*, WL-MS *leiŋ 'udek*, WL-PD *leiŋ 'udek*.

hill, ridge:

(1) WL-AB *dɔla*, WL-MS *dɔlu*.

(2) WL-MS *wɔtɔ 'wɛɔ?*, WL-PD *wɔtɔ*.

(3) WL-AD *woka*, WL-LWI *wɔka*.

(4) LH-SK *#volon*, CL-KK *'volo*, CL-KK *'volon*, SK-HEW *volon*.

hole:

(1) WL-AB *fəŋɔ*, WL-MS *wəŋɔ?*, WL-PD *wəŋɔ*.

(2) LH-KD *#vuok*, CL-LR *'vuok*, KD-LB *uɔ?*, WL-MS *nuɔ?*.

(3) LH *#uman*, CL-KK *u'ma*, CL-KK *u'man*, WL-LWI *'uman*.

horn:

- (1) WL-AB *u'haraj*, WL-BN *ru'ha:*, WL-MS *huar*, WL-PD *huar*.
 (2) LH-SK *#tara*, CL-KK *ta'ra*, SK-HEW *taran*, SK-TA *tara-ŋ*, WL-AD *ta'rã*, WL-LWI *'taraj*.

hot:

- (1) LH *#plati/kati*, CL-IL *kat*, CL-KK *ka'ti*, CL-KK *katiŋ*, CL-KK *kna'tiŋ*, CL-LK *kati*, CL-LP *kati*, CL-LR *kənati'ŋan*, CL-LWK *kati*, CL-LWT *kati*, CL-MR *kati*, CL-PN *katiŋən*, EL-LE *pati*, EL-LMT *pati*, WL-AB *pa'latiŋ*, WL-AD *'plate*, WL-BL *pəlate*, WL-BM *pəlate*, WL-BN *pa'latiŋ*, WL-BP *pəlate*, WL-BT *pəlate*, WL-DL *pəlate*, WL-HRW *pəlate*, WL-IA *pəlate*, WL-KO *pəlate*, WL-LG *pəlate*, WL-LH *pəlate*, WL-LMK *pəlate*, WL-LML *pəlate*, WL-LTB *pəlate*, WL-LWI *'pəlate*, WL-LWL *pəlate*, WL-MD *pəlate*, WL-ML *pəlate*, WL-MS *p'latiŋ*, WL-PD *p'latiŋ*, WL-PK *pate*, WL-RE *pəlate*, WL-TJ *pəlate*, WL-WB *pəlate*, WL-WK *pəlate*, WL-WT *pəlate*, WL-WW *pəlate*.
 (2) KD-LW *daja*, SK-HEW *dara*, SK-HEW *dara gahu?*.
 (3) SK-HEW *gahu*, SK-MM *gahu*, SK-TA *gahu*.
 (4) SK-TA *rou*, SK-TA *rou rada-ŋ*.

house:

- (1) PFL **ruma*, WL-AB *uma*, WL-BN *ʔuma*, WL-MS *uma*, WL-PD *uma*.
 (2) LH-KD **vetak* 'barn; house', KD-LB *we'ta?*.
 (3) WL-AD *'layɔ*, WL-LWI *'layɔ?*.
 (4) PFL **una* 'inside', CL-KK *'una*, CL-KK *u'nan*, CL-LR *'unan*.

how:

- (1) CL-KK *nao ga*, CL-LR *naboga*, KD-LB *nara 'bɔnɛ*, KD-LW *nara bɔnɛ*, WL-AB *namɔ'nanɔga*, WL-AD *nɜ'gaku*, WL-BN *na'mɔnɔnga*, WL-LWI *naʔaŋ'gaʔe*, WL-MS *nəmɔnga*, WL-MS *nəm'ga*, WL-PD *naməŋ'ga*.
 (2) SK-HEW *ganu upan*, SK-TA *ganupae*.

how much, how many:

- (1) PFL **pida*, CL-KK *'pira*, CL-LR *pira*, KD-LB *pie*, SK-HEW *pira*, WL-AB *pira*, WL-AD *'pira*, WL-BN *pir:a*, WL-LWI *'pira*, WL-MS *pira*, WL-PD *pira*.

hungry:

- (1) LH **maluv*, CL-KK *ma'luv*, CL-LR *'tei 'maluv*, WL-AB *tukarŋ 'malu*, WL-

AD 'malu, WL-LWI 'malu, WL-MS *tukaŋ* 'malu, WL-MS *tuka?* 'malu, WL-PD *tukaŋ* 'malu.

inside:

(1) PFL ***una**, SK-HEW *unen*, SK-MM *une*, WL-AB *ɔnɔŋ*, WL-AD 'ʔɔnʔeʔ, WL-BL *on*, WL-BM *onə*, WL-DL *onʔã*, WL-HRW *onʔã*, WL-KO *dəonã*, WL-LG *onə*, WL-LH *onã*, WL-LTB *onã*, WL-LWI 'ɔnɔ, WL-ML *onɔ*, WL-MS *ɔnɔŋ*, WL-PD *ɔnɔŋ*, WL-RE *onã*, WL-TJ *onə*, WL-WT *onʔã*, WL-WW *onã*.

(2) PFL ***laləm**, KD-LB *bɛ?* 'laleŋ.

(3) PFL ***bə-ləma**, EL-LE *ləmə*, KD-LB *bɛ?*, SK-HEW *ləman*, WL-BP *bəlomə*, WL-BT *bəlomã*, WL-IA *lomək*, WL-LMK *bəlomã*, WL-LML *belomã*, WL-LWL *bəlomo*, WL-MD *bəlomã*, WL-PK *ləmã*, WL-WB *bəlomə*, WL-WK *bəlome*.

(4) CL-IL *goməkən*, CL-KK *gomək*, CL-LK *gom*, CL-LP *gom*, CL-LWT *goməkən*, CL-MR *gomək*, CL-PN *goməkən*.

(5) CL-KK *dʒe ... or*, CL-LR *dʒe ... 'ora*, CL-LWK *or*.

itchy:

(1) PFL ***gatər**, CL-KK *gə'tək*, SK-HEW *gatar*, WL-AB *gate*, WL-AD 'gatsk, WL-LWI 'gatə, WL-MS *gate?*, WL-PD *gate*.

jackfruit:

(1) SK-HEW *nakat*, WL-AB *naŋka*, WL-MS *nakaŋ*, WL-PD *naka*.

(2) LH-KD #**kverak**, CL-KK *kve'rak*, CL-LR *kə'verak*, KD-LB *wɛ'ra?*, WL-AD 'kwɛrak, WL-LWI *kə'wɛra*.

knee:

(1) LH-KD #**kudul**, KD-LB 'ɔdul, KD-LW *udul*, WL-BN 'leiŋ 'kudul, WL-MS *lei 'kudul*, WL-MS *leiŋ 'kudul*, WL-MS *kudul*, WL-PD *kudul*.

(2) PFL ***lotur**, CL-KK *lo'tər*, CL-LR *lo'tor*, SK-HEW *tur*, SK-TA *βaʔi-ŋ tur*, SK-TA *tur*, WL-AD 'lotorɜt, WL-LWI 'lətər.

knife:

(1) WL-AB *duri*, WL-BN *du'ri*, WL-MS *duri?*, WL-PD *duri*.

(2) LH-KD #**me(r)it**, CL-KK *mə'rit*, KD-LB *mei?*.

(3) WL-AD *hepe?*, WL-LWI 'hepe?

leaf:

(1) LH-KD #lolo, CL-IL lolo, CL-KK l'lon, CL-KK lolon, CL-LK lolon, CL-LP lolo, CL-LR 'lolon, CL-LWK lolo, CL-LWT lolo, CL-MR lolo, CL-PN lolon, EL-LE lolō, EL-LMT lolō, KD-LB lolon, KD-LW ai lolon, KD-LW lolon, WL-AB lolon, WL-AD 'lolō, WL-BL lolo, WL-BM lolō, WL-BN lolon, WL-BT lolo, WL-DL lolo, WL-HRW lolo, WL-IA lolon, WL-KO lolo, WL-LH lolō, WL-LMK lolō, WL-LML lolo, WL-MD lolō, WL-ML lolo, WL-MS kaḏḏo 'lolon, WL-PD lolon, WL-PK lolon, WL-TJ lolo, WL-WB lolo, WL-WK lolo, WL-WT lolo, WL-WW lolo.

(2) LH-KD #ləpa 'leaf; sheet; lontar leaf', KD-LW lepa?, WL-BP ləpa, WL-LG lepaʔā, WL-LTB ləpā, WL-LWI 'ləpaŋ, WL-LWL ləpā, WL-RE ləpā.

(3) PFL *down, SK-HEW roun, SK-MM roun, SK-TA ʔrou, SK-TA ʔrou-ŋ.

left side:

(1) CL-IL səmekiŋən, CL-KK se'ke, CL-KK seke, CL-LK seke, CL-LP heke, CL-LR hε'ken, CL-LWK seke, CL-LWT həmekiŋən, CL-MR səmeki, CL-PN heken, WL-AB hεkē, WL-BL seki, WL-BN mekiŋ, WL-IA mekin, WL-LH meki, WL-LML meki, WL-LTB meki, WL-MD meki, WL-MS hekeŋ, WL-PD hekeŋ, WL-WK səmekiŋ.

(2) LH #səmekiŋ, WL-AD neke?, WL-AD 'nekĩ, WL-BM nekĩ, WL-BP nekĩ, WL-BT nekĩ, WL-DL nekĩ, WL-HRW nekiŋ, WL-KO nekin, WL-LG nekĩ, WL-LMK nekĩ, WL-LWI 'nekiŋ, WL-LWL nekĩ, WL-ML neki, WL-RE nekĩ, WL-TJ nekĩ, WL-WB nekĩ, WL-WT nekĩ, WL-WW nekĩ.

(3) PFL *viri, EL-LE wirĩ, EL-LMT wiri, KD-LB weri, KD-LW veri, SK-HEW virin, SK-HEW virin, SK-MM viri, SK-TA βiri, WL-PK wiri.

leg:

(1) LH-KD #lei 'foot; leg', CL-KK li, CL-LR li, KD-LB lei, WL-AB leiŋ, WL-AD 'leikət, WL-BN leiŋ, WL-LWI 'leiŋ, WL-MS leiŋ, WL-MS lei?, WL-PD leiŋ.

(2) PFL *vaʔi 'foot; leg', SK-HEW vaʔi, SK-MM vaʔin, SK-TA βaʔi, SK-TA βaʔi-ŋ.

light (weight):

(1) WL-AB beɾu, WL-MS beɾe, WL-PD beɾe.

(2) LH #kleak, CL-KK kle'ak, CL-LR kəlea'ŋen, KD-LB ʔa'haʔ, SK-HEW heak, WL-AD kə'liεʔ, WL-LWI kə'lea.

lightning:

(1) PFL ***k-silap**, CL-KK *ksi'lap*, CL-LR *'khillap*, KD-LB *kilaʔ*, KD-LB *nilaʔ*, WL-AB *hila bi'ba*, WL-AD *k3'hilat*, WL-LWI *kəɬə'ka kə'hila*, WL-MS *hilaʔ*, WL-PD *hila 'bilonj*.

lime:

(1) PFL ***ʔapur**, CL-KK *a'pər*, CL-LR *'apur*, KD-LB *apur*, SK-HEW *ʔapur*, WL-AB *apu*, WL-AD *ʔapuʔ*, WL-BN *ʔapu*, WL-LWI *'apuʔ*, WL-MS *apu*, WL-PD *apu*.

lips:

(1) PFL ***vivir**, KD-LB *wiwir*, SK-HEW *vivir*, WL-AB *fifiŋ*, WL-BN *fifiŋ*, WL-MS *wiwiŋ*, WL-MS *wiwiʔ*, WL-PD *wiwiŋ*.

(2) CL-LR *'nuhə lə'rit*, WL-AD *'nuhu'kət*, WL-LWI *'nuhuŋ*.

liver:

(1) PFL ***ate-n**, CL-LWK *kəbotik ate*, KD-LB *ha'ten*, SK-HEW *vaten*, SK-HEW *vaten*, SK-MM *vaʔten*, SK-TA *ʔbate*, WL-AB *a'tě*, WL-BM *atě*, WL-BN *hati*, WL-BP *atě*, WL-BT *atek*, WL-HRW *atek*, WL-IA *ate*, WL-LG *ate*, WL-LMK *ate*, WL-LTB *atě*, WL-LWI *'ateŋ*, WL-LWL *atě*, WL-ML *ate*, WL-MS *a'teŋ*, WL-PD *ateŋ*, WL-PK *ate*, WL-RE *ate*, WL-TJ *ate*, WL-WB *ate*, WL-WT *hati*, WL-WW *ate*.

(2) WL-AD *'mehi*, WL-DL *mehĩ*, WL-KO *mehĩ*.

(3) CL-IL *orasa*, WL-BL *onək*, WL-LH *onəkə*, WL-LML *onã*, WL-MD *onəkət*, WL-WK *one*.

(4) CL-KK *ɔ:r*, CL-KK *oren*, CL-LK *or*, CL-LP *oraha*, CL-LR *orən*, CL-LWT *oraha*, CL-MR *orasa*, CL-PN *oraj*.

(5) EL-LE *ũhã*, EL-LMT *huʔa*.

lizard:

(1) PFL ***təkek**, CL-KK *tə'kek*, CL-LR *tə'kek*, KD-LB *tɛ'kɛʔ*, SK-HEW *təke:*, WL-AB *ta'kɛ*, WL-AD *t3ke*, WL-LWI *'təkeʔ*, WL-MS *tə'k:ɛʔ*, WL-PD *ta'kɛ*.

long:

(1) WL-AB *bəla'hakaŋ*, WL-AB *bəla'hakaŋ*, WL-AD *b3lah3'a*, WL-AD *'b-laha*, WL-BM *bəlahə*, WL-BN *ba'lahə*, WL-BP *bəlahã*, WL-BT *bəlahã*, WL-DL *bəlahã*, WL-HRW *bəlahʔã*, WL-IA *bəlahakən*, WL-KO *bəlahʔã*, WL-LG *bəlahə*,

WL-LH *bəlahã*, WL-LMK *bəlahã*, WL-LML *blã*, WL-LTB *bəlahã*, WL-LWI *bə'laha*, WL-LWL *bəläha*, WL-MD *bəlahã*, WL-ML *bəlahɤ*, WL-MS *b'lahak*, WL-PD *b'lahak*, WL-PK *bahã*, WL-RE *bəlähã*, WL-TJ *bəlahã*, WL-WB *bəlahã*, WL-WK *bəlakã*, WL-WT *bəlahã*, WL-WW *bəlahã*.

(2) KD-LB *la'wa:*, KD-LW *lava dəkar*.

(3) PFL **doa* 'far; long', CL-KK *doak*, CL-KK *do'ak*, CL-LK *doak*, CL-LR *do'ak*, CL-LWK *doak*, CL-PN *doakən*, EL-LMT *doʔã*, KD-LW *ləla dɔaʔ*.

(4) LH-SK #*blavir* 'long; far', CL-IL *lafakən*, CL-LP *lafã*, CL-LWT *belamikən*, CL-MR *bəlamik*, EL-LE *bawã*, WL-BL *bəlawi*.

(5) SK-HEW *blo:n*, SK-MM *blo:n*, SK-TA *blo:ŋ*.

lontar palm:

(1) LH-SK **koli*, CL-KK *'koli*, CL-KK *'kolɛr*, CL-LR *ko'ler*, SK-HEW *koli*, WL-MS *koɔli* 'pukɔŋ, WL-PD *kɔli*.

louse in clothing:

(1) PFL **tuma*, CL-KK *kətu'ma*, CL-KK *kətu'maw*, CL-LR *kətu'mau*, WL-AD *ke'tumã*, WL-LWI *kə'tuma*.

louse in hair:

(1) PFL **kutu*, CL-IL *kutor*, CL-KK *kutu*, CL-KK *ku'tɔr*, CL-LR *ku'tor*, CL-LK *kutu*, CL-LP *kutodʒi*, CL-LWK *kutoru*, CL-LWT *kutodʒa*, CL-MR *kutoru*, CL-PN *kutoj*, EL-LE *uto*, EL-LMT *uto*, KD-LB *ʔutu*, KD-LW *utu*, SK-HEW *utu*, SK-HEW *ʔutu*, SK-MM *ʔutu*, SK-TA *ʔutu*, WL-AB *kuto*, WL-AD *'kutɔ*, WL-BL *kuto*, WL-BM *kuto*, WL-BN *ku'tɔ*, WL-BP *kuto*, WL-BT *kuto*, WL-DL *kuto*, WL-HRW *kuto*, WL-IA *kuto*, WL-KO *kuto*, WL-LG *kuto*, WL-LH *kuto*, WL-LMK *kuto*, WL-LML *kuto*, WL-LTB *kuto*, WL-LWI *'kuto*, WL-LWL *kuto*, WL-MD *kuto*, WL-ML *kuto*, WL-MS *kutɔ*, WL-PD *kuto*.

WL-PK *kuto*, WL-RE *kuto*, WL-TJ *kuto*, WL-WB *kuto*, WL-WK *kuto*, WL-WT *kuto*, WL-WW *kuto*,

lungs:

(1) LH #*vuda*, CL-KK *vu'dʒa*, CL-LR *'ore vu'dʒan*, WL-AB *furaŋ*, WL-AD *'wurhã*, WL-MS *wuraŋ*, WL-PD *wuraŋ*.

machete:

(1) LH-KD #*pada*, CL-KK *pe'dar*, CL-KK *pe'da*, CL-LR *'pada*, KD-LB *pe'daʔ*, WL-AB *sa'pada*, WL-AD *pada*, WL-BN *pada*, WL-MS *padaʔ*, WL-PD *pada*. From

Portuguese *espada* 'sword'.

man:

(1) PFL * (**kabe**) **laki**, CL-LR *la'kin*, CL-MR *analaki*, CL-PN *lakin*, SK-HEW *lai*, SK-HEW *laʔit*, SK-MM *laʔi*, SK-TA *ata laʔi*, WL-AB *ka'lake*, WL-AD *ʔama lake*, WL-BL *kəbailake*, WL-BM *amalake*, WL-BN *ka'lake*, WL-BP *bəlaki*, WL-BT *bəlakĩ*, WL-DL *amalake*, WL-HRW *amalake*, WL-IA *amalake*, WL-KO *amalake*, WL-LG *amalake*, WL-LH *bəlake*, WL-LMK *bəlakeʔ*, WL-LML *kəbaelake*, WL-LTB *bailaki*, WL-LWI *ina məla'ke*, WL-LWL *amalake*, WL-MD *kəbəlake*, WL-ML *kəblake*, WL-MS *k'lake*, WL-PD *k'lake*, WL-PK *kake*, WL-RE *bəlake*, WL-TJ *amalake*, WL-WB *bəlaki*, WL-WK *kəlake*, WL-WT *bəlake*, WL-WW *amalake*.

(2) KD-LB *anaʔ 'abε*, KD-LW *anaʔ aβε*.

(3) CL-IL *lamen*, CL-KK *lamen*, CL-KK *la'men*, CL-LK *lamedʒa*, CL-LP *analame*, CL-LWK *lame*, CL-LWT *ata lamen*.

(4) EL-LE *rəbã*, EL-LMT *reba*.

mango:

(1) PFL ***pau**, CL-KK *'pau*, CL-KK *'paor*, CL-LR *'paor*, KD-LB *peu*, SK-HEW *pau*, WL-AB *pao*, WL-AD *'pau*, WL-LWI *'pa:u*, WL-MS *pao*, WL-PD *pao*.

many:

(1) WL-MS *rasa*, WL-PD *rasa*.

(2) LH-KD **#rai**, CL-IL *raiŋa*, CL-KK *raik*, CL-KK *'raik*, CL-LK *rai*, CL-LP *rai*, CL-LR *'rɔi*, CL-LWK *rai*, CL-LWT *rai*, CL-MR *rai*, KD-LB *rai*, KD-LW *rai varan*.

(3) WL-AD *ʔajʔa*, WL-BL *adʒa*, WL-BM *aja*, WL-BP *adʒa*, WL-BT *ajʔã*, WL-DL *arʔã*, WL-HRW *ajʔã*, WL-IA *adʒa*, WL-KO *ajã*, WL-LG *ajaʔ*, WL-LH *ajʔã*, WL-LMK *adʒã*, WL-LML *adʒã*, WL-LTB *adʒã*, WL-LWI *'ajaʔ*, WL-MD *ajã*, WL-ML *adʒa*, WL-PK *ajaʔ*, WL-RE *aja*, WL-TJ *adʒã*, WL-WB *adʒã*, WL-WK *adʒã*, WL-WT *adʒaʔ*, WL-WW *ajaʔ*.

(3) WL-PD *lab:i*, WL-PD *lab:iŋ*, WL-BN *ʔata la'bi:*.

(4) SK-HEW *oher*, SK-HEW *oher*.

market:

(1) CL-KK *vul*, CL-LR *'vulan*, KD-LB *ul*, WL-AB *fu*, WL-AD *'wulɜ*, WL-LWI *'wuləŋ*, WL-MS *wu*, WL-PD *wu 'saluk*.

marungga:

(1) PFL **motoŋ*, CL-KK *mo'tɔŋ*, CL-KK *mo'tɔ*, CL-LR *'motoŋ*, KD-LB *mɔtɔŋ*, SK-HEW *motoŋ*, WL-AB *mɔtɔŋ*, WL-AD *'motõ*, WL-LWI *'mɔtɔŋ*, WL-MS *mɔtɔŋ*, WL-PD *mɔtɔŋ*.

mat:

(1) PFL **osan*, CL-KK *o'san*, CL-KK *o'sa*, CL-LR *'ohan*, KD-LW *ohan*, SK-HEW *ohaʔ*, WL-AB *ɔ'haŋ*, WL-AD *'ɔhãɜ*, WL-LWI *'ɔhaŋ*, WL-MS *ɔ'haŋ*, WL-MS *hoan*, WL-PD *ɔhaŋ*.

meat:

(1) PFL **isi-k* or **ihi-k*, KD-LW *ihij*, SK-TA *ih-iŋ*, WL-AB *i'hikiŋ*, WL-BM *ihik*, WL-BN *ʔihik*, WL-IA *ihin*, WL-KO *ihik*, WL-LG *ihik*, WL-LML *hik*, WL-LWL *ihik*, WL-MS *hik*, WL-PD *dagiŋ* *'ihi*, WL-RE *ihik*, WL-TJ *ihik*, WL-WT *ihik*.

(2) WL-AD *'umã*, WL-BT *umã*, WL-HRW *umã*.

(3) WL-BL *bənək*, WL-BP *mənəkẽ*, WL-LWI *'mənəkəŋ*.

(4) LH #*elam*, CL-IL *elam*, CL-KK *elam*, CL-KK *e'lam*, CL-LK *elam*, CL-LP *elam*, CL-LR *e'lam*, CL-LWK *elaŋ*, CL-LWT *elam*, CL-MR *eləm*, CL-PN *elam*, EL-LE *əlã*, EL-LMT *elã*, WL-LH *ela*, WL-LML *elã*, WL-MD *elã*, WL-WK *elã*.

(5) WL-LMK *dagiŋ*, WL-WB *dagiŋ*. From Indonesian/Malay *daging* 'meat'.

(6) LH-SK **(meiŋ)ʔətan*, SK-HEW *ətan*, SK-MM *ʔətan*, SK-TA *mei-ŋ* *ʔəta-ŋ*, WL-LTB *mətã*, WL-PK *mətã*, WL-WW *mətã*.

(7) SK-TA *ʔlui-ŋ* *klama-ŋ*, SK-TA *ʔlui-ŋ* *roto-ŋ*, SK-TA *ʔlui-ŋ*, WL-ML *nawək*.

milk:

(1) CL-LR *'susu*, SK-HEW *susu*, WL-AB *susu*, WL-BN *susu*, WL-LWI *'susu*, WL-MS *susu*, WL-PD *susu*.

(2) CL-KK *tu'sɔɜ*, CL-KK *tu'su*, KD-LB *tuʔu* *'wɛɛŋ*, WL-AD *'tuhɔ*.

millet:

(1) LH #*vətəm*, CL-KK *və'təm*, CL-LR *'vətəm*, KD-LB *wɛ'reʔ*, WL-LWI *'wətəŋ*.

monitor lizard:

(1) WL-AB *reha*.

(2) WL-MS *tɛto* *'damar*, WL-PD *təto* *'damar*.

(3) SK-HEW *oti*, WL-AD *'oteʔ*, WL-LWI *'ote*.

(4) CL-KK *bə'pap*, CL-LR *gə'pap*.

moon:

(1) PFL **vulan*, CL-KK *vu'lan*, CL-LR *'vulan dʒe kə'ləm*, KD-LB *wu'la*, KD-LW *ula lɔɔ*, SK-HEW *vulan*, SK-TA *βula-ŋ*, WL-AB *fulaŋ*, WL-AD *'wulã*, WL-BN *fulaŋ*, WL-LWI *'wula*, WL-MS *wulaŋ*, WL-PD *wulaŋ*.

morning:

- (1) WL-AB *fu'ra fu'ra*, WL-MS *wuraʔ*, WL-PD *wura*.
 (2) WL-AD *g'hula*, WL-AD *hə'gula*, WL-LWI *'guləŋ*.
 (3) CL-KK *bə'en*, CL-LR *bae'nan*.

mosquito:

(1) PFL **k<n>əpuŋ/*həpuŋ*, KD-LW (Samely 2013) *hepun*, SK-KR (Pareira and Lewis 1998) *həpuŋ*, WL-AB *kapuŋ*, WL-LWI *kə'nəpuŋ*, WL-MS *kap:uŋ*, WL-PD *kap:uŋ*.

- (2) CL-KK *kəmə'ŋe*, CL-KK *kəmə'ŋeŋ*, CL-LR *kəmə'ŋeŋ*.

mother:

(1) PFL **ina*, CL-IL *ina*, CL-KK *ina*, CL-KK *'ina*, CL-LP *ina*, CL-LR *ina*, CL-LWK *ina*, CL-LWT *ina*, CL-MR *ina*, CL-PN *inan*, EL-LE *ina*, EL-LMT *ina*, KD-LB *ʔine*, SK-HEW *ina*, SK-MM *ina*, SK-TA *ina*, WL-AB *inaŋ*, WL-AD *'ina*, WL-BL *ina*, WL-BN *ʔina*, WL-DL *ina*, WL-HRW *ina*, WL-IA *ina*, WL-KO *ina*, WL-LH *inakə*, WL-MD *ina*, WL-ML *ina*, WL-MS *ina*, WL-MS *inaŋ*, WL-MS *nina*, WL-PD *nina*, WL-WK *ina*, WL-WW *inaʔ*.

(2) CL-LK *əma*, SK-HEW *əma*, WL-BM *əma*, WL-BP *əmaʔ*, WL-BT *əma*, WL-LG *əmaʔ*, WL-LMK *əma*, WL-LML *əma*, WL-LTB *əma*, WL-LWI *'ə'maʔ*, WL-LWL *əma*, WL-PK *əmaʔ*, WL-RE *əma*, WL-TJ *əma*, WL-WB *əmaʔ*, WL-WT *əma*.

mountain:

- (1) WL-BN *foto*, WL-MS *wəɔ b'ələk*, WL-PD *wəɔ 'bələ*.

(2) PFL **ili*, CL-IL *iler*, CL-KK *ile*, CL-KK *i'ler*, CL-KK *i'li*, CL-LK *iler*, CL-LP *ile*, CL-LR *'iler*, CL-LWK *ileru*, CL-LWT *ileru*, CL-MR *ileru*, CL-PN *iler*, EL-LE *ile*, EL-LMT *ile*, KD-LB *ili*, KD-LW *ili*, SK-HEW *ilin*, SK-HEW *ilin*, SK-TA *ili-n*, WL-AD *ʔile*, WL-BL *ile*, WL-BM *ile*, WL-BP *ile*, WL-BT *ile*, WL-DL *ile*, WL-HRW *ile*, WL-IA *ile*, WL-KO *ile*, WL-LG *ile*, WL-LH *ile*, WL-LMK *ile*, WL-LML *ile*, WL-LTB *ile*, WL-LWI *ile*, WL-LWL *ile*, WL-MD *ile*, WL-ML *ile*, WL-PK *ile*, WL-RE *ile*, WL-TJ *ile*, WL-WB *ile*, WL-WK *ile*, WL-WT *ile*, WL-WW *ile*.

(3) LH-SK #volon 'hill; ridge', SK-TA *βolo*, SK-TA *βolo-n*.

mouth:

(1) PFL **nusu*, CL-KK *nu:s*, CL-KK *nusən*, CL-LK *nus*, CL-LP *nuha*, CL-LR *nu'hə*, CL-LWK *nusəga*, CL-MR *nusa*, CL-PN *nuhag*, EL-LE *nuhe*, EL-LMT *nuhe*, KD-LB *nunu*, KD-LW *nunu vɔvɔ*, WL-BL *nūk*, WL-LH *nuhukə*, WL-MD *nuhu*, WL-MS *nuhuŋ*, WL-MS *nuhu?*.

(3) PFL **vava*, CL-IL *fəfasa*, CL-LWT *fəfaha*, SK-HEW *va:*, SK-HEW *va:*, SK-MM *va:n*, WL-AB *fə'faj*, WL-AD *'wəwa'kət*, WL-BM *wəwa*, WL-BN *fəfaj*, WL-BP *wəwat*, WL-BT *wəwã*, WL-DL *wəwat*, WL-HRW *wəwal*, WL-IA *wəwak*, WL-KO *wəwak*, WL-LG *wəwat*, WL-LMK *wəwã*, WL-LML *fəfã*, WL-LTB *wəwa*, WL-LWI *'wəwaŋ*, WL-LWL *wewa*, WL-ML *wəwa*, WL-PD *wə'waŋ*, WL-PK *wəwa*, WL-RE *wəwa*, WL-TJ *wəwak*, WL-WB *wəwã*, WL-WK *wəwã*, WL-WT *wəwã*, WL-WW *wəwak*.

name:

(1) PFL **nadan*, CL-IL *naḍʒanəs*, CL-KK *naḍʒan*, CL-KK *'naḍʒan*, CL-LK *naḍʒan*, CL-LP *naḍʒan*, CL-LR *'najan*, CL-LWK *najan*, CL-LWT *naḍʒan*, CL-MR *naḍʒa*, CL-PN *najan*, EL-LE *nara*, EL-LMT *narã*, KD-LB *naja*, KD-LW *naja*, SK-HEW *naran*, SK-HEW *naraŋ*, SK-MM *naran*, SK-TA *nara*, SK-TA *nara-ŋ*, WL-AB *naraŋ*, WL-AD *'narã*, WL-BL *naraŋ*, WL-BM *narã*, WL-BN *naraŋ*, WL-BP *narã*, WL-BT *narã*, WL-DL *narã*, WL-HRW *narã*, WL-IA *naran*, WL-KO *narã*, WL-LG *narã*, WL-LH *narã*, WL-LMK *narã*, WL-LML *naraŋ*, WL-LTB *narã*, WL-LWI *'naraŋ*, WL-LWL *narã*, WL-MD *narã*, WL-ML *nara*, WL-MS *naraŋ*, WL-MS *nara?*, WL-PD *naraŋ*, WL-PK *narã*, WL-RE *narã*, WL-TJ *narã*, WL-WB *narã*, WL-WK *nara*, WL-WT *narã*, WL-WW *narã*.

narrow:

(1) PFL **kiput*, CL-IL *kənipotən*, CL-KK *kənipot*, CL-KK *kni'pət*, CL-LK *kənipot*, CL-LR *hənipotən*, CL-LR *kənipotən*, CL-LWK *kənipot*, CL-MR *kipo*, CL-PN *kənipotən*, KD-LB *i'pe?*, WL-AB *kipe*, WL-AD *'keni?*, WL-BN *ki'p:ɛ*, WL-DL *kipu?*, WL-LG *kəripu*, WL-LWI *'kəne?*, WL-MD *kənipu*.

(2) WL-MS *kupɔ?*, WL-PD *kupo*.

(3) CL-LP *hipe*, CL-LWT *hipet*, WL-BL *hipet*, WL-BM *kəhipuk*, WL-LML *hipet*, WL-WK *hipet*.

(4) LH-KD #*tudak*, keda-samely *tuda?*, WL-BT *tudak*, WL-HRW *tənud?ã*, WL-KO *tudək*, WL-LH *tudak*, WL-LMK *kətudak*, WL-ML *tudak*, WL-TJ *tudək*,

WL-WT *tudak*, WL-WW *tudak*.

(5) LH-SK #*taker*, SK-HEW *taker*, SK-MM *taker*, WL-IA *takəl*, WL-LTB *kətəkə*, WL-LWL *kətəker*, WL-PK *təkə*, WL-RE *kətəkə*, WL-WB *kətəkə*.

(6) EL-LE *ʔuhu*, EL-LMT *uhu*.

navel:

(1) PFL **pusər*, CL-KK *kəpu'sər*, CL-LR *kəpu'hər*, KD-LB *pu'he:*, SK-HEW *puher*, WL-AB *kəpu'hərɔŋ*, WL-AD *kə'puhu'rət*, WL-BN *ka'pu'hər*, WL-LWI *kəpu'hur*, WL-MS *puhər*, WL-PD *puhor*.

near:

(1) LH-KD #*dahe-k*, CL-IL *daekən*, CL-KK *daek*, CL-KK *da'ek*, CL-LK *daek*, CL-LP *dae*, CL-LR *dae'kən*, CL-LWK *dae*, CL-LWT *dae*, CL-MR *daek*, CL-PN *daekən*, EL-LE *dahe*, EL-LMT *dahe*, KD-LB *de'hiʔ*, WL-AB *dahe*, WL-AD *'dahe*, WL-BL *dae*, WL-BM *dahe*, WL-BN *dah:ɛ*, WL-BP *dahe*, WL-BT *dahʔẽ*, WL-DL *dahe*, WL-HRW *dahe*, WL-IA *dahe*, WL-KO *dahʔẽ*, WL-LG *dahe*, WL-LH *dahe*, WL-LMK *dahẽ*, WL-LML *dae*, WL-LTB *dãhe*, WL-LWI *'dahe*, WL-LWL *dahẽ*, WL-MD *dahe*, WL-ML *dahe*, WL-MS *daheʔ*, WL-PD *dahe*, WL-PK *dahe*, WL-RE *dahẽ*, WL-TJ *dahe*, WL-WB *dahe*, WL-WK *dahe*, WL-WT *dahe*, WL-WW *dahe*.

(2) SK-HEW *roʔo*, SK-HEW *roʔo*, SK-MM *roʔon*, SK-TA *groʔo*, SK-TA *roʔo*, SK-TA *roʔo-ŋ*.

neck:

(1) LH-KD #*vuli*, CL-IL *fulesa*, CL-KK *vu'le*, CL-KK *wulen*, CL-LK *wule*, CL-LP *fule*, CL-LR *'vule*, CL-LWK *wule*, CL-LWT *fuleha*, CL-MR *fule*, CL-PN *wuleg*, EL-LE *wulĩ*, EL-LMT *wuli*, KD-LW *ɥli*, WL-AB *fuliŋ*, WL-AD *'wuli'kət*, WL-BL *wulik*, WL-BM *wuli*, WL-BN *fuliŋ*, WL-BP *wuli*, WL-BT *wuli*, WL-DL *wuliŋ*, WL-HRW *wulĩ*, WL-IA *wuli*, WL-KO *wulik*, WL-LG *wulĩ*, WL-LH *wulikə*, WL-LMK *wulĩ*, WL-LML *fuli*, WL-LTB *wuli*, WL-LWI *'wuliŋ*, WL-LWL *wuli*, WL-MD *wulik*, WL-ML *fuli*, WL-MS *wuliŋ*, WL-PD *wuliŋ*, WL-PK *wulĩ*, WL-RE *wulĩ*, WL-TJ *wuli*, WL-WB *wulik*, WL-WK *wuli*, WL-WT *wulĩ*, WL-WW *wuli*.

(2) KD-LB *bɔŋan*, KD-LW *bɔŋan aduŋ*.

(3) SK-HEW *eru*, SK-HEW *əru*, SK-TA *ʔəru-ŋ*.

(4) SK-MM *boʔir*, SK-TA *boʔir*.

necklace:

- (1) WL-AB *manje*, WL-MS *manjer*, WL-MS *manjer* 'bɔta, WL-PD *manjer*.
 (2) LH **niləŋ*, CL-KK *ni'ləŋ*, CL-LR 'niŋal, WL-AD 'nilɜ, WL-LWI 'niləŋ.

needle:

- (1) LH-KD #*batul*, KD-LB *ba'tur*, WL-AB *batu*, WL-MS *batul*, WL-PD *batul*.
 (2) CL-KK 'lusir, CL-LR 'luhir, SK-HEW *luhir*, WL-AD 'luhi, WL-LWI 'luhi.

new:

- (1) PFL **vəru*, CL-IL *wərun*, CL-KK *və'rɔn*, CL-KK *wərun*, CL-LK *wərun*, CL-LP *fərun*, CL-LR *una və'run*, CL-LWK *wəru*, CL-LWT *fərudʒa*, CL-MR *wəru*, CL-PN *wərun*, EL-LE *wərũ*, EL-LMT *wərũ*, KD-LB *wɛ'run*, KD-LW *vərun taʔɛn*, SK-HEW *vərun*, SK-HEW *vərun*, SK-MM *vərun*, SK-TA *βəru*, WL-AD 'wuʔũ, WL-BL *wu*, WL-BM *wuʔũ*, WL-BN *funɔ*, WL-BP *wuʔũ*, WL-BT *wuʔũ*, WL-DL *wuʔũ*, WL-HRW *wuʔũ*, WL-IA *wun*, WL-KO *wuʔũ*, WL-LG *wuʔũ*, WL-LMK *wuʔũ*, WL-LML *fu*, WL-LTB *wuʔũ*, WL-LWI 'wuʔuŋ, WL-LWL *wuʔũ*, WL-MD *wuʔũ*, WL-ML *wu*, WL-PK *wuʔu*, WL-RE *wuʔũ*, WL-TJ *wuʔũ*, WL-WB *wuʔũ*, WL-WK *wuʔũ*, WL-WT *wuʔũ*, WL-WW *wu*.
 (2) WL-AB *uma* 'funɔŋ, WL-MS *uma* 'wunɔŋ, WL-PD *uma* 'wunɔŋ, WL-PD *wunɔŋ*.

night:

- (1) LH-SK #*dzəma*, CL-KK *dzəma* 'time unit', CL-LWK *rəma*, EL-LMT *rəma*, SK-HEW *rema* 'time unit', WL-AB *ma'reŋ*, WL-AD 'rɜmaʔ, WL-BL *rəma*, WL-BM *rəma*, WL-BN *ma'reŋ*, WL-BT *rəmʔã*, WL-DL *rəmã*, WL-HRW *rəmã*, WL-IA *rəma*, WL-KO *rəmãʔ*, WL-LG *rəmã*, WL-LH *rəma*, WL-LMK *rəmã*, WL-LML *remã*, WL-LWI *rə'maʔaŋ*, WL-MD *rəma*, WL-ML *rəma*, WL-MS *mə'r:ɛŋ*, WL-PD *ma'r:ɛŋ*, WL-WT *rəmã*, WL-WW *rəmã*.
 (2) KD-LB *u'bɛn*, KD-LW *ubɛn*.
 (3) WL-BP *noko*, WL-LTB *nokõ*, WL-LWL *nokõ*, WL-PK *noko*, WL-RE *noko*, WL-TJ *noko*, WL-WB *nokõ*.
 (4) CL-IL *buk*, CL-KK *bu:k*, CL-KK *bukən*, CL-LK *buk*, CL-LP *buk*, CL-LR *bu'kənən*, CL-LWT *buk*, CL-MR *bukɤ*, CL-PN *bukən*, SK-TA ʔβau-ŋ.
 (5) LH-SK #*guman*, EL-LE *ugũ*, SK-HEW *guman*, SK-MM *guman*, SK-TA *guma-ŋ*, WL-WK *baʔu*.

nine:

(1) PFL *siva, CL-KK 'siva, CL-LR 'hiva, SK-HEW hiva, WL-AB hifa, WL-AD hiwa, WL-BN hif:a, WL-LWI 'hiwa, WL-MS hiwa, WL-PD hiwa.

no, not:

(1) WL-AB lahe, WL-AD 'halaʔ, WL-BN la'h:ε, WL-MS lahe, WL-MS laheʔ, WL-PD lahe.

(2) PFL *ta(ke), CL-KK 'ta, CL-KK 'tak, CL-KK 'take, CL-LR ta'ken, WL-LWI 'take.

(3) SK-HEW eʔon, SK-TA eʔo-ŋ.

nose:

(1) PFL *(n)idun, CL-IL nidʒunʔəs, CL-KK ni'dʒu, CL-KK nidʒuga, CL-LK nidʒu, CL-LP nidʒun, CL-LR ni'dʒu, CL-LWK nidʒuga, CL-LWT nidʒun, CL-MR nidʒu, CL-PN nidʒug, EL-LE nirũ, EL-LMT iru, KD-LB niŋ, KD-LW niŋ, SK-HEW iru, SK-HEW iʔru, SK-MM irun, SK-TA iʔru-ŋ, WL-AB nirun, WL-AD 'iru'nɔt, WL-BL nirun, WL-BM irũ, WL-BN iʔruŋ, WL-BP iru, WL-BT irũ, WL-DL iʔruŋ, WL-HRW irun, WL-IA iʔruŋ, WL-KO iʔruŋ, WL-LG irut, WL-LH irunə, WL-LMK irũ, WL-LML nirun, WL-LTB irũ, WL-LWI 'iruŋ, WL-LWL iru, WL-MD iru, WL-ML nirun, WL-MS iʔruŋ, WL-MS iʔruʔ, WL-PD iʔruŋ, WL-PD nirun, WL-PK irũ, WL-RE irũ, WL-TJ irũ, WL-WB iru, WL-WK nirun, WL-WT irũ, WL-WW iru.

old (for people):

(1) PFL *m-tuʔa, CL-KK tu'ana, CL-LP tənərahən, CL-LR 'tua 'magu, CL-MR tənəras, CL-PN tuanakən, KD-LB tua 'lahar, SK-HEW duan, WL-AD tɔ'nuʔe, WL-BM tənue, WL-BP tənue, WL-BT tənueʔe, WL-DL tənue, WL-HRW tənueʔe, WL-IA nuhen, WL-KO tənueʔe, WL-LG tənue, WL-LMK tənueʔe, WL-LTB tua, WL-LWI 'tənueŋ, WL-LWL tənueʔ, WL-PD ɔraj 'tua, WL-PD tua, WL-PK duʔã, WL-RE tənueʔe, WL-TJ tənue, WL-WB tənue, WL-WT tənueʔe, WL-WW tənueʔe.

(2) LH #magun, CL-IL magun, CL-KK magun, CL-KK ma'guna, CL-LK magun, CL-LR 'tua 'magu, CL-LWK magu, CL-LWT magudʒa, EL-LE magũ, EL-LMT magũ, WL-BL magu, WL-LH magũ, WL-LML magu, WL-MD magũ, WL-ML magu, WL-WK magu.

(3) SK-HEW blupur, SK-MM blupur.

one:

(1) LH #toʔu, CL-KK *tune*, CL-KK *tu*, CL-LR *tu*, WL-AB *tou*, WL-AD 'toʔu, WL-BN *tɔː*, WL-LWI 'toʔu, WL-MS *tou*, WL-PD *tou*.

(2) KD-LB ʔu'deʔ, KD-LW *udeʔ*.

(3) PFL *əha 'one; alone', SK-HEW *ha*, SK-TA *ha*.

one hundred:

(1) CL-KK 'ratu *tu*, CL-LR *rat*, KD-LB *ratuʔ ʔudeʔ*, WL-AB *ratu*, WL-AD *tɜ'ratu*, WL-BN *ratu*, WL-LWI *ratu 'toʔu*, WL-LWI 'təratu, WL-MS *ratu*, WL-PD *ratu*.

other:

(1) LH #gesak, CL-IL *geakən*, CL-KK *gesak*, CL-LK *gesak*, CL-LP *gehak*, CL-LWK *gesak*, CL-LWT *gehakən*, EL-LMT *gehʔã*, WL-BL *geak*, WL-BM *gehã*, WL-BP *geha*, WL-LG *gehʔã*, WL-LH *geha*, WL-LMK *geha*, WL-LML *geak*, WL-LTB *gehã*, WL-LWL *gehã*, WL-MD *gehʔã*, WL-ML *gehak*, WL-TJ *gehã*, WL-WB *gehã*, WL-WK *gehak*.

(2) WL-BT *ikər*, WL-DL *ikən*, WL-HRW *ikã*, WL-KO *ikʔã*, WL-RE *ikər*, WL-WT *ikər*, WL-WW *ikər*.

(3) LH-SK #pehan, EL-LE *pehã*, SK-HEW *pehan*.

(4) CL-MR *wisa*, CL-PN *wiha*.

oven:

(1) LH-SK #likat, CL-KK *li'kat*, CL-LR 'likat, SK-HEW *liʔat*, WL-AB *lika*, WL-AD 'likat, WL-LWI 'lika, WL-MS *kada 'likaʔ*, WL-MS *likaʔ*, WL-PD *lika*.

palm of hand:

(1) WL-AB *liman ʔonɔŋ*, WL-AD 'limakɜt ʔonʔeʔ, WL-LWI *liman ʔonəʔonɔŋ*, WL-MS *lima ʔonɔŋ*, WL-MS *liman ʔonɔŋ*, WL-PD *liman ʔonɔŋ*.

(2) LH-SK #əpak 'palm of hand; footprint', CL-KK *li'ma ə'pak*, SK-HEW *lima əpak*.

papaya:

(1) WL-AB *kasi*, WL-PD *kasi*.

(2) WL-MS *sela*, WL-PD *sela*, WL-PD *tela*.

(3) LH-KD #payam, CL-KK *pa'dʒam*, CL-LR *pa'jam*, KD-LB *wa'ja*, WL-AD *pa'jã*, WL-LWI *wua 'pajanɔŋ*.

penis:

- (1) WL-AB *la'hakaŋ*, WL-LWI *'laha*, WL-PD *lahak*.
 (2) PFL **uti*, CL-KK *'uti*, CL-LR *u'ti*, KD-LB *wu'ti*, SK-HEW *uti*, WL-AD *'ute*, WL-MS *utiŋ*.

person:

- (1) PFL **ata*, CL-LK *ate*, SK-HEW *ata*, WL-AB *ata*, WL-AD *'ata*, WL-BM *atã*, WL-BN *ʔata*, WL-MS *ata*, WL-PD *ata*, WL-TJ *ata*.
 (2) WL-MS *mə'sia*, WL-PD *mə'sia*. From Indonesian/Malay *manusia* 'human; person', ultimately from Sanskrit.
 (3) LH-KD #(ata) *dikə-n*, CL-IL *atadikəd̪ʒa*, CL-KK *'ata di'kən*, CL-KK *atadikən*, CL-LP *atadikəd̪ʒ*, CL-LR *ata dikən*, CL-LWT *atadikəd̪ʒa*, CL-MR *atadikə*, EL-LE *atadijã*, EL-LMT *ədiʔã*, KD-LB *ate 'diʔen*, WL-BL *atadik*, WL-BP *atadikə*, WL-BT *atadikə*, WL-DL *atadikã*, WL-HRW *atadikã*, WL-IA *atadikən*, WL-KO *atadikʔã*, WL-LG *atadikã*, WL-LH *dikã*, WL-LMK *atadikã*, WL-LML *atadikã*, WL-LTB *atadikã*, WL-LWI *ata dikə'ʔəŋ*, WL-LWL *atadikẽ*, WL-MD *ədikã*, WL-ML *atadikɤ*, WL-PK *atadikẽ*, WL-RE *atadikã*, WL-WB *atadikã*, WL-WK *atadikẽ*, WL-WT *atadikã*, WL-WW *atadike*.
 (4) SK-HEW *atabia*, SK-HEW *biʔan*, SK-MM *biʔan*.

pig:

- (1) PFL **vavi*, CL-KK *vavi*, CL-KK *va'vɛr*, CL-LR *va'vɛr*, KD-LB *wavi*, SK-HEW *vavi*, WL-AB *fafɛ*, WL-AD *'wawe*, WL-LWI *'wawe*, WL-MS *wawe*, WL-PD *wawe*.

price:

- (1) PFL **veli* 'price; bride price; expensive; buy', CL-KK *ve'li*, CL-LR *'velin*, KD-LB *welin*, SK-HEW *velin*, WL-AB *feliŋ*, WL-AD *'welĩ*, WL-LWI *'weliŋ*, WL-MS *weliŋ*, WL-PD *weliŋ*.

rack above hearth:

- (1) WL-AB *huruj*, WL-PD *huruj*.
 (2) WL-MS *herɛʔ*, WL-PD *herɛ*.
 (3) KD-LB *rak*, WL-LWI *'ra:k*.
 (4) LH #*magar*, CL-KK *ma'gar*, WL-AD *maga*.

rain:

(1) PFL ***udan**, CL-IL *udžan*, CL-KK *udžan*, CL-KK *u'džan*, CL-LK *udžan*, CL-LP *udžan*, CL-LR *'udžan*, CL-LWT *udžan*, CL-MR *udžau*, CL-PN *udžan*, EL-LE *urã*, EL-LMT *urã*, KD-LB *u'ja*, KD-LW *ujã*, SK-HEW *uran*, SK-MM *uran*, SK-TA *ura-n*, WL-AB *uraj*, WL-AD *'uraj*, WL-BL *ura*, WL-BM *urã*, WL-BP *urã*, WL-BT *urã*, WL-DL *urut*, WL-HRW *urã*, WL-IA *uran*, WL-KO *urut*, WL-LG *urã*, WL-LH *urã*, WL-LMK *urã*, WL-LML *urã*, WL-LTB *urã*, WL-LWI *'uraj*, WL-LWL *urã*, WL-MD *urã*, WL-ML *urɔ*, WL-MS *uraj* *'tuɲ*, WL-MS *uraj*, WL-PD *uraj*, WL-PK *urã*, WL-RE *urã*, WL-TJ *urã*, WL-WB *urã*, WL-WK *ura*, WL-WT *urã*, WL-WW *urã*.

rat:

(1) LH #**kromi**, CL-KK *kro'mi*, CL-KK *kro'mer*, WL-AB *ka'more*, WL-AD *'krome*, WL-BN *ka'more*, WL-LWI *kə'rɔme*, WL-MS *more*, WL-PD *k'mɔre*, WL-PD *mɔre*.

(2) KD-LB *tiu*, SK-HEW *teʔu*.

rattan:

(1) PFL ***uay**, CL-KK *u'adž*, CL-KK *u'a*, CL-LR *'uar*, KD-LB *wua'wei*, SK-HEW *ue*, WL-AB *ufi*, WL-AD *'ua*, WL-LWI *u'wajəɲ*, WL-MS *uwe*, WL-PD *ue*.

(2) WL-MS *lua*, WL-PD *lua*.

(3) KD-LB *wei*, KD-LB *wua'wei*.

red:

(1) PFL ***meran**, CL-IL *merajən*, CL-KK *meran*, CL-KK *me'ran*, CL-LK *meran*, CL-LP *meranu*, CL-LR *me'ran*, CL-LWK *mera*, CL-LWT *meran*, CL-MR *merau*, CL-PN *meran*, EL-LE *mera*, EL-LMT *meã*, SK-HEW *meran*, SK-MM *meraʔ*, SK-TA *mera-ɲ*, SK-TA *merak*, WL-AB *meaj*, WL-AD *'mɜʔã*, WL-BL *mea*, WL-BM *meʔa*, WL-BN *meã*, WL-BP *meʔa*, WL-BT *meʔã*, WL-DL *meʔã*, WL-HRW *meʔã*, WL-IA *mean*, WL-KO *meʔã*, WL-LG *meʔã*, WL-LH *meʔa*, WL-LMK *meã*, WL-LML *meã*, WL-LTB *meʔã*, WL-LWI *'meʔa*, WL-LWL *meʔã*, WL-MD *meʔã*, WL-ML *meɔ*, WL-MS *meaj*, WL-PD *meaj*, WL-PK *meʔã*, WL-RE *meʔã*, WL-TJ *meʔa*, WL-WB *meʔã*, WL-WK *meʔã*, WL-WT *meʔã*, WL-WW *meʔã*.

(2) KD-LB *pu'tuʔ*, KD-LW *putuʔ*.

rice ear bug:

(1) PFL ***kəmeruɲ**, CL-KK *kəmə'rɔɲ*, CL-KK *kəmə'rɔ*, CL-LR *kəmə'roɲ*,

WL-AB *ka'meruŋ*, WL-AD *kəmerū'date*, WL-LWI *kə'meruŋ*, WL-MS *meruŋ b'lurɔŋ*, WL-MS *meruŋ 'gihi*, WL-MS *meruŋ*, WL-PD *k'meruŋ*.

rice grains:

(1) CL-LR *kə'luok 'vuliran*, SK-HEW *vulir*, WL-LWI *'tahaŋ 'wuliŋ*, WL-MS *pare 'wuliŋ*, WL-PD *pare 'wuliŋ*.

(2) PFL **padi*, WL-AB *pari mu'leŋ*, WL-MS *pare 'wuliŋ*, WL-PD *pare 'wuliŋ*.

rice plant:

(1) PFL **padi*, WL-AB *pari*, WL-MS *pare 'nawaŋ*, WL-PD *pare*.

(2) CL-KK *li'sɔ*, CL-KK *li'sɔr*.

rich:

(1) CL-KK *kova'san*, CL-LR *kəva'han*, WL-AB *ka'fasaŋ*, WL-AD *kə'wasa*, WL-MS *k'wasas*, WL-PD *k'wasas*. From Indonesian/Malay *kuasa* 'power', ultimately from Sanskrit.

(2) SK-HEW *bisa?*, WL-MS *bisaŋ*. Probably from Indonesian/Malay *bisa* 'can'.

(3) KD-LB *kaja*, WL-LWI *'kaja*. From Indonesian/Malay *kaya* 'rich'.

ridge:

(1) PFL *(*v*)*uvuŋ*, CL-KK *u'vɔŋ*, CL-LR *'uvuŋ*, SK-HEW *puvun*, WL-AB *mou'fufuŋ*, WL-LWI *'uvuŋ*.

(2) WL-MS *b'lokɔŋ*, WL-PD *b'lokɔŋ*.

(3) LH-KD #*volar*, CL-LR *'volar*, KD-LB *wɔlar*, WL-AD *'wola?*.

right side:

(1) PFL **dikə* 'good; right; person', WL-AB *dikē*, WL-BN *di'kē*, WL-MS *dikeŋ*, WL-PD *dikeŋ*.

(2) PFL **vanan*, CL-IL *fanan*, CL-KK *va'na*, CL-KK *wana*, CL-LK *wana*, CL-LP *fanan*, CL-LR *va'nan*, CL-LWK *wanã*, CL-LWT *fanan*, CL-MR *wana*, CL-PN *wanan*, EL-LE *wana*, EL-LMT *wana*, KD-LB *wana*, KD-LW *vana*, SK-HEW *vanan*, SK-HEW *vanan*, SK-MM *vana*, SK-TA *βana*, WL-AD *'wana*, WL-BL *wana*, WL-BM *wana*, WL-BP *wana*, WL-BT *wanã*, WL-DL *wanã*, WL-HRW *wanã*, WL-IA *wanan*, WL-KO *wana*, WL-LG *wanã*, WL-LH *wana*, WL-LMK *wanã*, WL-LML *fana*, WL-LTB *wana*, WL-LWI *'wanaŋ*, WL-LWL *wana*, WL-MD *wanã*, WL-ML *wɔnɔ*, WL-PK *wana*, WL-RE *wana*, WL-TJ *wana*, WL-WB *wana*, WL-

WK *wana*, WL-WT *wanã*, WL-WW *wanan*.

ripe:

(1) PFL ***m-tasak**, CL-KK *ta'sak*, CL-LR *taha'kən*, KD-LB *ta'ʔen*, SK-HEW *dahaʔ*, WL-AB *taha*, WL-AD *'tahak*, WL-BN *tah:a*, WL-LWI *'taha*, WL-MS *ta-hak*, WL-PD *ta'hak*.

(2) CL-IL *fai*, CL-KK *waer*, CL-LK *waer*, CL-LP *fai*, CL-LR *'vei ra'jan*, CL-LWT *fai mənəjaken*, CL-MR *wai*, CL-PN *wai*, KD-LB *wei*, SK-HEW *vair ba*, SK-TA *βair ba*, WL-BP *wai ba*, WL-BT *wai maʔã*, WL-DL *wai*, WL-HRW *wai*, WL-IA *wai nahan*, WL-KO *wai maʔã*, WL-LMK *waimənaʔaŋ*, WL-ML *wai*, WL-WW *wai*.

river, stream:

(1) PFL *(**n**)**ubak**, CL-KK *ə'bak*, CL-LWK *baku*, KD-LW *ʏbeʔ*, SK-KR *nuba*.

(2) SK-HEW *vair ba*, SK-TA *naŋa ba*, SK-TA *βair ba*.

(3) WL-AD *suje*, WL-BM *suje*, WL-LG *suje*, WL-LML *suje*, WL-LWL *suje*, WL-PK *suje*, WL-RE *suje*, WL-WT *suje*. From Indonesian/Malay *sun-gai* 'river'.

(3) LH #(**a**)**luŋu**, EL-LE *luŋu*, EL-LMT *luŋu*, WL-AB *alu*, WL-BL *luŋu*, WL-LH *luŋu*, WL-LTB *luŋu*, WL-LWI *'luŋuŋ*, WL-MD *luŋu*, WL-MS *aluk*, WL-PD *alu*, WL-TJ *luŋu*, WL-WB *lũ wãi*, WL-WK *luŋu*.

road:

(1) PFL ***lalan**, CL-IL *lala*, CL-KK *la'la*, CL-KK *lalan*, CL-KK *la'lan*, CL-LK *lalan*, CL-LP *lalan*, CL-LR *'lalan*, CL-LWK *lalanu*, CL-LWT *lalanu*, CL-MR *lalan*, CL-PN *lalan*, EL-LE *laran*, EL-LMT *larã*, KD-LB *lala*, SK-HEW *larã*, SK-HEW *laran*, SK-MM *lalan*, SK-TA *lala-ŋ*, WL-AB *laraŋ*, WL-AD *rarã*, WL-BL *lara*, WL-BM *larã*, WL-BP *rarã*, WL-BT *rarã*, WL-DL *rarã*, WL-HRW *rarã*, WL-IA *laran*, WL-KO *rarã*, WL-LG *laran*, WL-LH *laran*, WL-LMK *rarã*, WL-LML *larã*, WL-LTB *larã*, WL-LWI *'laraŋ*, WL-LWL *rarã*, WL-MD *larã*, WL-ML *larã*, WL-MS *laraŋ*, WL-PD *laraŋ*, WL-PK *larã*, WL-RE *larã*, WL-TJ *larã*, WL-WB *larã*, WL-WK *lara*, WL-WT *rarã*, WL-WW *larã*.

(2) WL-BN *tɔ:r*, WL-MS *tɔr*.

root:

(1) PFL ***ramut**, CL-IL *ramut*, CL-KK *ramut*, CL-KK *ra'mut*, CL-LK *ra-mut*, CL-LP *ramut*, CL-LR *'ramut*, CL-LWT *ramut*, CL-MR *ramut*, CL-PN *ra-*

mutən, EL-LE *ramu*, EL-LMT *ramutã*, KD-LB *ramu?*, KD-LW *ramu?*, SK-HEW *ramut*, SK-MM *ramuť*, SK-TA *ʔramut*, WL-AB *ram'bukuy*, WL-AD *ʔamut*, WL-BL *ramut*, WL-BM *ramut*, WL-BN *ramuk*, WL-BP *amut*, WL-BT *əmut*, WL-DL *ramut*, WL-HRW *əmut*, WL-IA *ramut*, WL-KO *amut*, WL-LG *ramuk*, WL-LH *amutə*, WL-LMK *amutã*, WL-LML *ramut*, WL-LTB *ramu*, WL-LWI *'ramu*, WL-LWL *amut*, WL-MD *amut*, WL-ML *ramut*, WL-MS *ramuk*, WL-PD *ramu*, WL-PK *ramut*, WL-RE *amut*, WL-TJ *ramut*, WL-WB *amut*, WL-WK *amut*, WL-WT *ramut*, WL-WW *amut*.

rope:

(1) PFL **tali*, CL-KK *ta'ler*, CL-KK *ta'li*, CL-KK *tale*, CL-LWK *tali*, CL-PN *tali*, EL-LE *tale*, EL-LMT *tale*, SK-HEW *tali*, SK-MM *ťali*, SK-TA *tali*, WL-AB *tale*, WL-AD *tale?*, WL-BL *tale*, WL-BM *tale*, WL-BN *tale*, WL-BP *tale*, WL-BT *tale*, WL-DL *tale*, WL-HRW *tale*, WL-IA *tale*, WL-KO *tale?*, WL-LG *tale*, WL-LH *tale*, WL-LMK *tale*, WL-LML *tale*, WL-LTB *tale*, WL-LWI *tale*, WL-LWL *tale*, WL-MD *tale*, WL-ML *tale*, WL-MS *tale*, WL-PD *tale*, WL-PK *tale*, WL-RE *tale*, WL-TJ *tale*, WL-WB *tale*, WL-WK *tale*, WL-WT *tale?*, WL-WW *tale*.

(2) LH-KD *#vadək*, CL-IL *fədək*, CL-LP *fədek*, CL-LR *və'dək*, CL-LWT *fədəku*, CL-MR *fədək*, KD-LB *wa'de?*, KD-LW *vade?*.

rotten:

(1) CL-KK *kəda'tək*, WL-AD *'datɜ*, WL-BN *date*.

(2) CL-IL *fuk*, CL-KK *wauk*, CL-LP *fuk*, CL-LWK *wauk*, CL-MR *fukʁ*, CL-PN *wauk*, SK-HEW *vauk*, SK-HEW *vauk*, SK-TA *βau-ŋ*, SK-TA *βauk*, WL-AB *fou*, WL-BL *wau*, WL-BM *wău*, WL-BP *wău*, WL-BT *wău*, WL-DL *wău*, WL-IA *waun*, WL-KO *wău*, WL-LG *waũ*, WL-LH *waũ*, WL-LMK *waũ*, WL-LML *fău*, WL-LWI *'wauŋ*, WL-LWL *wău*, WL-MD *wău*, WL-ML *wɔu*, WL-MS *wou?*, WL-PD *wau*, WL-PK *wău*, WL-RE *wău*, WL-TJ *wău*, WL-WB *wău*, WL-WK *wau*, WL-WT *wău*, WL-WW *wău*.

(3) CL-LWT *buru*, KD-LB *bu'ru?*, KD-LW *buru? balan*.

(4) EL-LE *weu*, EL-LMT *we?u*.

sail:

(1) PFL **layar*, CL-KK *ladʒar*, CL-LR *'lajar*, KD-LB *lajar*, SK-HEW *lajar*, WL-AB *ladʒa*, WL-AD *laja?*, WL-LWI *'laja?*, WL-MS *ladʒa*, WL-PD *ladʒa*.

salt:

(1) PFL *hira, CL-IL *irar*, CL-KK *i'ra*, CL-KK *irar*, CL-KK *i'rar*, CL-LK *irar*, CL-LP *irar*, CL-LR *i'rar*, CL-LWK *iraru*, CL-LWT *iraru*, CL-MR *iraru*, CL-PN *irar*, EL-LMT *hira*, WL-AB *sia*, WL-AD *'siʔa*, WL-BL *sia*, WL-BM *siʔa*, WL-BN *siʔa*, WL-BP *siʔa*, WL-BT *siʔa*, WL-DL *siʔa*, WL-HRW *siʔa*, WL-IA *sia*, WL-KO *siʔa*, WL-LG *siʔa*, WL-LH *sia*, WL-LMK *siʔa*, WL-LML *sia*, WL-LTB *siʔa*, WL-LWI *'siʔa*, WL-LWL *siʔa*, WL-MD *siʔa*, WL-ML *sia*, WL-MS *sia*, WL-PD *sia*, WL-PK *siʔa*, WL-RE *siʔa*, WL-TJ *siʔa*, WL-WB *siʔa*, WL-WK *siʔa*, WL-WT *siʔa*, WL-WW *sia*. WL forms with initial *s* are loans. See §5.2.3.2 for a discussion.

(2) LH-KD *taʔu, EL-LE *taʔo*, KD-LB *te'ʔu*, KD-LW *tæʔu apur*.

(3) SK-HEW *hini*, SK-MM *hini*, SK-TA *hini bru*, SK-TA *hini*.

salty:

(1) PFL *m-pədu, CL-KK *pəne'dʒʊk*, CL-LR *məɖʒu'kən*, CL-LR *pɔi pɔi*, KD-LB *pɛ'ju*, SK-HEW *bəru*, WL-AB *parɔ*, WL-AD *'pəro*, WL-LWI *'pəro*, WL-MS *pə'rɔɔ*, WL-PD *parɔɔ*.

same:

(1) PFL *sama, CL-KK *sna'maŋ*, CL-LR *'hama*, KD-LB *hama*, SK-HEW *hama*, WL-AB *hama*, WL-AD *'hama*, WL-LWI *'hama*, WL-MS *hama*, WL-PD *hama*.

sand:

(1) WL-AB *ta'n:iŋ*, WL-MS *tə'n:i*.

(2) PFL *ənay, KD-LW *ɣnɛ*, SK-HEW *ne:n*, SK-MM *ne:*, SK-TA *ne:*, WL-MS *a'n:e*, WL-MS *ə'n:i*, WL-PD *an:e*, WL-PD *ane*.

(3) LH-KD #bota(n), CL-KK *'bota*, CL-KK *botan*, CL-KK *'botan*, CL-LK *botan*, CL-LP *bota*, CL-LR *'botan*, CL-MR *bota*, CL-PN *botan*, EL-LE *botã*, EL-LMT *botã*, KD-LW *buta*, WL-LML *botã*, WL-ML *botɔ*.

(4) PFL *vura, CL-IL *fərar*, CL-LWK *wəraru*, CL-LWT *fəradʒa*, KD-LB *ʔerah*, SK-TA *ne βura*, WL-AD *'wəra*, WL-BL *wəra*, WL-BM *wəra*, WL-BP *wəra*, WL-BT *wəra*, WL-DL *wəra*, WL-HRW *wəra*, WL-IA *wəra*, WL-KO *wəra*, WL-LG *wəra*, WL-LH *wəra*, WL-LMK *wəra*, WL-LTB *wəra*, WL-LWI *'wəra*, WL-LWL *wəra*, WL-MD *wəra*, WL-PK *wəra*, WL-RE *wəra*, WL-TJ *wəra*, WL-WB *wəra*, WL-WK *wəra*, WL-WT *wəra*, WL-WW *wəra*.

(5) SK-HEW *əda*, SK-HEW *ʔəda*.

sarong for women:

(1) WL-AB *ka'fate*, WL-AD *kə'wata*, WL-LWI *kə'watə*, WL-MS *k'wate?*, WL-PD *k'wate*.

(2) CL-KK *kre'ət*, CL-KK *kre'mət*.

scabies:

(1) WL-AB *katu*, WL-MS *katu wə'ne*, WL-MS *katu*, WL-PD *katu*.

(2) LH-KD *#kəburak*, CL-KK *kə'bura*, CL-LR *kəpu'rok*, KD-LB *pu'rə?*, WL-AD *kə'bu?u*, WL-LWI *kə'bu?u*.

scared:

(1) PFL **takut*, CL-KK *takut*, CL-LK *takut*, CL-LR *tə'nakutən*, CL-LWK *takutga*, CL-PN *takut*, EL-LMT *ta?u*, KD-LB *tau?*, WL-AB *taku*, WL-BM *taku*, WL-BN *taku*, WL-BP *takut*, WL-LG *taku*, WL-LH *takut*, WL-LTB *taku?*, WL-LWI *'taku*, WL-LWL *takut*, WL-MD *takut*, WL-ML *taku*, WL-MS *taku?*, WL-PD *takuk*, WL-PK *tapuk*, WL-WB *takut*.

(2) WL-AD *'so?ot*, WL-BT *so?ot*, WL-DL *so?ot*, WL-HRW *so?ot*, WL-IA *so?ot*, WL-KO *so?ot*, WL-LMK *so?o*, WL-RE *so?ot*, WL-TJ *so?ot*, WL-WT *so?o*, WL-WW *so?ot*.

(3) CL-IL *kəruitəs*, CL-LP *keruit*, CL-LWT *kəruit*, CL-MR *kəruī*, WL-BL *kəruī*, WL-LML *kəruit*, WL-WK *kəruī*.

(4) SK-HEW *bərau*, SK-HEW *blau?*, SK-MM *blau*.

scorpion:

(1) KD-LB *ela*, WL-AB *ala*, WL-MS *ala*, WL-PD *ala*.

(2) LH-SK *#kəmekot*, SK-HEW *mekot*, WL-AD *'mekət*, WL-LWI *kə'meko*.

(3) CL-KK *i'lot*, CL-LR *i'lot*.

sea:

(1) PFL **tahik*, EL-LE *tahi*, EL-LMT *tahi*, KD-LB *ta'hi?*, KD-LW *tahi?əjin*, SK-HEW *tahi*, SK-HEW *tahi?*, SK-MM *ṭahi*, SK-TA *tahi*, WL-AB *tahi*, WL-BM *tahik*, WL-BP *tahik*, WL-BT *tahik*, WL-DL *tahik*, WL-HRW *tahik*, WL-IA *tahi*, WL-KO *tahi*, WL-LG *tahik*, WL-LH *tahi*, WL-LMK *tahik*, WL-LTB *tahi*, WL-LWI *tahi*, WL-LWL *tahik*, WL-MD *tahi*, WL-MS *tahi?*, WL-PD *tahi*, WL-PK *tahik*, WL-RE *tahik*, WL-TJ *tahik*, WL-WB *tahik*, WL-WT *tahik*, WL-WW *tahik*.

(2) CL-IL *lefar*, CL-LP *lefar*, CL-LWT *lefa*, CL-MR *lewa*, KD-LW *leva*, WL-AD *lewa?*, WL-LML *lefa*, WL-ML *lewa*.

(3) PFL *(l)oyor ‘wave; sea’, CL-KK *lo'dʒor*, CL-KK *lodʒor*, CL-LK *lodʒor*, CL-LR *'lojor*, CL-LWK *lojoru*, CL-PN *lojor*, WL-BL *odʒok*, WL-WK *odʒok*.

seed:

(1) LH-KD #*kuluk*, CL-IL *kuluk*, CL-KK *kuluk*, CL-LK *kuluk*, CL-LP *kuluk*, CL-LWK *kuluk*, CL-MR *kuluk*, CL-PN *kulukən*, EL-LE *ʔulü*, KD-LB *ʔu'luʔ*, KD-LW *uluʔ*, WL-AB *ku'lukuŋ*, WL-AD *'kulʔü*, WL-BL *kulu*, WL-BM *kuluʔ*, WL-BN *kuluŋ*, WL-BP *kuluk*, WL-BT *kulʔü*, WL-DL *kulʔü*, WL-HRW *kəlulü*, WL-IA *ku-luk*, WL-KO *kulʔü*, WL-LG *kulʔü*, WL-LH *kulü*, WL-LMK *kulʔü*, WL-LTB *kulu*, WL-LWI *'kuluŋ*, WL-MD *kulʔü*, WL-ML *kənuluk*, WL-MS *kuluŋ*, WL-PD *kuluŋ*, WL-PK *kuluʔü*, WL-TJ *kuluʔ*, WL-WB *kənuluʔ*, WL-WK *kulʔu*, WL-WT *kulʔü*, WL-WW *kəluʔü*.

(2) PFL **luri* ‘bone’, CL-KK *ri'uk*, CL-LR *'riuk*, CL-LWT *riuk*.

(3) EL-LMT *upu*, WL-LML *upu*.

(4) SK-HEW *ihin*, SK-TA *βini*.

seven:

(1) PFL **pitu*, CL-KK *'pito*, CL-LR *'pito*, KD-LB *pitu*, SK-HEW *pitu*, WL-AB *pito*, WL-AD *'pito*, WL-BN *pitɔ*, WL-LWI *'pitɔ*, WL-MS *pito*, WL-PD *pito*.

shark:

(1) PFL **iu*, CL-KK *i'u*, CL-KK *i'ɔr*, CL-LR *'iɔr*, KD-LB *iʔa hiiu*, SK-HEW *hiiu*, WL-AB *idʒo*, WL-AD *'hio*, WL-LWI *'io*, WL-MS *io*, WL-PD *io*.

sharp:

(1) LH #*bərəkət/dəkət*, CL-IL *bərək*, CL-KK *rəkət*, CL-KK *rə'kət*, CL-LK *bərəkət*, CL-LP *bərəkət*, CL-LR *rəkə'tən*, CL-LWK *rəkət*, CL-LWT *bərəkət*, CL-MR *rəkət*, CL-PN *rəkət*, EL-LMT *raʔe*, WL-AB *dakɛ*, WL-AD *'dəkət*, WL-BL *rəkət*, WL-BM *bərəkət*, WL-BN *dakɛ*, WL-BP *dəkət*, WL-BT *dəkət*, WL-HRW *bərəkət*, WL-IA *dəkət*, WL-KO *dəkət*, WL-LG *bərekə*, WL-LH *rəkətə*, WL-LMK *bərəkət*, WL-LML *bərəkət*, WL-LWI *'bərəkə*, WL-LWL *dəkət*, WL-MD *bərəkət*, WL-ML *rəkə*, WL-MS *də'ke 'tɔrak*, WL-MS *də'kek*, WL-PD *də'kek*, WL-RE *dəkət*, WL-TJ *dəkət*, WL-WB *dəkət*, WL-WK *rəkət*, WL-WT *dəkət*, WL-WW *bərəkət*.

(2) PFL **m-tidəm*, KD-LB *dəʔeʔ*, KD-LW *dəjæʔ*, SK-HEW *diran*, SK-MM *diran*, SK-TA *dira-ŋ*.

(3) WL-LTB *gã*, WL-PK *gã*.

shirt:

- (1) WL-AB *kəndʒɔ*, WL-MS *kəndɔ*, WL-PD *kəndɔ*.
 (2) PFL **labur*, CL-KK *la'bur*, CL-LR *la'bur*, KD-LB *labor*, SK-HEW *labur*, WL-AD *'labu*, WL-LWI *'labu*.

short:

- (1) LH-KD #*məkul*, CL-LR *mə'kul*, KD-LB *mɛ'ku?*, WL-AB *maku*, WL-AB *'maku*, WL-BN *makɔ*, WL-MS *mə'ku*, WL-PD *ma'ku*.
 (2) WL-AB *fəro*, WL-MS *wɛrɔ?*, WL-PD *wero*.
 (3) WL-AD *kɜʔrũ*, WL-BL *kəru*, WL-BP *kərũ*, WL-BT *kərʔũ*, WL-HRW *kərʔũ*, WL-IA *kərukən*, WL-KO *kərʔũ*, WL-LH *kərʔu*, WL-LML *kəru*, WL-LWL *kəru*, WL-RE *kəru*, WL-WB *karu*, WL-WW *kəru*.
 (4) WL-DL *kəsũ*, WL-LG *kəsɔ*, WL-LMK *kəsũ*, WL-LTB *pəsɔ*, WL-MD *kəsʔu*, WL-WK *kəsʔu*, WL-WT *kəsʔũ*.
 (5) CL-IL *pəɖʒokən*, CL-LP *pəɖʒokən*, CL-LR *pa'dʒok*, CL-LWT *pəɖʒokən*, CL-MR *pəɖʒok*, CL-PN *pəɖʒokən*.
 (6) CL-KK *mə'gɔk*, CL-KK *məgok*, CL-LK *məgok*, CL-LWK *məgok*.
 (7) CL-KK *le'rek*, CL-LR *ata'dikə lere'kən*, WL-LWI *'ləɛ?*, WL-PK *lerʔɛ*.
 (8) LH-SK #*buʔu*, EL-LE *boʔɔ*, SK-HEW *buʔu*, SK-HEW *buʔu*.
 (9) WL-LWI *'kənu*, WL-ML *kənəpɔ*.
 (10) SK-HEW *buluk*, SK-HEW *buluk*, SK-MM *buluk*, SK-TA *buluk*.

shoulder:

- (1) WL-AB *hanɛ*, WL-AD *'hana'nɜt*, WL-LWI *'hanəŋ*, WL-MS *hanəŋ*, WL-PD *hanəŋ*.
 (2) PFL **kpali-k* / **kwali-k*, CL-KK *kva'lek*, CL-LR *'kvalek*, KD-LB *wali*, SK-HEW *palik*.

shy, ashamed:

- (1) PFL **meya*, SK-HEW *mean*, WL-AB *manian*, WL-AD *'mia*, WL-LWI *'mia*, WL-MS *m'nian*, WL-PD *m'nian*.
 (2) CL-KK *kru'it*, CL-LR *kru'it*.

sick, painful:

- (1) LH-SK #*blara*, SK-HEW *blara?*, WL-AB *bə'lara*, WL-AD *bɜ'rərʔa*, WL-AD *'brara*, WL-MS *b'larəŋ*, WL-PD *b'lara*.
 (2) WL-AB *pa'latin*, WL-PD *wə'k:in p'latin*.

six:

(1) PFL ***ənəm**, CL-KK *ə'nəm*, CL-LR *ə'nam*, KD-LB *enɛŋ*, SK-HEW *əna*, WL-AB *namuŋ*, WL-AD *'namu*, WL-BN *nam:u*, WL-LWI *'nəmuŋ*, WL-MS *nə'mu*, WL-PD *na'm:u*.

skin:

(1) PFL ***kamak** 'skin; bark (tree)', CL-IL *kamakəs*, CL-LP *kama*, CL-LR *ka'mak*, CL-LWT *kamakha*, CL-MR *kamak*, CL-PN *kamak*, KD-LB *ʔa'maʔ*, KD-LW *amaʔ*, SK-KR *ʔama*, WL-AD *'kama'kət*, WL-BL *kama*, WL-BM *kama*, WL-BN *kamaŋ*, WL-BT *kamʔã*, WL-DL *kamʔa*, WL-HRW *kamakət*, WL-IA *kamak*, WL-KO *kamʔa*, WL-LG *kama*, WL-LMK *kamʔa*, WL-LML *kāmã*, WL-LWL *kamaʔ*, WL-MD *kama*, WL-ML *kəmɤ*, WL-PD *kamak*, WL-RE *kamaʔ*, WL-TJ *kama*, WL-WB *kama*, WL-WK *kama*, WL-WW *kamʔa*.

(2) PFL ***kulit**, EL-LE *uli*, EL-LMT *ulitã*, SK-HEW *ulit*, SK-HEW *ʔulit*, SK-MM *ulit̚*, SK-TA *ʔuli-t*, WL-AB *ku'likij*, WL-BP *kulit*, WL-LH *kulit*, WL-LTB *kuli*, WL-LWI *'kuli*, WL-MS *kulik*, WL-PD *kulik*, WL-PK *kulit*, WL-WT *kulit*.

(3) CL-KK *ku'da*, CL-KK *kudan*, CL-LK *kudan*, CL-LWK *kuda*.

(4) KD-LB *piki*, WL-MS *kipek*, WL-PD *kipek*.

skinny:

(1) LH-SK ***krogoŋ**, CL-KK *krɔ'gɔŋ*, SK-HEW *rugʉʔ*, WL-LWI *kə'rogɔŋ*.

sky:

(1) PFL ***kələm**, CL-IL *kələm*, CL-KK *kələm*, CL-KK *kə'ləm*, CL-LK *kələm*, CL-LP *kələm*, CL-LR *'kələm 'leva*, CL-LWK *kələm*, CL-LWT *kələmu*, CL-MR *kələm*, EL-LE *əlẽ*, EL-LMT *əle*, KD-LB *ʔɛ'leŋ*, KD-LW *ælæŋ*, SK-HEW *kələŋ*, WL-AD *'kəlɜ*, WL-BL *kələm*, WL-BM *kəlã*, WL-BP *kəlã*, WL-BT *kəlã*, WL-DL *kələŋ*, WL-HRW *kəlã*, WL-IA *kələm*, WL-KO *kəlɤ*, WL-LG *kəlã*, WL-LH *kələm*, WL-LMK *kələŋ*, WL-LML *kelã*, WL-LTB *kəlã*, WL-LWI *'kələŋ*, WL-LWL *kəle*, WL-MD *kəlã*, WL-ML *kəlɤ*, WL-PD *kal:əŋ*, WL-PK *kulã*, WL-RE *kəlã*, WL-TJ *kəlã*, WL-WB *kəlẽ*, WL-WK *kələm*, WL-WT *kəlã*, WL-WW *kəlã*.

(2) WL-BN *laŋi*, WL-MS *laŋi*.

small:

(1) PFL ***anak** 'small; child', CL-IL *anakən*, CL-KK *anak*, CL-KK *a'nak*, CL-LR *ana'kən*, CL-LWK *anak*, CL-LWT *anak*, EL-LMT *kanaʔã*, SK-HEW *anak*, SK-TA *anak*, WL-BL *ana*, WL-BN *an:əŋ*, WL-MS *anaŋ*, WL-PD *anaŋ*, WL-RE

anã.

- (2) KD-LW *kulɛŋ kai?*, WL-AB *kae*.
- (3) WL-MS *kihu*, WL-PD *kihu*.
- (4) WL-LTB *kərē*, WL-ML *kəre*, WL-PD *kar:i*.
- (5) EL-LE *kana*, WL-AD *'keni*, WL-BM *kəni*, WL-BP *kəne*, WL-HRW *keni*, WL-KO *keni*, WL-LG *kəne*, WL-LML *keni*, WL-LWI *'kənɛ*, WL-LWL *kəni*, WL-TJ *kəni*, WL-WB *kəne*, WL-WW *kene*.
- (6) SK-MM *kəŋtik*, SK-TA *kəsik*, WL-BT *kəsĩ*, WL-DL *kusʔã*, WL-LH *kesi*, WL-MD *kesi*, WL-PK *kəsĩʔ*.
- (7) WL-LMK *kəda*, WL-WK *kudu*, WL-WT *kedi*.
- (8) CL-LP *kəlurkən*, CL-PN *kəlurəkən*.
- (9) CL-LK *bənut*, CL-MR *bəluit*.

smoke:

- (1) LH-SK *#nuhi*, CL-LR *ap 'nunan*, CL-PN *nunan*, EL-LE *nuhẽ*, EL-LMT *nuhẽ*, SK-HEW *nuhi*, SK-HEW *nuhin*, SK-MM *nuhin*, SK-TA *nuhi-n*, WL-AB *pə'nuhuŋ*, WL-AD *'nuhũ*, WL-BM *nuhu*, WL-BP *nuhũ*, WL-BT *nuhũ*, WL-DL *nuhũ*, WL-HRW *nuhu*, WL-IA *bənu*, WL-KO *nuhũ*, WL-LG *nuhũ*, WL-LH *nuhu*, WL-LMK *nuhuŋ*, WL-LWI *nuhũŋ*, WL-LWL *huhũ*, WL-MD *nuhu*, WL-ML *nuhu*, WL-MS *p'nuhuŋ*, WL-PD *pə'nuhuŋ*, WL-PK *nuhã*, WL-RE *nuhu*, WL-TJ *nuhu*, WL-WB *nuhu*, WL-WT *nuhũ*, WL-WW *nuhu*.
- (2) KD-LB *rubu*, KD-LW *rubu*.
- (3) CL-IL *sənəgurən*, CL-KK *sənəgur*, CL-KK *snə'gur*, CL-LK *sənəgur*, CL-LP *negur*, CL-LWK *sənəgur*, CL-LWT *hənəgur*, CL-MR *sənəgur*, WL-BL *sənəgur*, WL-LML *sənəgur*, WL-WK *sənəgur*.

smooth:

- (1) PFL **kəmelu*, CL-KK *gəmə'ls*, CL-LR *kə'mel kə'mel*, CL-LR *kəmelu'tən*, KD-LB *mɛ'luʔ*, SK-HEW *mɛlur*, WL-AB *ha'melu*, WL-AD *kə'melut*, WL-MS *mɛluk*, WL-PD *mɛluk*.
- (2) WL-LWI *kə'lɔhɔ*, WL-MS *b'lɔhɔʔ*.

snake:

- (1) PFL **ular*, CL-IL *ular*, CL-KK *ula*, CL-KK *u'la*, CL-KK *u'lar*, CL-LK *ular*, CL-LP *ularu*, CL-LR *u'lar*, CL-LWK *ularu*, CL-LWT *ularu*, CL-MR *ularu*, CL-PN *ulaj*, EL-LE *ula*, EL-LMT *ula*, KD-LB *u'lar*, KD-LW *ɯlar pupun*, SK-HEW *ular*, SK-MM *ular*, SK-TA *ular*, WL-AB *ula*, WL-AD *'ʔulaʔ*, WL-BL *ula*, WL-BM

ula?, WL-BN *ʔu'la*, WL-BP *ula?*, WL-BT *ula*, WL-DL *ula?*, WL-HRW *ula?*, WL-IA *ula*, WL-KO *ula?*, WL-LG *ula*, WL-LH *ula*, WL-LMK *ula*, WL-LML *ula*, WL-LTB *ula?*, WL-LWI *'ula?*, WL-LWL *ula*, WL-MD *ula*, WL-ML *ula*, WL-MS *ula*, WL-PD *ula*, WL-PK *ula*, WL-RE *ula*, WL-TJ *ula?*, WL-WB *ula*, WL-WK *ula*, WL-WT *ula?*, WL-WW *ula*.

soil:

- (1) SK-HEW *tana*, WL-AD *tana*, WL-LWI *'tana*, WL-MS *tana*, WL-PD *tana*.
- (2) PFL **ənay* 'sand; soil', CL-KK *ə'na*, CL-KK *ə'nan*, CL-LR *'anai*.

sole of foot:

- (1) LH-SK #*əpak* 'palm of hand; footprint', CL-KK *li ə'pak*, CL-LR *ə'pak*, SK-HEW *vaʔi əpak*.
- (2) WL-AB *leij 'əncəŋ*, WL-AD *leikət 'əncəŋ*, WL-MS *lei 'əncəŋ*, WL-MS *leij 'əncəŋ*, WL-PD *leij 'əncəŋ*.

some:

- (1) CL-IL *bopira*, CL-KK *uapira*, CL-LK *uapira*, CL-LP *bopira*, CL-LWK *pira*, CL-LWT *bopira*, CL-MR *pira*, CL-PN *dapira*, EL-LE *pirəpira*, EL-LMT *pirəpira*, WL-BL *bopira*, WL-BM *bopira*, WL-BP *bopira*, WL-BT *bopira*, WL-DL *bopira*, WL-HRW *bopira*, WL-IA *pira pira*, WL-KO *pira pira*, WL-LG *bopira*, WL-LH *pira*, WL-LMK *bopira*, WL-LML *bopira*, WL-LTB *bopirã*, WL-LWL *bopira*, WL-MD *pira pira*, WL-ML *pira pira*, WL-PK *bopira*, WL-RE *bopira*, WL-TJ *bopira*, WL-WB *bopira*, WL-WK *pirapira*, WL-WT *bopira*, WL-WW *bopira*.

sorcerer:

- (1) LH #*m<an>akap*, CL-KK *'makap 'taŋən*, WL-AB *ma 'naka*, WL-MS *m 'naka?*, WL-PD *m 'naka*.
- (2) CL-LR *'svaŋgi*, WL-PD *su 'waŋgi*.
- (3) LH-KD #*molan*, CL-KK *molan*, KD-LB *molan*, WL-LWI *'molan*.

sour:

- (1) PFL **g-nilu-k*, CL-KK *kni'luk*, KD-LB *ki'ru*, SK-HEW *niluk*, WL-AB *gilo*, WL-AD *'gilo*, WL-LWI *'gilo*, WL-MS *gilon*, WL-PD *gilon*, WL-PD *gilo*.
- (2) WL-MS *rapan*, WL-PD *rapan*.

spear:

- (1) LH-KD #*gala(r)*, CL-KK *ga'la*, CL-KK *ga'lar*, CL-LR *'galar*, KD-LB *kala*,

WL-AD *gala?*, WL-LWI *'gala*.

spider:

- (1) WL-AB *ka'raŋga*, WL-LWI *'kɔbɔ*, WL-MS *kay'ga 'ruka?*
 (2) PFL **kalaka*, CL-KK *lakbo'rit*, CL-LR *'laka bo'rit*, SK-HEW *kuku raka*,
 WL-PD *k'lake*.

star:

- (1) PFL **mətala*, EL-LE *malã*, EL-LMT *malã*, KD-LB *ma'le*, KD-LW *malæ manu?*, KD-LW *malæ tala*, SK-HEW *dala*, SK-MM *dala*, SK-TA *dala*, WL-AB *ta'mala*, WL-BL *təmala*, WL-BM *pətala*, WL-BN *ta'mala*, WL-BP *pətala*, WL-HRW *pətala*, WL-IA *təmala*, WL-LG *pətala*, WL-LWI *'pətala*, WL-LWL *pətala*, WL-MS *mala*, WL-PD *tə'mala*, WL-PK *tala?*, WL-TJ *kətala*, WL-WB *pətala*, WL-WW *pətala*.
 (2) WL-AD *ʔetəp*, WL-BT *epət*, WL-DL *etəp*, WL-KO *e-təp*, WL-LMK *e təp*, WL-MD *etəp*, WL-WT *etəp*.
 (3) CL-IL *tonar*, CL-KK *tona*, CL-KK *to'na*, CL-KK *to'nar*, CL-LK *tonar*, CL-LP *tonaru*, CL-LR *'tonar*, CL-LWK *tonaru*, CL-LWT *tonadʒa*, CL-MR *tona*, CL-PN *tonar*, WL-LH *tona*, WL-LML *tona*, WL-WK *tona*.
 (4) WL-LTB *bəlia*, WL-ML *bəlia*, WL-RE *bəlia*.

stem:

- (1) LH *#puka*, CL-IL *puka*, CL-KK *puk*, CL-LK *punak*, CL-LWK *puk*, CL-LWT *puk*, CL-MR *puk*, WL-BM *pukã*, WL-BP *pukən*, WL-BT *pukã*, WL-IA *pukən*, WL-LG *pukã*, WL-LH *pukã*, WL-LMK *pukã*, WL-LML *pukã*, WL-LTB *pukã*, WL-LWL *pukẽ*, WL-MD *pukã*, WL-ML *pukə*, WL-PK *pukã*, WL-RE *pukã*, WL-TJ *pukə*, WL-WB *pukə*, WL-WT *pukã*, WL-WW *pukã*.
 (2) LH-KD *#tava 'stem; grow'*, CL-LP *tafa*, CL-PN *tawan*, EL-LE *nawa*, EL-LMT *nawã*, WL-BL *tawa*, WL-WK *tawa*.
 (3) WL-DL *əpʔã*, WL-HRW *əpã*, WL-KO *əpʔã*.

stomach, belly:

- (1) PFL **tai*, CL-KK *tai kəbotin*, CL-KK *tai*, CL-LK *taisa orasa*, CL-LK *taisa*, CL-LP *tai kəboti*, CL-LWK *taiga*, CL-LWK *taik*, CL-MR *taisa*, CL-PN *tai oraj*, CL-PN *taih*, SK-HEW *tai dula*, SK-MM *taʔin*, SK-TA *taʔi-ŋ*, WL-AB *teiŋ*, WL-AD *taĩ*, WL-AD *taim*, WL-BL *taik kaluŋ*, WL-BN *teiŋ*, WL-BP *tăikəboti*, WL-BT *taik*, WL-BT *taik lorã*, WL-HRW *taik*, WL-IA *tain onən*, WL-IA *taik*, WL-KO *taik*,

WL-LG *tāi*, WL-LH *taikət onã*, WL-LH *taik*, WL-LMK *taĩ*, WL-LTB *tāi kənaluy*, WL-LWL *taikəboti*, WL-MD *taik*, WL-MD *taik kalunã*, WL-ML *təi*, WL-PD *taij*, WL-TJ *tāi kəboti*, WL-WB *tāikəboti*, WL-WK *taik*, WL-WK *taik kaluy*, WL-WT *taĩ*, WL-WW *tāikəboti*, WL-WW *tāij*.

(2) CL-IL *tukasa usər*, CL-IL *tukasa*, CL-LP *tuka*, CL-LWT *tukaha*, CL-LWT *tukaha kebotiha*, WL-MS *tukaŋ*, WL-MS *tuka?*, WL-PD *tukaŋ*.

(3) LH-KD *#(kə)boti*, CL-KK *taikəbotin*, CL-LP *taikəboti*, CL-LR *kəbo'tin*, CL-LWT *tukaha kebotiha*, CL-MR *kəbotisa*, EL-LE *botih*, EL-LMT *boti*, EL-LMT *lau boti*, KD-LB *bə'tin*, KD-LW *bətin*, WL-BM *kəboti*, WL-BP *tāikəboti*, WL-DL *kəboti*, WL-KO *kəbotik*, WL-LMK *kəboti*, WL-LML *kəboti*, WL-LWL *taikəboti*, WL-ML *kəboti*, WL-RE *kəboti*, WL-TJ *tāi kəboti*, WL-WB *tāikəboti*, WL-WT *kəbotik*, WL-WW *tāikəboti*.

(4) WL-BP *ale*, WL-LWL *ale*, WL-RE *ale*, WL-TJ *ale*, WL-WB *alek*.

(5) SK-HEW *dula*, SK-HEW *taidula*, WL-BM *dulak*, WL-PK *dulak onoʔo*, WL-PK *dulak*.

(6) WL-BL *luwuk*, WL-LG *luwu*, WL-LML *lufu*, WL-LTB *luwu*, WL-LWI *'luwuy*.

(7) WL-BT *taik lorã*, WL-HRW *lorã*.

(8) EL-LE *ũhã lau*, EL-LMT *lau boti*.

(9) WL-BL *taik kaluy*, WL-LTB *tāi kənaluy*, WL-MD *taik kalunã*, WL-WK *taik kaluy*.

stone:

(1) PFL **vatu*, CL-IL *fator*, CL-KK *va'tər*, CL-KK *'vatu*, CL-KK *watu*, CL-LK *wator*, CL-LP *fatoru*, CL-LR *'vator*, CL-LWK *watoru*, CL-MR *wato*, CL-PN *wator*, EL-LE *wato*, EL-LMT *wato*, KD-LB *wa?*, KD-LW *va? ai*, KD-LW *va? era*, KD-LW *va? luli*, KD-LW *va? nima?*, SK-HEW *vatu*, SK-HEW *vatu*, SK-MM *vaʔu*, SK-TA *βatu*, WL-AB *fato*, WL-AD *wato*, WL-BL *wato*, WL-BM *wato*, WL-BN *fato*, WL-BP *wato*, WL-BT *wato*, WL-DL *wato*, WL-HRW *wato*, WL-IA *wato*, WL-KO *wato*, WL-LG *wato*, WL-LH *wato*, WL-LMK *wato*, WL-LML *fato*, WL-LTB *wato*, WL-LWI *'wato*, WL-LWL *wato*, WL-MD *wato*, WL-ML *wato*, WL-MS *wato*, WL-PD *wato*, WL-PK *wato*, WL-RE *wato*, WL-TJ *wato*, WL-WB *wato*, WL-WK *wato*, WL-WT *wato*, WL-WW *wato*.

(2) WL-PD *nəhi*, WL-PD *mə'nəhi*.

storage house:

(1) LH-KD *#kebaŋ*, KD-LB *ʔebaŋ*, WL-LWI *'kebaŋ*.

(2) CL-KK *ve'tak*, CL-LR *'vetak*.

straight:

(1) PFL **mulur*, CL-IL *mulurən*, CL-KK *mo'lur*, CL-KK *mulur*, CL-LP *mulurən*, CL-LR *mulu'rən*, CL-LWK *mulur*, CL-LWT *mulu*, CL-MR *mulur*, CL-PN *mulurən*, EL-LMT *mulülū*, KD-LB *ʔolər*, KD-LW *ɔlər ælər*, SK-TA *dəlor*, WL-AB *molo*, WL-BL *mulu*, WL-BN *malləŋ*, WL-IA *mūr*, WL-LH *murã*, WL-LML *mūlu*, WL-MD *murã*, WL-ML *mulu*, WL-MS *mələŋ*, WL-PD *moloŋ*, WL-PK *lorʔã*, WL-WK *mulu*.

(2) WL-AD *kə'lohɔ*, WL-BT *kəlohok*, WL-DL *kəloho*, WL-HRW *kəloho*, WL-KO *kəlohʔã*, WL-LMK *kəloho*, WL-LTB *kəlodo*, WL-TJ *kəlo*, WL-WT *kəloho*, WL-WW *kəloho*.

(3) LH #*mopa*, EL-LE *mohpa*, WL-BM *mopa*, WL-BP *mopa*, WL-LG *mo-paʔ*, WL-LWI *'mopaʔ*, WL-LWL *mopa*, WL-RE *mopã*, WL-WB *mopa*.

sugar palm:

(1) LH-KD #*pola*, KD-LB *pɔ'le*, WL-AD *'polaʔ*, WL-MS *pɔla 'pukəŋ*, WL-PD *pɔla*.

(2) LH-KD #*kebol*, CL-KK *'kebol*, CL-LR *'kebol*, WL-LWI *'kebo*.

sugarcane:

(1) PFL **təvu*, CL-KK *tə'vu*, CL-KK *tə'vɔr*, CL-LR *tə'vor*, KD-LB *tə'hu*, SK-HEW *təvu*, WL-AB *ta'fɔ*, WL-AD *'təwɔ*, WL-LWI *'təwɔ*, WL-MS *tə'wɔ*, WL-PD *tawɔ*.

sun:

(1) PFL **lədav* 'sun; day', CL-IL *lədʒaf*, CL-LP *lədʒa*, CL-LWK *lədʒawu*, CL-LWT *lədʒaf*, CL-MR *lədʒa*, EL-LE *ləra*, KD-LB *lɔ'jɔ*, KD-LW *lɔjɔ matan*, KD-LW *lɔjɔ*, SK-HEW *ləro*, SK-TA *ləro*, WL-AB *la'ra*, WL-AD *'rɜra*, WL-BL *ləra*, WL-BM *ləra*, WL-BN *la'ra*, WL-BP *rəra*, WL-BT *rəra*, WL-DL *ləra*, WL-HRW *ləra*, WL-IA *ləra*, WL-KO *rəra*, WL-LG *ləra*, WL-LH *ləra*, WL-LMK *rəra*, WL-LML *ləra*, WL-LTB *ləra*, WL-LWI *'ləra:*, WL-LWL *rəra*, WL-MD *ləra*, WL-ML *ləra*, WL-MS *lə'r:a*, WL-PD *lə'r:a*, WL-PK *ləra*, WL-RE *ləra*, WL-TJ *ləra*, WL-WB *ləra*, WL-WK *ləra*, WL-WT *rəra*, WL-WW *ləra*.

(2) LH #*luvak*, CL-KK *lu'vak*, CL-KK *luwak*, CL-LK *luwak*, CL-LR *'luvak* *lə'dʒau*, CL-PN *luwak*, EL-LMT *luwa*.

sweat:

(1) WL-AB *ha'name*, WL-MS *nə'mək*, WL-PD *nam:ε*.

(2) PFL **(k)rəvun*, CL-KK *krə'vun*, CL-LR *kəṛə'vun*, KD-LB *re'wun*, SK-HEW *rəvu*, WL-AD *'ʔɜwũ*, WL-LWI *'əwuŋ*.

sweet:

(1) WL-AB *tajge*, WL-MS *tə'gε*, WL-PD *ta'g:ε*, WL-PD *taj'gε*.

(2) PFL **səru-k*, CL-KK *snə'ruk*, CL-LR *hənəru'kən*, CL-LR *hər hər*, KD-LB *ε'ru*, SK-HEW *hure?*, WL-LWI *'hərə*.

tail:

(1) PFL **ikur*, CL-IL *ikur*, CL-KK *ikur*, CL-KK *i'kur*, CL-LK *ikur*, CL-LP *ikur*, CL-LR *i'kur*, CL-LWK *ikur*, CL-LWT *ikur*, CL-MR *ikurɤ*, CL-PN *ikurən*, EL-LE *ʔiʔu*, EL-LMT *iʔũ*, SK-HEW *iʔur*, SK-HEW *iʔur*, SK-MM *iʔur*, SK-TA *iʔur*, WL-AB *ikukun*, WL-AD *'ʔikʔũ*, WL-BL *iku*, WL-BM *iku*, WL-BN *ikuk*, WL-BP *ikũ*, WL-BT *iku*, WL-DL *ikũ*, WL-HRW *ikuk*, WL-IA *ikuk*, WL-KO *ikũ*, WL-LG *ikũ*, WL-LH *ikũ*, WL-LMK *iku*, WL-LML *iku*, WL-LTB *kiku*, WL-LWI *iku'ʔuŋ*, WL-LWL *ikũ*, WL-MD *iku*, WL-ML *iku*, WL-MS *ikuk*, WL-PD *ikuk*, WL-PK *kikuʔũ*, WL-RE *iku*, WL-TJ *iku*, WL-WB *iku*, WL-WK *iku*, WL-WT *ikũ*, WL-WW *ikũ*.

(2) KD-LB *ʔebɔ*, KD-LW *ebɔ*.

tall:

(1) LH **blolo/golo*, CL-KK *gɔ'lɔk*, CL-LR *ata'dikə golo'kən*, WL-AB *ba'lɔlɔ*, WL-AD *blɔʔ'lɔ*, WL-LWI *bə'lola*, WL-MS *b'lɔlɔk*, WL-PD *b'lɔlɔk*.

tear:

(1) PFL **luu*, KD-LB *lun*, SK-HEW *lu*, WL-AB *lou*, WL-AD *'lõ:*, WL-LWI *'louŋ*, WL-MS *matan* *'louŋ*, WL-MS *lou*, WL-MS *louŋ*, WL-PD *matan* *'louŋ*.

(2) CL-KK *'mata* *'vaen*, CL-LR *mat vaer*, WL-LWI *mata* *'waʔiŋ*.

teeth:

(1) WL-AB *ulɔŋ*, WL-BN *ʔulɔŋ*, WL-MS *ulɔŋ*, WL-MS *ulɔ?*, WL-PD *ulɔŋ*.

(2) KD-LB *ʔawar*, KD-LW *avar*.

(3) PFL **ipə-*, CL-KK *pai*, CL-KK *pa'i*, CL-LK *paisa*, CL-LR *pai*, CL-LWK *paiga*, CL-MR *ipã*, CL-PN *paig*, EL-LE *ihpe*, EL-LMT *ipe*, WL-AD *'ipɜ?*, WL-BL *ipək*, WL-BM *ipə*, WL-BP *ipət*, WL-BT *ipə*, WL-DL *ipɤ*, WL-HRW *ipək*, WL-IA *ipək*, WL-KO *ipək*, WL-LG *ipə*, WL-LH *ipəkə*, WL-LMK *ipã*, WL-LML *ipã*, WL-

LTB *ipə*, WL-LWI *'ipəʔ*, WL-LWL *ipə*, WL-MD *ipək*, WL-ML *ipə*, WL-PK *ipək*, WL-RE *ipə*, WL-TJ *ipə*, WL-WB *ipə*, WL-WK *ipě*, WL-WT *ipʔã*, WL-WW *ipə*.

(4) CL-IL *kiakəs*, CL-LP *kiak*, CL-LWT *kiakha*.

(5) SK-HEW *niu*, SK-HEW *niʔu*, SK-MM *niun*, SK-TA *niu-ŋ*.

ten:

(1) WL-AB *kartou*, WL-BN *kartɔ*, WL-MS *kər'tou*, WL-PD *kə'rtou*. First part is an Alor-Pantar loan. See §9.4.4 for a discussion.

(2) PFL **s-pulu*, CL-KK *'spulo*, CL-LR *he'pulo*, KD-LB *pulu*, SK-HEW *pulu*, WL-AD *pulo*, WL-LWI *'pulɔ*.

testicles:

(1) PFL **lahe-k*, CL-KK *lɛ:k*, CL-LR *'leke 'teluk*, KD-LB *la'hɛʔ*, WL-AB *la'hakəŋ ku'lukuy*, WL-AD *'lahak*, WL-MS *lahak*, WL-MS *lahak 'kuluŋ*, WL-PD *lahak 'kuluŋ*.

(2) WL-AB *la'hakəŋ ku'lukuy*, WL-LWI *kulu'ʔuy*, WL-MS *lahak 'kuluŋ*, WL-PD *lahak 'kuluŋ*.

(3) CL-LR *'leke 'teluk*, SK-HEW *tɛɔ*.

thatch for roofing:

(1) LH #*luʔo*, CL-KK *lu'ɔr*, CL-KK *lu'ɔ*, WL-AB *luo*, WL-AD *'luʔo*, WL-LWI *'luʔɔ*, WL-MS *luo*, WL-PD *takar 'luo*, WL-PD *luo*.

thatched roof:

(1) LH #*tnakar*, CL-KK *na'kər*, CL-LR *'nakar*, WL-AB *luo 'taka*, WL-AD *'nakɜ*, WL-MS *takar*, WL-PD *takar*, WL-PD *tə'nakar*.

(2) CL-KK *u'vu*, CL-KK *u'vuy*.

thick:

(1) PFL **m-kapal*, SK-HEW *apar*, SK-MM *apar*, SK-TA *ʔapar*, KD-LB *kapal*, KD-LW *kapal vahɔʔ*, WL-AB *gapa*, WL-BN *gapa*, WL-MS *gapal*, WL-PD *gapal*.

(2) LH #*prəvak*, CL-KK *pərəwak*, CL-KK *prə'vak*, CL-LK *pəfərak*, CL-LR *prə'vak*, CL-PN *pərəwakən*, EL-LE *wəra*, EL-LMT *prəwa*, WL-AD *pɜ'wɜrã*, WL-BT *pəwərã*, WL-DL *pəwərã*, WL-HRW *pəwərʔe*, WL-IA *fəran*, WL-KO *pəwərã*, WL-LH *pəwərã*, WL-LMK *pəwərʔě*, WL-LML *pəfəre*, WL-LTB *kəwara*, WL-RE *pəwərě*, WL-WB *pəwəre*, WL-WT *pəwəre*, WL-WW *pəwəre*, WL-ML *pəwəwɜ*.

(3) WL-BM *bəsiʔ*, WL-LG *bəsiʔ*, WL-LWI *bə'siʔ*, WL-LWL *bəsĩʔ*, WL-TJ *bəsiʔ*.

(4) LH #*pənəŋe*, CL-IL *pənəŋenən*, CL-LP *pəŋe*, CL-LWT *pəŋe*, CL-MR *pənəŋe*, WL-BL *pəŋe*, WL-MD *pənəŋe*, WL-WK *pəŋe*.

(5) CL-LWK *təbalɤ*, WL-PK *təbal*.

thin:

(1) PFL **m-nipih-i*, CL-IL *mipi*, CL-KK *mip'iv*, CL-KK *mipiw*, CL-LK *mipi*, CL-LP *mipi*, CL-LR *mipiu*, CL-LWK *mipiwu*, CL-LWT *mənipifən*, CL-PN *mipiwən*, EL-LE *mipi*, EL-LMT *mipih*, KD-LB *mipi*, KD-LW *mipi tɛʔul*, WL-AB *mə'nipi*, WL-AD *mə'nipi*, WL-BL *nipi*, WL-BM *mənipi*, WL-BP *mənipi*, WL-BT *məniphĩ*, WL-DL *mənipi*, WL-HRW *məniphĩ*, WL-IA *nipi*, WL-KO *məniph*, WL-LG *mənipi*, WL-LH *bənipi*, WL-LMK *mənipi*, WL-LML *mənipi*, WL-LTB *mənipi*, WL-LWI *mə'nipi*, WL-LWL *mənipĩ*, WL-MD *bənipi*, WL-ML *bənipi*, WL-MS *nipi*, WL-MS *nipĩ*, WL-PD *mə'nipi*, WL-PD *nipi*, WL-PK *nipi*, WL-RE *mənipĩ*, WL-TJ *nipi*, WL-WB *mənipi*, WL-WK *nipi*, WL-WT *mənipi*, WL-WW *mənipi*.

(2) SK-HEW *bleler*, SK-HEW *leler*, SK-MM *beler*, SK-TA *blelär*.

thousand:

(1) PFL **rivu*/**ribu*, CL-KK *'ribu tu*, CL-LR *rib tu*, KD-LB *ribu*, SK-HEW *rivu*, WL-AB *ribu*, WL-AD *ri:bu*, WL-LWI *ribu 'toʔu*, WL-MS *ribu*, WL-PD *ribu*. See §6.2.4 for a discussion.

thread:

(1) CL-KK *ka'pək*, CL-LR *'kapha*, WL-MS *kapɔ*, WL-PD *kapo*.

(2) KD-LB *bə'naŋ*, SK-HEW *bənaŋ*, WL-AB *ba'n:əŋ*. From Indonesian/Malay *benang* 'thread'.

three:

(1) PFL **təlu*, CL-KK *tə'lu*, CL-LR *təl*, KD-LB *telu*, KD-LW *təlu*, SK-HEW *təlu*, SK-TA *təlu*, WL-AB *tal:ɔ*, WL-AD *'telo*, WL-BN *tal:au*, WL-LWI *'tɛ:ɔ*, WL-MS *tə'l:ɔ*, WL-PD *tal:ɔ*.

thunder:

(1) WL-AB *hila*, WL-MS *hilaʔ brə'kaŋ*, WL-PD *hila*.

(2) LH-SK #*kleka*, CL-KK *kələ'gor*, CL-LR *'kləkə*, SK-HEW *kleka liar*, WL-AD *kɜ'ɜka*.

tinea:

(1) PFL ***m-panau**, SK-HEW *na:u?*, WL-AB *manau*, WL-AD *mɜ'naɔ*, WL-LWI *'mau*, WL-MS *nao*, WL-PD *m'nao*, WL-PD *nao*.

to accuse:

- (1) PFL ***tuɖu**, WL-AB *turuŋ*, WL-MS *turuŋ 'apa*, WL-PD *turuŋ 'tumaŋ*.
- (2) LH #**əvan**, CL-KK *ə'van*, WL-AD *ʔɜ'wa?*.
- (3) SK-HEW *rudu*, WL-LWI *pə'lainj*.
- (4) WL-MS *hode k'dire?*, WL-PD *hɔde*.

to answer:

- (1) LH #**tapan**, CL-KK *ta'pan*, WL-AD *'tapã*, WL-LWI *'tapaŋ*.
- (2) LH-KD #**soloi**, CL-LR *'holoi*, KD-LB *hɔlɔ*.

to ask question:

- (1) LH KD #**kəter**, KD-LW *kete(r)*, WL-AB *gate*, WL-LWI *'gəta*, WL-MS *gə'ter*, WL-PD *ga'ter*.
- (2) PFL ***dasan** 'ask; report', CL-LR *'daha*, KD-LB *da'haŋ*, SK-HEW *dahan* 'report', WL-AD *'dahã*.

to bark:

- (1) LH-KD #**bovoŋ**, CL-KK *bo'vo*, KD-LB *bɔ'wɔŋ*, WL-LWI *'bowo*, WL-MS *bowɔŋ*, WL-PD *bɔwɔŋ*.

to bathe:

- (1) LH-KD #**həbu**, CL-KK *lə'bo*, CL-LR *lə'bo*, KD-LB *hebu*, WL-AB *habɔ*, WL-AD *'habo*, WL-BN *hab:ɔ*, WL-LWI *'həbɔ*, WL-MS *ha'b:ɔ*, WL-MS *habɔ*, WL-PD *hab:ɔ*.

to be silent:

- (1) WL-MS *rama?*, WL-PD *rama*.
- (2) CL-KK *kə'pae*, WL-AD *'paha*.
- (3) CL-KK *mə'li*, CL-LR *məl məl*.

to bite:

- (1) PFL ***giki**, CL-IL *gik*, CL-KK *'gike*, CL-KK *giki*, CL-KK *'giki*, CL-LK *giki*, CL-LP *giki*, CL-LR *'gike*, CL-LWK *giki*, CL-LWT *gike*, CL-MR *giki*, CL-PN *gik*,

EL-LE *giʔi*, EL-LMT *giʔi*, KD-LB *kiʔi*, KD-LW *kiʔi avar*, KD-LW *kiʔi ka*, SK-MM *giʔi*, SK-TA *ʔiʔi*, WL-AB *gaki*, WL-AD *'gike*, WL-BL *gike*, WL-BM *gike*, WL-BN *gaki*, WL-BP *gike*, WL-BT *gike*, WL-DL *gike*, WL-HRW *gike*, WL-IA *gike*, WL-KO *gike*, WL-LG *gike*, WL-LH *gike*, WL-LMK *gike*, WL-LML *goki*, WL-LWI *'gike*, WL-LWL *gike*, WL-MD *gike*, WL-ML *gike*, WL-MS *gakiʔ*, WL-PD *gaki*, WL-PK *gike*, WL-RE *gike*, WL-TJ *gike*, WL-WB *gike*, WL-WK *gike*, WL-WT *gike*, WL-WW *gike*.

(2) LH-SK **#baka**, SK-HEW *baka*, SK-HEW *vaka*, WL-LWI *'baka*.

to blow:

(1) EL-LE *pui*, EL-LMT *pu*, KD-LB *puiʔ*, WL-AB *pui*, WL-AD *'puit*, WL-BN *pui*, WL-IA *puit*, WL-KO *puit*, WL-LMK *pui*, WL-MD *puit*, WL-MS *puiʔk*, WL-PD *pui*, WL-WK *puis*.

(2) WL-BM *bu*, WL-BP *bu*, WL-BT *bu*, WL-LG *bu*, WL-LTB *bu*, WL-LWI *bu*, WL-LWL *bu*, WL-ML *bu*, WL-RE *kəbuha*, WL-TJ *bu*, WL-WB *bu*, WL-WW *bu*.

(3) LH **#dihe**, CL-IL *die*, CL-KK *did̥ʒi*, CL-LK *did̥ʒi*, CL-LWT *dihe*, WL-BL *did̥i*, WL-DL *dihe*, WL-HRW *dihẽ*, WL-LH *dihe*, WL-LML *die*, WL-PK *dihẽ*, WL-WT *dihẽ*.

(4) LH-KD **#puur**, CL-KK *pu:r*, CL-LR *purə'həŋ*, CL-LWK *pur*, CL-MR *purəs*, CL-PN *pura*.

(5) PFL ***prupi/plupi**, SK-HEW *pərupi*, SK-HEW *plupi*, SK-MM *prupi*.

to breathe:

(1) CL-KK *nai*, CL-LWT *ər naiha*, CL-MR *ərə nai*, SK-TA *aiŋ*, WL-AB *nami* *'nihij*, WL-AD *'nəri'naĩ*, WL-AD *'nəri'naik*, WL-BT *nəri nahij*, WL-DL *nəri nahij*, WL-HRW *nəri nahij*, WL-KO *nəri naik*, WL-LMK *nərināi*, WL-LML *ərā nāi*, WL-PD *naŋ 'nahij*, WL-RE *ər nāi*, WL-WT *nəri nahij*, WL-WW *nərināi*.

(2) KD-LB *pəhɔʔ*, KD-LW *ipɔ*, WL-MS *hɔpaŋ*.

(3) LH **#esari nai**, CL-IL *ər naisa*, CL-KK *esa nain*, CL-LK *esa nai*, CL-LR *'ehe 'nai*, CL-LWK *sənaiga*, CL-PN *ehe naig*, EL-LMT *ha nei*, WL-BL *hea naik*, WL-BM *harĩ*, WL-BP *harinaĩ*, WL-IA *həŋə naikə*, WL-LG *hari nait*, WL-LH *eha naikə*, WL-LTB *hari nāi*, WL-LWI *'hariŋ*, WL-LWL *harināi*, WL-MD *eha naikə*, WL-ML *eha nai*, WL-TJ *hari nāi*, WL-WB *harināi*.

(4) SK-HEW *bo:ʔ*, WL-PK *boʔ*.

to burn, to clear land:

(1) PFL ***tuno** 'burn; grill', CL-KK *'tunɔ*, WL-BN *tun:ɔ*, WL-HRW *tuno*, WL-LMK *tuno*, WL-LTB *tuno*, WL-MD *tuno*, WL-ML *tuno*, WL-PK *tuno*, WL-RE *tuno*, WL-TJ *tuno*, WL-WW *tuno*.

(2) WL-AB *saru 'ekaj*, WL-AD *'səru?*, WL-BL *səru*, WL-BM *səru*, WL-BP *səru*, WL-BT *səru*, WL-DL *səru*, WL-IA *səru*, WL-KO *səru*, WL-LG *səru*, WL-LH *səru*, WL-LWI *se'ru?*, WL-LWL *səru*, WL-PD *sa'ru*, WL-PD *sa'ru 'ekaj*, WL-WB *səru*, WL-WK *səru*, WL-WT *səru*.

(3) LH-KD ***papi**, CL-IL *papi*, CL-KK *papi*, CL-KK *pa'pi*, CL-LK *papi*, CL-LP *papi*, CL-LR *'papi 'ole*, CL-LWK *papi*, CL-LWT *papi*, KD-LB *pa?*, KD-LW *pa? bəɔ?*, WL-LML *papi*, WL-MS *papij*, WL-PD *papij*.

(4) PFL ***deket**, EL-LE *de?e*, EL-LMT *da?e*.

(5) SK-HEW *holo*, SK-TA *holo*, SK-TA *olor*.

to buy:

(1) LH #**hopi**, CL-KK *'opi*, CL-LR *op*, WL-AB *hɔpɛ*, WL-AD *'hope*, WL-LWI *'hope*, WL-MS *hɔpɛ*, WL-PD *hɔpɛ*.

to chew:

(1) PFL ***mama?**, CL-KK *ma'mɔt*, KD-LB *ma'mɛ?*, SK-HEW *ma?ma?*, WL-AB *mame*, WL-MS *mame*, WL-PD *mame*.

(2) CL-LR *ma'hap*, WL-AD *'bahak*, WL-LWI *'baha*.

to choose:

(1) PFL ***pili?**, CL-KK *'pili*, CL-KK *'pile*, CL-LR *'pile*, KD-LB *pil*, SK-HEW *li?i*, WL-AB *pilɛ*, WL-AD *'pile*, WL-LWI *'pile?*, WL-MS *pilɛ*, WL-PD *pile*.

to climb:

(1) LH-KD #**gedi** 'go up; ascend', KD-LB *kɛu*, WL-AB *gere*, WL-AD *'gere*, WL-LWI *'gere*, WL-MS *gere*, WL-PD *gere*.

(2) CL-KK *tu'kar*, CL-LR *'tuka*.

(3) PFL ***hakay** 'ascend', CL-KK *akadz*, KD-LW *a?*, SK-HEW *ha?e*, WL-AD *haka*.

to close:

(1) WL-MS *tera?*, WL-PD *tera*.

(2) LH-KD #**letu?**, KD-LB *le'tu?*, WL-AD *'lɜ:tu?*, WL-LWI *'lətu?*.

(3) LH-SK #*ihere*, CL-LR *ihε'raŋ*, SK-HEW *hære?*.

to come:

(1) FL #*bəCo(l)*, CL-IL *bəsol*, CL-LR *'bedʒɔ*, CL-LWT *bəsol*, CL-MR *bəsol*, EL-LMT *bəso*, WL-AB *beta*, WL-AD *'beto*, WL-BL *bəsol*, WL-BN *beta*, WL-BT *beto*, WL-DL *beto*, WL-HRW *beto*, WL-IA *beto*, WL-KO *beto*, WL-LMK *beto*, WL-LML *bəso*, WL-ML *bəso*, WL-MS *beta*, WL-MS *betaŋ*, WL-PD *beta*, WL-WB *beto*, WL-WK *bəsol*, WL-WT *beto*, WL-WW *beto*.

(2) SK-HEW *səga*, WL-BM *səga*, WL-BP *səga*, WL-LG *səga*, WL-LTB *səga*, WL-LWI *'səga*, WL-PK *səgat*, WL-RE *səga*, WL-TJ *səga*.

(3) LH #*oli*, CL-KK *ɔ'li*, CL-KK *oli*, CL-PN *oliŋ*, WL-LWL *oli*.

(4) WL-LH *lohe*, WL-MD *lohe*.

(5) CL-LWK *haka*, EL-LE *hoʔe*.

(6) CL-LP *kəfi*, CL-LP *məfi*.

(7) PFL **mai*, CL-LK *mene*, SK-HEW *mai*, SK-HEW *main*, SK-MM *main*, SK-TA *mai*.

to cook:

(1) WL-AB *dakaŋ* 'apa, WL-BN *dakaŋ*, WL-MS *dakaŋ*, WL-PD *dakaŋ*.

(2) LH-KD #(g)*iu*, CL-IL *giu*, CL-LK *giu*, CL-LP *giu*, CL-LWK *giu*, CL-LWT *giu*, CL-MR *giu*, CL-PN *giu*, KD-LB *iu*, KD-LB *iu* 'daʔa.

(3) WL-AD *'dena*, WL-DL *dena*, WL-KO *dena*, WL-LML *denã*, WL-WT *dena*.

(4) LH #*bisu*, EL-LE *bihi*, EL-LMT *bihi*, WL-BL *bio*, WL-BM *bihi*, WL-BP *bihi*, WL-BT *bihi*, WL-HRW *bihi*, WL-IA *bihi*, WL-LG *bihi*, WL-LH *bihi*, WL-LMK *bihi*, WL-LTB *bihi*, WL-LWI *'bihi*, WL-LWL *bihi*, WL-MD *bihi*, WL-ML *bihi*, WL-PK *bihi*, WL-RE *biso*, WL-TJ *bihi*, WL-WB *bihi*, WL-WK *bihi*, WL-WW *bihi*.

(5) CL-KK *lo'lu*, CL-KK *loluy*, CL-LR *'lolu*.

(6) SK-HEW *huna*, SK-HEW *hunaʔ*, SK-MM *huna*.

to count:

(1) SK-TA *reke-ŋ*, WL-AB *rekeŋ*, WL-BN *rekiŋ*, WL-MS *rekeŋ*, WL-PD *rekeŋ*.
From Dutch *rekenen* 'calculate'.

(2) KD-LB *ka'reʔ*, KD-LW *kareʔ kataʔ*.

(3) LH-SK #*gasik*, CL-KK *ga'sip*, CL-LR *'gahip*, SK-HEW *gasik*, SK-TA *api-ŋ gasik*, SK-TA *gasik*, WL-AD *'gasik*, WL-LWI *'gasi*.

(4) KD-LW *kare? kata?*, SK-TA *gata kira-ŋ*.

to crouch:

(1) WL-AB *tɔbɔ 'tae*, WL-MS *tɔbɔ 'tɛlɛŋ*.

WL-MS *tɔbɔ 'tɛlɛŋ*, (2) SK-HEW *bɔkɔŋ*, WL-PD *sɔrɔ 'bɔkɔŋ*.

(3) WL-LWI *'luguŋ*, WL-PD *lugu*.

(4) CL-KK *sva 'kat*, CL-LR *həva 'kat*.

to cry:

(1) PFL **tani*, EL-LE *tani*, EL-LMT *tani*, SK-HEW *dani*, SK-HEW *tani*, SK-MM *ṭani*, WL-AD *'tani*, WL-BL *tani*, WL-BM *tani*, WL-BN *tanĩ*, WL-BP *tani*, WL-BT *tani*, WL-DL *tani*, WL-HRW *tani*, WL-IA *tani*, WL-KO *tani*, WL-LG *tani*, WL-LH *tani*, WL-LMK *tani*, WL-LML *tani*, WL-LWI *'taniŋ*, WL-LWL *tani*, WL-MD *tani*, WL-ML *tani*, WL-MS *taniŋ*, WL-MS *tani?*, WL-PD *tani*, WL-PD *taniŋ*, WL-PK *taniŋ*, WL-RE *tani*, WL-TJ *tani*, WL-WB *tani*, WL-WK *tani*, WL-WT *tani*, WL-WW *tani*.

(2) CL-IL *əra*, CL-KK *ə'ra*, CL-KK *əraw*, CL-LK *əra*, CL-LWT *əra*.

(3) CL-LR *'giar*, CL-LWK *giar*, CL-PN *giar*.

(4) CL-LP *giləŋ*, CL-MR *gilɔ*.

to cut:

(1) WL-AB *pakɔu*, WL-BN *pakɔu*, WL-PD *pə'kɔu*.

(2) KD-LB *pɛ'ri?*, KD-LW *pæri? pi?*.

(3) LH-KD #*belu*, CL-IL *bel*, CL-KK *belu*, CL-KK *'belu*, CL-LK *belu*, CL-LP *belu*, CL-LR *'belo*, CL-LWK *belu*, CL-LWT *belu*, CL-MR *belu*, CL-PN *bel*, EL-LE *belu*, EL-LMT *belo*, KD-LW *beluŋ*, WL-AD *'belo*, WL-BL *belo*, WL-BM *belo*, WL-BP *belo*, WL-BT *belo*, WL-DL *belo*, WL-HRW *belo*, WL-IA *belo*, WL-KO *belo*, WL-LG *belo*, WL-LH *belo*, WL-LMK *belo*, WL-LML *belo*, WL-LTB *belo*, WL-LWI *'belo*, WL-LWL *belo*, WL-MD *belo*, WL-ML *belo*, WL-PK *belo*, WL-RE *belo*, WL-TJ *belo*, WL-WB *belo*, WL-WK *belo*, WL-WT *belo*, WL-WW *belo*.

(4) PFL **pati*, KD-LW *pate iɦu*, SK-HEW *bati*, SK-HEW *pati*, SK-MM *paṭi*, SK-TA *patar pati*, WL-LWI *'pata*.

(5) SK-TA *poro*, SK-TA *poʔok*, WL-AD *'poro?*.

(6) KD-LB *ta'di?*, WL-MS *tə'rri*, WL-PD *ta'rri*.

(7) CL-KK *ba'kat*, CL-LR *bakə'təŋ*.

to dance:

(1) LH-KD #sedu, CL-KK sedu, CL-LR he'du, CL-PN hedu, KD-LB he'duy.

(2) WL-AD 'ni:g3, WL-HRW nigə, WL-KO nigə.

(3) FL #soka, CL-IL soka, CL-KK soka, CL-KK 'soka, CL-LK soka, CL-LP soka, CL-LWK soka, CL-LWT soka, CL-MR soka, EL-LE sokə, EL-LMT sokə, SK-HEW soka, WL-BL soka, WL-BM soka, WL-BP soka, WL-BT soka, WL-IA soka, WL-LG soka, WL-LH soka, WL-LMK soka, WL-LML soka, WL-LTB soka, WL-LWI 'soka, WL-LWL soka, WL-MD soka, WL-ML soka, WL-PK soka, WL-RE soka, WL-TJ soka, WL-WB soka, WL-WK soka, WL-WT soka, WL-WW soka.

to die:

(1) PFL *matay, CL-IL matad̥ʒa, CL-KK matad̥ʒ, CL-KK matek, CL-LK matad̥ʒ, CL-LP matad̥ʒ, CL-LR ma'tei, CL-LWT matad̥ʒ, CL-MR matad̥ʒ, CL-PN matejəg, EL-LE mata, EL-LMT mata, KD-LB mate, SK-HEW mate, SK-MM matən, SK-TA mate, WL-AB mati, WL-AD 'mata:, WL-BL mate, WL-BM mata, WL-BN mate, WL-BP mata, WL-BT mate, WL-DL matē, WL-HRW mata, WL-IA matad̥ʒ, WL-KO mata, WL-LG mata, WL-LH matad̥ʒ, WL-LMK mata, WL-LML mata, WL-LTB mata, WL-LWI 'mata, WL-LWL mata, WL-MD mataja, WL-ML mate, WL-MS mate, WL-PD mate, WL-PK matĩ, WL-RE mata, WL-TJ mata, WL-WB mata, WL-WK matad̥ʒ, WL-WT mata, WL-WW mate.

to dig:

(1) PFL *gali, CL-IL gali, CL-KK ga'li, CL-KK galiŋ, CL-LK gali, CL-LP gali, CL-LR gal, CL-LWK gali, CL-LWT gali, CL-MR gali, CL-PN gal, EL-LE gəli, EL-LMT gali, KD-LB kæl, KD-LW kæl kəɾɔʔ, KD-LW kaliŋ kəɾɔʔ, SK-HEW ʒli, SK-HEW gali, SK-HEW ʔali, SK-MM ʔali, SK-TA gali, WL-AB galiŋ, WL-BL gəli, WL-BN gali, WL-BP gəli, WL-LG gali, WL-LH gali, WL-LWL gəli, WL-MD gali, WL-ML gali, WL-MS galiŋ, WL-PD galiŋ, WL-PK galiŋ, WL-TJ gəli, WL-WB gəli.

(2) PFL *gurit, CL-LR 'gurit, KD-LW koroʔ, SK-TA goʔi, WL-LML gui, WL-LTB guʔi, WL-LWI 'guʔi, WL-PD gui, WL-RE guĩ.

(3) WL-AD 'baʔe, WL-BT baʔe, WL-DL baʔe, WL-HRW baʔe, WL-KO baʔe, WL-LMK bake, WL-WK gaʔe, WL-WT baʔe, WL-WW baʔe.

(4) KD-LW kæl kəɾɔʔ, KD-LW kaliŋ kəɾɔʔ, SK-TA gorak.

to divide:

(1) PFL *bagi, SK-HEW bige, WL-AB bage, WL-MS bage, WL-PD bage.

(2) LH #vekan, CL-KK ve'ka, CL-LR ve'kan, WL-AD 'weka, WL-LWI 'wekan,

WL-MS *wekaŋ*.

to dream:

(1) LH #*turu*, CL-KK *tu'rən*, CL-LR *'turən*, WL-AB *nureŋ*, WL-AD *t3'uru*, WL-LWI *'turəŋ*, WL-MS *nurɔŋ*, WL-PD *nurɔŋ*.

to drink:

(1) PFL *-*inu*, CL-IL *ten*, CL-IL *men*, CL-KK *-enu*, CL-KK *menu*, CL-LK *menu*, CL-LP *menu*, CL-LP *tenu*, CL-LR *mən*, CL-LWK *menu*, CL-LWT *tenu*, CL-MR *tenu*, CL-PN *ken*, CL-PN *men*, EL-LE *enu*, EL-LE *menu*, EL-LMT *enu*, EL-LMT *menu*, KD-LB *min*, KD-LW *in a*, SK-HEW *-inu*, SK-HEW *tinu*, SK-MM *tinu*, SK-TA *ʔ-inu*, WL-AB *kenuŋ*, WL-AD *'kenu*, WL-AD *'menu*, WL-AD *'nenu*, WL-BL *tenu*, WL-BM *tenu*, WL-BP *tenu*, WL-BT *tenu*, WL-DL *tenu*, WL-HRW *kenu*, WL-IA *tenu*, WL-KO *kenu*, WL-LG *tenu*, WL-LH *kenu*, WL-LMK *tenu*, WL-LML *tenu*, WL-LTB *tenu*, WL-LWI *'tenuŋ*, WL-LWL *tenu*, WL-MD *kenu*, WL-MD *menu*, WL-ML *tenu*, WL-MS *go 'kenuŋ*, WL-MS *kenuŋ*, WL-PD *menuŋ*, WL-PK *tenu*, WL-RE *tenu*, WL-TJ *tenu*, WL-WB *tenu*, WL-WK *nenu*, WL-WT *menu*, WL-WW *tenu*.

to eat:

(1) PFL *-*kaan*, CL-IL *ka*, CL-IL *ga*, CL-KK *ga*, CL-KK *ga:*, CL-KK *ka*, CL-KK *ka:*, CL-LK *ga*, CL-LK *ka*, CL-LP *ga*, CL-LP *ka*, CL-LR *ga*, CL-LR *ka*, CL-LWK *ga*, CL-LWK *ka*, CL-LWT *ga*, CL-LWT *pa*, CL-MR *ka*, CL-PN *ga*, CL-PN *ka*, EL-LE *a*, EL-LE *ga*, EL-LMT *a*, EL-LMT *ga*, KD-LB *ka:*, KD-LW *a*, SK-HEW *ea*, SK-HEW *ʔa:*, SK-MM *ʔa:*, SK-TA *ʔoa*, WL-AB *kaŋ 'apa*, WL-AD *'goe kã*, WL-BL *ka*, WL-BM *təkã*, WL-BP *go*, WL-BP *kã*, WL-BT *go*, WL-BT *ka*, WL-DL *təka*, WL-HRW *go*, WL-HRW *ka*, WL-IA *go*, WL-IA *ka*, WL-KO *ka*, WL-LG *təkã*, WL-LH *kã*, WL-LMK *go*, WL-LMK *kã*, WL-LML *go*, WL-LML *kã*, WL-LTB *təkã*, WL-LWI *'təkaŋ*, WL-LWL *go*, WL-LWL *kã*, WL-MD *go*, WL-MD *kã*, WL-ML *go*, WL-ML *kɤ*, WL-MS *kaŋ*, WL-PD *gaŋ*, WL-PD *kaŋ*, WL-PK *go*, WL-PK *kã*, WL-RE *go*, WL-RE *kã*, WL-TJ *təkã*, WL-WB *go*, WL-WB *kã*, WL-WK *ga*, WL-WK *ka*, WL-WT *go*, WL-WW *go*, WL-WW *kã*.

to fall from above:

(1) PFL *-*lodoŋ*, KD-LW *lɔduŋ*, SK-HEW *lodoŋ*, WL-MS *gɔkal 'lɔdoŋ*, WL-PD *gɔkal 'lɔdoŋ*.

(2) EL-LE *moru*, KD-LB *mo'ruʔ*, KD-LW *mɔruʔ hɔkɔ*.

- (3) KD-LW *hubaʔ həkɔ*, KD-LW *hubaʔ kɔliʔ*.
- (4) KD-LW *hubaʔ həkɔ*, KD-LW *mɔruʔ həkɔ*.
- (5) WL-BM *dəkak*, WL-BT *dəkak*, WL-IA *dəka*, WL-LG *dəka*, WL-LWI *'dəka*, WL-PK *dəkak*, WL-RE *dəka*, WL-TJ *dəkak*, WL-WB *dəkak*.
- (6) LH #*ləŋat*, CL-IL *ləŋata*, CL-KK *lə'ŋat*, CL-LK *ləŋat*, CL-LWK *ləŋat*, WL-BL *ləŋat*, WL-LH *ləŋa*, WL-LMK *ləŋa*, WL-MD *ləŋat*, WL-ML *ləŋa*.
- (7) CL-KK *pe'al*, EL-LMT *goli*.
- (8) CL-LR *pə'hol d̥ʒe 'lodo*, CL-PN *pahol*.
- (9) CL-MR *gəta*, WL-WK *gətak*.
- (10) SK-MM *ʔəla*, SK-TA *ʔela*.

to fall over:

- (1) KD-LB *hubaʔ*, WL-AB *hoba*, WL-AD *'hobat*, WL-MS *hobak*, WL-PD *hɔbak*.
- (2) CL-LR *'tobaŋ*, WL-AD *'tobaŋ*, WL-LWI *'tobaŋ*.
- (3) LH-KD #*gokal*, CL-KK *gokal*, CL-LWT *gokala*, KD-LW *kɔʔal*, WL-AB *gɔka 'iti 'loloŋ*, WL-AB *gɔka*, WL-AD *'gɔka*, WL-BN *gɔka*, WL-DL *goka*, WL-HRW *goka*, WL-KO *goka*, WL-LML *goka*, WL-LTB *goka*, WL-MS *gɔkal 'lɔdɔŋ*, WL-PD *gɔkal*, WL-PD *gɔkal 'lɔdɔŋ*, WL-WT *goka*, WL-WW *goka*.

to fight:

- (1) WL-BN *ka'laĩ 'fak:ĩ*, WL-PD *k'lainʔ wə'ki*.
- (2) PFL **pə-vunu*, CL-IL *puno*, CL-KK *puno*, CL-KK *'punu*, CL-LK *punu*, CL-LP *pun*, CL-LR *'punu vak'hi*, CL-LWK *puno*, CL-LWT *punod̥ʒ*, CL-MR *puno*, CL-PN *punu*, KG-LW *unu*, SK-HEW *punu viʔin*, SK-TA *punu βiʔi-ŋ*, WL-AD *pə'wuno*, WL-BL *uno*, WL-BT *pəwuno*, WL-DL *pəwuno*, WL-HRW *pəwuno*, WL-IA *uno*, WL-KO *pəwuno*, WL-LH *uno*, WL-LMK *pəwuno*, WL-LML *uno*, WL-LTB *kəwuno*, WL-MD *uno*, WL-ML *pəwuno*, WL-WK *unho*, WL-WT *pəwuno*, WL-WW *pəwuno*.
- (3) LH #*gəni*, EL-LMT *gəni*, WL-BM *gəni*, WL-BP *gəni*, WL-LG *gəni*, WL-LWI *'gəniŋ*, WL-LWL *gəni*, WL-RE *gəni*, WL-TJ *gəni*, WL-WB *gəni*.
- (4) WL-MS *boa wə'ki*, WL-PD *boa wə'ki*.
- (5) KD-LB *ʔawe 'wɛʔ*, KD-LW *avɛ unu*.

to float:

- (1) WL-AB *nepi*, WL-BN *nepi*, WL-MS *nepik*, WL-PD *nepik*.
- (2) PFL **baɔ*, CL-IL *baɔŋa*, CL-LK *baɔ*, CL-LP *baɔŋa*, CL-LR *'baɔ*, CL-LWT *baɔŋa*, CL-MR *baɔ*, CL-PN *baɔ*, EL-LE *bɔɔ*, EL-LMT *baɔ*, KD-LB *baɔ*, KD-LW

baɔ tata, SK-HEW *bau lele*, WL-AD 'baɔʔ, WL-BL *baɔ*, WL-BM *baɔ*, WL-BP *baɔ*, WL-BT *baɔ*, WL-DL *baɔ*, WL-HRW *baɔ*, WL-IA *baɔ*, WL-KO *baɔ*, WL-LG *baɔ*, WL-LH *baʔo*, WL-LMK *baɔ*, WL-LML *baɔ*, WL-LTB *baɔ*, WL-LWI 'baɔ, WL-LWL *baɔ*, WL-MD *baɔ*, WL-ML *baɔ*, WL-PK *baɔ*, WL-RE *baɔ*, WL-TJ *baɔ*, WL-WB *baɔ*, WL-WK *baɔ*, WL-WT *baɔ*, WL-WW *baɔ*.

(3) CL-KK *teda*, CL-KK 'teda.

(4) CL-LWK *nebon*, SK-HEW *nebon*.

to flow:

(1) WL-AB *barɛ*, WL-PD *bə'r:ɛŋ*.

(2) LH-SK #*baɲak*, CL-IL *baɲ*, CL-KK *baɲak*, CL-KK *ba'ɲak*, CL-LK *baɲak*, CL-LP *baɲ*, CL-LR 'baɲak, CL-LWK *baɲak*, CL-LWT *baɲ*, CL-MR *baɲɤ*, CL-PN *ba*, EL-LE *bã*, EL-LMT *bã*, SK-HEW *baɪ*, SK-MM *baɪ*, SK-TA *baɪ*, WL-AD *baʔ*, WL-BM *ba*, WL-BP *ba*, WL-BT *baʔ*, WL-DL *baʔ*, WL-HRW *baʔ*, WL-IA *baɲ*, WL-LG *ba*, WL-LMK *ba*, WL-LML *bã*, WL-LTB *ba*, WL-LWI *baɪ*, WL-LWL *ba*, WL-MS *wai 'baɲ*, WL-PK *ba*, WL-RE *ba*, WL-TJ *ba*, WL-WB *ba*, WL-WK *baɲak*, WL-WT *ba'ʔ*, WL-WW *ba*.

(3) KD-LB *kawaɲ*, KD-LW *kavaɲ kɔɲ*.

(4) WL-BL *golo*, WL-LH *golo*, WL-MD *golo*.

to fly:

(1) LH-KD #*bəka*, CL-IL *bək*, CL-KK *bəka*, CL-KK *bə'ka*, CL-LK *bəka*, CL-LP *bəkə*, CL-LR 'bəkə, CL-LWK *bəka*, CL-LWT *bək*, CL-MR *bəkɤ*, CL-PN *bəka*, EL-LE *bɔʔʔ*, EL-LMT *baʔã*, KD-LB *baʔa*, KD-LW *baʔa baæʔ*, WL-AB *baka*, WL-AD 'bɜka, WL-BL *bəka*, WL-BM *bəka*, WL-BN *bak:ɲ*, WL-BP *bəka*, WL-BT *bəka*, WL-DL *bəka*, WL-HRW *bəka*, WL-IA *bəka*, WL-KO *bəka*, WL-LG *bəka*, WL-LH *bəka*, WL-LMK *bəka*, WL-LML *bəka*, WL-LTB *bəka*, WL-LWI 'bəka, WL-LWL *bəka*, WL-MD *bəka*, WL-ML *bəka*, WL-MS *bə'kaɲ*, WL-PD *bək:ɲ*, WL-PK *bəka*, WL-RE *bəka*, WL-TJ *bəka*, WL-WB *bəka*, WL-WK *bəka*, WL-WT *bəka*, WL-WW *bəka*.

(2) SK-HEW *horo*, SK-MM *horo*, SK-TA *horo*.

to fold:

(1) PFL **ləpət*, CL-KK *lə'pət*, CL-LR 'ləpət, KD-LB *lə'piʔ*, SK-HEW *ləpet*, WL-AB *lapɛ*, WL-AD 'ləpɜt, WL-LWI 'ləpək, WL-MS *lə'pɛʔ*, WL-PD *la'p:ɛ*.

to give:

(1) LH-KD #**neʔi**, CL-LK *ne*, CL-LWK *ne*, KD-LW *ni*, WL-AB *nẽ*, WL-AD *'neĩ*, WL-BN *neĩ*, WL-BP *nei*, WL-BT *nei*, WL-DL *nei*, WL-HRW *nei*, WL-KO *nei*, WL-LMK *neʔi*, WL-LWI *'neiŋ*, WL-LWL *nei*, WL-ML *ni*, WL-MS *nij*, WL-PD *neiŋ*, WL-PD *nij*, WL-TJ *nei*, WL-WB *neʔi*, WL-WT *nei*.

(2) FL #**soroŋ**, CL-LP *soro*, CL-LWT *soro*, EL-LE *sorõ*, EL-LMT *sorõ*, KD-LB *sɔɔŋ*, KD-LW *sɔɔŋ nateŋ*, WL-BL *soro*, WL-IA *soro*, WL-LG *sorõ*, WL-LH *soro*, WL-LML *soro*, WL-LTB *sorõ*, WL-LWI *'sɔɔŋ*, WL-MD *sorõ*, WL-PK *sorõ*, WL-RE *sorõ*, WL-WK *sorõ*, WL-WW *sorõ*.

(3) PFL ***bəli**, SK-HEW *bəli*, SK-TA *βəli*.

(4) CL-IL *beŋa*, CL-KK *be*, CL-KK *be:*, CL-LR *beŋ*, CL-PN *beŋ*.

(5) CL-KK *noto*, CL-MR *noto*.

to go:

(1) LH-SK #-**ai**, WL-AB *mei*, WL-AD *'naʔi*, WL-LWI *'naʔi*, WL-MS *kaiʔ*, WL-PD *rahi*.

(2) PFL ***panav** 'walk', CL-KK *'pana*, CL-LR *pa'nav*, KD-LB *pan*, SK-HEW *pano*.

to grasp:

(1) LH #**pəhen** 'grasp; hold', CL-LR *pəhe'naŋ*, WL-AB *paha*, WL-AD *'pehẽ*, WL-LWI *'pəheŋ*, WL-PD *paha*.

(2) WL-MS *kə'kuŋ*, WL-MS *mə'kuŋ*, WL-MS *mə'kuŋ*, WL-MS *mə'kuŋ*, WL-MS *nə'kuŋ*, WL-MS *rə'kuŋ*, WL-MS *tə'kuŋ*.

(3) SK-HEW *pligo*, WL-PD *p'rageŋ*.

to grow:

(1) LH-KD #**tava** 'stem; grow', CL-LR *'tava*, KD-LB *tawe*, WL-AB *tafa*, WL-AD *'tawã*, WL-LWI *'tawa*, WL-MS *tawaŋ*, WL-PD *tawaŋ*.

to hear:

(1) PFL ***dəŋər**, CL-IL *dəŋ*, CL-KK *daŋar*, CL-KK *'dəŋər*, CL-LK *dəŋ*, CL-LP *dəŋ*, CL-LR *'dəŋə*, CL-LWK *daŋər*, CL-LWT *dəŋ*, CL-MR *dəŋə*, CL-PN *dəŋa*, EL-LE *dəŋe*, EL-LMT *dəŋe*, KD-LB *dəŋər*, KD-LW *dəŋer bæiŋ*, SK-HEW *rəna*, SK-TA *rəna*, WL-AB *daŋa*, WL-AD *'dəŋəʔ*, WL-BL *dəŋ*, WL-BM *weŋe*, WL-BN *daŋ:a*, WL-DL *dəŋɔ*, WL-HRW *dəŋɔ*, WL-IA *dəŋa*, WL-KO *dəŋə*, WL-LG *weŋe*, WL-LH *dəŋa*, WL-LMK *dəŋa*, WL-LML *dəŋa*, WL-LWI *'weŋe*, WL-MD *dəŋa*,

WL-ML *dəŋa*, WL-MS *də'ŋa*, WL-PD *dəŋa*, WL-WK *dəŋə*, WL-WT *dəŋa*, WL-WW *weŋe*.

(2) KD-LW *baŋær*, WL-BP *bãi*, WL-BT *bãi*, WL-LTB *bãi*, WL-LWL *bãi*, WL-PK *bãi*, WL-RE *bãi*, WL-TJ *bãi*, WL-WB *bãi*.

(3) SK-HEW *di?in*, SK-HEW *di?iŋ*.

to help:

(1) WL-AB *sambɔ*, WL-MS *sambɔ*, WL-PD *sambɔ*.

(2) LH-KD #*pohiŋ*, CL-KK *'pɔi*, CL-LR *'poɛŋ*, KD-LB *po'hiŋ*.

to hide:

(1) WL-AB *dafu*, WL-MS *də'wuk*, WL-PD *dawu*.

(2) LH #*soroŋ*, CL-KK *so'ɾɔ*, CL-LR *ho'roŋ*, WL-AD *'horõ*, WL-LWI *'hɔɾɔŋ*.

to hit:

(1) WL-AB *bɛhɛ*, WL-BN *bɛhɛ*, WL-PD *bɛhɛ*.

(2) LH #*dani* 'hit (drum)', CL-KK *dani* 'hit drum', WL-AB *danɛ*, WL-MS *danɛ*.

(3) WL-MS *boa?*, WL-PD *boa*.

(4) PFL **paluk*, CL-KK *pa'luk*, KD-LB *pa'lu?*, KD-LW *palu?*.

to hold:

(1) LH #*pehen* 'grasp; hold', CL-LR *pehe'naŋ*, EL-LE *pehe*, EL-LMT *pehe*, WL-AB *paha*, WL-AD *'pehẽ*, WL-BM *pehe*, WL-BN *pa'ha:*, WL-BP *pehe*, WL-BT *pehẽ*, WL-DL *pehẽ*, WL-HRW *pehe*, WL-IA *pehe*, WL-KO *pehẽ*, WL-LG *pehẽ*, WL-LH *pehe*, WL-LMK *pẽhe*, WL-LML *pe*, WL-LTB *pehẽ*, WL-LWI *'peheŋ*, WL-LWL *pehẽ*, WL-ML *pehe*, WL-MS *paha*, WL-PD *paha*, WL-RE *pehẽ*, WL-TJ *pehe*, WL-WB *pehe*, WL-WK *pehe*, WL-WT *pehẽ*, WL-WW *pehẽ*.

(2) LH #*soga*, CL-KK *so'ga*, CL-KK *sogan*, CL-LK *soga*, CL-LWK *soga*, WL-BL *soga*.

(3) CL-IL *pida*, CL-LP *pida*, CL-LWT *pid*, CL-MR *pidɔ*.

(4) CL-PN *wetenəŋ*, WL-MD *wetet*.

(5) SK-MM *doe*, SK-TA *toe*.

(6) WL-MS *gena* 'ekaŋ, WL-PD *gena* 'ekaŋ.

(7) EL-LE *deru*, KD-LB *de'ruŋ*, KD-LW *dəruŋ dəŋ*, WL-LMK *dore*.

to hunt:

(1) LH #*bati*, CL-IL *bati*, CL-KK *ba'ti*, CL-KK *batin*, CL-LK *bati*, CL-LP *bati*, CL-LR *ba'ti*, CL-LWK *bati*, CL-LWT *bati*, WL-AD *'batĩ*, WL-BL *bati*, WL-BP *hewābati*, WL-BT *batĩ*, WL-DL *batij*, WL-HRW *batĩ*, WL-IA *batijən*, WL-KO *bati*, WL-LH *batĩ*, WL-LML *bati*, WL-RE *batĩ*, WL-WT *bat?*, WL-WW *batĩ*.

(2) CL-PN *welok*, EL-LMT *welo*, WL-BM *pəreha*, WL-LG *pəreha*, WL-LWI *pə'reha?*, WL-MD *welok*, WL-ML *welo*.

(3) WL-LTB *lərakə*, WL-PK *ləra?*.

(4) WL-BP *hewābati*, WL-LWL *hewa*, WL-TJ *hewa*, WL-WB *hewā*.

(5) CL-MR *apu*, WL-WK *hapō*.

(6) SK-HEW *rakan*, SK-MM *ʔou rakan*, SK-TA *raka-ŋ*.

(7) CL-LR *'undaŋ*, KD-LB *ʔundaŋ*, WL-MS *ɔndaŋ*, WL-PD *ɔnda*.

(8) CL-KK *'mali*, WL-AD *'marĩ*, WL-LWI *'mariŋ*, WL-PD *mariŋ 'ata*.

(9) CL-KK *sə'dʒan*, SK-HEW *həɾɔn*.

to kill:

(1) WL-AB *bunɔ*, WL-BN *bunɔ*, WL-MS *bunɔ*, WL-PD *bunɔ*.

(2) LH-KD #*belu*, CL-IL *bel*, CL-KK *belu*, CL-LP *belu*, CL-LWK *belu*, EL-LE *belu*, EL-LMT *belu*, KD-LB *bel*, KD-LW *bæloʔ ba?*, KD-LW *bæloʔ behar*, KD-LW *bæloʔ ɔhɔr*, WL-AD *'belo*, WL-BM *belo*, WL-BP *belo*, WL-BT *belo*, WL-DL *belo*, WL-HRW *belo*, WL-IA *belo*, WL-KO *belo*, WL-LG *belo*, WL-LH *belo*, WL-LMK *belo*, WL-LWI *'belo*, WL-LWL *belo*, WL-PK *belo*, WL-RE *belo*, WL-TJ *belo*, WL-WB *belo*, WL-WT *belo*, WL-WW *belo*.

(3) CL-LK *təbadʒak*, CL-LWT *təbadʒak*, CL-MR *təbadʒak*, WL-BL *təbadʒak*, WL-LML *təbadʒak*, WL-WK *təbadʒak*.

(4) CL-KK *-ar ma'tadʒa*, CL-LR *mə'raŋ na 'matei*, CL-PN *həbal*, WL-ML *səbəle*.

(5) SK-HEW *pati*, SK-TA *pati mate*.

(6) SK-HEW *robak*, SK-MM *roba*.

to know:

(1) CL-IL *kin*, CL-IL *min*, CL-LP *kin*, CL-LP *min*, CL-LWT *toi*, CL-MR *tina*, EL-LMT *moi*, EL-LMT *ʔoi*, KD-LW *ui*, WL-AB *koiŋ*, WL-AD *'koi*, WL-AD *'noi*, WL-BL *toi*, WL-BM *koi*, WL-BM *moi*, WL-BP *koi*, WL-BP *moi*, WL-BT *koi*, WL-BT *moi*, WL-DL *koi*, WL-DL *moi*, WL-HRW *koi*, WL-HRW *moi*, WL-IA *toi*, WL-KO *koi*, WL-KO *moi*, WL-LG *koi*, WL-LG *moi*, WL-LH *koi*, WL-LMK *koi*, WL-LMK *moi*, WL-LML *toi*, WL-LTB *koi*, WL-LTB *moi*, WL-LWI *nɔi*, WL-LWI *tɔi*,

WL-LWL *koi*, WL-LWL *moi*, WL-MD *koi*, WL-ML *toi*, WL-MS *koiŋ*, WL-PD *koiŋ*, WL-PD *moiŋ*, WL-PD *noiŋ*, WL-PK *koi*, WL-PK *moi*, WL-RE *koi*, WL-RE *moi*, WL-TJ *koi*, WL-TJ *moi*, WL-WB *koi*, WL-WB *moi*, WL-WK *koi*, WL-WK *moi*, WL-WT *koi*, WL-WT *moi*, WL-WW *koi*, WL-WW *moi*.

(2) LH-SK **#(raʔi) -etən**, CL-KK *-etən*, CL-KK *ketən*, CL-LK *metəna*, CL-LR *meʔtən*, CL-LWK *ketən*, CL-PN *ketan*, CL-PN *metan*, SK-HEW *raitan*, SK-HEW *raʔitan*, SK-MM *raʔinʔan*, SK-TA *raʔinta-ŋ*.

to laugh:

(1) LH **#gekay**, CL-IL *gekaj*, CL-KK *geʔka*, CL-KK *gekadʒ*, CL-LK *geka*, CL-LP *gekaj*, CL-LR *geʔkai*, CL-LWK *geka*, CL-LWT *gekadʒ*, CL-MR *gekadʒɤ*, CL-PN *gekaj*, EL-LE *geʔa*, EL-LMT *geʔa*, WL-AB *geki*, WL-AD *ʔgeka*, WL-BL *gekadʒ*, WL-BM *geka*, WL-BN *geki*, WL-BP *geka*, WL-BT *geka*, WL-DL *geka*, WL-HRW *geka*, WL-IA *gekat*, WL-KO *geka*, WL-LG *geka*, WL-LH *gekaj*, WL-LMK *geka*, WL-LML *geka*, WL-LTB *geka*, WL-LWI *ʔgeka*, WL-LWL *geka*, WL-MD *geka*, WL-ML *geka*, WL-MS *geki*, WL-MS *gekiŋ*, WL-MS *gekiʔ*, WL-PD *geke*, WL-PK *gekat*, WL-RE *geka*, WL-TJ *geka*, WL-WB *geka*, WL-WK *gekadʒ*, WL-WT *geka*, WL-WW *geka*.

(2) PFL ***tave**, KD-LB *tawe*, KD-LW *tave hekɔ*, SK-HEW *to:*, SK-MM *to:*, SK-TA *to:*.

to let go:

(1) LH **#loka**, CL-KK *ʔloka*, CL-LR *ʔlokaŋ*, WL-AD *ʔloʔok*, WL-LWI *ʔloʔo*.

(2) LH **#loit**, CL-KK *ʔoit*, WL-LWI *ʔio*.

to lie down:

(1) SK-HEW *duʔe labe*, SK-HEW *tuʔe*, SK-TA *tuʔə*, WL-BN *turu*, WL-LWI *ʔturu*, WL-MD *turu*, WL-PD *turu*, WL-WB *tobāʔturu*.

(2) CL-LWK *gəle*, CL-PN *galek*, WL-AB *pɔna* *ʔpɔna*, WL-MS *pɔna*, WL-PD *pɔna* *ʔaleŋ*.

(3) WL-BM *pola*, WL-KO *pola*, WL-LG *polar*.

(4) LH **#pawaŋ** **ʔlie (position for things)**, CL-IL *pafaŋa*, CL-KK *paʔva*, CL-KK *pawan*, CL-LK *pawa*, CL-LWT *pafaŋa*, WL-BL *pawaŋ*, WL-LML *pafã*, WL-ML *pawaŋ*.

(5) WL-DL *pədenək*, WL-LMK *ledã*, WL-LTB *ledãna*, WL-WW *pədenək*.

(6) CL-LR *ʔeva*, EL-LE *ewa*, EL-LMT *ewã*, WL-IA *ewan*, WL-LH *ewa*.

(7) CL-LP *toba*, CL-MR *tobana*, WL-WB *tobāʔturu*, WL-WK *toba*.

(8) WL-RE *gola*, WL-TJ *epa gola*.

to open:

(1) PFL **bukat*, CL-KK *bukat*, CL-LR *bukə'təŋ*, KD-LB *bu'kaʔ*, WL-AB *buka*, WL-AD *'buka*, WL-LWI *'bukaʔ*.

(2) WL-MS *leka*, WL-PD *leka*.

to plant rice:

(1) PFL **mula* 'plant', CL-KK *mul*, CL-LR *'mula kə'luok*, WL-AB *mula 'pari*, WL-AD *'mula*, WL-LWI *mula 'tahaʔaŋ*, WL-MS *mula 'pare*, WL-PD *mula 'pare*.

to plant yam:

(1) PFL **mula* 'plant', CL-KK *mul*, CL-LR *'mula 'hurar*, SK-HEW *mula*, WL-AB *mula 'kura 'dʒafa*, WL-LWI *mula 'uwe*, WL-MS *mula 'kura*, WL-PD *mula 'kura*.

(2) LH-SK #*pahat*, SK-HEW *pahe*, SK-HEW *paʔat*, WL-AD *'pahaʔ*.

to play:

(1) WL-AB *ka'naku*, WL-BM *gənaku*, WL-BN *ka'noku*, WL-BP *gənaku*, WL-LG *gənaku*, WL-LWI *gənə'ku*, WL-LWL *gənaku*, WL-MS *knə'ku*, WL-PD *k'naku*, WL-RE *gənəkuk*, WL-TJ *gənaku*, WL-WB *gənaku*.

(2) KD-LB *huan*, KD-LW *huan æbæl*.

(3) LH #*glasa*, CL-IL *gəlaŋ*, CL-KK *gəlasa*, CL-KK *'glasa*, CL-LK *gəlasa*, CL-LP *gəlahan*, CL-LWK *gəlasa*, CL-LWT *gəlah*, WL-AD *g'lahəʔ*, WL-BL *gəla*, WL-BT *gəlahə*, WL-DL *gəlahə*, WL-HRW *gəlahə*, WL-KO *gəlahə*, WL-LH *gəlahə*, WL-LMK *gəlahə*, WL-LML *gəla*, WL-ML *gəlahə*, WL-WK *gəlahə*, WL-WT *gəlahə*, WL-WW *gəlahə*.

(4) CL-LR *kə'raŋa*, CL-PN *kəraŋa*.

(5) SK-HEW *ləbe*, SK-HEW *ləbeʔ*, SK-TA *ʔləbe*.

to pound:

(1) PFL **bayu*, CL-KK *'badʒu*, CL-LR *ba'dʒo*, KD-LB *bae*, SK-HEW *bai*, WL-AB *badʒo*, WL-AD *'badʒo*, WL-LWI *'bajo*, WL-MS *badʒo*, WL-PD *badʒo*.

to promise:

(1) LH #*nakin*, CL-LR *nə'kaŋ*, WL-AB *nakin*, WL-PD *nak:iŋ*.

(2) KD-LB *d̥ʒand̥ʒi*, WL-LWI *'d̥ʒand̥ʒi*. From Indonesian/Malay *janji* 'promise'.

(3) LH #*baya*, WL-AD *'baja*.

to pull:

(1) WL-AB *tar:ε*, WL-BN *tar:ε*, WL-MS *tε'r:ε?*, WL-PD *ta'r:ε*.

(2) LH #*vidu*, CL-IL *fido*, CL-KK *'vidu*, CL-LP *fidu*, EL-LE *ido*, WL-DL *wada*, WL-PD *wider*, WL-RE *widu*.

(3) KD-LB *be?*, KD-LW *dera beke*, KD-LW *be?*.

(4) WL-AD *'gehã*, WL-BM *gehẽ*, WL-BP *gehã*, WL-HRW *gehã*, WL-LG *gehã*, WL-LTB *geha?*, WL-LWI *'gehaŋ*, WL-MD *gihã*, WL-PK *gehã*, WL-TJ *gehe*, WL-WB *gehã*, WL-WW *geã*.

(5) LH-SK #*tubu*, CL-MR *tubu*, SK-TA *toβa*, WL-AD *'tubo*, WL-BL *tubo*, WL-BT *tubo*, WL-IA *tubo*, WL-KO *tubo*, WL-LH *tubo*, WL-LMK *tubo*, WL-ML *tubo*, WL-WK *tubo*, WL-WT *tubo*.

(6) CL-KK *tugu*, CL-LK *tugu*, CL-LR *'tugo*, CL-LWK *tugu*, CL-PN *tugo*.

(7) CL-LWT *gəfea*, WL-LML *gəfea*.

(8) LH-SK #*gide*, EL-LMT *gidə*, SK-HEW *ide*, SK-TA *gide*.

to push:

(1) KD-LB *obaŋ*, KD-LW *ɔbaŋ*, WL-AB *tobaŋ*, WL-BN *tobaŋ*, WL-MS *tobaŋ*, WL-PD *tobaŋ*.

(2) KD-LW *uru?*, WL-AB *uruŋ*.

(3) WL-AD *'tojõ*, WL-BT *tojõ*, WL-HRW *tojõ*, WL-LWI *'tojo*, WL-LWL *tođʒõ*, WL-WB *tođʒõ*, WL-WW *tojõ*.

(4) LH-KD #*uduk/uruk*, CL-IL *uduka*, CL-KK *uduk*, CL-KK *u'duk*, CL-LK *uduk*, CL-LR *'uduk*, CL-LWK *uduk*, CL-MR *udur*, CL-PN *udukaŋ*, EL-LE *udu*, EL-LMT *udũ*, WL-BL *uduk*, WL-BM *odo*, WL-BP *odo*, WL-LG *odo*, WL-LH *uduk*, WL-LML *odo*, WL-LML *uruk*, WL-LTB *odo*, WL-MD *uduk*, WL-ML *udu*, WL-TJ *odo*, WL-WK *udur*.

(5) SK-HEW *sogor*, SK-HEW *sogor*, WL-DL *osõ*, WL-KO *osõ*, WL-PK *sogor*, WL-WT *osõ*.

(6) CL-LP *ogor*, CL-LWT *bogor*, WL-RE *bogõ*.

(7) CL-LR *'gehiŋ*, WL-AD *ge'hik*, WL-LWI *'gehi?*, WL-MS *ɔno* *'gehiŋ*, WL-PD *gehiŋ*.

to return:

(1) PFL **gəvalik*, SK-HEW *gəvali*, WL-BP *gəwali*, WL-IA *balik*, WL-LMK *balok*, WL-LTB *gəwali*, WL-RE *gəwalik*, WL-WT *balok*, WL-WW *balok*.

(2) LH #*tue*, CL-IL *tue*, CL-KK *tue*, CL-LK *tue*, CL-LP *tue*, CL-LWK *tue*, CL-LWT *tuel*, CL-MR *tue*, CL-PN *tue*, EL-LMT *tue*, WL-BL *tue*, WL-BT *tue*, WL-DL *tuẽ*, WL-HRW *tuẽ*, WL-LH *tue*, WL-LML *tue*, WL-LWL *tuse*, WL-MD *tue*, WL-ML *tue*, WL-PK *tue?*, WL-WB *tue*, WL-WK *tue*.

(3) EL-LE *peko*, WL-KO *pekot*, WL-LG *pekot*.

to rub:

(1) WL-AB *dəhɔ*, WL-BN *də'hɔ*, WL-MS *dəhɔ?*, WL-PD *dəhɔ*.

(2) CL-LR *poho'kəŋ*, KD-LB *pə'hɔ?*.

(3) LH-KD #*doru* 'rub; wipe', CL-KK *doru*, CL-KK 'doru, CL-LK *doru*, CL-LWK *doru*, CL-PN *dore*, EL-LE *doru*, EL-LMT *doru*, KD-LW *doru*, WL-AD *do'ru?*, WL-BL *doru*, WL-BM *doru*, WL-BP *doru*, WL-BT *doruk*, WL-DL *doruk*, WL-HRW *doruk*, WL-IA *doru*, WL-LG *doruk*, WL-LH *doru*, WL-LML *doru*, WL-LTB *doru*, WL-LWI 'doru?', WL-LWL *doru*, WL-MD *doru*, WL-ML *doru*, WL-PK *doru*, WL-TJ *doru*, WL-WW *doru*.

(4) CL-IL *odo*, CL-LP *godonj*, CL-LWT *odo*, CL-MR *godo*, WL-WK *godo*.

(5) WL-KO *pusak*, WL-RE *puʔũ*.

(6) SK-HEW *berosok*, SK-HEW *blosok*, SK-TA *bloso*, SK-TA *boho*, SK-TA *moho*, WL-LMK *oso*, WL-WB *rosuk*, WL-WT *oso*.

to rule:

(1) WL-AB *gahij*, WL-MS *sɔrak 'sɔgaj*.

to run:

(1) PFL **plari/kari*, CL-KK 'kari, CL-LR *kar*, SK-HEW *plari*, WL-AB *palae*, WL-AD *pa'laʔe*, WL-BN *pa'lae*, WL-LWI *pə'laʔe*, WL-MS *p'laeŋ*, WL-PD *p'lae*.

to sail:

(1) LH #*bua*, CL-LR 'bua 'laja, WL-AB *bua*, WL-AD *bua:*, WL-LWI *bua 'təna*, WL-MS *bua?*, WL-PD *bua*.

to say:

(1) LH-SK #*madi*, CL-LK *maik*, CL-LWK *maik*, CL-LWT *madʒi*, SK-HEW *marĩ*, SK-HEW *marin*, WL-BL *mari*, WL-BM *marĩ*, WL-BN *ma'riũ*, WL-BP *mari*,

WL-BT *marĩ*, WL-DL *marĩ*, WL-HRW *marĩ*, WL-KO *mari*, WL-LG *marĩ*, WL-LH *mari*, WL-LMK *mari*, WL-LML *mari*, WL-LTB *mari*, WL-LWL *marĩ*, WL-ML *mxi*, WL-MS *mariŋ*, WL-PD *mariŋ*, WL-RE *marĩ*, WL-TJ *mari*, WL-WB *marĩ*, WL-WT *marĩ*, WL-WW *marĩ*.

(2) KD-LB *tutu?*, KD-LB *tutu?* 'nanaŋ, SK-MM *tuṭur*, SK-TA *tutur bəta*, WL-AB *tutu*, WL-LWI 'tutu?, WL-WK *tutu*.

(3) WL-IA *koda*, WL-LWI 'kɔda.

(4) CL-IL *gas*, CL-MR *gasi*, WL-MD *gasip*.

(5) CL-KK *pə'nua*, CL-LR *pə'nua*.

(6) CL-KK *tədək*, CL-LP *tədək*, CL-LR *tə'dək*, EL-LE *təde*, EL-LMT *təde*.

(7) SK-HEW *həron*, SK-TA *həro-ŋ*, WL-PK *hərō*.

(8) SK-TA *tutur bəta*, SK-TA *βeta*.

to scratch:

(1) PFL **garu*, CL-IL *rag*, CL-KK *kəragu*, CL-KK *kə'ragu*, CL-LK *ragu*, CL-LP *ragu*, CL-LR *kə'ragum*, CL-LWK *kəragum*, CL-LWT *ragu*, CL-MR *ragu*, CL-PN *kərag*, EL-LE *ragu*, EL-LMT *ragu*, KD-LB *karɔ*, KD-LW *karɔ*, SK-HEW *garu*, SK-HEW *ʔaro*, SK-MM *ʔaro*, SK-TA *kəru*, WL-AB *gou*, WL-AD 'ragu?, WL-BL *ragu*, WL-BM *ragu*, WL-BN *gaɔ*, WL-BP *ragu*, WL-BT *ragu*, WL-DL *ragu*, WL-HRW *ragu*, WL-IA *gau*, WL-KO *ragu*, WL-LG *ragu*, WL-LH *ragu*, WL-LMK *ragu*, WL-LML *rago*, WL-LTB *ragu*, WL-LWI 'ragu?, WL-LWL *ragu*, WL-MD *ragu*, WL-ML *ragu*, WL-MS *gau*, WL-PD *gau*, WL-PK *gaʔu*, WL-RE *ragu*, WL-TJ *ragu*, WL-WB *ragu*, WL-WK *ragu*, WL-WT *garu*, WL-WW *garuk*.

to search, to look for, to hunt for:

(1) WL-AB *gena*, WL-BN *gena*, WL-MS *gena*, WL-PD *gena*.

(2) KD-LB *ha'ba*, WL-AD 'sɜba?, WL-LWI 'seba?.

to see:

(1) CL-IL *sure*, KD-LW *bɔra? hera?*, WL-AB *seru*, WL-BN *seru*, WL-MS *seru?*, WL-PD *seru*.

(2) KD-LB *nui*, KD-LW *ui*, WL-BL *koi*, WL-TJ *koi*, WL-WB *koi*.

(3) KD-LW *bɔra? hera?*, KD-LW *ɛbeŋ bɔra?*.

(4) WL-AD 'hulɜ, WL-BT *hulǎ*, WL-DL *hulǎ*, WL-HRW *hulǎ*, WL-KO *hulə*, WL-LH *hulǎ*, WL-WK *hule*, WL-WW *hulǎ*.

(5) WL-BM *təda*, WL-BP *təda*, WL-LG *təda*, WL-LMK *təda*, WL-LWL *təda*, WL-RE *təda*, WL-WT *təda?*.

- (6) CL-KK *du*, CL-LK *du*.
- (7) CL-LWK *bele*, CL-MR *bəlele*, CL-PN *bəlele*, EL-LMT *bəlele*, WL-MD *bəlele*, WL-ML *bele*, WL-PK *bələr*.
- (8) SK-HEW *tejer*, SK-HEW *təjer*, WL-LWI *təjə*.
- (9) CL-LP *tene*, CL-LWT *tene*.
- (10) SK-HEW *ʔitaʔ*, SK-TA *ʔita*.
- (11) SK-MM *niʔa*, SK-TA *niʔa*.

to sell:

- (1) WL-AB *pa'nahanj*, WL-MS *p'nahanj*, WL-PD *p'nahanj*.
- (2) CL-KK *du'rum*, CL-LR *du'rum*, KD-LB *du'ruŋ*, WL-AD *'duʔũ*, WL-LWI *'duʔuŋ*.

to sew:

- (1) LH-KD *#saur*, CL-IL *saur*, CL-KK *saur*, CL-LK *saur*, CL-LP *haur*, CL-LR *hau'rəŋ*, CL-LWK *saur*, CL-LWT *haur*, CL-MR *saur*, CL-PN *hau*, EL-LE *heu*, EL-LMT *heu*, KD-LB *he'woŋ*, WL-AD *'hau*, WL-BL *haur*, WL-BM *hău*, WL-BN *haur*, WL-BP *hawu*, WL-BT *hau*, WL-DL *hau*, WL-HRW *hau*, WL-IA *hau*, WL-KO *hau*, WL-LG *hawurã*, WL-LH *haur*, WL-LMK *hawu*, WL-LML *hau*, WL-LTB *hau*, WL-LWL *hawu*, WL-MD *haur*, WL-ML *hau*, WL-MS *haur*, WL-PD *haur*, WL-PK *hau*, WL-RE *hau*, WL-WB *hau*, WL-WK *haur*, WL-WT *hau*, WL-WW *hau*.
- (2) WL-AB *həru 'apa*, WL-PD *horu*.
- (3) PFL **daru*, WL-LWI *'dauŋ*, WL-TJ *dau*.
- (4) PFL **daʔit*, SK-HEW *raʔit*, SK-HEW *raʔit*, SK-MM *raʔit̚*, SK-TA *raʔit*.

to shine:

- (1) WL-AB *saru*, WL-MS *sə'ruʔ*.
- (2) KD-LB *ɔlɔr*, KD-LB *duʔ 'ɔlɔr*, WL-AB *bɔlo*, WL-MS *bɔlɔr*, WL-PD *bolɔr*.
- (3) WL-MS *tutuŋ*, WL-PD *tutuŋ*.
- (4) CL-KK *va'ra*, SK-HEW *baraʔ*.
- (5) CL-KK *va'ləka*, CL-LR *valə'kən*.
- (6) CL-LR *'talek 'kay*, CL-LR *'talek*.
- (7) CL-LR *'hodiŋ*, WL-LWI *'hode neʔa*, WL-LWI *'hode*.
- (8) WL-AB *tutɛ*, WL-MS *tutɛ*, WL-PD *tutɛ*.

to shoo, to chase away, to send away:

(1) PFL *sika, CL-KK *si'ka*, WL-AD *'sika*.

(2) SK-HEW *bəve*, SK-HEW *pəve*.

to shoot:

(1) FL #pasak, CL-IL *pasak*, CL-KK *pasak*, CL-KK *pa'sak*, CL-LK *pasak*, CL-LP *pasak*, CL-LR *pa'hak*, CL-LWK *pasak*, CL-LWT *pasak*, CL-MR *pasak*, CL-PN *pahak*, EL-LE *pasa*, EL-LMT *pasa*, SK-HEW *pasak*, WL-AB *pasa*, WL-AD *'pasak*, WL-BL *pasak*, WL-BM *pasak*, WL-BN *pasa*, WL-BP *pasak*, WL-BT *pasak*, WL-DL *pasak*, WL-HRW *pasak*, WL-IA *pasak*, WL-KO *pasak*, WL-LG *pasak*, WL-LH *pasa*, WL-LMK *pasak*, WL-LML *pasa*, WL-LTB *pasa*, WL-LWI *'pasak*, WL-LWL *pasak*, WL-MD *pasak*, WL-ML *pasa*, WL-MS *pasa?*, WL-PD *pasa*, WL-PK *pasak*, WL-RE *pasak*, WL-TJ *pasak*, WL-WB *pasak*, WL-WK *pasak*, WL-WT *pasak*, WL-WW *pasak*.

to sing:

(1) WL-AB *lian*, WL-BN *bɔtɛ 'lian*, WL-MS *kan 'lian*, WL-MS *nan 'lian*, WL-PD *kan 'lian*.

(2) SK-HEW *de:daŋ*, WL-AD *'dedã*.

(3) WL-BM *nani*, WL-BP *nani*, WL-KO *nani*, WL-LG *nani*, WL-LMK *nani*, WL-LTB *nani*, WL-LWI *'njanji*, WL-LWL *nani*, WL-MD *nani*, WL-ML *nani*, WL-PK *nani*, WL-TJ *nani*, WL-WT *nani*. From Indonesian/Malay *nyanyi* 'sing'.

(4) CL-IL *kantar*, CL-KK *kantar*, CL-KK *kan'tar*, CL-LK *kantar*, CL-LP *kantar*, CL-LR *kan'tar*, CL-LWK *kantar*, CL-LWT *kantar*, CL-MR *kantar*, CL-PN *kantar*, EL-LE *kata*, EL-LMT *katərã*, SK-HEW *kantar*, SK-TA *kantar*, WL-BL *kantar*, WL-BT *kantar*, WL-DL *kantar*, WL-HRW *kantar*, WL-IA *kantar*, WL-LH *kantar*, WL-LML *kantar*, WL-LWI *'kantar*, WL-RE *kantar*, WL-WB *kantar*, WL-WK *kantar*, WL-WW *kantar*. From Portuguese *cantar* 'sing'.

to sit:

(1) LH-KD #tobe, CL-IL *tobe*, CL-KK *tobe*, CL-KK *'tobe*, CL-LK *tobe*, CL-LP *tobo*, CL-LR *'tobe*, CL-LWK *tobe*, CL-LWT *tobe*, CL-MR *tobe*, CL-PN *tobe*, EL-LE *tobo*, EL-LMT *tobo*, WL-AB *tobɔ*, WL-AD *'tobo*, WL-BL *tobo*, WL-BM *tobo*, WL-BN *tobɔ*, WL-BP *tobo*, WL-BT *tobo*, WL-DL *tobo*, WL-HRW *tobo*, WL-IA *tobo*, WL-KO *tobo*, WL-LG *tobo*, WL-LH *tobo*, WL-LMK *tobo*, WL-LML *tobo*, WL-LTB *tobo*, WL-LWI *'tobɔ*, WL-LWL *tobo*, WL-MD *tobo*, WL-ML *tobo*, WL-MS *tobɔ*, WL-PD *tobɔ*, WL-PK *tobo*, WL-RE *tobo*, WL-TJ *tobo*, WL-WB *tobo*,

WL-WK *tobo*, WL-WT *tobo*, WL-WW *tobo*.

(2) KD-LB *te'beʔ*, KD-LW *tɛbɛʔ*.

(3) SK-HEW *dəri*, SK-HEW *təri*, SK-MM *təri*, SK-TA *təri*.

to skewer:

(1) KD-LB *tuhuʔ*, SK-HEW *tuhuk*, WL-AB *tuhu*, WL-AD *'tuhuk*, WL-MS *tuhuʔ*, WL-PD *tuhu*.

(2) LH-KD *#təguʔ*, CL-KK *tə'gu*, KD-LW *tukoʔ* 'stab', WL-AD *'təguʔ*.

to slap:

(1) KD-LB *le'paʔ*, WL-AB *lapa*, WL-MS *lə'pa*, WL-PD *lə'p:a* 'kapuŋ, WL-PD *lə'p:a*.

(2) CL-KK *tə'pa*, CL-LR *tə'pa*, WL-AD *'təpak*, WL-AD *'təpɔ̃*, WL-LWI *'təpa*.

to sleep:

(1) PFL **tudu*, EL-LE *turu*, EL-LMT *turʔo*, WL-AB *turu*, WL-AD *'turu*, WL-BL *turu*, WL-BM *turu*, WL-BN *turu*, WL-BP *turu*, WL-BT *turu*, WL-DL *turu*, WL-HRW *turuk*, WL-IA *turu*, WL-KO *turu*, WL-LG *turu*, WL-LH *turu*, WL-LMK *turu*, WL-LML *turu*, WL-LTB *turu*, WL-LWI *'turu*, WL-LWL *turu*, WL-MD *turu*, WL-ML *turu*, WL-MS *turu*, WL-PD *turu*, WL-PK *turu*, WL-RE *turu*, WL-TJ *turu*, WL-WB *turu*, WL-WK *turu*, WL-WT *turu*, WL-WW *turu*.

(2) LH-SK *#tuʔay*, CL-KK *ktuadz*, KD-LB *tɛʔɛl*, KD-LW *təʔæl tɛbɛʔ*, SK-HEW *duʔe*, SK-HEW *tuʔe*, SK-MM *tuʔe*, SK-TA *tuʔə*.

(3) CL-IL *gəle*, CL-KK *gəle*, CL-KK *gə'le*, CL-LK *gəle*, CL-LP *gəle*, CL-LR *gə'le*, CL-LWK *gəle*, CL-LWT *gəle*, CL-MR *gəle*, CL-PN *galek*.

to smell:

(1) LH-SK *#si(n)on*, SK-TA *sino-ŋ*, WL-BM *siõ*, WL-BN *siõ*, WL-BP *siõ*, WL-LG *siõ*, WL-LMK *sio*, WL-LTB *siõ*, WL-LWI *'siõŋ*, WL-LWL *siõ*, WL-PK *siõ*, WL-RE *siõ*, WL-WB *siõ*, WL-WT *sioro*, WL-WW *sio*.

(2) LH-KD *#(m)are*, EL-LMT *are*, WL-AB *mare*, WL-PD *mare*.

(3) KD-LB *naɛʔ*, KD-LW *naɛʔ pɔvɔn*.

(4) WL-AD *pənu'huk*, WL-BT *pənuk*, WL-DL *pənuhuk*, WL-HRW *pənuhuk*, WL-IA *nuhuk*, WL-KO *pənuhuk*, WL-LML *pənukã*.

(5) LH *#sodam*, CL-IL *sodam*, CL-KK *sɔ'dam*, CL-KK *sodam*, CL-LK *sodam*, CL-LP *hodam*, CL-LR *ho'dam*, CL-LWK *sodam*, CL-LWT *hodam*, CL-MR *sodam*, CL-PN *hodam*, EL-LE *hodã*, WL-BL *hodam*, WL-LH *hodã*, WL-MD *hodã*, WL-

ML *hodɤ*, WL-WK *hodã*.

(6) SK-HEW *movo*, SK-HEW *movoʔ*, SK-MM *movo*.

to speak:

(1) CL-KK *pətu'tosa*, KD-LB *tutuʔ*, KD-LB *tutuʔ 'nanaŋ*, WL-AB *tutu*, WL-BN *tutu*, WL-LWI *'tutuʔ*, WL-MS *tutu 'apa*, WL-MS *tutu*, WL-PD *tutu*.

(2) SK-HEW *ulun*, WL-AD *pɜ'ʔulũ*.

to spit:

(1) SK-TA *ilur*, SK-TA *nilur*, WL-AB *uli*, WL-BN *buhzu 'ill:u*, WL-MS *ilu*, WL-PD *pitɔŋ 'ilu*.

(2) WL-MS *pitɔŋ*, WL-PD *pitɔŋ*, WL-PD *pitɔŋ 'ilu*.

(3) PFL ***m-potaŋ**, KD-LB *pɔ'taʔ*, KD-LW *pɔtaʔ mijuʔ*, SK-TA *bota-ŋ*, WL-TJ *putã*.

(4) LH-KD **#təmidu**, CL-IL *təmi*, CL-KK *təmidʒu*, CL-KK *tə'midʒu*, CL-LK *təmidʒu*, CL-LP *təmidʒu*, CL-LR *tə'midʒɔ*, CL-LWK *təmidʒu*, CL-LWT *təmidʒu*, CL-MR *təmidʒu*, CL-PN *təmidʒ*, KD-LW *ijuʔ*, KD-LW *mijuʔ*, KD-LW *pɔtaʔ mijuʔ*, WL-BL *təmidʒo*, WL-LML *təmiro*, WL-WK *təmidʒo*.

(5) LH **#pəriŋo**, EL-LE *piru*, EL-LMT *piru*, WL-AD *'prino*, WL-BM *pəriŋo*, WL-BP *pəriŋo*, WL-BT *pəriŋo*, WL-DL *pəriŋo*, WL-HRW *pəriŋo*, WL-IA *pino*, WL-KO *pəriŋo*, WL-LG *pəriŋo*, WL-LH *priŋo*, WL-LMK *pəriŋo*, WL-LTB *pino*, WL-LWI *pə'riŋo*, WL-LWL *pəriŋo*, WL-MD *pəriŋo*, WL-ML *pəriŋo*, WL-PK *pino*, WL-RE *pəriŋo*, WL-WB *pəriŋo*, WL-WT *pəriŋo*, WL-WW *pəriŋo*.

(6) SK-HEW *niru*, SK-MM *niru*.

to split:

(1) LH-KD **#beta**, KD-LB *bɛ'tɛ*, KD-LW *baʔ*, WL-AB *bat:a*, WL-BN *bat:a*, WL-IA *bata*, WL-PD *bə't:a*.

(2) WL-MS *peaʔ*, WL-PD *pea*.

(3) CL-LR *lə'kaŋ*, SK-HEW *pleka*, SK-HEW *plika*, SK-MM *ləka*, SK-TA *kəlak*, WL-AD *'lɜkaʔ*, WL-BL *ləka*, WL-LMK *ləka*, WL-LML *ləka*, WL-LWI *'likaʔ*, WL-WK *ləka*, WL-WT *ləka*, WL-WW *ləka*.

(4) WL-BM *tika*, WL-LG *tika*, WL-TJ *tika*.

(5) WL-BT *hiwək*, WL-DL *hiwək*, WL-HRW *hiwak*, WL-KO *hiwək*.

(6) CL-IL *legata*, CL-KK *ləga*, CL-KK *lə'ga*, CL-LK *ləga*, CL-LP *ləgat*, CL-LWK *ləga*, CL-LWT *ləgat*, CL-MR *legat*, CL-PN *ləga*, EL-LE *ləgo*, EL-LMT *ləgə*, KD-LW *likɔʔ mapəʔ*, WL-LTB *ləga*, WL-MD *ləga*.

(7) WL-LH *gika*, WL-ML *gika*, WL-WB *giah*.

(8) WL-BP *bəlah*, WL-LWL *sidak*, WL-PK *bəla*, WL-RE *bəla*.

to squeeze:

(1) LH-SK #*pəmek*, SK-MM *pəme*, WL-AB *pəme*, WL-MS *pəme?*, WL-PD *pəme*.

(2) CL-LP *ikut*, KD-LB *iku?*, WL-BP *ikut*, WL-LMK *iku*, WL-LWL *ikut*, WL-TJ *ikut*.

(3) LH #*piʔuk*, CL-IL *piuka*, CL-KK *piuk*, CL-KK *'piuk*, CL-LK *piuk*, CL-LR *'piuŋ*, CL-LWK *piuk*, CL-LWT *piuk*, CL-PN *piuŋ*, EL-LE *piu*, EL-LMT *piu?*, WL-AD *'piʔuk*, WL-BL *piuk*, WL-BM *piʔuk*, WL-BT *peʔut*, WL-DL *piʔuk*, WL-HRW *piʔuk*, WL-IA *piuk*, WL-KO *piʔuk*, WL-LG *piʔu*, WL-LH *piuk*, WL-LML *piuk*, WL-LTB *piʔu*, WL-LWI *'piʔu*, WL-MD *piuk*, WL-ML *piu*, WL-PK *piʔuk*, WL-RE *piʔuk*, WL-WB *piʔuk*, WL-WK *piʔuk*, WL-WW *piʔuk*.

(4) SK-HEW *bəra*, SK-HEW *pəra*, SK-TA *pəra*, WL-WT *pərəs*.

to stab:

(1) WL-AB *sika*, WL-BN *sika*, WL-MS *sika?*, WL-PD *sika*.

(2) KD-LB *tuhu?*, KD-LW *tuhu?* *bakɔ?*, WL-PD *tuhu*.

(3) KD-LW *bəlo?* *ba?*, KD-LW *bəlo?* *bəhar*, KD-LW *bəlo?* *ɔhɔr*.

(4) LH #*səgat*, CL-LR *həgə'təŋ*, CL-PN *həgat*, WL-AD *'səgat*.

(5) LH-KD #*tubak*, CL-IL *tubak*, CL-KK *tubak*, CL-KK *tu'bak*, CL-LP *tubak*, CL-LWK *tubak*, CL-LWT *tubak*, CL-MR *tubak*, EL-LE *tuba*, EL-LMT *tuba*, WL-BL *tubak*, WL-BM *tubak*, WL-BP *tubak*, WL-BT *tubak*, WL-DL *tubak*, WL-HRW *tubak*, WL-IA *tuba*, WL-KO *tubak*, WL-LG *tubak*, WL-LH *tubak*, WL-LMK *tubak*, WL-LML *tuba*, WL-LTB *tuba*, WL-LWL *tubak*, WL-MD *tubak*, WL-ML *tuba*, WL-PK *tubak*, WL-RE *tubak*, WL-TJ *tubak*, WL-WB *tubak*, WL-WK *tubak*, WL-WT *tubak*, WL-WW *tubak*.

(6) SK-TA *təgu*, WL-LWI *'dəgo?*.

(7) LH-SK #*robak*, SK-HEW *robak*, SK-TA *robak*, SK-TA *rohuk*, WL-LWL *robak*.

to stand:

(1) WL-AB *tide*, WL-BN *tid:ɛ*, WL-MS *tide*, WL-PD *tide*.

(2) KD-LB *mader*, KD-LW *mader* *hədaŋ*, KD-LW *mader* *təbe?*.

(3) PFL **diri*, CL-IL *dir*, CL-KK *diri*, CL-KK *'diri*, CL-LK *diri*, CL-LP *diri*, CL-LR *dir*, CL-LWK *diri*, CL-LWT *diri*, CL-MR *diri*, CL-PN *dir*, EL-LE *diri*, EL-

LMT *diri*, WL-AD *'de?i*, WL-BL *dei*, WL-BM *de?i*, WL-BP *de?i*, WL-BT *de?i*, WL-DL *de?i*, WL-HRW *dei*, WL-IA *dei*, WL-KO *de?i*, WL-LG *de?i*, WL-LH *de?i*, WL-LMK *de?i*, WL-LML *dei*, WL-LTB *de?i*, WL-LWI *de'iy*, WL-LWL *de?i*, WL-MD *de?i*, WL-ML *dei*, WL-PK *de?i*, WL-RE *de?i*, WL-TJ *de?i*, WL-WB *de?i*, WL-WK *de?i*, WL-WT *de?i*, WL-WW *de?i*.

(4) PFL **kəda*, SK-HEW *gəra*, SK-HEW *gəra*, SK-HEW *ʔəra*, SK-MM *ʔəra*, SK-TA *gəra*.

to steal:

(1) PFL **t<əm>akav*, CL-KK *ta'kav*, CL-LR *ta'kav*, KD-LB *mə'ʔə*, WL-AB *ta'maku*, WL-LWI *təmakə*, WL-MS *maka*, WL-PD *maka*, WL-PD *t'maka*.

(2) CL-KK *la'vit*, WL-AD *'lā:*.

to suck:

(1) KD-LB *du'məʔ*, KD-LW *duməʔ*, WL-AB *dumə*, WL-BN *dumə*, WL-PD *da'm:e*, WL-PD *dumə*.

(2) WL-AD *'isəp*, WL-LWI *'isə*.

(3) CL-IL *tus*, CL-KK *tusu*, CL-LK *tusu*, CL-LP *tuhu*, CL-LWK *tusu*, CL-LWT *tuho*, CL-MR *tuso*, CL-PN *tuhuy*, EL-LE *tuhu*, EL-LMT *tuhu*, SK-HEW *vuho*, WL-BL *tuo*, WL-BM *tuho*, WL-BP *tuho*, WL-BT *tuho*, WL-DL *tuhu*, WL-HRW *tuho*, WL-IA *tuho*, WL-KO *tuhu*, WL-LG *tuho*, WL-LH *tuho*, WL-LMK *tuho*, WL-LML *tu*, WL-LTB *tuho*, WL-LWL *tuho*, WL-MD *tuho*, WL-ML *tuho*, WL-PK *tuho*, WL-RE *tuho*, WL-TJ *tuho*, WL-WB *tuho*, WL-WK *tuho*, WL-WT *tuho*, WL-WW *tuho*.

(4) CL-LR *məm'hi*, CL-LR *məmi'həŋ*.

to swell:

(1) SK-HEW *bovoʔ*, WL-AB *baʔ*, WL-BN *baʔ*.

(2) WL-AB *ihĩ*, WL-MS *hiŋ*, WL-MS *ihiy*, WL-PD *ihiy*.

(3) KD-LB *babaŋ*, WL-PD *baŋ*.

(4) CL-IL *kəbarasa*, CL-LP *kəbaraha*, CL-LWT *kəbarah*, WL-AD *'baʔa*, WL-BM *baʔa*, WL-BP *baʔa*, WL-BT *baʔa*, WL-DL *baʔa*, WL-HRW *baʔa*, WL-IA *bā*, WL-KO *baʔa*, WL-LG *baʔa*, WL-LMK *baʔa*, WL-LML *kəbaras*, WL-LTB *baʔa*, WL-LWI *'baʔa*, WL-LWL *baʔa*, WL-ML *kebaras*, WL-RE *baʔa*, WL-TJ *baʔa*, WL-WB *baʔa*, WL-WT *baʔa*, WL-WW *baʔ*. (5) CL-KK *sənobad̥z̥*, CL-KK *sno'bad̥z̥*, CL-LK *sobad̥z̥*, CL-LR *kə'hobɔi*, CL-LWK *sobad̥z̥*, CL-PN *kəhobaj*, WL-BL *hobad̥z̥*.

(6) EL-LMT *hogo*, WL-LH *hogor*, WL-MD *nogor*.

(7) CL-MR *kəbetɔ*, WL-WK *kəbətə*.

(8) SK-HEW *bovo*, SK-MM *bovo*, SK-TA *boβo*, WL-PK *bowo*.

to swim:

(1) PFL **naŋi*, CL-IL *naŋ*, CL-KK *naŋe*, CL-KK 'naŋe, CL-LP *naŋi*, CL-LR *naŋ*, CL-LWK *naŋe*, CL-LWT *naŋi*, CL-MR *naŋe*, CL-PN *naŋ*, EL-LE *naŋi*, EL-LMT *naŋi*, KD-LB *naŋi*, KD-LW *naŋi ɟjaŋ*, SK-HEW *nani*, SK-MM *nani*, SK-TA *nani*, WL-AB *naŋɛ*, WL-AD 'naŋe, WL-BL *naŋe*, WL-BM *naŋe*, WL-BN *naŋɛ*, WL-BP *naŋe*, WL-BT *naŋe*, WL-DL *naŋe*, WL-HRW *naŋe*, WL-IA *naŋe*, WL-KO *naŋe*, WL-LG *naŋe*, WL-LH *naŋe*, WL-LMK *naŋe*, WL-LML *naŋe*, WL-LTB *naŋe*, WL-LWI 'naŋe, WL-LWL *naŋe*, WL-MD *naŋe*, WL-ML *naŋe*, WL-MS *naŋɛ*, WL-PD *naŋe*, WL-PK *naŋe*, WL-RE *naŋe*, WL-TJ *naŋe*, WL-WK *naŋe*, WL-WT *naŋe*, WL-WW *naŋe*.

(2) CL-KK 'dulo, CL-LK *dulo*.

to tell:

(1) CL-KK 'tutu, CL-LR *tut*, KD-LB *tutu?* 'nanaŋ, WL-AB *tutu*, WL-AD 'tu-tu?, WL-LWI 'tutu?, WL-MS *tutu 'mariŋ*, WL-PD *tutu*.

to think:

(1) KD-LB *pikir*, WL-BN *pikir*. From Indonesian/Malay *pikir* 'think'.

(2) CL-PN *gənuan*, WL-AB *ga'nokuŋ*, WL-MS *g'nuku?*, WL-PD *g'nuku*.

(3) LH #*petən*, CL-IL *petən*, CL-KK *petən*, CL-KK 'petən, CL-LK *petən*, CL-LWK *petən*, CL-LWT *petən*, CL-MR *petə*, WL-AD 'petə, WL-BL *pet*, WL-BT *petən*, WL-DL *petəŋ*, WL-HRW *petən*, WL-IA *petən*, WL-KO *petə*, WL-LH *petən*, WL-LMK *petā*, WL-LML *petā*, WL-MD *petā*, WL-ML *petɔ*, WL-RE *petə*, WL-WK *petā*, WL-WT *petā*, WL-WW *petən*.

(4) LH-SK #*hukut*, EL-LE *huã*, EL-LMT *hu?ə*, SK-MM *hu:k*, SK-TA *hu:k*, WL-BP *hukut*, WL-LTB *hu*, WL-LWL *hukut*, WL-TJ *ukut*, WL-WB *ukut*.

(5) WL-BM *nəkət*, WL-LG *nəkət*, WL-LWI 'nəkə, WL-PK *nəkət*.

(6) SK-HEW *sugo*, SK-HEW *sugo*.

to throw:

(1) WL-AB *deiŋ*, WL-BN *deĩ*, WL-MS *dei?*, WL-PD *dei*, WL-PK *tubak*.

(2) KD-LB *ka'ka?*, KD-LW *kaka?*.

(3) KD-LW *bə?*, WL-AD *ge'ba?*.

(4) KD-LW *tida?*, KD-LW *vida?*.

(5) WL-AD *ga'hak*, WL-DL *gahak*, WL-IA *gaha*, WL-LMK *gahak*, WL-WT *gahak*, WL-WW *gahak*.

(6) EL-LE *geba*, EL-LMT *gebə*, SK-TA *geba*, WL-BL *geba*, WL-BP *geba*, WL-BT *geba*, WL-KO *geba*, WL-LH *geba*, WL-LML *geba*, WL-LTB *geba*, WL-LWL *geba*, WL-MD *geba*, WL-RE *geba*, WL-WB *geba*, WL-WK *geba*.

(7) WL-BM *wərə*, WL-LG *wərǎ*, WL-LWI *'wərəŋ*.

(8) CL-KK *batu*, CL-KK *'batu*, CL-LK *batu*, CL-LR *'bato*.

(9) CL-LR *dəkaŋ*, CL-LWK *dəka*, CL-PN *dəkaŋ*, WL-ML *dəka*, WL-TJ *dəka*.

(10) SK-TA *βelo*, WL-LWI *'waleŋ*.

(11) CL-LP *pədi*, CL-LWT *pədi*, CL-MR *pədi*.

(12) SK-HEW *dəgu*, SK-HEW *təgu*.

(13) SK-MM *loga*, SK-TA *roga*.

to tie:

(1) CL-IL *pui*, CL-KK *pu'dʒu*, CL-LR *'pui*, CL-LWK *pui*, CL-MR *pui*, CL-PN *puinəŋ*, WL-AB *pɔhi*, WL-AD *'puĩ*, WL-BL *pui*, WL-BM *puĩ*, WL-BN *pɔ'hi*, WL-BP *puĩ*, WL-BT *puĩ*, WL-DL *puin*, WL-HRW *puĩ*, WL-IA *pui*, WL-LG *puĩ*, WL-LH *pui*, WL-LTB *püi*, WL-LWI *'puin*, WL-LWL *püĩ*, WL-ML *pui*, WL-MS *pɔhi*, WL-PD *pɔhi*, WL-PK *püi*, WL-RE *püĩ*, WL-TJ *pui*, WL-WB *püĩ*, WL-WK *pui*, WL-WW *püi*.

(2) LH-KD *#soŋa*, CL-KK *soŋ*, CL-LK *soŋ*, CL-LP *hoŋ*, CL-LWT *hoŋ*, EL-LMT *hoŋe*, KD-LB *ho'ŋen*, KD-LW *hɔŋen*, WL-LML *hoŋǎ*, WL-MD *hoŋǎ*, WL-PD *hɔŋɔŋ*.

(3) WL-KO *wido*, WL-LMK *wido*, WL-WT *wido*.

(4) SK-HEW *mataʔ*, SK-HEW *mataʔ*, SK-MM *maṯa*.

to vomit:

(1) EL-LE *mu*, KD-LW *mutɛ ʒkaʔ*, SK-HEW *mutaʔ*, SK-HEW *mutaʔ*, SK-MM *muṯa*, SK-TA *muta*, WL-AB *muta*, WL-BN *muta*, WL-BP *muta*, WL-BT *muta*, WL-DL *muta*, WL-IA *muta*, WL-KO *muta*, WL-LG *muta*, WL-LH *mutakə*, WL-LMK *muta*, WL-LML *muta*, WL-LTB *muta*, WL-LWI *'mutaʔ*, WL-LWL *muta*, WL-MS *mutaŋ*, WL-MS *mutaʔ*, WL-PD *muta*, WL-PD *mutaŋ*, WL-PK *muta*, WL-RE *muta*, WL-TJ *muta*, WL-WB *muta*, WL-WT *muta*, WL-WW *muta*. (2) KD-LB *ɔ'kaʔ*, KD-LW *ʒkaʔ*, KD-LW *mutɛ ʒkaʔ*.

(3) SK-TA *doʔuk*, WL-AD *kɜ'doʔuk*.

(4) LH-KD *#luaŋ*, CL-IL *luaŋa*, CL-KK *lu'a*, CL-KK *luaŋ*, CL-LK *lua*, CL-LP *lua*, CL-LR *'luaŋ*, CL-LWK *lua*, CL-LWT *luaŋa*, CL-MR *lua*, CL-PN *luag*, EL-

LMT *luã*, WL-BL *lua*, WL-MD *luã*, WL-ML *lua*, WL-WK *lua*.

to wait:

- (1) WL-AB *baiŋ*, WL-MS *baiŋ*, WL-PD *baiŋ*.
- (2) WL-AD 'mĩã, WL-LWI 'niaŋ, WL-MS *miaŋ*, WL-PD *kiaŋ*, WL-PD *miaŋ*, WL-PD *riaŋ*, WL-PD *tiaŋ*.
- (3) CL-KK *kide*, CL-KK *kli'de*.

to wake someone up:

- (1) WL-BN *guo* 'baũ, WL-PD *pui* 'bauŋ.
- WL-PD *guo* 'bauŋ, (2) WL-MS *pui*, WL-PD *pui*, WL-PD *pui* 'bauŋ.
- (3) LH #rio, CL-KK 'riɔ, CL-LR 'riɔ, WL-AD 'ʔio.
- (4) WL-LWI *neiŋ* 'hogo, WL-LWI *sogo* 'hogo.

to wake up:

- (1) PFL **baŋun*, CL-KK 'baŋu, CL-LR *baŋ*, WL-AB *tide*, WL-MS *baŋuŋ*, WL-MS *baŋu*, WL-PD *baŋu*.
- (2) PFL **hogo*, KD-LB *hɔkɔ*, SK-HEW *hogo*, WL-AD 'hogo, WL-LWI 'hogo.

to walk:

- (1) PFL **pana(v)*, CL-KK *pana*, CL-KK *panav-*, CL-KK 'pana mə'ra lim, CL-LP *pana*, CL-LR 'panə, CL-LWK *pana*, CL-LWT *pana*, CL-MR *pana*, EL-LE *pã*, EL-LMT *pana*, KD-LB *pan*, KD-LW *pan ledɔ*, SK-HEW *pano*, SK-MM *pano*, SK-TA *pano*, WL-AB *pana* 'leiŋ, WL-AD 'pana, WL-BL *pana*, WL-BM *pana*, WL-BN *pana*, WL-BP *pana*, WL-BT *pana*, WL-DL *pana*, WL-HRW *pana*, WL-IA *pana*, WL-KO *pana*, WL-LG *pana*, WL-LH *pana*, WL-LMK *pana*, WL-LML *pana*, WL-LTB *pana*, WL-LWI 'pana, WL-LWL *pana*, WL-MD *pana*, WL-ML *pana*, WL-MS *pana* 'lei, WL-MS *pana* 'leiŋ, WL-PD *pana*, WL-PK *pana*, WL-RE *pana*, WL-TJ *pana*, WL-WB *pana*, WL-WK *pana*, WL-WT *pana*, WL-WW *pana*.
- (2) CL-IL *labi*, CL-LK *labi*.

to wash:

- (1) WL-BN *lamiŋ*, WL-MS *lamiŋ*, WL-PD *lamiŋ*.
- (2) WL-AB *beme* 'apa, WL-BN *beme*, WL-MS *bema*, WL-PD *bema*.
- (3) WL-BN *hue*, WL-MS *hue*, WL-PD *hue*.
- (4) PFL **basa*, CL-IL *basa*, CL-KK *basa*, CL-KK *ba'sa*, CL-LK *basa*, CL-LP *baha*, CL-LR 'bahaŋ, CL-LWK *basa*, CL-LWT *baha*, CL-PN *baha*, EL-LE *baha*,

EL-LMT *baha*, WL-AD *'baha*, WL-BM *baha*, WL-BP *baha*, WL-BT *baha*, WL-DL *baha*, WL-HRW *baha*, WL-IA *baha*, WL-KO *baha*, WL-LG *baha*, WL-LH *baha*, WL-LMK *baha*, WL-LML *ba*, WL-LTB *baha*, WL-LWI *'baha*, WL-MD *baha*, WL-ML *baha*, WL-PK *baha*, WL-RE *baha*, WL-TJ *baha*, WL-WB *baha*, WL-WT *baha*, WL-WW *bak*.

(5) WL-BL *səsa*, WL-WK *sesa*.

(6) CL-KK *pu:s*, CL-MR *pusʁ*, WL-LML *pu*.

(7) LH-SK #*puʔu*, SK-HEW *puʔu*, WL-LWI *'puʔu*, WL-LWL *puʔũ*.

(8) SK-HEW *bopo*, SK-HEW *popo*, SK-TA *popo*.

to weave:

(1) PFL **tani*, CL-KK *'tani*, CL-LR *'pano tan*, KD-LB *nekeʔ 'tan*, WL-AB *tane 'apa*, WL-AD *nekʔt tane*, WL-AD *tane*, WL-LWI *'tane*, WL-MS *tane*, WL-PD *tane*.

to wipe:

(1) PFL **hapu*, CL-IL *hapus*, CL-LK *apus*, CL-LWK *apus*, CL-LWT *apudʒa*, CL-MR *apudʒʁ*, EL-LE *hapu*, EL-LMT *hapu*, SK-HEW *hapus*, WL-AB *hapɔ*, WL-BN *hapɔ*, WL-LH *hapu*, WL-LML *hapu*, WL-ML *hapu*, WL-MS *hapɔ*, WL-PD *hapo*, WL-PK *hapũ*, WL-WK *hapus*.

(2) LH-KD #*doru* 'wipe; rub', KD-LW *dɔru*, WL-PD *dɔhɔ*.

(3) CL-LP *hok*, CL-LR *poho'kəj*, CL-PN *pohok*, KD-LB *pɔ'hɔʔ*, KD-LW *pɔ'hɔʔ*, WL-AD *'bɔhɔʔ*, WL-BT *bohok*, WL-DL *bohok*, WL-HRW *bohok*, WL-IA *boho*, WL-KO *bohok*, WL-MD *tohok*.

(4) WL-BP *teʔu*, WL-LWI *tɛ'ruʔ*.

(5) LH-KD #*hamu* 'wipe; sweep', KD-LW *hamur*, WL-BM *hamu*, WL-LG *hamu*, WL-LTB *hamu*, WL-LWL *hamu*, WL-TJ *hamu*, WL-WB *hamu*, WL-WW *hamo*.

(6) CL-KK *sɔ:k*, CL-KK *soka*, WL-LMK *oso*, WL-WT *oso*.

to yawn:

(1) WL-AB *panoe*, WL-MS *p'noek*, WL-PD *p'noe*.

(2) PFL **muav*, CL-KK *kəmu'av*, CL-LR *kə'moav*, KD-LB *moa*, SK-HEW *moaʔ*, WL-AD *pʔ'moa:*, WL-LWI *pə'moa*.

tongue:

(1) LH-KD #*ebel*, CL-IL *efeləs*, CL-KK *e'vɛl*, CL-KK *ewel*, CL-LK *ewel*, CL-

LP *efel*, CL-LR *e'vel*, CL-LWK *ewel*, CL-LWT *efel*, CL-MR *efel*, CL-PN *ewelag*, EL-LE *ebe*, EL-LMT *eblə*, KD-LB *ebel*, KD-LW *æbæł*, WL-AB *fɛ'fɛłɛŋ*, WL-AD *'ɜwɜ'rɜt*, WL-BL *ewel*, WL-BM *wewel*, WL-BN *fɛfɛł*, WL-BP *wewe*, WL-BT *wewe*, WL-DL *wewen*, WL-HRW *wewer*, WL-IA *ewel*, WL-KO *wewejək*, WL-LG *wewel*, WL-LH *ewel*, WL-LMK *ewel*, WL-LML *efel*, WL-LTB *wewe*, WL-LWI *'wewe'łəŋ*, WL-LWL *wewe*, WL-MD *ewelə*, WL-ML *ewel*, WL-MS *wewel*, WL-PD *wewel*, WL-PK *wewel*, WL-RE *wewel*, WL-TJ *wewel*, WL-WB *wewel*, WL-WK *ewel*, WL-WT *wewer*, WL-WW *wewer*.

(2) SK-HEW *ma:*, SK-MM *ma:n*, SK-TA *ma:*, SK-TA *ma:ŋ*.

tree:

(1) PFL **kayu* 'tree; wood', CL-IL *kad̥ʒor*, CL-KK *'kad̥ʒor*, CL-KK *kad̥ʒu*, CL-KK *'kad̥ʒu*, CL-LK *kad̥ʒor*, CL-LP *kad̥ʒu*, CL-LR *'kajor*, CL-LWK *kajoru*, CL-LWT *kad̥ʒoru*, CL-MR *kad̥ʒu*, CL-PN *kajo*, KD-LB *ʔai*, KD-LW *ai vaʔ*, KD-LW *kain laraŋ*, SK-HEW *ai*, SK-HEW *ʔai*, SK-TA *ʔai*, WL-AB *kad̥ʒo*, WL-AD *'kajo*, WL-BL *kad̥ʒo*, WL-BM *kajo*, WL-BN *kad̥ʒu 'lələŋ*, WL-BP *kad̥ʒo*, WL-BT *kajo*, WL-DL *karo*, WL-HRW *kajo*, WL-IA *kajo*, WL-KO *kajo*, WL-LG *kajo*, WL-LH *kajo*, WL-LMK *kad̥ʒo*, WL-LML *kad̥ʒo*, WL-LTB *kad̥ʒopukã*, WL-LWL *kad̥ʒo*, WL-MD *kajo*, WL-ML *kad̥ʒo*, WL-MS *kad̥ʒo 'pukəŋ*, WL-PD *kad̥ʒo 'pukəŋ*, WL-PK *kaju*, WL-RE *kajo*, WL-TJ *kad̥ʒo*, WL-WB *kad̥ʒo*, WL-WK *kad̥ʒo*, WL-WT *kad̥ʒo*, WL-WW *kajo*.

(2) EL-LE *əso*, EL-LMT *əso*.

(3) WL-LTB *kad̥ʒopukã*, WL-LWI *'pukəŋ*.

trousers:

(1) LH-KD #*deko*, CL-KK *dekor*, CL-KK *deko*, CL-LR *kə'deko* 'doak, KD-LB *deko 'lawan*, WL-AB *deko 'leiŋ bəla'hakaŋ*, WL-AD *'deko'blahã*, WL-MS *deko 'leiŋ b'lahak*, WL-PD *deko 'leiŋ b'lahak*.

turtle:

(1) PFL **kera*, CL-KK *'kear*, CL-KK *'kea*, CL-LR *ke'rar*, KD-LB *ʔe're*, SK-HEW *ʔera*, WL-AB *kea*, WL-AD *'keʔa*, WL-LWI *'keʔa*, WL-MS *kea*, WL-PD *kea*.

two:

(1) PFL **d̥ʒua*, CL-KK *'d̥ʒua*, CL-LR *'d̥ʒua*, KD-LB *su'we*, KD-LW *suɛ*, SK-HEW *rua*, SK-TA *rua*, WL-AB *rua*, WL-AD *'rua*, WL-BN *rua*, WL-LWI *'rua*, WL-MS *rua*, WL-PD *rua*.

uncooked rice:

- (1) WL-AB *pari*, WL-MS *wata* 'pare, WL-MS *pare*, WL-PD *pare*.
- (2) LH-KD #(d)anen, CL-KK 'liso da'nen, KD-LB ?a'nen.
- (3) LH #kluok, CL-LR kə'luok 'nalatən, CL-LR kə'luok.

vagina:

- (1) PFL *puki, CL-LR pu'ki, WL-AB *pukij*, WL-MS *pukij*, WL-MS *puki?*, WL-PD *pukij*.
- (2) LH #məna, CL-KK mə'na, WL-AD 'mənə, WL-LWI 'mena.

vein:

- (1) PFL *alis 'tendon', CL-KK a'lis, CL-LR *alhin*, KD-LB a'li, SK-HEW a:li, WL-AD a:l'hĩ, WL-LWI *ali'hij*, WL-MS *alij*.

village:

- (1) LH-KD #ləvu, CL-KK lə'vu, CL-KK lə'vɔr, CL-LR 'lavor, KD-LB *lewu*, WL-AB *lafɔ*, WL-AD lə'wɔ:, WL-LWI 'ləwɔ, WL-MS *ləwɔ*, WL-PD *lawɔ*.

voice:

- (1) WL-AB *alaŋ*, WL-MS *ala?*, WL-MS *alaŋ*, WL-PD *alaŋ*.
- (2) LH #raaŋ, CL-KK ra:ŋ, CL-LR 'raŋa, WL-AD 'rã:, WL-LWI 'ra:ŋ.

waist:

- (1) LH-KD #ale, CL-KK a'le, CL-LR 'ale, KD-LB a'len, WL-AB *alɛ*, WL-AD kɜ'lɜgɜrɜt, WL-MS *alɛŋ*, WL-MS *alɛ?*, WL-PD *alɛŋ*.

wall:

- (1) PFL *gəbi/gnəbin, CL-KK kənə'bin, CL-KK kənə'ber, CL-LR kəna'ber, KD-LB nɛ'bi?, SK-HEW *gəbi*, WL-AB *ga'nabɛŋ*, WL-AD 'kənɜ'bi?, WL-LWI kə'nəbi, WL-MS *gə'b:e*, WL-MS *keba*, WL-PD *g'nab:ɛŋ*.

water:

- (1) PFL *vaʔir, CL-IL *faer*, CL-KK va'er, CL-KK 'vai, CL-KK *waer*, CL-LK *waer*, CL-LP *faer*, CL-LR *vaer*, CL-LWK *waeru*, CL-LWT *faeru*, CL-MR *waeru*, CL-PN *waer*, EL-LE *wae*, EL-LMT *wae*, KD-LB *wei*, KD-LW *væi ai*, SK-HEW *vair*, SK-HEW *vair*, SK-MM *vai*, SK-TA *βair*, WL-AB *fei*, WL-AD *wai*, WL-BL

wai, WL-BM *wai*, WL-BP *wai*, WL-BT *wai*, WL-DL *wai*, WL-HRW *wai*, WL-IA *wai*, WL-KO *wai*, WL-LG *wai*, WL-LH *waʔi*, WL-LMK *wai*, WL-LML *fai*, WL-LTB *waiʔ*, WL-LWI *wai*, WL-LWL *wai*, WL-MD *waʔi*, WL-ML *wai*, WL-MS *wai*, WL-PD *wai*, WL-PK *wai*, WL-RE *wai*, WL-TJ *wai*, WL-WB *wai*, WL-WK *wai*, WL-WT *wai*, WL-WW *wai*.

wave:

(1) PFL *(1)oyor ‘wave; sea’, CL-KK ɔ’d͡ʒɔk, KD-LB ɔ’jɔʔ, WL-AB ɔd͡ʒɔ, WL-LWI ɔjo, WL-MS ɔd͡ʒɔʔ, WL-PD ɔd͡ʒɔ.

(2) WL-AB *umba*, WL-MS *ɔmbaʔ*.

wet:

(1) WL-AB *ɛɛ*, WL-BN *ɛɛ*, WL-MS *ɛɛk*, WL-PD *ɛɛk*.

(2) KD-LB *ri’taʔ*, KD-LW *ritaʔ dɔrɔʔ*.

(3) KD-LW *dɔrɔʔ ɤerɔʔ*, KD-LW *ritaʔ dɔrɔʔ*.

(4) WL-AD *’namʔɛ*, WL-BT *nəme*, WL-DL *nəme*, WL-HRW *nəme*, WL-IA *nəmaŋ*, WL-KO *nəmʔe*, WL-LH *nəme*, WL-LMK *nəmaŋ*, WL-MD *nəmʔe*, WL-WT *nəmʔe*, WL-WW *nəme*.

(5) WL-BM *dəme*, WL-BP *kədəmu*, WL-LG *dəmə*, WL-LWI *də’mɛŋ*, WL-LWL *kədəmu*, WL-PK *dəman*, WL-RE *dəme*, WL-TJ *dəma*, WL-WB *kədəmu*.

(6) CL-IL *sənəbar*, CL-KK *sənəbar*, CL-KK *sənə’bar*, CL-LK *səbar*, CL-LP *həba*, CL-LR *hənəba’rən*, CL-LWK *sənəbar*, CL-LWT *həba*, CL-MR *sənəbar*, CL-PN *hənəbarən*, WL-BL *səba*, WL-LML *sənəbe*, WL-ML *sənəbe*, WL-WK *səba*.

(7) EL-LMT *weɿ*, WL-LTB *kəwaʔi*.

(8) SK-HEW *gəma*, SK-HEW *gəmaʔ*, SK-MM *gəma*, SK-TA *gəma*.

whale:

(1) WL-AB *ikaŋ ka’laru*, WL-MS *k’laruʔ*, WL-PD *k’laru*.

(2) KD-LB *iʔa paus*, SK-HEW *iʔan paus*, WL-LWI *’paus*.

(3) CL-KK *tə’mus*, CL-LR *təm’hu*.

what:

(1) PFL **apa*, CL-IL *kapa*, CL-LP *kapa*, CL-LR *a’pɔi*, CL-PN *apaj*, EL-LE *ape*, KD-LB *a’pe*, KD-LW *ape*, SK-HEW *apa*, SK-HEW *apar*, SK-MM *apa*, WL-AB *pai*, WL-BN *pei*, WL-MS *apai*, WL-MS *pai*, WL-PD *pai*.

(2) CL-LK *aleka*, CL-LWT *alaka*, CL-MR *alaku*, WL-AD *’aku*, WL-BL *alaku*, WL-BM *a*, WL-BP *a*, WL-BT *aku*, WL-DL *aku*, WL-HRW *aku*, WL-IA *aku*, WL-

KO *aku*, WL-LG *a*, WL-LH *a*, WL-LMK *aku*, WL-LML *alaka*, WL-LTB *a*, WL-LWI *'a*, WL-ML *ale*, WL-PK *a*, WL-TJ *a*, WL-WB *a*, WL-WK *alaku*, WL-WT *aku*, WL-WW *aku*.

(3) PFL ***an**, CL-KK *an*, WL-LWL *ã*, WL-RE *ã*.

(4) EL-LMT *naʔa*, WL-MD *naʔã*.

when:

(1) CL-IL *d̥zəm pira*, CL-KK *'d̥zəm 'pira*, CL-KK *d̥zəm pira*, CL-LK *d̥zəm pira*, CL-LP *d̥zəmpira*, CL-LR *'d̥zəma 'pira*, CL-LWK *d̥zəm pira*, CL-LWT *d̥zəma pira*, CL-MR *d̥zəma pira*, EL-LE *erepirə*, EL-LMT *arapira*, KD-LB *wɛŋ 'pie*, KD-LW *ɛɛ*, KD-LW *væŋ pie*, SK-HEW *rəma pira*, SK-HEW *rəmapira*, SK-MM *rema pira*, WL-AB *ɛrə 'pira*, WL-AD *'ʔɜrɜ'pira*, WL-BL *arapira*, WL-BM *arəmpira*, WL-BP *arəmpira*, WL-BT *ər pira*, WL-DL *ərən pira*, WL-HRW *ərəm pira*, WL-IA *arampira*, WL-KO *ər pira*, WL-LG *arəmpira*, WL-LH *arapira*, WL-LMK *arəm pira*, WL-LML *arapira*, WL-LWI *ərən 'pira*, WL-LWL *arapira*, WL-MD *ar-apira*, WL-ML *arapira*, WL-MS *ər'pira*, WL-PD *ari 'pira*, WL-PK *arəmpira*, WL-RE *arəmpira*, WL-TJ *arəmpira*, WL-WB *arəmpira*, WL-WK *ar pira*, WL-WT *ərəm pira*, WL-WW *arəmpira*.

(2) SK-TA *nora ləro-ŋ pae*, SK-TA *nora oras*, SK-TA *nora pae*.

where:

(1) CL-IL *kətəga*, CL-KK *gane*, CL-KK *nagane*, CL-KK *ga*, CL-LK *nanaga*, CL-LP *kətega*, CL-LR *mai gaʔ*, CL-LR *no gə*, CL-LR *tamo nəgaʔ*, CL-LWK *maga*, CL-LWT *ketega*, CL-MR *maga*, CL-PN *nanaga*, EL-LMT *nəgaʔe*, WL-AB *ɔɔ 'n-an̩ga*, WL-AB *nan̩ga*, WL-AB *nan̩ga 'd̥zafa*, WL-AD *d̥z'gaku*, WL-BL *nəga*, WL-BM *gaʔe*, WL-BN *na'ŋga 'ɔɔ*, WL-BP *təga*, WL-BT *dəga*, WL-HRW *təga*, WL-IA *təga*, WL-KO *dəga*, WL-LG *gaʔe*, WL-LH *təga*, WL-LMK *dəga*, WL-LML *diga*, WL-LWI *'gaʔe*, WL-LWL *təga*, WL-MD *təga*, WL-ML *nəga*, WL-MS *r'ga*, WL-MS *r'ga 'naŋ*, WL-PD *ɔrga*, WL-PD *ər'ga*, WL-PK *gaʔe*, WL-RE *diga*, WL-TJ *təgaʔe*, WL-WB *təga*, WL-WK *dəga*, WL-WT *digaku*, WL-WW *təga*.

(2) KD-LB *d̥iʔ 'dita*, KD-LW *denita*, KD-LW *dita*.

(3) SK-HEW *upa*, SK-MM *epae*.

white:

(1) PFL ***budaʔ**, CL-IL *bid̥zakən*, CL-KK *bud̥zak*, CL-KK *bu'd̥zak*, CL-LK *bud̥zak*, CL-LP *bud̥za*, CL-LR *bu'd̥zak*, CL-LWK *bud̥zak*, CL-LWT *bud̥za*, CL-MR *bud̥zakɾ*, CL-PN *bud̥zakən*, EL-LE *bur̥ʔ*, EL-LMT *bur̥ʔã*, KD-LB *bu'jaʔ*, KD-

LW *buja? tape*, SK-HEW *bura?*, SK-HEW *bura?*, SK-MM *bura*, SK-TA *bura*, SK-TA *bura bara*, WL-AB *bura*, WL-AD *'bura?*, WL-BL *bura*, WL-BM *burã?*, WL-BN *bur:a*, WL-BP *bura?*, WL-BT *burã*, WL-DL *bur?ã*, WL-HRW *burã*, WL-IA *burakən*, WL-KO *bur?ã*, WL-LG *bura?*, WL-LH *bur?ã*, WL-LMK *bur?ã*, WL-LML *burã*, WL-LTB *burã*, WL-LWI *'bura*, WL-LWL *burã*, WL-MD *bur?a*, WL-ML *burɤ*, WL-MS *burak*, WL-PD *burak*, WL-PK *bura?ã*, WL-RE *burã*, WL-TJ *bura?*, WL-WB *burã*, WL-WK *bura*, WL-WT *bur?ã*, WL-WW *bura*.

who:

- (1) WL-AB *fiaru*, WL-BN *hafa*, WL-MS *haru*.
- (2) WL-MS *hewai'aru*, WL-MS *hewai*, WL-PD *hewai*.
- (3) KD-LB *sio*, KD-LW *sio*.
- (4) LH *#henaku*, CL-IL *kena*, CL-KK *e'nak*, CL-KK *naku*, CL-LK *enak*, CL-LP *kena*, CL-LR *e'na*, CL-LWK *enaku*, CL-LWT *kena*, CL-MR *henaku*, CL-PN *ena*, EL-LE *he?e*, EL-LMT *he*, WL-AD *'heku*, WL-BL *heku*, WL-BM *hege*, WL-BP *hege*, WL-BT *hege*, WL-DL *heku*, WL-HRW *heku*, WL-IA *heku*, WL-KO *heku*, WL-LG *hege*, WL-LH *heku*, WL-LMK *heku*, WL-LML *heku*, WL-LTB *hege*, WL-LWI *'hege*, WL-LWL *hege*, WL-MD *heku*, WL-PK *hege*, WL-RE *hege*, WL-TJ *hege*, WL-WB *hege*, WL-WK *heku*, WL-WT *hege*, WL-WW *hege*.
- (5) SK-HEW *hai*, SK-HEW *hainimu*, SK-MM *hai*, SK-TA *hai*.

wide:

- (1) WL-AB *pa'leŋa*, WL-BL *pəlagad̪ʒ*, WL-MS *p'leŋaŋ*, WL-PD *p'lengaŋ*.
- (2) CL-IL *pəlafat*, CL-LP *pəlafatən*, CL-LWT *pəlafə*, CL-MR *pəlawat*, WL-LML *pəlafat*, WL-PD *p'lawa?*.
- (3) EL-LE *bəle*, EL-LMT *pəlo?e*, SK-HEW *pəloeŋ*, WL-AD *'bele?*, WL-BM *belã*, WL-BP *bəlapa*, WL-BT *pəlelet*, WL-DL *bel?ã*, WL-HRW *pəlelet*, WL-IA *pəloen*, WL-KO *pəlelet*, WL-LG *belə*, WL-LWI *'belə*, WL-LWL *belẽ*, WL-MD *belã*, WL-ML *belɤ*, WL-PK *bel?ã*, WL-RE *pəlo?ẽ*, WL-TJ *belə*, WL-WB *bəlã*, WL-WK *bel?ẽ*, WL-WW *pəlelet*.
- (4) LH-SK *#ledan*, CL-KK *ple'dan*, SK-KR *lədaŋ* 'wide open room'.
- (5) CL-KK *kedak*, CL-LK *kedak*, CL-LR *keda'kən*, CL-PN *kedakən*.
- (6) SK-MM *pəŋaŋ*, SK-TA *peŋa-ŋ*.

wife:

- (1) PFL **hava* 'spouse', CL-IL *kəfae*, CL-KK *a'van bi'ne*, CL-KK *awan*, CL-LK *awan*, CL-LR *kəveinan*, CL-LWK *awaga bine*, EL-LMT *hawã*, WL-AB *ka'fae*,

WL-BM *kəwae*, WL-BN *ka'fae*, WL-BP *kəwae*, WL-BT *kəwae*, WL-DL *kəwae*, WL-HRW *kəwahe*, WL-LG *kəwae*, WL-LMK *bərkəwahe*, WL-LML *kəfae*, WL-LTB *kəwae*, WL-LWI *kə'wae*, WL-LWL *kəwae*, WL-MD *kəwae*, WL-ML *kəwae*, WL-MS *k'wae* 'nara, WL-MS *k'wae*, WL-PD *k'wae*, WL-PK *vahe*, WL-RE *kəwāe*, WL-TJ *kəwae*, WL-WB *kəwae*, WL-WK *kəwae*, WL-WW *kəwae*.

(2) KD-LB *we?* 'rian, KD-LW *væ?* rian.

(3) WL-AD *hāi*, WL-BL *ha*, WL-IA *hān*, WL-KO *hak*, WL-LH *hakā*, WL-WT *hā*.

(4) CL-LP *gərapu*, CL-LWT *gerapu*, CL-MR *gərapu*, EL-LE *bare*.

(5) SK-HEW *du?a*, SK-HEW *du?at*, SK-MM *du?a*.

(6) SK-TA *ata du?a* *βai*, SK-TA *βai*.

wind:

(1) PFL **aŋin*, CL-IL *aŋi*, CL-KK *a'ŋi*, CL-KK *aŋin*, CL-KK *a'ŋin*, CL-LK *aŋin*, CL-LP *aŋin*, CL-LR *'aŋin*, CL-LWK *aŋinu*, CL-LWT *aŋinu*, CL-MR *aŋi*, CL-PN *aŋin*, EL-LE *aŋin*, EL-LMT *aŋi*, KD-LB *a'ŋin*, KD-LW *aŋin vɛŋin*, SK-HEW *anin*, SK-HEW *aŋin*, SK-MM *anin*, SK-TA *ani-n*, WL-AB *aŋĩ*, WL-AD *'aŋĩ*, WL-BL *aŋi*, WL-BM *aŋi*, WL-BN *aŋi*, WL-BP *aŋi*, WL-BT *aŋi*, WL-DL *aŋi*, WL-HRW *aŋi*, WL-IA *aŋin*, WL-KO *aŋi*, WL-LG *aŋi*, WL-LH *aŋi*, WL-LMK *aŋi*, WL-LML *aŋi*, WL-LTB *aŋi*, WL-LWI *'aŋin*, WL-LWL *aŋi*, WL-MD *aŋi*, WL-ML *aŋi*, WL-MS *aŋiŋ*, WL-PD *aŋiŋ*, WL-RE *aŋi*, WL-TJ *aŋi*, WL-WB *aŋi*, WL-WK *aŋi*, WL-WT *aŋi*, WL-WW *aŋi*.

wing:

(1) PFL **kapik*, CL-IL *kəpik*, CL-KK *kəpik*, CL-KK *kə'pik*, CL-LK *kəpik*, CL-LP *kəpik*, CL-LR *kə'pik*, CL-LWK *kəpik*, CL-LWT *kəpik*, CL-MR *kəpik*, CL-PN *kəpikən*, EL-LE *api*, EL-LMT *api?ĩ*, KD-LB *a'pi?*, KD-LW *api?*, SK-HEW *kapik*, SK-MM *kəpik*, SK-TA *kəpik*, WL-AB *ka'pikiŋ*, WL-AD *'kəp?ĩ*, WL-BL *kəpi*, WL-BM *kəpi?*, WL-BP *kəpi?*, WL-BT *kəp?ĩ*, WL-DL *kəpiŋ*, WL-HRW *kəpik*, WL-IA *kəpikən*, WL-KO *kəp?i*, WL-LG *kəp?ĩ*, WL-LH *kəpi*, WL-LMK *kəpĩ*, WL-LML *kəpik*, WL-LTB *kəpi?*, WL-LWI *kə'pi?iŋ*, WL-LWL *kəpi?*, WL-MD *kəpi*, WL-ML *kəpik*, WL-MS *kapik*, WL-PD *kapik*, WL-PK *kəpi?ĩ*, WL-RE *kəpi*, WL-TJ *kəpi?*, WL-WB *kəpi?*, WL-WK *kəpi?*, WL-WT *kəpi*, WL-WW *kəpik*.

with:

PFL **nora*, CL-IL *na*, CL-KK *ko*, CL-KK *mo*, CL-KK *no*, CL-LK *no*, CL-LP *no*, CL-LR *no*, CL-LWK *no*, CL-PN *no*, EL-LMT *mo*, EL-LMT *ʔo*, KD-LB *no're*,

KD-LW *nɔrɛ*, SK-HEW *-ora*, SK-HEW *nora*, WL-AB *nɔŋ*, WL-AD *nɔʔ*, WL-BL *no*, WL-BM *no*, WL-BN *nɔ̃:*, WL-BP *nɔ̃*, WL-BT *noʔo*, WL-DL *noʔo*, WL-HRW *noʔo*, WL-IA *no*, WL-KO *noʔɔ̃*, WL-LG *noʔɔ̃*, WL-LH *kã*, WL-LH *mã*, WL-LMK *noʔɔ̃*, WL-LML *nã*, WL-LTB *noʔɔ̃*, WL-LWI *'nɔʔɔŋ*, WL-LWL *noʔɔ̃*, WL-MD *ko*, WL-MD *mo*, WL-ML *nɔ̃*, WL-MS *naŋ*, WL-PD *naŋ*, WL-PK *noʔɔ̃*, WL-RE *noʔɔ̃*, WL-TJ *noʔo*, WL-WB *nã*, WL-WK *no*, WL-WT *noʔɔ̃*, WL-WW *noʔɔ̃*.

(2) CL-LWT *fɛ*, CL-MR *we*.

woman:

(1) PFL **vai*, SK-TA *ata βai*, WL-AB *ka'fae*, WL-AD *'ina wae*, WL-BL *kəbarawae*, WL-BM *inawae*, WL-BN *ka'fae*, WL-BP *bərwaĩ*, WL-BT *bərwaie*, WL-DL *inawae*, WL-HRW *inawae*, WL-IA *inawae*, WL-KO *inawae*, WL-LG *inawae*, WL-LH *bərwaie*, WL-LMK *bərkəwae*, WL-LML *kəbarafae*, WL-LTB *kəbarawāi*, WL-LWI *ina 'wai*, WL-LWL *inawae*, WL-MD *bərwaie*, WL-ML *kəbarawae*, WL-MS *k'wae*, WL-PD *k'wae*, WL-PK *vae*, WL-RE *bərkəwāe*, WL-TJ *inawae*, WL-WB *bərkəwāe*, WL-WK *kəwae*, WL-WT *bərkəwae*, WL-WW *kawae*.

(2) KD-LW *inaʔ*, KD-LW *inɔ̃*, KD-LW *inɛ*, WL-AD *'ina wae*, WL-DL *inawae*, WL-HRW *inawae*, WL-IA *inawae*, WL-KO *inawae*, WL-LG *inawae*, WL-LWI *ina 'wai*, WL-LWL *inawae*, WL-TJ *inawae*.

(3) CL-IL *gərapor*, CL-LP *gərapod̃ʒa*, CL-LWT *gərapod̃ʒa*, CL-MR *gəraporu*.

(4) PFL **binay* 'female; sister', CL-KK *binen*, CL-KK *bi'nɛn*, CL-LK *bined̃ʒa*, CL-LWK *bine*.

(5) CL-LR *ronen*, CL-PN *ronen*.

(6) SK-HEW *duʔa*, SK-HEW *duʔat*, SK-MM *duʔa*.

(7) EL-LE *bare*, EL-LMT *bare*.

wood:

(1) PFL **kayu* 'wood; tree', CL-KK *'kad̃ʒu*, CL-KK *'kad̃ʒɔr*, CL-LR *kaj 'tavan*, KD-LB *ʔai*, SK-HEW *ʔai*, WL-AB *kad̃ʒɔ*, WL-AD *'kajo*, WL-LWI *'kajo*, WL-MS *kad̃ʒɔ*, WL-PD *kad̃ʒɔ*.

word:

(1) WL-AB *kirɛ*, WL-MS *k'dirɛʔ*, WL-PD *k'dirɛ*.

(2) WL-AD *'koda*, WL-LWI *'koda*.

wound:

(1) LH #ugadak, CL-KK *ga'dʒak*, CL-LR *ga'jak*, WL-AB *uka*, WL-MS *ugar*, WL-MS *ukak*, WL-PD *ugar*, WL-PD *uka*.

wrong:

(1) PFL *sala, CL-KK *sna'lak*, CL-LR *bənəla'ɣən*, SK-HEW *hala*, WL-AB *hala*, WL-AD *'nalã*, WL-LWI *'nalaŋ*, WL-MS *hala*, WL-MS *halaŋ*, WL-PD *halaŋ*.

year:

(1) PFL *tuun, CL-IL *tunən*, CL-KK *tu:n*, CL-KK *tunən*, CL-LK *tunən*, CL-LP *tun*, CL-LR *'tunan*, CL-LWK *tunən*, CL-LWT *tun*, CL-MR *tunɤ*, CL-PN *tunanən*, EL-LE *tuŋ*, EL-LMT *tũ*, KD-LB *tun*, KD-LW *tun*, WL-AB *tuŋ*, WL-AD *tũ:*, WL-BL *tuŋ*, WL-BM *sũ*, WL-BN *tu:ŋ*, WL-BP *sũ*, WL-BT *tun*, WL-DL *tuŋ*, WL-HRW *tuŋ*, WL-IA *tun*, WL-KO *tũn*, WL-LG *sũ*, WL-LH *tunə*, WL-LMK *tuŋ*, WL-LML *tuŋ*, WL-LTB *sũ*, WL-LWI *'su:ŋ*, WL-LWL *sũ*, WL-MD *tuŋ*, WL-ML *tu*, WL-MS *nuŋ*, WL-MS *nuŋ 'nuŋ*, WL-MS *tuŋ*, WL-PD *tuŋ*, WL-PK *sũ*, WL-RE *tũ*, WL-TJ *sũ*, WL-WB *sũ*, WL-WK *tuŋ*, WL-WT *tun*, WL-WW *tũ*.

(2) SK-HEW *livan*, SK-HEW *livan*, SK-MM *livan*, SK-TA *?liβa-ŋ*.

yellow:

(1) LH-KD #kumas, CL-IL *kumasən*, CL-KK *kumas*, CL-KK *ku'mas*, CL-LK *kumas*, CL-LP *kənumahan*, CL-LR *kuma'hən*, CL-LWK *kumas*, CL-LWT *kuma*, CL-MR *kumas*, CL-PN *kumahən*, EL-LE *umɔ*, EL-LMT *uməhã*, KD-LB *ʔu'man*, KD-LW *uma kateʔ*, WL-AB *kumɔŋ*, WL-AD *'kumhã*, WL-BL *kuma*, WL-BM *kuma*, WL-BN *kumõ:*, WL-BT *kumha*, WL-DL *kuma*, WL-HRW *kumha*, WL-IA *kuma*, WL-KO *kumha*, WL-LG *kumã*, WL-LH *kuma*, WL-LMK *kumha*, WL-LML *kumã*, WL-LTB *kumã*, WL-LWI *'kuma*, WL-LWL *kumã*, WL-MD *kumhã*, WL-ML *kumɤ*, WL-MS *kumɔŋ*, WL-PD *kumɔŋ*, WL-RE *kumã*, WL-TJ *kuma*, WL-WB *kumã*, WL-WK *kuma*, WL-WT *kumã*, WL-WW *kuma*.

(2) SK-HEW *heret*, SK-MM *heret*, SK-TA *herə-t*, WL-PK *heret*.

yesterday:

(1) LH-KD #evian, KD-LB *ʔɛ'wɪn*, WL-AB *fiaŋ*, WL-AD *'wia*, WL-LWI *'wia*, WL-MS *wiaŋ*, WL-PD *wiaŋ*.

(2) CL-KK *nəgə'ro*, CL-LR *garo*.

younger sibling:

(1) PFL *vadi, CL-IL *wad̥zi*, CL-KK *va'd̥zin*, CL-KK *wad̥zi*, CL-LP *wad̥zin*, CL-LR *va'd̥zi*, CL-LWK *wad̥ziga*, CL-LWT *fad̥zin*, CL-MR *wad̥zi*, CL-PN *wad̥zig*, EL-LE *wari*, EL-LMT *ari*, KD-LB *ʔa'riʔ*, SK-HEW *vari*, WL-AB *arij*, WL-AD *'arim*, WL-BL *ari*, WL-BM *ariʔ*, WL-BN *ʔarĩ*, WL-BT *arik*, WL-DL *arik*, WL-HRW *arik*, WL-IA *ari*, WL-KO *arik*, WL-LG *ariʔ*, WL-LH *arikə*, WL-LMK *arĩ*, WL-LML *arik*, WL-LTB *ari*, WL-LWI *'arij*, WL-MD *ari*, WL-ML *ari*, WL-MS *arij*, WL-PD *arij*, WL-PK *arĩ*, WL-RE *arĩ*, WL-WK *ari*, WL-WT *arik*, WL-WW *arĩ*.

(2) WL-MS *kau*, WL-PD *kau*.

(3) WL-MS *ul:a*, WL-PD *ula*.

APPENDIX C

Orthographies and glosses from other sources

C.1 List of orthographic representations

Table C.1 shows a list of languages from other sources and the original transcription conventions used in these sources. Only sounds that are represented in varying ways across sources are listed. Thus, the phonemes listed per language do not represent the complete phoneme inventory. The exact realization of some phonemes can vary. The phoneme /v/ has also been described as [ʋ] or [w] or [β] for certain languages. The phoneme /e/ has also been described as [ɛ] and /o/ has also been described as [ɔ]. But no phonemic contrast has been described for those variants. Vowel length is represented by doubling the vowel symbol.

Table C.1: Transcription conventions in this dissertation and their equivalents in other sources.

IPA symbol	?	ŋ	ɸ	v	w	dʒ	z	j	ə	e/ɛ	æ	ɐ/a	..	ɓ	d'	f
Orthography in this thesis	'	ng	f	w	w	j	z	y	e	é	è	a	..	ɓ	d'	f
Amarasi (Edwards 2016a)	ʔ	ŋ	-	-	-	dʒ	-	-	-	e	-	a	-	-	-	-
Blagar (Steinhauer 2014)	ʔ	ŋ	-	-	-	-	j	y	-	e	-	a	-	ɓ	-	-
Bunaq (Schapper 2009)	q															
CL-Atadei Painara (Krauß 2016:119)	-	ng	-	w	-	j	-	y	ě	é	-	a	-	-	-	-
Dela (Rote) (Tamelan 2007)	'	ng	-	-	-	-	-	-	-	e	-	a	-	-	-	-
Termanu (Rote) (Jonker 1908)	?	ng	-	w	-	-	-	j	-	e	è	a	-	-	-	-
Fataluku (van Engelenhoven 2009)	— no information on orthography in this publication —															
Hawu (Walker 1982)	'	ŋ	-	w	-	j	-	-	ə	e	-	a	-	ɓ'	d'	j'
Helong (Balle 2017)	ʔ	ŋ	-	-	-	-	-	-	-	e	-	a	-	-	-	-
Kaera (Klamer 2014b)	-	ng	-	v	w	-	-	y	-	ɛ	-	a	-	-	-	-
Kamera (Klamer 1998)	-	ng	-	-	w	j	-	y	-	e	-	a	-	ɓ	d	-
Makalero (Huber 2011)	ʔ	-	-	v	-	-	-	j	-	e	-	a	-	-	-	-
Makasae (Huber 2008)	ʔ	-	-	-	w	-	-	-	-	e	-	a	-	-	-	-
Mambae (Grimes et al. 2014)	-	-	-	-	-	-	-	-	-	e	-	a	-	-	-	-
Mangarai (Semiun 1993)	— no information on orthography in this publication —															
KD-Leuwayang (Samely 1991a:56)	q	ng	-	w	-	-	-	y	-	e	è	a	>	-	-	-
Keo (Baird 2002:64-66)	'	ng	-	w	-	j	-	-	e	é	-	a	-	-	-	-
SK (Arndt 1931:3-4)	-	ng	-	w	-	dz	-	-	e	é	-	a	c	-	-	-

IPA symbol	?	ŋ	ϕ	v	w	dʒ	z	j	ə	e/ɛ	æ	ɐ/ɑ	..	ʙ	d	f
Orthography in this thesis	'	ng	f	w	w	j	z	y	e	é	è	a	..	ʙ	d	f
SK-Hewa (Fricke 2014a:12, 19)	'	ng	-	w	-	j	-	-	ê	e	-	a	-	-	-	-
SK (Rosen 1986)	?	ng	-	w̥	-	j̥	-	-	ě	e	-	a	-	-	-	-
SK-Krowe (Lewis and Grimes 1995)	?	ŋ	-	β	-	j̥	-	-	ə	e	-	a	-	-	-	-
SK-Krowe (Pareira and Lewis 1998)	'	ng	-	w	-	j	-	-	e	é	-	a	-	-	-	-
Teiwa (Klamer 2010)	'	ng	f	v	-	-	-	y	-	e	-	a	-	-	-	-
Tetun (van Klinken 1999)	'	-	-	w	-	-	-	-	-	e	-	a	-	-	-	-
WL (Arndt 1937:4-5)	'	ng	-	w	-	dz	-	j	e	é	-	a	-	-	-	-
WL-Alorese (Klamer 2011:30)	'	ng	-	-	-	j	-	y	-	e	-	a	-	-	-	-
WL-Lamalera (Keraf 1978a:xvi)	-	ng	-	-	-	j	-	-	e	é	-	a	-	-	-	-
WL-Lewoingu (Nishiyama and Kelen 2007:9)	'	ng	-	w	-	-	-	y	v	e	-	a	-	-	-	-
WL-Lewolema (Pampus 1999:17-18)	'	ng	-	w	-	j	-	-	ə	e	-	a	-	-	-	-
WL-Lewotobi (Nagaya 2011:91)	?	ŋ	-	w	-	j	-	y	ə	e	-	a	-	-	-	-
WL-Solor (Kroon 2016:35-36)	'	ng	-	w	-	j	-	y	e	é	-	a	-	-	-	-
Western Pantar (Holton 2014:27,35)	'	ng	-	-	-	-	-	y	-	é	-	a	-	-	-	-

C.2 List of glosses from other sources

This appendix contains a list of glosses from other sources that are different to the conventions of this dissertation. Therefore, in the examples in Part III, these glosses have been adapted to the conventions in this work. The table below can be used to retrieve the original glosses. Glosses which follow the same conventions in the original sources as in this dissertation are not listed in this table. A list of glosses used in this dissertation is found in the list of abbreviations on page xix. More information on glossing conventions is found in §7.3.

Table C.2: Glosses from other sources

Language and source	This thesis	Original	Meaning
Amarasi (Edwards 2016a)	3SG.POSS	3SG.GEN/3GEN	third person singular genitive
	DIST	2DEM	demonstrative ‘near’ ad- dressee
	1SG.POSS	3PL/1GEN	third person plural and first person genitive
CL-Atadei Painara (Krauß 2016)	3SG	3SG.PROCL	third person singular proclitic pronoun
	sidewards	inside.CIRC	circumfix meaning ‘in- side’
	NAME	[name spelled out]	a personal name
Dela (Rote) (Tamelan 2007)	NMLZ	Nom	nominaliser
Hawu (Walker 1982)	CLF	COUNT	classifier
	PROX	DEM2sg	a place in the immediate vicinity of the speaker

Language and source	This thesis	Original	Meaning
	PROX	DEM1sg	a specified point near the speaker
Helong (Balle 2017)	1PL.EXCL	1pex	first person plural exclusive
	NEG	not	negation
Kambera (Klamer 1998)	3SG.POSS	3SG	third person singular genitive
	1SG.NOM	1SN	first person singular nominative
	NAME	[name spelled out]	a personal name
KD-Leuwayang (Samely 1991a)	2SG.POSS	poss-2s	2nd person singular possessor
	1SG.POSS	poss-1s	first person singular possessor
	2SG.POSS	poss-2s	second person singular possessor
	3SG.POSS	poss-3s	third person singular possessor
	3SG	pp-3s	third person singular pronoun

Language and source	This thesis	Original	Meaning
Keo (Baird 2002)	1sg	pp-1s	first person singular pronoun
	1sg upwards	1s deic	first person singular deictic particle (that points to a higher location)
	downwards	deic	deictic particle (that points to a lower location)
	sidewards	deic	deictic particle (that points to a sideward location)
	hillwards	deic	deictic particle (that points to a hillward location)
	seawards	deic	deictic particle (that points to a seaward location)
	LOC	on/in	locative preposition
	CLF	CL	classifier

Language and source	This thesis	Original	Meaning
	NAME	[name spelled out]	a personal or place name
Makalero (Huber 2011)			
Makasae (Huber 2008)	1sg.POSS	1SG:POSS	first person singular pos- sessor
	CLF	CLASS	classifier
	PROX	DEM/be.here	proximate (demonstrat- ive)
Manggarai (Semium 1993)	POSS	GEN.	genitive; possessive
	1sg	1s (bound) / I (free)	first person singular
	3sg	3s	third person singular
	NAME	[name spelled out]	a personal name
	CLF	class.	classifier
Roma (Steven 1991)	exist	there_is	existential predicate
	3sg	3s	third person singular
Sika (several varieties) (Rosen 1986)	1PL.INCL	we incl.	first person plural includ- ive

Language and source	This thesis	Original	Meaning
Sika (Lewis and Grimes 1995)	1PL.INCL-go-1PL.INCL	direc. pt.	directional particle
	3SG-go	direc. pt.	directional particle
	3SG	she/he	third person singular
Sika (Lewis and Grimes 1995)	POSS	GEN	genitive
	NAME	[name spelled out]	a personal name
	LOW	DOWN	in downward direction
	upwards hillwards CLF	HIGH HILL CLASS	in upward direction in hillward direction classifier
Teiwa (Klamer 2010)	3SG	3s	third person singular
Termanu (Jonker 1908)	POSS	GEN	possessive
Tetun (Fehan) (van Klinken 1999)	POSS	POS	possessive
	3SG	3S	third person singular
	NAME	[name spelled out]	a personal name

Language and source	This thesis	Original	Meaning
WL-Lewotobi (Nagaya 2011)	NMLZ	NMZ	nominalizer
	DIST	DEM.DIS	distal demonstrative
	NAME	[name spelled out]	a personal name
WL-Lewoingu (Nishiyama and Kelen 2007)	DIST	at	distal demonstrative
	DIST	on	distal demonstrative
	3SG.POSS	his/her/3sg	third person singular pronoun
	NAME	[initials]	a personal name
WL-Solor (Kroon 2016)	LOC	prep	locative preposition
	DIST	there	distal demonstrative
	3SG.POSS	his	free third person singular pronoun
	1PL.INCL	1PL(inc)	first person plural inclusive
	go.up	ascend	to go in upward direction

Bibliography

- Abdurachman, Paramita. 2008. *Bunga angin Portugis di Nusantara: jejak-jejak kebudayaan Portugis di Indonesia*. Jakarta: LIPI Press & Asosiasi Persahabatan dan Kerjasama Indonesia-Portugal dan Yayasan Obor Indonesia.
- Aikhenvald, Alexandra, and Robert Dixon, ed. 2006. *Serial verb constructions: A cross-linguistic typology*. Oxford: Oxford University Press.
- Akoli, Marcelinus. 2010. Lamaholot dialect variations. Master thesis, Monash University, Melbourne.
- Antonius, Porat, and Abdul Ruskhan. 1997. *Struktur bahasa Ngadha dialek Rongga*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
- Aoki, Eriko, and Satoshi Nakagawa. 1993. Endenese-English Dictionary (unpublished). In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Aritonang, Jan, and Karel Steenbrink, ed. 2008. *A history of Christianity in Indonesia*. Leiden: Brill.
- Arka, Wayan, Fransiscus Seda, Antonius Gelang, Yohanes Nani, and Ivan Ture. 2007. A Rongga-English dictionary with an English-Rongga finderlist. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University

- Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Arndt, Paul. 1931. *Grammatik der Sika-Sprache*. Ende: Arnoldus.
- Arndt, Paul. 1937. *Grammatik der Solor-Sprache*. Ende: Arnoldus.
- Arndt, Paul. 1938. Demon und Padzi, die feindlichen Brueder des Solor-Archipels. *Anthropos* XXXIII:1–58.
- Arndt, Paul. 1940. *Soziale Verhältnisse auf Ost-Flores, Adonare und Solor*. Münster: Aschendorffsche Verlagsbuchhandlung.
- Arndt, Paul. 1951. *Religion auf Ostflores, Adonare und Solor*. Wien: Mission-druckerei St. Gabriel.
- Arndt, Paul. 1961. Wörterbuch der Ngadhasprache. Published in *Anthropos*. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Ataladjar, Thomas. 2015. *Pai hone tala la, pai wane tele pia lame lusi lako: dari tanah Nubanara menuju tanah misi*. Jakarta: Koker.
- Badan Pusat Statistik Kabupaten Lembata. 2018. *Kecamatan Atadei dalam angka 2018*. Lewoleba: Badan Pusat Statistik Kabupaten Lembata.
- Baird, Louise. 2002. A grammar of Keo. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Balle, Misriani. 2017. Types of reduplication in Helong, an Austronesian language in Eastern Indonesia. Doctoral Dissertation, Payap University, Chiang Mai.
- Barnes, Robert. 1974. *Kédang: A study of the collective thought of an eastern Indonesian people*. Oxford: Clarendon Press.

- Barnes, Robert. 1982. The Majapahit dependency Galiyao. *Bijdragen tot de Taal-, Land- en Volkenkunde* 138:407–412.
- Barnes, Robert. 1987. Weaving and non-weaving among the Lamaholot. *Indonesia Circle* 42:16–31.
- Barnes, Robert. 1996. *Sea hunters of Indonesia: Fishers and weavers of Lamalera*. Oxford: Clarendon Press.
- Barnes, Robert. 2013. *Excursions into Eastern Indonesia: Essays on History and Social Life*. New Haven: Yale Univ Southeast Asia Studies.
- Barnes, Ruth. 1989. *The ikat textiles of Lamalera: A study of an Eastern Indonesian weaving tradition*. Leiden: Brill.
- Barnes, Ruth. 2005. Moving between cultures. Textiles as a source of innovation in Kedang, eastern Indonesia. In *Textiles in Indian Ocean societies*, ed. Ruth Barnes, 150–162. London: RoutledgeCurzon.
- Bellwood, Peter. 2017. *First islanders: Prehistory and human migration in Island Southeast Asia*. Hoboken: John Wiley & Sons.
- Blevins, Juliette. 1995. The syllable in phonological theory. In *The handbook of phonological theory*, ed. John Goldsmith, 206–244. Cambridge: Blackwill.
- Blevins, Juliette. 1999. Untangling Leti infixation. *Oceanic Linguistics* 38:383–403.
- Blevins, Juliette. 2008. Consonant Epenthesis: Natural and Unnatural Histories. In *Linguistic Universals and Language Change*, ed. Jeff Good, 79–107. Oxford University Press.
- Blust, Robert. 1993. Central and Central-Eastern Malayo-Polynesian. *Oceanic Linguistics* 32:241–293.
- Blust, Robert. 2008. Is there a Bima-Sumba subgroup? *Oceanic Linguistics* 47:45–113.
- Blust, Robert. 2009. The Position of the Languages of Eastern Indonesia: A Reply to Donohue and Grimes. *Oceanic Linguistics* 48:36–77.

- Blust, Robert. 2013. *The Austronesian languages*. Canberra: Pacific Linguistics.
- Blust, Robert, and Stephen Trussel. 2010. The Austronesian Comparative Dictionary: Web Edition. <http://www.trussel2.com/acd/>.
- Bolscher, H. 1982. *Tati Lalang Sara Sikka*. Maumere: Lela.
- Brandes, Jan. 1884. *Bijdrage tot de vergelijkende klankleer der Westersche afdeling van de Maleisch-Polynesische taalfamilie*. Utrecht: Van de Weijer.
- Campbell, Lyle, and William Poser. 2008. How to show languages are related: The methods. In *Language Classification: History and Method*, ed. Lyle Campbell and William Poser, 162–223. Cambridge: Cambridge University Press.
- Croft, William. 1991. The Evolution of Negation. *Journal of Linguistics* 27:1–27.
- Culhane, Kirsten. 2019. Consonant insertions: A synchronic and diachronic account. Honours thesis, Australian National University, Canberra.
- Dixon, Robert. 1982. *Where have All the Adjectives Gone?, And Other Essays in Semantics and Syntax*. Berlin: Mouton de Gruyter.
- Donohue, Mark. 2003. Palu'e Dictionary. Electronic Edition 1.0. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Donohue, Mark. 2007. The Papuan Language of Tambora. *Oceanic Linguistics* 46:520–537.
- Donohue, Mark, and Charles Grimes. 2008. Yet More on the Position of the Languages of Eastern Indonesia and East Timor. *Oceanic Linguistics* 47:114–158.
- Doyle, Matthew. 2010. Internal divisions of the Flores-Lembata subgroup of Central Malayo-Polynesian. Master thesis, Leiden University, Leiden.

- Dryer, Matthew, and Martin Haspelmath. 2013. *WALS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/>.
- Edwards, Owen. 2016a. Metathesis and Unmetathesis: Parallelism and Complementarity in Amarasi, Timor. Doctoral Dissertation, Australian National University, Canberra.
- Edwards, Owen. 2016b. Parallel Sound Correspondences in Uab Meto. *Oceanic Linguistics* 55:52–86.
- Edwards, Owen. 2018a. Preliminary report on Funai Helong. *NUSA: Linguistic studies of languages in and around Indonesia* 65:1–27.
- Edwards, Owen. 2018b. Top-Down Historical Phonology of Rote-Meto. *Journal of the Southeast Asian Linguistics Society* 11:63–90.
- Edwards, Owen. 2019. Reintroducing Welaun. *Oceanic Linguistics* 58:31–58.
- Edwards, Owen. in prep. Rote-Meto comparative dictionary.
- Elias, Alexander. 2017a. The internal diversity of Lamaholot. Manuscript, Leiden University, Leiden.
- Elias, Alexander. 2017b. Lio Wordlist. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Elias, Alexander. 2017c. Nage Wordlist recorded in Ondorea Barat. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Elias, Alexander. 2017d. Wordlist of Oja Nga'o. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.

- Elias, Alexander. 2017e. Wordlist of Watumite Nga'o. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.modeling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Elias, Alexander. 2018. Lio and the Central Flores languages. Master thesis, Leiden University, Leiden.
- Ewing, Michael, and Marian Klamer, ed. 2010a. *East Nusantara: Typological and areal analyses*. Canberra: Pacific Linguistics.
- Ewing, Michael, and Marian Klamer. 2010b. The languages of East Nusantara: An introduction. In *East Nusantara: Typological and areal analyses*, ed. Michael Ewing and Marian Klamer, 1–24. Canberra: Pacific Linguistics.
- Fedden, Sebastian, Dunstan Brown, and Greville Corbett. 2010. Conditions on pronominal marking: A set of 42 video stimuli for field elicitation. University of Surrey. <http://dx.doi.org/10.15126/SMG.25/1>.
- Fernandez, Inyo. 1977. Bahasa Lamaholot Ile Mandiri. Bachelor thesis, Universitas Gadjah Mada, Yogyakarta.
- Fernandez, Inyo. 1996. *Relasi historis kekerabatan bahasa Flores: Kajian linguistik historis komparatif terhadap sembilan bahasa di Flores*. Ende: Nusa Indah.
- Fogaça, Helem Andressade de Oliveira. 2017. O ecossistema fundamental da língua Mambae: Aspectos endoecológicos e exoecológicos de uma língua Austronésia de Timor-Leste. Doctoral dissertation, Universidade de Brasília, Brasília.
- François, Alexandre, and Siva Kalyan. 2018. Freeing the Comparative Method from the tree model: A framework for Historical Glottometry. In *Let's talk about trees: Tackling Problems in Representing Phylogenetic Relationships among Languages*, ed. Ritsuko Kikusawa and Lawrence Reid, 59–89. Osaka: National Museum of Ethnology.
- Fricke, Hanna. 2014a. *Topics in the grammar of Hewa: A variety of Sika in Eastern Flores, Indonesia*. Munich: Lincom Europe.

- Fricke, Hanna. 2014b. Topics in the grammar of Hewa. Published at Lincom. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klammer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Fricke, Hanna. 2015a. Field notes on Lamaholot-Kalikasa. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klammer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Fricke, Hanna. 2015b. Field notes on Lamaholot-Lerek. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klammer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Fricke, Hanna. 2017. The rise of clause-final negation in Flores-Lembata, Eastern Indonesia. *Linguistics in the Netherlands* 34:47–62.
- Fricke, Hanna. 2019. Central Lembata Corpus. Nijmegen: The Language Archive, Max Planck Institute for Psycholinguistics. <https://hdl.handle.net/1839/c12bb123-1c4b-47cf-8d21-f185daf0557e>.
- Fricke, Hanna, Marian Klammer, and Yunus Sulistyono. 2019. Hewa Collection (2013-2015). Nijmegen: The Language Archive, Max Planck Institute for Psycholinguistics. <https://hdl.handle.net/1839/851c8532-4a32-47b1-aa3b-62dd91cecla4>.
- Grangé, Philippe. 2015a. The expression of possession in some languages of the eastern Lesser Sunda islands. *Masyarakat Linguistik Indonesia* 33:35–51.
- Grangé, Philippe. 2015b. The Lamaholot dialect chain (East Flores, Indonesia). In *Language change in Austronesian languages: Papers from 12-ICAL*, ed. Malcolm Ross and Wayan Arka. Canberra: Asia-Pacific Linguistics.
- Grimes, Charles, Bernadus Lado, Thomas Ly, and Simon Tari. 2008. Lii Hawu (Sabu) Online Dictionary. Kupang: UBB-GMIT. In *LexiRumah 2.0.0.*

- (2019), ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.modeling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Grimes, Charles, Carlos Marçal, and Paolino Viana Fereira. 2014. *Introductory Dictionary of Mambae (Same)*. Kupang: Unit Bahasa & Budaya, Kupang (UBB) and Australian Society for Indigenous Languages (AuSIL).
- Hägerdal, Hans. 2012. *Lords of the land, lords of the sea: Conflict and adaptation in early colonial Timor, 1600 - 1800*. Leiden: KITLV Press.
- Heine, Bernd, and Tania Kuteva. 2003. On contact-induced grammaticalization. *Studies in Language* 27:529–572.
- Heine, Bernd, and Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.
- Heston, Tyler. 2015. The segmental and suprasegmental phonology of Fataluku. Doctoral Dissertation, University of Hawai'i.
- Himmelman, Nikolaus. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In *The Austronesian languages of Asia and Madagascar*, 110–181. London: Routledge.
- Himmelman, Nikolaus, John Bowden, Maurício Belo, John Hajek, Alex Tilman, and Alex Freitas. 2006. Waima'a Lexical database. <https://hdl.handle.net/1839/00-0000-0000-0008-3458-8>.
- Holm, John. 2000. *An introduction to pidgin and creoles*. Cambridge: Cambridge University Press.
- Holton, Gary. 2014. Western Pantar. In *The Papuan Languages of Timor, Alor and Pantar: Volume 1*, ed. Antoinette Schapper, 23–95. Berlin: Mouton de Gruyter.
- Holton, Gary, and Marian Klamer. 2017. The Papuan languages of East Nusantara and the Bird's Head. In *The Languages and Linguistics of the New Guinea Area: A Comprehensive Guide*, ed. Bill Palmer, 569–640. Berlin: Mouton de Gruyter.

- Holton, Gary, Marian Klamer, František Kratochvíl, Laura Robinson, and Antoinette Schapper. 2012. The historical relations of the Papuan languages of Alor and Pantar. *Oceanic Linguistics* 51:86–122.
- Holton, Gary, and Laura Robinson. 2017a. The internal history of the Alor-Pantar language family. In *The Alor-Pantar languages: History and typology*, ed. Marian Klamer, 49–91. Berlin: Language Science Press.
- Holton, Gary, and Laura Robinson. 2017b. The linguistic position of the Timor-Alor-Pantar languages. In *The Alor-Pantar languages: History and typology*, ed. Marian Klamer, 147–191. Berlin: Language Science Press.
- Huber, Juliette. 2008. *First steps towards a grammar of Makasae : A language of East Timor*. Munich: Lincom Europa.
- Huber, Juliette. 2011. A grammar of Makalero: A Papuan language of East Timor. Doctoral dissertation, Leiden University, Leiden.
- Hudjashov, Georgi, Tatiana Karafet, Daniel Lawson, Sean Downey, Olga Savina, Herawati Sudoyo, Stephen Lansing, Michael Hammer, and Murray Cox. 2017. Complex Patterns of Admixture across the Indonesian Archipelago. *Molecular Biology and Evolution* 34:2439–2452.
- Hyman, Larry. 2001. Fieldwork as a state of mind. In *Linguistic Fieldwork*, ed. Paul Newman and Martha Ratliff, 15–33. Cambridge: Cambridge University Press.
- Ismail, Mansyur. 1985. Kamus Bima-Indonesia. Published at Pusat Pembinaan dan Pengembangan Bahasa Departemen Pendidikan dan Kebudayaan Jakarta. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Jacob, June, and Charles Grimes. 2011. Aspect and directionality in Kupang Malay serial verb constructions: Calquing on the grammars of substrate languages. In *Creoles, their substrates, and language typology*, ed. Claire Lefebvre, 337–366. Amsterdam: John Benjamins.
- Jonker, J. 1908. *Rottineesch-Hollandsch woordenboek*. Leiden: Brill.

- Jonker, J. 1915. *Rottineesche spraakkunst*. Leiden: Brill.
- Kaiping, Gereon, Owen Edwards, and Marian Klamer. 2019. LexiRumah 2.0.0. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Kaiping, Gereon, and Marian Klamer. 2018. LexiRumah: An online lexical database of the Lesser Sunda Islands. *PloS one* 13:e0205250.
- Kaiping, Gereon, and Marian Klamer. to appeara. CultureRumah. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/culturerrumah/>.
- Kaiping, Gereon, and Marian Klamer. to appearb. Subgrouping the Timor-Alor-Pantar languages using systematic Bayesian inference. *Language Dynamics and Change*.
- Karafet, Tatiana, Brian Hallmark, Murray Cox, Herawati Sudoyo, Sean Downey, Stephen Lansing, and Michael Hammer. 2010. Major East–West Division Underlies Y Chromosome Stratification across Indonesia. *Molecular Biology and Evolution* 27:1833–1844.
- Keraf, Gregorius. 1978a. *Morfologi Dialek Lamalera*. Ende: Arnoldus.
- Keraf, Gregorius. 1978b. Morfologi Dialek Lamalera. Doctoral dissertation at Universitas Indonesia. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Klamer, Marian. 1998. *A grammar of Kambara*. Berlin: Mouton de Gruyter.
- Klamer, Marian. 2000. How report verbs become quote markers and complementisers. *Lingua* 110:69–98.
- Klamer, Marian. 2002. Typical features of Austronesian languages in Central/Eastern Indonesia. *Oceanic Linguistics* 41:363–383.

- Klamer, Marian. 2008. The semantics of semantic alignment in eastern Indonesia. In *The typology of semantic alignment*, ed. Mark Donohue and Søren Wichmann, 221–251. Oxford: Oxford University Press.
- Klamer, Marian. 2010. *A grammar of Teiwa*. Berlin: Mouton de Gruyter.
- Klamer, Marian. 2011. *A short grammar of Alorese (Austronesian)*. Munich: Lincom Europe.
- Klamer, Marian. 2012a. Papuan-Austronesian language contact: Alorese from an areal perspective. In *Melanesian languages on the Edge of Asia: Challenges for the 21st Century*, ed. Marian Klamer and Nicholas Evans, 72–108. Honolulu: University of Hawai'i Press.
- Klamer, Marian. 2012b. Tours of the past through the present of eastern Indonesia. *Language Documentation & Conservation Special Publication* 6:54–63.
- Klamer, Marian. 2014a. The history of numeral classifiers in Teiwa (Papuan). In *Number - Constructions and semantics: Case studies from Africa, Amazonia, India and Oceania*, ed. Anne Storch and Gerrit Dimmendaal, 135–166. Amsterdam: John Benjamins.
- Klamer, Marian. 2014b. Kaera. In *The Papuan Languages of Timor, Alor and Pantar: Volume 1*, ed. Antoinette Schapper, 97–146. Berlin: Mouton de Gruyter.
- Klamer, Marian. 2014c. Numeral classifiers in the Papuan languages of Alor and Pantar: A comparative perspective. In *Number and quantity in East Nusantara: Papers from I2-ICAL*, ed. Marian Klamer and František Kratochvíl, 103–122. Canberra: Asia-Pacific Linguistics.
- Klamer, Marian. 2015a. Field notes on Hewa. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.modeling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Klamer, Marian. 2015b. Field notes on Kedang. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden:

- Leiden University Centre for Linguistics. <https://lexirumah.modeling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Klamer, Marian. 2015c. Field notes on Lamaholot-Adonara. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.modeling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Klamer, Marian. 2015d. Field notes on Lamaholot-Lewoingu. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.modeling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Klamer, Marian, ed. 2017. *The Alor-Pantar languages: History and typology*. Berlin: Language Science Press.
- Klamer, Marian. 2019. The dispersal of Austronesian languages in Island South East Asia: Current findings and debates. *Language and Linguistics Compass* 13:1–26.
- Klamer, Marian. to appear. From Lamaholot to Alorese: Morphological loss in adult language contact. In *Austronesian Undressed: How and why languages become isolating*, ed. David Gil and Antoinette Schapper. Amsterdam: John Benjamins.
- Klamer, Marian, Ger Reesink, and Miriam van Staden. 2008. East Nusantara as a linguistic area. In *From linguistic areas to areal linguistics*, ed. Pieter Muysken, Werner Abraham, and Michael Noonan, 95–149. Amsterdam: John Benjamins.
- Klamer, Marian, Antoinette Schapper, and Greville Corbett. 2017. Plural number words in the Alor-Pantar languages. In *The Alor-Pantar languages: History and typology*, ed. Marian Klamer, 365–403. Berlin: Language Science Press.
- Kohl, Karl-Heinz. 1998. *Der Tod der Reisjungfrau: Mythen, Kulte und Allianzen in einer ostindonesischen Lokalkultur*. Stuttgart: Kohlhammer.

- Kratochvíl, František. 2007. A grammar of Abui: A Papuan language of Alor. Doctoral dissertation, Leiden University, Leiden.
- Kratochvíl, František. 2014. Sawila. In *The Papuan Languages of Timor, Alor and Pantar: Volume 1*, ed. Antoinette Schapper, 351–438. Berlin: Mouton de Gruyter.
- Krauße, Daniel. 2016. A brief grammar of the Eastern Atadei language of Lembata, Indonesia. *Linguistik Indonesia*, 34:113–128.
- Kroon, Yosep. 2016. A grammar of Solor Lamaholot: A language of Flores, Eastern Indonesia. Doctoral dissertation, University of Adelaide, Adelaide.
- Le Roux, C. 1928. *De Elcano's tocht door den Timorarchipel met Magalhaes' schip 'Victoria'*. Weltevreden: Kolff.
- Levinson, Stephen. 2003. *Space in language and cognition: Explorations in cognitive diversity*. Cambridge: Cambridge University Press.
- Lewis, Douglas. 1988. *People of the source: The social and ceremonial order of Tana Wai Brama on Flores*. Dordrecht: Foris.
- Lewis, Douglas. 1995. Sika wordlist. Published in the Comparative Austronesian Dictionary by Tryon. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Lewis, Douglas. 1998. The tyranny of the text: Oral tradition and the power of writing in Sikka and Tana 'Ai, Flores. *Bijdragen tot de Taal-, Land-en Volkenkunde* 154:457–477.
- Lewis, Douglas. 2006. Ritual and reflexes of lost sovereignty in Sikka: A regency of Flores in eastern Indonesia. *Bijdragen tot de Taal-, Land-en Volkenkunde* 162:306–335.
- Lewis, Douglas, and Charles Grimes. 1995. Sika. In *Comparative Austronesian Dictionary*, ed. Darrell Tryon and Malcolm Ross, 601–609. Berlin: Mouton de Gruyter.

- Max Planck Institute for Evolutionary Anthropology. 2015. The Leipzig Glossing Rules. <https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf>.
- Michaelis, Susanne. 2008. *Roots of creole structures weighing the contribution of substrates and superstrates*. Amsterdam: John Benjamins.
- Michels, Marc. 2017. Western Lamaholot. A cross-dialectal grammar sketch. Master thesis, Leiden University, Leiden.
- Moro, Francesca. 2016a. Field notes on Alorese-Alor-Basar. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Moro, Francesca. 2016b. Field notes on Alorese-Baranusa. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Moro, Francesca. 2016c. Field notes on Alorese-Munaseli. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Moro, Francesca. 2016d. Field notes on Alorese-Pandai. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Moro, Francesca. 2018. The Plural Word hire in Alorese: Contact-Induced Change from Neighboring Alor-Pantar Languages. *Oceanic Linguistics* 57:22.
- Moro, Francesca. 2019. Loss of Morphology in Alorese (Austronesian): Simplification in Adult Language Contact. *Journal of Language Contact* 12:378–403.

- Moro, Francesca, and Hanna Fricke. submitted. Contact-induced change in Alorese give-constructions. *Oceanic Linguistics*.
- Muysken, Pieter. 2008. Code-switching and code-mixing. In *Functional Categories*, 163–176. Cambridge: Cambridge University Press.
- Muysken, Pieter. 2010. Scenarios for language contact. In *The Handbook of Language Contact*, ed. Raymond Hickey. Malden: Wiley-Blackwell.
- Nagaya, Naonori. 2011. The Lamaholot language of Eastern Indonesia. Doctoral dissertation, Rice University, Houston.
- Nishiyama, Kunio, and Herman Kelen. 2007. *A grammar of Lamaholot, Eastern Indonesia: The morphology and syntax of the Lewoingu dialect*. Munich: Lincom Europe.
- Onvlee, L. 1984. *Kamberaas (Oost-Soembaas) - Nederlands woordenboek met Nederlands - Kamberaas register*. Dordrecht: Foris.
- Paauw, Scott. 2008. The Malay contact varieties of Eastern Indonesia: A typological comparison. Doctoral dissertation, State University of New York, Buffalo.
- Pampus, Karl-Heinz. 1999. *Koda kiwā: Dreisprachiges Wörterbuch der Lamaholot-Sprache, Dialekt von Lewolema*. Stuttgart: Deutsche Morgenländische Gesellschaft Steiner.
- Pareira, Mandalangi, and Douglas Lewis. 1998. *Kamus sara Sikka bahasa Indonesia*. Ende: Nusa Indah.
- Parker, Stephen. 2002. Quantifying the sonority hierarchy. Doctoral dissertation, University of Massachusetts, Amherst.
- Pawley, Andrew. 2005. *Papuan pasts: Cultural, linguistic and biological histories of Papuan-speaking peoples*. Canberra: Pacific Linguistics.
- Payne, Thomas. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.

- Pigeaud, Theodore. 1960. *Java in the 14th century: a study in cultural history: the Nāgara-Kĕrtāgama by Rakawi, prapañca of Majapahit, 1365 A.D. II: Notes on the texts and the translations (3rd edition)*. The Hague: Nijhoff.
- Pigeaud, Theodore. 1962. *Java in the 14th century: A study in cultural history: The Nāgara-Kĕrtāgama by Rakawi, prapañca of Majapahit, 1365 A.D. IV: Commentaries and recapitulation (3rd edition)*. The Hague: Nijhoff.
- Rappoport, Dana. 2014. Songs and sorrow in Tanjung Bunga: Music and the myth of the origin of rice (Lamaholot, Flores, Indonesia). *Bijdragen tot de Taal-, Land- en Volkenkunde* 170:215–249.
- Reesink, Ger. 2002. Clause-final negation: Structure and interpretation. *Functions of Language* 9:239–268.
- Reesink, Ger, and Michael Dunn. 2018. Contact phenomena in Austronesian and Papuan languages. In *The languages and linguistics of the New Guinea area: A comprehensive guide*, ed. Bill Palmer, 939–983. Berlin: Mouton de Gruyter.
- Robbeets, Martine, and Alexander Savelyev. 2017. *Language Dispersal Beyond Farming*. Amsterdam: John Benjamins.
- Robinson, Laura. 2015. The Alor-Pantar (Papuan) languages and Austronesian contact in East Nusantara. In *Language change in Austronesian languages: Papers from I2-ICAL*, ed. Malcolm Ross and Wayan Arka, 19–33. Canberra: Asia-Pacific Linguistics.
- Robinson, Laura C., and John Haan. 2014. Adang. In *The Papuan Languages of Timor, Alor and Pantar: Volume 1*, ed. Antoinette Schapper, 221–283. Berlin: Mouton de Gruyter.
- Rosen, Joan. 1986. Phonemes, verb classes and personal endings in Maumere. *NUSA: Linguistic studies of languages in and around Indonesia* 25:39–69.
- Ross, Malcolm. 1992. The Sound of Proto-Austronesian: An Outsider's View of the Formosan Evidence. *Oceanic Linguistics* 31:23–64.

- Ross, Malcolm. 2013. Diagnosing contact processes from their outcomes: The importance of life stages. *Journal of Language Contact* 6:5–47.
- Samely, Ursula. 1991a. *Kedang (Eastern Indonesia): Some aspects of its grammar*. Hamburg: H. Buske.
- Samely, Ursula. 1991b. Kedang (Eastern Indonesia): Some aspects of its grammar. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.model-ling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Samely, Ursula, and Robert Barnes. 2013. *A dictionary of the Kedang language: Kedang-Indonesian-English*. Leiden: Brill.
- Sapir, Edward. 1916. *Time perspective in aboriginal American culture: A study in method*. Ottawa: Government Printing Bureau.
- Sassoon, Galit. 2013. *Vagueness, Gradability and Typicality: The interpretation of adjectives and nouns*. Leiden: Brill.
- Schapper, Antoinette. 2009. *Bunaq: A Papuan language of central Timor*. Doctoral dissertation, Australian National University, Canberra.
- Schapper, Antoinette. 2014a. Kamang. In *The Papuan Languages of Timor, Alor and Pantar: Volume 1*, ed. Antoinette Schapper, 285–350. Berlin: Mouton de Gruyter.
- Schapper, Antoinette, ed. 2014b. *The Papuan Languages of Timor, Alor and Pantar: Volume 1*. Berlin: Mouton de Gruyter.
- Schapper, Antoinette. 2015. Wallacea, a linguistic area. *Archipel* 90:99–151.
- Schapper, Antoinette. 2017. Elevation in the spatial deictic systems of Alor-Pantar languages. In *The Alor-Pantar languages: History and typology*, ed. Marian Klamer, 239–277. Berlin: Language Science Press.
- Schapper, Antoinette, and Harald Hammarström. 2013. Innovative Numerals in Malayo-Polynesian Languages outside of Oceania. *Oceanic Linguistics* 52:423–456.

- Schapper, Antoinette, and Rachel Hendery. 2014. Wersing. In *The Papuan Languages of Timor, Alor and Pantar: Volume 1*, ed. Antoinette Schapper, 439–503. Berlin: Mouton de Gruyter.
- Schapper, Antoinette, Juliette Huber, and Aone van Engelenhoven. 2012. The historical relation of the Papuan languages of Timor and Kisar. *Language & Linguistics in Melanesia: Journal of the Linguistic Society of Papua New Guinea* Special Issue:192–240.
- Schapper, Antoinette, Juliette Huber, and Aone van Engelenhoven. 2017. The relatedness of Timor-Kisar and Alor-Pantar languages: A preliminary demonstration. In *The Alor-Pantar languages: History and typology*, ed. Marian Klamer, 91–147. Berlin: Language Science Press.
- Schapper, Antoinette, and Marian Klamer. 2017. Numeral systems in the Alor-Pantar languages. In *The Alor-Pantar languages: History and typology*, ed. Marian Klamer, 277–329. Berlin: Language Science Press.
- Semiun, Agustinus. 1993. The basic grammar of Manggarai: Kempo subdialect. Doctoral dissertation, La Trobe University, Bundoora.
- Shopen, Timothy. 2007. *Language Typology and Syntactic Description: Complex Constructions (Volume II)*. Cambridge: Cambridge University Press.
- Sneddon, James. 2003. *The Indonesian language: its history and role in modern society*. Sydney: UNSW Press.
- Steinhauer, Hein. 1991. Malay in east Indonesia: The case of Lantuka (Flores). *Pacific Linguistics. Series A. Occasional Papers*. Canberra 81:177–195.
- Steinhauer, Hein. 2014. Blagar. In *The Papuan Languages of Timor, Alor and Pantar: Volume 1*, ed. Antoinette Schapper, 147–219. Berlin: Mouton de Gruyter.
- Steven, Lee. 1991. The phonology of Roma, an Austronesian language of eastern Indonesia. Master thesis, University of Texas, Arlington.
- Stevens, Alan, and Aluzius Schmidgall-Tellings. 2004. *A comprehensive Indonesian-English dictionary*. Athens: Ohio University Press.

- Stokhof, W. 1975. *Preliminary notes on the Alor and Pantar languages (East Indonesia)*. Canberra: Pacific Linguistics.
- Tadmor, Uri, Martin Haspelmath, and Bradley Taylor. 2010. Borrowability and the notion of basic vocabulary. *Diachronica* 27:226–246.
- Tamelan, Thersia. 2007. An overview of Dela-Oenale verb morphology: A language spoken in western Rote. Manuscript, Artha Wacana Christian University, Kupang.
- Thomason, Sarah. 2001. *Language contact*. Edinburgh: Edinburgh University Press.
- Thomason, Sarah, and Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- van der Auwera, Johan. 2009. The Jespersen Cycles. In *Cyclical change*, ed. Elly van Gelderen, 35–71. Amsterdam: John Benjamins.
- van der Auwera, Johan, and Denies Du Mon. 2015. Negation, Linguistics of. In *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*, ed. James Wright, 409–414. Oxford: Elsevier.
- van der Auwera, Johan, and Frens Vossen. 2015. Negatives between Chamic and Bahnaric. *Journal of the Southeast Asian Linguistics Society* 8:24–38.
- van der Auwera, Johan, and Frens Vossen. 2016. Jespersen cycles in the Mayan, Quechuan and Maipurean languages. In *Cyclical change continued*, ed. Elly van Gelderen. Amsterdam: John Benjamins.
- van Engelenhoven, Aone. 2009. On derivational processes in Fataluku, a non- Austronesian language in East-Timor. In *The Linguistics of Endangered Languages: Contributions to Morphology and Morpho-Syntax*, ed. Leo Wetzels, 333–362. Utrecht: LOT.
- van Gelderen, Elly. 2008. Negative cycles. *Linguistic Typology* 12:195–243.
- van Klinken, Catherina. 1999. *A grammar of the Fehan dialect of Tetun. An Austronesian language of West Timor*. Canberra: Pacific Linguistics.

- Vatter, Ernst. 1932. *Ata Kiwan: unbekannte Bergvölker im tropischen Holland*. Leipzig: Bibliographisches Institut.
- Veloso, Alexandre. 2016. A grammar sketch of Naueti, a language of East Timor. Master thesis, Leiden University, Leiden.
- Verheijen, Jilis. 1967. Kamus Manggarai: Manggarai-Indonesia & Indonesia-Manggarai. Published at Nijhoff in The Hague. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.modeling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Verheijen, Jilis. 1982. Komodo: Het Eiland, Het Volk en De Taal. Published at Nijhoff in The Hague. In *LexiRumah 2.0.0. (2019)*, ed. Gereon Kaiping, Owen Edwards, and Marian Klamer. Leiden: Leiden University Centre for Linguistics. <https://lexirumah.modeling.eu/lexirumah/>. DOI <https://doi.org/10.5281/zenodo.1164782>.
- Vossen, Frens. 2016. On the typology of the Jespersen Cycles. Doctoral dissertation, Universiteit van Antwerpen, Antwerpen.
- Vossen, Frens, and Johan van der Auwera. 2014. The Jespersen Cycles Seen from Austronesian. In *The diachrony of negation*, ed. Maj-Britt Mosegaard Hansen and Jacqueline Visconti, 47–83. Amsterdam: John Benjamins.
- Walker, Alan. 1982. *A Grammar of Sawu*. Jakarta: NUSA.
- Wellfelt, Emilie. 2016. *Historyscapes in Alor: Approaching indigenous histories in eastern Indonesia*. Växjö: Linnaeus University.
- Wierzbicka, Anna. 1988. *The Semantics of Grammar*. Amsterdam: John Benjamins.
- Wittenburg, Peter, Hennie Brugman, Albert Russel, Alex Klassmann, and Han Sloetjes. 2006. ELAN: A Professional Framework for Multimodality Research. *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation* 1556–1559.

Wolff, John. 2010. *Proto-Austronesian phonology with glossary: Volume I*. Ithaca: Cornell Southeast Asia Program Publications.

Wu, Jiang. 2017. Plural words in Austronesian languages: Typology and History. Master thesis, Leiden University, Leiden.

English summary

The overall aim of this dissertation is to reconstruct the history of the Flores-Lembata languages including traces of contact-induced change. The Flores-Lembata languages are a lower-level subgroup within the Austronesian language family spoken in eastern Indonesia.

Part I of this dissertation fills a gap in the documentation of the Flores-Lembata languages by providing a descriptive grammar of the language Central Lembata. This description is the first extensive description of a language belonging to the Central Lamaholot subgroup of Flores-Lembata. Part II concerns the history of the phonology and the lexicon of the Flores-Lembata languages with the aim of providing evidence, on the one hand, for inherited Austronesian vocabulary in the Flores-Lembata languages, and on the other hand, for the presence of a non-Austronesian lexical substrate. Part III examines morpho-syntactic features and their history of contact. Eight features of the Flores-Lembata languages which are atypical for Austronesian languages are described and evaluated on their potential of being the result of contact with non-Austronesian languages of the area.

This dissertation shows that the Flores-Lembata languages are, genealogically, closely related to the Austronesian languages located to the west of Flores-Lembata. The Flores-Lembata languages form an innovation-defined subgroup together with their western neighbours on Flores, as well as further Austronesian languages on the islands of Sumba and of Bima (based on lenition of Proto-Malayo-Polynesian (PMP) *b > w). The lexicon of individual Flores-Lembata subgroups contains a considerable component of vocabulary that cannot be reconstructed to an Austronesian ancestor (up

to 50% in individual languages). This vocabulary most likely entered the languages due to contact with now extinct non-Austronesian languages. Typologically, especially concerning morpho-syntax, the Flores-Lembata languages are mixed. In addition to inherited Austronesian features, these languages also share features with their eastern neighbours of the non-Austronesian Timor-Alor-Pantar family. In the morpho-syntax of the Flores-Lembata languages, a number of contact-induced innovations are found. The nominal domain appears to be affected in particular. On the one hand, the word order in the noun phrase, concerning the position of the nominal possessor, the numeral and the locative noun, differs from typical Austronesian languages. On the other hand, new semantic distinctions, such as an alienability distinction in possessive constructions or the plural number marking on nouns, become grammaticalized in a subset of the Flores-Lembata languages. Also the clausal syntax underwent changes. In contrast to the inherited pre-predicate negation and the SVO word order, some of the Flores-Lembata languages have innovated clause-final negation and all Flores-Lembata languages have innovated a set of clause-final deictic motion verbs. The word order features of the noun phrase and two of the deictic motion verbs can be reconstructed to Proto-Flores-Lembata. All other innovated features appear to have entered the languages after the break-up of Flores-Lembata into subgroups. All these innovated features in the Flores-Lembata languages are also attested in the non-Austronesian Alor-Pantar languages, which are spoken on two adjacent islands to the east, and these features can be reconstructed to Proto-Alor-Pantar.

Combining lexical and typological evidence, I propose that the Flores-Lembata languages have been in contact with languages typologically similar to the Alor-Pantar languages since the time of Proto-Flores-Lembata until after the break-up of the family into subgroups. Some subgroups have gained more non-Austronesian features than others. This suggests that there was variation in length and intensity of contact among the subgroups.

Nederlandse samenvatting

Het centrale doel van deze dissertatie is om de geschiedenis van de Flores-Lembatatalen te reconstrueren. Hierbij wordt ook gekeken naar taalverandering door taalcontact. De Flores-Lembatatalen vormen een subgroep van de Austronesische taalfamilie en worden gesproken in Oost-Indonesië.

Deel I van deze dissertatie vult de bestaande documentatie van de Flores-Lembatatalen aan door een beschrijvende grammatica van Centraal-Lembata te presenteren. Dit is de eerste uitgebreide beschrijving van een taal die behoort tot de Centraal-Lamaholotsubgroep van Flores-Lembata. Deel II behandelt de geschiedenis van zowel de fonologie als het lexicon van de Flores-Lembatatalen. Dit deel laat zien dat de Flores-Lembatatalen niet alleen geërfd Austronesisch vocabulaire bevatten, maar ook een niet-Austronesisch lexicaal substraat. Deel III onderzoekt morfosyntactische kenmerken en hun geschiedenis met betrekking tot taalcontact. Er worden acht kenmerken van de Flores-Lembatatalen beschreven die atypisch zijn voor Austronesische talen. Ook wordt geëvalueerd of het mogelijk is dat deze kenmerken voort zijn gekomen uit taalcontact met niet-Austronesische talen uit de regio.

Deze dissertatie toont aan dat de Flores-Lembatatalen genealogisch gezien nauw verwant zijn aan de Austronesische talen ten westen van Flores-Lembata. De Flores-Lembatatalen vormen samen met hun westerburen op Flores en andere Austronesische talen op de eilanden Sumba en Bima een subgroep op basis van innovatie (lenitie van Proto-Maleis-Polynesisch (PMP) *b > w). Het lexicon van de individuele Flores-Lembatataalgroepen bevat een aanzienlijk deel (tot wel 50% in afzonderlijke talen) dat niet herleid

kan worden tot een Austronesische voorouder. Dit vocabulaire is hoogstwaarschijnlijk geïntroduceerd onder invloed van taalcontact met niet-Austronesische talen die nu uitgestorven zijn.

Typologisch gezien, vooral waar het gaat om morfosyntaxis, zijn de Flores-Lembatatalen gemengd van aard. Naast geërfde Austronesische eigenschappen zijn er ook kenmerken die ze delen met hun oosterburen van de niet-Austronesische Timor-Alor-Pantarfamilie. In de Flores-Lembatatalen zijn er diverse morfosyntactische innovaties aan te wijzen die het gevolg zijn van taalcontact. Het nominale domein lijkt in dezen het meest beïnvloed te zijn. Enerzijds verschilt de woordvolgorde in de nominale constituent (NP) met die van typisch Austronesische talen; met name wat betreft de nominal bezitter, het telwoord en het locatieve naamwoord. Anderzijds zijn er nieuwe semantische categorieën gegrammaticaliseerd in een subset van de Flores-Lembatatalen: vervreemdbaar en onvervreemdbaar bezit wordt van elkaar onderscheiden, en het verschil tussen enkelvoud en meervoud wordt gemarkeerd op naamwoorden. Ook de syntaxis op zinsniveau heeft veranderingen ondergaan. In tegenstelling tot de geërfde situatie waarin de negatie aan het predicaat voorafgaat en het werkwoord tussen het subject en het object in staat (SVO-volgorde), hebben sommige Flores-Lembatatalen zinsfinale negatie ontwikkeld en hebben alle Flores-Lembatatalen een set van zinsfinale deictische werkwoorden van beweging gekregen. De woordvolgorde in de nominale constituent en twee van de deictische werkwoorden van beweging kunnen worden herleid tot Proto-Flores-Lembata. Alle andere innovaties lijken na de opsplitsing van Flores-Lembata plaatsgevonden te hebben. Al deze innovatieve kenmerken in de Flores-Lembatatalen zijn ook geattesteerd in de niet-Austronesische Alor-Pantartalen, die op twee aangrenzende eilanden in het oosten gesproken worden; deze kenmerken kunnen worden gereconstrueerd voor Proto-Alor-Pantar.

Op basis van een combinatie van lexicaal en typologisch bewijsmateriaal stel ik voor dat de Flores-Lembatatalen in contact moeten hebben gestaan met talen die typologisch gezien sterk lijken op de Alor-Pantartalen. Dit contact moet zowel voor als na de opsplitsing van Proto-Flores-Lembata plaats hebben gevonden. Sommige subgroepen van Flores-Lembatatalen hebben meer niet-Austronesische kenmerken verworven dan andere. Dit suggereert dat het taalcontact tussen de subgroepen varieerde in duur en intensiteit.

Ringkasan dalam bahasa Indonesia

Disertasi ini disusun dengan tujuan untuk merekonstruksi sejarah kelompok bahasa-bahasa Flores-Lembata, termasuk menelusuri jejak-jejak perubahan yang disebabkan oleh kontak bahasa. Bahasa-bahasa Flores-Lembata di Indonesia timur adalah sub-kelompok bahasa yang berada pada tingkat lebih rendah dalam rumpun bahasa Austronesia.

Bagian I disertasi ini menambahkan dokumentasi bahasa-bahasa Flores-Lembata dengan menyiapkan suatu deskripsi tata bahasa dari bahasa *Central Lembata* (bahasa Atadei Demon). Deskripsi tata bahasa ini adalah yang pertama kali mendeskripsikan secara luas salah satu bahasa yang termasuk dalam sub-kelompok Lamaholot Tengah di dalam bahasa-bahasa Flores-Lembata. Bagian II disertasi ini membahas fonologi historis dan leksikon bahasa-bahasa Flores-Lembata. Di satu sisi, pembahasan ini bertujuan untuk memberikan bukti keberadaan kosakata Austronesia yang diwariskan dalam bahasa-bahasa Flores-Lembata. Namun, di sisi lain, pembahasan ini juga membuktikan kehadiran substratum leksikal dari bahasa-bahasa non-Austronesia di dalam bahasa-bahasa Flores-Lembata. Sementara itu, bagian III disertasi ini mengkaji fitur-fitur morfo-sintaksis dan riwayat kontak bahasa dari fitur-fitur tersebut. Terdapat delapan fitur dalam bahasa-bahasa Flores-Lembata yang dianggap tidak lazim untuk bahasa-bahasa Austronesia pada umumnya. Fitur-fitur tersebut dideskripsikan dan dievaluasi potensinya sebagai hasil dari pertemuannya dengan bahasa-bahasa non-Austronesia di wilayah Kepulauan Flores-Lembata.

Disertasi ini menunjukkan bahwa bahasa-bahasa Flores-Lembata, secara silsilah, berkerabat erat dengan bahasa Austronesia yang terletak di

sebelah barat Kepulauan Flores-Lembata. Bahasa-bahasa Flores-Lembata dengan kerabatnya di sebelah barat, yaitu di Flores, serta kerabatnya yang lain, yaitu bahasa-bahasa yang dituturkan di Pulau Sumba dan Bima, membentuk sub-kelompok berdasarkan inovasi bersama (lenisi Proto-Malayo-Polinesia (PMP) *b>w). Namun, leksikon masing-masing bahasa di sub-kelompok Flores-Lembata berisi komponen kosakata yang tidak dapat direkonstruksi ke dalam suatu bahasa moyang Austronesia (persentase kekerabatan hingga 50% dalam beberapa bahasa). Kosakata ini kemungkinan besar masuk ke bahasa-bahasa Flores-Lembata karena kontak dengan bahasa-bahasa non-Austronesia yang sekarang sudah punah.

Secara tipologis, khususnya menyangkut fitur-fitur morfo-sintaksis, bahasa-bahasa Flores-Lembata adalah hasil percampuran. Selain ada fitur-fitur Austronesia yang diwarisi, bahasa-bahasa Flores-Lembata juga memiliki fitur-fitur yang mirip dengan tetangga bahasa-bahasa non-Austronesia, yaitu bahasa-bahasa Timor-Alor-Pantar. Di dalam aspek morfo-sintaksis bahasa-bahasa Flores-Lembata, ditemukan sejumlah inovasi yang disebabkan oleh kontak dengan bahasa-bahasa non-Austronesia. Dari aspek-aspek morfo-sintaksis ini, domain nominal adalah yang mendapat pengaruh. Di satu sisi, urutan kata dalam frasa benda, seperti pada penentuan letak *nominal possessor*, numeral dan kata benda lokatif, berbeda dengan urutan kata dalam bahasa Austronesia pada umumnya. Di sisi lain, pembeda semantik yang baru, seperti pembeda *alienability* dalam konstruksi posesif atau penanda plural dalam kata benda, berubah menjadi pembeda gramatikal di beberapa bahasa (tidak terbatas pada satu sub-kelompok). Selain itu, aspek sintaksis klausal dalam bahasa-bahasa Flores-Lembata juga telah mengalami perubahan. Bertolak belakang dengan negasi pra-predikat dan urutan kata SVO yang diwariskan dari nenek moyang bahasa Austronesia, beberapa bahasa-bahasa Flores-Lembata telah menginovasi satu set kata kerja *deictic motion* yang berada di akhir kalimat. Fitur-fitur urutan kata pada frasa benda dan dua dari kata kerja *deictic motion* dapat direkonstruksi ke dalam Proto bahasa Flores-Lembata. Bentuk-bentuk inovasi lainnya tampaknya baru digunakan di dalam bahasa-bahasa Flores-Lembata setelah perpecahan bahasa-bahasa Flores-Lembata menjadi beberapa sub-sub-kelompok. Semua fitur-fitur inovatif dalam bahasa-bahasa Flores-Lembata juga telah terbukti ada dalam bahasa-bahasa non-Austronesia, Timor-Alor-Pantar yang dituturkan di dua pulau berdekatan di sebelah timur, dan fitur-fitur itu bisa direkonstruksi ke dalam Proto bahasa Timor-Alor-Pantar.

Dengan menggabungkan bukti leksikal dan tipologis, disertasi ini mengusulkan bahwa bahasa-bahasa Flores-Lembata telah mengalami kontak dengan bahasa-bahasa yang secara tipologis mirip dengan bahasa-bahasa Alor-Pantar sejak zaman Proto-Flores-Lembata sampai setelah pecahnya keluarga bahasa Flores-Lembata ke dalam sub-sub-kelompok. Beberapa subkelompok memiliki fitur non-Austronesia lebih banyak daripada sub-kelompok yang lain. Hal ini menunjukkan bahwa terdapat variasi dalam hal panjang-pendek dan intensitas kontak di antara sub-sub-kelompok tersebut.

Curriculum Vitae

Hanna Fricke was born in Winsen an der Luhe, Germany in 1987. After graduating from high school at the Ellenrieder Gymnasium in Konstanz in 2007, she studied a Bachelor of Arts in Southeast Asian Studies at the University of Hamburg. She focussed on Indonesian language and culture in her major and took a minor in French linguistics. In 2010, she spent one semester of her studies at the Gadjah Mada University in Yogyakarta, Indonesia. In 2012, she continued her linguistic studies in the Research Master program at the Leiden University Center for Linguistics (LUCL). For her Master thesis, she conducted field research in the village of Hewa in eastern Indonesia and compiled a description of several grammatical aspects of the Hewa language (published as Fricke 2014a). After obtaining her Master's degree, she worked as a teacher of German as a foreign language in Berlin. In 2014, she started her PhD research in Leiden, as part of the NWO funded VICI project *Reconstructing the past through languages of the present: The Lesser Sunda Islands* led by Prof. dr. Marian Klamer (project number: 277-70-012). She is currently a research assistant in the same project.