

CAMBRIDGE GRAMMATICAL DESCRIPTIONS

Editors: R. M. W. Dixon, Keren Rice

This series is devoted to the publication of comprehensive descriptive grammars of languages that have not previously been documented, and that have interesting and unusual characteristics which will expand our understanding of human language in all its diversity and challenge the limits of current linguistic theory. Some of these languages are spoken by only a small number of people and can be considered endangered.

Each grammar briefly introduces the society in which the language is spoken, and covers the key areas of phonology, morphology and syntax, together with typological and historical considerations. In each case, a sample text or texts in the language are provided, with full gloss and translation. A glossary of basic vocabulary is also included. The series aims to provide theoretical linguists in the various subdisciplines with reliable data and analysis which will provide a permanent and invaluable set of source materials.

Previously published in this series

David E. Watters *A Grammar of Kham*
0 521 81245 3

Alexandra Y. Aikhenvald *A Grammar of Tariana*
0 521 82664 0

A GRAMMAR OF SEMELAI

Nicole Kruspe

Research Centre for Linguistic Typology
La Trobe University, Melbourne



CB1
1980

for my daughters Amelia and Ruby Goss

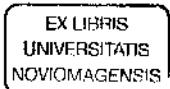
PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge, CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

<http://www.cambridge.org>

© Nicole Kruspe 2004

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



First published 2004

Printed in the United Kingdom at the University Press, Cambridge

Typeface Times 10.5/13.5 pt System L^AT_EX 2_E [TB]

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

ISBN 0 521 81497 9 hardback

Contents

Tables and figures xiv

Preface xvi

Acknowledgements xviii

Abbreviations and conventions xx

Maps xxiii

1 Semelai 1

1.1 Linguistic type 1

1.2 The Aslian languages 10

1.2.1 Internal subgroupings of the Aslian languages 10

1.2.2 Geographic distribution 11

1.2.3 The term Aslian 12

1.2.4 Aslian: an historical overview 12

1.2.5 Linguistic affiliation of Aslian with Mon-Khmer 14

1.2.6 Prehistory of the Aslian languages 16

1.3 The Semelai and their language 18

1.3.1 Ethnonym 19

1.3.2 Prior documentation of the Semelai 19

1.3.3 The Semelai at Tasek Bera 21

1.3.4 Current linguistic situation 28

1.4 The Orang Asli 29

1.4.1 Semang, Senoi and Melayu Asli 30

2 Phonology and phonotactics 32

2.1 Inventory of phonemes 32

2.1.1 Consonants 32

2.1.2 Vowels 36

2.2 The syllable 39

2.2.1 Syllable structure 39

2.2.2 Stress 40

2.2.3 Syllabification 40

2.3 The distribution of phonemes 48

2.3.1 Onset phonemes 48

2.3.2 Coda phonemes 51

2.3.3 The nucleus 52

2.4 Malay loanword phonology 55	5.3 Derivational morphology 110
2.5 Notes on comparative phonology of Aslian languages 58	5.3.1 Imperfective +C+ 'IMPERF' 111
2.5.1 Consonants 58	5.3.2 Imperfective mN- 'IMPERF' 115
2.5.2 Vowels 59	5.3.3 Middle voice b(r)- 'MID' 117
2.5.3 Syllable structure 59	5.3.4 Collective br<>an 'TOG' 123
3 Morphology 61	5.3.5 The causative 124
3.1 Preliminaries 61	5.3.6 Valency increasing -i? 'APPL', 'ITER' 135
3.1.1 Terminology 61	5.3.7 Happenstance t(r)- 'HAPP' 140
3.1.2 Affixation and cliticisation 63	5.3.8 Comparative +ra?+ 'COMP' 146
3.2 Affixation 64	5.3.9 Intensive 'INTNS' 149
3.2.1 Overview of forms and processes 64	5.4 Derivation: nominal to verb 150
3.2.2 Nonconcatenative morphology 69	5.4.1 Coda copy 150
3.2.3 Concatenative morphology 82	5.4.2 Performative mN- 'PERFM' 152
3.2.4 Non-productive lexicalisation patterns 85	5.4.3 Middle voice b(r)- 'HAVE/USE' 153
3.3 Cliticisation 87	5.4.4 Causative 'EQUIP' 155
3.3.1 Pronominal proclitics 88	5.5 Verbal clitics 156
3.3.2 Irrealis ma= 'IRR' 89	5.5.1 The pronominal proclitic 157
3.3.3 Agentive/causal la= 'A'/'BCS' 89	5.5.2 Irrealis ma= 'IRR' 161
3.3.4 The clitic hn 90	5.5.3 Imminent aspect ga= 'IMM' 163
3.3.5 The proclitic mæ= 91	5.6 The modal category 164
3.3.6 Clitics to free pronouns 92	5.6.1 Necessity 165
3.3.7 Name ?i= 'NM' 93	5.6.2 Possibility 168
4 Word classes 94	6 Pronouns: personal, ignorative, and demonstrative 170
4.1 Introductory remarks 94	6.1 Personal pronouns 170
4.2 Open classes 97	6.1.1 Semantic distinctions in the pronominal system 170
4.2.1 Nominals 97	6.1.2 Functions of the free personal pronouns 172
4.2.2 Verbs 100	6.1.3 Determiner ?a= 'DET' 174
4.2.3 Expressives 102	6.1.4 Possessive dæ= - dde= 'OF' 174
4.3 Closed classes 102	6.1.5 Possessor focus pæ= 'POC' 175
4.3.1 Prepositions 102	6.1.6 dr? 'sell' 175
4.3.2 Adverbs 103	6.2 Ignoratives 176
4.3.3 Auxiliaries and supplemental verbs 103	6.2.1 Distinctions in the ignorative class 176
4.3.4 Existential and ascriptive predicates 103	6.2.2 THING 179
4.3.5 Negators 103	6.2.3 PROPERTY; REASON 181
4.3.6 Connectives 103	6.2.4 PERSON 185
4.3.7 Interjections 104	6.2.5 PLACE 187
5 The verb 105	6.2.6 MANNER 190
5.1 The structure of the verbal word 105	6.2.7 QUANTITY 190
5.2 Verb classification 106	6.2.8 TIME 191
5.2.1 Verbal sub-classes 107	6.3 Demonstratives 192
	6.3.1 The demonstratives 192
	6.3.2 Adnominal and pronominal functions 194
	6.3.3 The major functions of demonstratives 198

7	The noun phrase 202	9	Grammatical relations, constituent order and coding strategies 247
7.1	The internal structure of the noun phrase 202	9.1	Grammatical relations 247
7.2	Functions of the noun phrase 203	9.1.1	Core relations 247
7.3	The quantifier phrase 204	9.1.2	Obliques 249
7.3.1	Numerals 205	9.1.3	Syntactic behaviour of core relations 251
7.3.2	Classifiers 206	9.2	Constituent order 253
7.3.3	Measure nouns 208	9.2.1	Core arguments 254
7.3.4	Non-numeral quantifiers 210	9.2.2	Obliques 256
7.4	The attributive phrase 211	9.2.3	Fronting: the clause initial position 256
7.4.1	The locative prepositional phrase 211	9.3	Coding strategies 258
7.4.2	The relative clause 211	9.3.1	The transitive clause 258
7.4.3	The verb 211	9.3.2	The intransitive and non-verbal clauses 264
7.5	The associative phrase 213	9.3.3	A-marking and the intransitive clause 265
7.5.1	Possessive enclitic =hn '3POSS' 214	9.4	Zero anaphora 266
7.5.2	Functions of the associative phrase 214	10	Basic clauses 269
7.5.3	Locative nominals in the associative phrase 215	10.1	Identificational, ascriptive and existential clauses 269
7.5.4	The associative phrase and compounds 216	10.1.1	Clauses without overt predators 269
7.6	Nominal derivation 218	10.1.2	Clauses with overt predators 273
7.6.1	Noun to noun derivation 218	10.2	Irrealis constructions 278
7.6.2	Deverbal nouns 220	10.2.1	Formal features of an irrealis clause 279
8	Prepositions and the prepositional phrase 227	10.2.2	Epistemic modality 281
8.1	The prepositions 227	10.2.3	Deontic modality 293
8.2	The structure of the prepositional phrase 228	10.3	Adverbial modifiers of the basic clause 296
8.3	Functions of the prepositional phrase 229	10.3.1	Lexical adverbs 296
8.3.1	Peripheral expressions 229	10.3.2	Nominal adjuncts 312
8.3.2	Predicative function 229	10.3.3	Verbal adjuncts 315
8.3.3	Attributive function 230	10.4	Negation 316
8.3.4	Core function 230	10.4.1	Internal negation da? 'NEG' 316
8.4	The neutral locative 231	10.4.2	Metalinguistic negation 319
8.4.1	Peripheral functions 231	10.5	Interrogatives 324
8.4.2	Core function 232	10.5.1	Polar interrogatives 325
8.5	Deictic locatives 233	10.5.2	Information interrogatives 328
8.5.1	Location and orientation 234	10.6	Imperatives 331
8.5.2	Distributed reading 236	10.6.1	Formal features of an imperative clause 331
8.5.3	Peripheral functions 237	10.6.2	The positive imperative 332
8.5.4	Core function 238	10.6.3	The negative imperative 336
8.6	Directionals 238	11	Complex clauses 339
8.6.1	Directionals and the deictic locatives 240	11.1	Relative clauses 339
8.7	Source 241	11.1.1	Relative clauses formed with me= 'REL' 340
8.8	The instrumental/comitative 243	11.1.2	The juxtaposed relative clause 346
8.9	The possessive 245	11.2	Complement clauses 346
		11.2.1	Paratactic clause complements 348

11.2.2 Reduced complement clauses	351	14.4.3 c<nn>man sma? pe? b-knøn 'The narrative of the person who had three children'	433
11.2.3 Desiderative and achievement predicates	357	14.4.4 c<nn>man sma? k ^b oy 'The narrative of the person who was a head'	439
11.3 Serial verb constructions	359	 	
11.3.1 General features of serial verb constructions	361	 	
11.3.2 Types of serial verb construction	363	 	
11.4 Concatenated clauses	373	 	
11.4.1 Semantic relationships in concatenated clauses	374	 	
11.4.2 The syntax of the concatenated clause	376	 	
11.5 Connective clauses	380	 	
11.5.1 General features of connectives	380	 	
11.5.2 A note on enclitic =hn 'CONN'	381	 	
11.5.3 The adverbial connective construction	382	 	
11.5.4 Discourse connectives	391	 	
12 Expressives	396	 	
12.1 The semantics of expressives	397	 	
12.2 The syntax of expressives	399	 	
12.3 Notes on the phonology and morphology of expressives	401	 	
13 The quotative marker, interjections and discourse clitics	403	 	
13.1 The quotative marker	403	 	
13.1.1 The syntactic distribution of k ^b laŋ 'QUOTE'	405	 	
13.2 Interjections	407	 	
13.2.1 Expressions of negative attitude	407	 	
13.2.2 Interjections with imperative force	408	 	
13.2.3 Other interjections	409	 	
13.3 Discourse clitics	410	 	
13.3.1 Emphatic =cø? 'EM'	411	 	
13.3.2 Speaker conclusion =søn 'SC'	414	 	
13.3.3 Clitic =ja 'CL'	417	 	
13.3.4 Factual =pa 'FACT'	419	 	
14 Texts	420	 	
14.1 ttikyan 'Riddles'	420	 	
14.2 Recollections	424	 	
14.2.1 Establishing the identity of ga=tɔ?ɛ?	424	 	
14.2.2 gatɔ?ɛ? and her elder brother	425	 	
14.3 Then and now	426	 	
14.4 Traditional narratives	429	 	
14.4.1 c<nn>man sma? ɬadi?bradi? rom romgasi 'The narrative of the siblings and romgasi'	429	 	
14.4.2 sma? ki=ca podɔŋ 'The person the tiger ate'	432	 	
		Vocabulary	448
		References	466
		Index	477

Tables and figures

Tables

- 1.1 The semantic domains of nouns in the avoidance speech style 8
- 1.2 The semantic domains of verbs in the avoidance speech style 9
- 1.3 Comparative South-eastern Sakai forms 20
- 1.4 Kinship reference and address terms 24
- 1.5 Orang Asli population figures 29
- 2.1 Consonant phonemes 33
- 2.2 Onset and coda phonemes 48
- 2.3 Distribution of vowel phonemes 53
- 2.4 Phonemic inventory of Aslian languages 58
- 2.5 Inventory of standard Aslian vowel phonemes 59
- 4.1 Lexical, semantic and pragmatic categories 94
- 5.1 Derivational morphemes 110
- 5.2 The range of functions of coda copy 110
- 5.3 Middle marking 118
- 5.4 Distribution of causative affixes 127
- 5.5 Summary of verbal derivation strategies 150
- 6.1 The pronominal paradigm 171
- 6.2 Ignoratives: primary ontological categories 177
- 6.3 Ignoratives: derived categories 177
- 6.4 Derived demonstratives 193
- 7.1 The numeral system 205
- 7.2 Lexical nominalisation processes 221
- 9.1 Role marking 258
- 9.2 Available expressions of core arguments 267
- 10.1 Non-verbal clause types 269
- 10.2 Interrogative pronouns 329
- 11.1 Manner serial verbs 363
- 11.2 Associated motion serial verb constructions 368
- 11.3 Summary of connective conditional constructions 386

Figures

- 1.1 Genetic affiliation of the Aslian languages 11
- 1.2 Blagden's Aslian language groupings 13

- 1.3 Schebesta's Aslian language classification 13
- 2.1 Phonemic oral and nasal vowels 37
- 2.2 The internal structure of the syllable 39
- 2.3 Derivation 1 40
- 2.4 Derivation 2 41
- 5.1 The structure of the verbal word 106
- 5.2 The imperfective 111
- 5.3 The decausative 119
- 5.4 The two place causative construction 125
- 5.5 The three place causative construction 126
- 5.6 The one participant auto-causative 131
- 5.7 The two participant auto-causative 131
- 5.8 Directional raising applicatives 136
- 5.9 Argument frame of the happenstance causative 145
- 7.1 The internal structure of the noun phrase 202
- 8.1 Deictic locatives 234

Preface

This is a description of Semelai, an Aslian language spoken by some 4,000 people in Peninsular Malaysia. The present book is a revision of my doctoral dissertation submitted to the University of Melbourne in 1999. The study is based on a corpus of data collected during an initial period of fieldwork undertaken from July 1990 to June 1991 at Tasek Bera, in the state of Pahang. Insights from recent visits in 2000-1, whilst a postdoctoral student at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, have also been included. Revisions have been completed during a post-doctoral fellowship at the Research Centre for Linguistic Typology, La Trobe University, Melbourne, Australia.

First and foremost, this book represents an initial step in the documentation of a unique and complex language, in an area which to date has been largely neglected by linguists.

My main consideration in writing this grammar was to provide a lucid and accessible account of Semelai for linguists of all theoretical persuasions. The language is described in terms of Basic Linguistic Theory. I have sought to treat equally all aspects of the grammar, examining form and meaning and providing semantic motivations for the various distinctions made within it. A generous number of examples and texts are included to illustrate the discussion and to create a reference work, which will be of use to descriptivists and typologists alike. The discussion of the grammar is contextualised with reference to other languages and comparative notes are included where possible. The analysis is supported by references to literature on issues of theoretical and typological relevance.

When I first set out to write this description, the meagre resources available only hinted at what was to come. Each advance was – reflecting Diffloth's sentiments with regard to writing a description of Jah Hut – 'a step in the unknown' (1976c). This book provides a model of grammatical description for future researchers in this region. In addition, it will benefit typologists by providing empirical data for an area under-represented in cross-linguistic studies.

With this book I hope to draw attention to the Aslian languages and advance our knowledge not only of Aslian, but also of Mon-Khmer linguistics in general. Tucked away at the end of a vast continent, Semelai and the other Aslian languages represent an outlier of Austroasiatic, the oldest phylum present in mainland Southeast Asia. They hold an important place for our understanding of the linguistic typology of the region and the study of Southeast Asian prehistory.

A few points are required on the method of data collection. The technique employed was to work monolingually with a cross-section of the community, collecting narratives

in a range of genres. I worked on a daily basis with *normah het*, who became not only my principal informant, but also my friend. Her willingness to accept me into her family circle and allow me to participate in their daily life consolidated my knowledge of the Semelai language. Not only did *normah* provide me with language data, she also introduced me to her wider family and encouraged them to assist me in the collection of lexical and textual materials. Textual materials were also provided by the following people: *?amay isa* – whose prolific knowledge of traditional narratives, and passion for relating them, despite her age and ill-health, was astounding; *pa?yah jahudi* provided traditional narratives; *normah*'s mother *kop*, in addition to traditional narratives, recounted events from her childhood and the Communist Emergency; the late *brahim* of Kampong Jelawat, *galenah*, her eldest son *kassim* and the late *nari*, also had much to offer.

Text collection was usually undertaken in a group setting, at the suggestion of the narrator, who felt that an interactive audience was a necessary and important component of the performative retelling of traditional narratives. Audience interaction ranged from interjections and non-verbal gestures, to direct comments and questions. This context of spontaneous conversation proved useful for data collection. Direct elicitation of language materials was kept to a minimum, employed only to clarify phenomena in the texts, or to work through infrequent or poorly understood structures. Unless otherwise noted, all examples in this grammar come from the collected texts. A selection of texts is found in §14.

The texts were transcribed and annotated primarily with the assistance of *normah*. These sessions took place in the afternoons when groups of people gather on each others' verandahs. These group settings, which included people of all ages from different settlements, provided a forum of spontaneous and often lively debate which greatly increased my understanding of the language. In my translation and commentaries I have tried to capture some of the spirit of these engagements and the culture of the Semelai.

The reader should note that examples are numbered separately within each chapter; cross-references to examples are only made within individual chapters.

Semelai is not a written language. The orthography used here is phonemic, and follows Benjamin (1976b) and Diffloth (1976c) employing the standard IPA symbols with the following exceptions: IPA /j/, /y/ and /s/ are replaced by /y/, /j/ and /s/, and /m/, /n/ and /ŋ/ are represented as the digraphs /hm/, /hn/ and /hn/.

The initial letter of Semelai personal names and place names is not capitalised in accordance with IPA conventions.

The language/ethnic names of some Aslian groups are not standardised. I have maintained the spelling found in the original document. The following variants are noted: *Bcsisi/Mah Meri*; *Jahai/Jehai*; *Kensi/Kensiw*; *Kentaq Bong/Kintaq Bong* and *Semaq Beri/Semoq Beri*.

Acknowledgements

Without the acceptance of the Semelai community this work would not have been possible. It is out of respect for them that I persisted with this research. It is impossible to acknowledge every one who assisted in teaching me Semelai, for every exchange provided a contribution to my knowledge of their language. I am especially indebted to my primary informant normah net for her friendship; to her extended family, and in particular her mother kɔp, for their hospitality and assistance in my endeavour; to my neighbours at Kampong Gau and their extended families, borhan pipin, maʔpah krdr, the late mntri? ʔabas, the late nari and noriah nari. I would also like to thank ʔamir wadi for the hours he spent with me in the early days; the people at Kampong Jelawat who always welcomed my visits; the residents of Kampong Putat paʔŋah jahudi, ʔamay isa and jimpang and layon who spent many hours enthusiastically retelling oral narratives for my benefit, and matnor sujin, ʔon, tatey, raizan maris, rohana ʔinolan, shamsuddin gosi and the late brahim. I am also indebted to Bah Tony Williams-Hunt who generously found time to provide me with an introduction to the community, and to ʔatim padot and family for making their home available during my initial field trip.

This research would not have been possible without the permission of the Government of Malaysia and the State of Pahang, to whom I am grateful. I would also like to thank the Jabatan Hal Ehwal Orang Asli, in particular Mohd. Sayed Abdullah, and Puan Munirah Abd. Manan, Economic Planning Unit, The Prime Minister's Department. Thanks are also due to Professor Asmah Haji Omar, Universiti Malaya, and Professors Hood Salleh and Shamsul Amri Baharuddin, Universiti Kebangsaan Malaysia, for their assistance and support in Malaysia.

My principal supervisor, Nicholas Evans, first suggested the Asian languages as an area for research. Nick's enthusiasm, challenging criticisms and perceptive observations were invaluable in shaping this work, enabling me to finally find order amidst chaos.

I have greatly appreciated the contributions of Mark Durie, Bill McGregor and Jean Mulder, who took over supervision at various periods over the years. I have also benefited from discussions with Geoffrey Benjamin, James T. Collins, Gerard Diffloth, Ilia Peiros and Niclas Burenhult, and from the considered comments and encouragement of my examiners, the series editors Bob Dixon and Keren Rice and the anonymous reviewer from Cambridge University Press. I would also like to thank Sasha Aikhenvald for her support and encouragement.

Formatting was initially done by Erich Round, before Adam Bowles took over. Adam's valuable suggestions and attention to detail in preparing the manuscript are greatly appreciated. The maps were drawn by Chandra Jayasuriya.

Financial support during my candidature came from an Australian Postgraduate Research Award; fieldwork was made possible by a generous award from the Alma Hansen Travelling Scholarship and a Departmental Travel grant, and a Luce fellowship and a tuition fellowship from the University of Hawai'i at Manoa allowed me to undertake tuition in Malay at the Southeast Asian Summer Institute in 1989.

The Max Planck Institute for Evolutionary Anthropology, the Harbison-Higginbotham Research Scholarship and the Research Centre for Linguistic Typology, La Trobe University, provided generous financial and institutional support at various stages of this project.

To my friends and family who have provided me with moral and practical support over the years, I am deeply grateful. Too many years and too many people have passed by for me to thank every one, but particular thanks go to Alec Coupe, Catriona Hyslop, Andrew Ingram and Janet Sharp for their encouragement and practical assistance in the final stages, and to Anthony Jukes and Tonya Stebbins.

Abbreviations and conventions

The following lists comprise the major abbreviations and conventions employed throughout the thesis. Occasionally some abbreviations which only have local relevance are introduced where necessity dictates, e.g. BEN 'Benefactor' in §9. Due to their limited application they are not listed here.

Some general conventions:

*	non-occurring or reconstructed form
~	'in free variation with'
→	'is realised as, becomes'
←	'is derived from'
=	'is equivalent to'

The abbreviations and conventions are listed according to domains.

Phonological conventions:

	phonemic representation	UR	underlying representation
//	underlying representation	Nu	nucleus
///	phonetic representation	On	onset
[]	checked release	R	rhyme
:	length	Co	coda
·	stress	C	consonant
.	syllable boundary	V	vowel
σ	syllable	N	nasal
σ _R	reduced syllable	IPA	International Phonetic Association
ω	word		

Morphological conventions:

✓	root
*✓	synchronously non-occurring formative
[]	morpheme template
++	morpheme boundary: underspecified affixes
-	morpheme boundary: prefixes and suffixes
<>	morpheme boundary: infixes and the circumfix
=	clitic boundary
:	separates meaning elements in a portmanteau morpheme
μ	morpheme

Grammatical relations:

A	agent-like argument of a transitive verb
O	patient-like argument of a transitive verb
S	single argument of an intransitive verb
IO	indirect object
Subj	subject, A, S or the subject of an NVC

Personal pronouns and kin terms:

I	first person
1&2	first person inclusive
2	second person
3	third person
aug	augmented
f	familiar
min	minimal
pl	plural
sg	singular
EB	elder male sibling, or parent's sibling's offspring
EZ	elder female sibling, or parent's sibling's offspring
UA	unidentified agent
YS	younger sibling, or parent's sibling's offspring

Derivational affixes:

APPL	applicative	LSR	light syllable reduplication
BE	to be a number	MCAUS	manipulative causative
CAUS	causative	MID	middle voice
COMP	comparative	NMZ	nominalisation
DEM	demonstrative	PERFM	perform
ENM	enumerator	SRC	source
EQUIP	equip with, provide	RDP	reduplication, specific to nominals
HAPP	happenstance	TOG	collective
HAVE	possessive	UNIT	a unit, or measure
IMPERF	imperfective		nominalisation
INDIV	individuation	USE	utilise
INTNS	intensive	XS	excessive agent/performer
ITER	iterative		

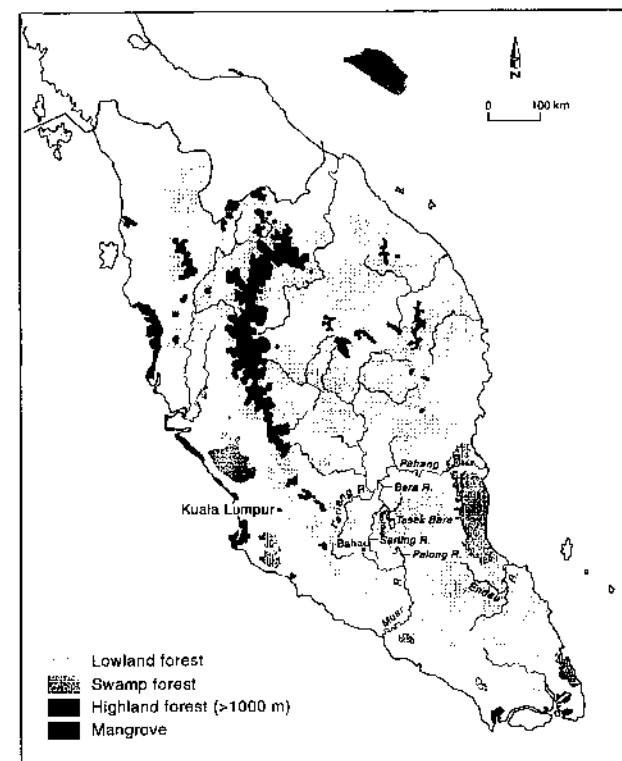
Clitics:

ABS	absolutive	IMM	imminent aspect
ATTN	attention	IMP	imperative
AUG	augmented	IRR	irrealis
BCS	because	NM	name
CL	clitic	OF	possessive

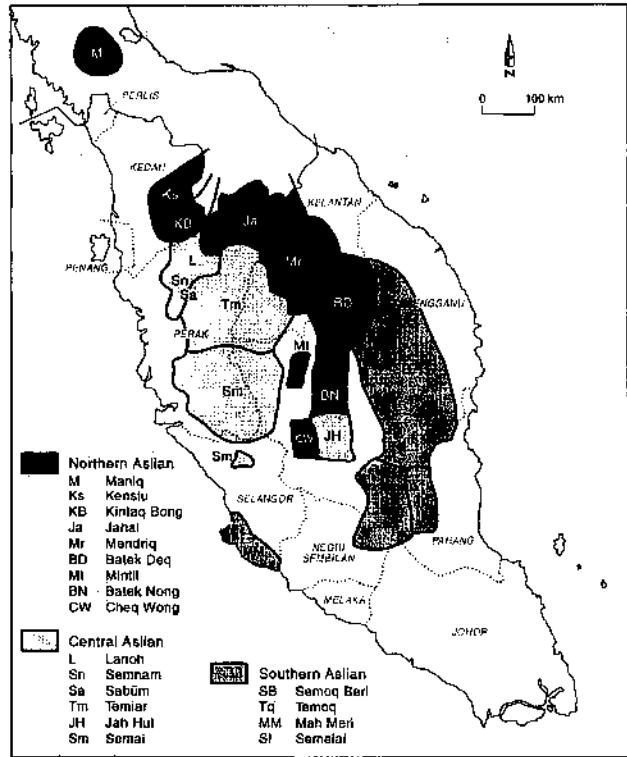
CONN	connective	PFOC	possessor focus
DET	determiner	REL	relative proclitic
EM	emphatic	SC	speaker conclusion
FACT	factual		

Other interlinear glosses and abbreviations:

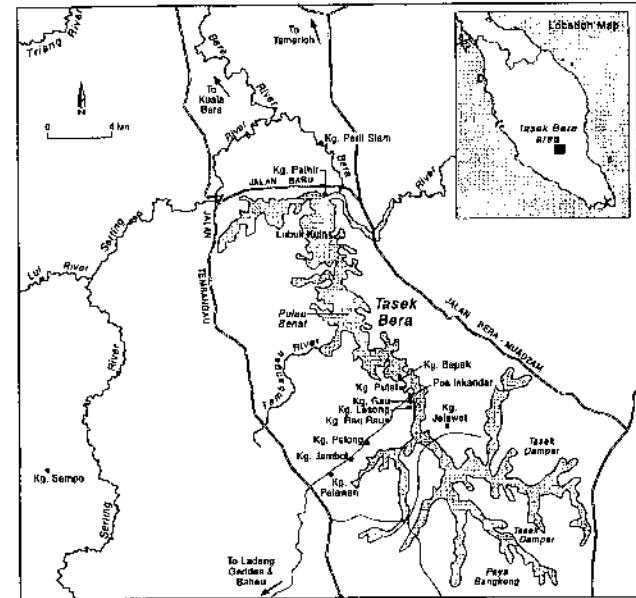
AFF	affirmative	NP	noun phrase
ALT	interrogative tag expressing alternative	NVC P	nonverbal clause preposition
AT	locative preposition	POSS	possessor
CLF	classifier	PP	prepositional phrase
EXCL	exclamation, interjection	PRED	predicate
EXIST	existential predicate	Q	interrogative tag expressing doubt
HORT	hortative		
LOC	locative	TO:spec	directional preposition with direction specified
N	noun	TO:unspec	directional preposition with direction unspecified
NEG	negator	V	verb
NEG:IMP	negative imperative	VP	verb phrase
NO	external negator		
NOT	external negator (nominal)		
(av.)	lexical item from the avoidance speech style		
[top]	toponym		
[name]	personal name		
sp.	species		
[HES]	hesitation		



Map 1 The Malay Peninsula



Map 2 Approximate distribution of Aslian languages



Map 3 Tasek Bera and environs

1 Semelai

Semelai is an Aslian language belonging to the Mon-Khmer division of Austroasiatic. It is spoken by approximately 4,103 people in Peninsular Malaysia around the shores of Tasek Bera, along the banks of the Bera, Teriang and Serting rivers in south-west Pahang and north-west Negeri Sembilan states, and the Muar River in north-west shore (see Maps).¹

As with all Aslian languages, there is no written tradition, or recorded history, other than oral based.

1.1 Linguistic type

Semelai exhibits many of the typological characteristics of a mainland Southeast Asian language, as well as incorporating aspects of Austronesian, notably from sustained contact with Malay. Whilst there has been considerable impact from Malay, particularly at the lexical level, Semelai remains a distinctly mainland Austroasiatic language, but at the same time raises questions with respect to our understanding of the typology of this region.

A. Phonology Semelai has a rich phonemic inventory. There are thirty-two consonant phonemes, including a series of voiceless nasals which have not been recorded previously in an Aslian language; a series of glottalised sonorants, and twenty vowel phonemes: ten oral vowels and ten phonemically nasal counterparts (§2.1).

The maximal canonical syllable is [CV(C)]_σ. Phonemic contrasts of both consonants and vowels are richest in the final syllable; for consonants these contrasts are maximised in the onset of the final syllable.

Words have the structure: $\omega \rightarrow (\sigma_R)^n (\sigma)^1 \sigma$, where: $n \leq 2$. Reading from the right, σ is the final syllable – word stress is always on the final syllable, and there is no secondary stress, σ is the penultimate syllable and σ_R is the prepenultimate syllable. The $,$ indicates significantly reduced phonotactic possibilities for both consonants and vowels, including the characteristic underspecification of vowels associated with this position. Any oral vowel phoneme has the potential to occur as the phonetic realisation of an underspecified vowel in a given environment.

The minimal word is monosyllabic, e.g. /t̪ʰi/ 'hand', the maximal word is tetrasyllabic: /k.r.wan.ceŋ/ [kəruwanceŋ] 'coral snake sp.'

B. Morphology Semelai is an isolating language with agglutinating features. Typically for the Aslian family, but at variance with many Mon-Khmer languages, Semelai has a complex morphological system, containing a rich inventory of prefixes,

¹ The Semoq Beri do not inhabit this area (pace Parkin 1991: 55), but an area in north eastern Pahang.

infixes, suffixes and a circumfix. Some have their origins firmly in Austroasiatic (e.g. the nominalising affixes $+n^2$ and $\langle m \rangle$ (NMZ)), others have been acquired through language contact with varieties of Malay (e.g. $br-$ (MID) 'middle voice', and $tr-$ (HAPP) 'happenstance'), see §3.2. Semelai has distinctive word classes (see C below and §4); the typical function of derivational morphology is to change the word class or subclass of the root, e.g. from verb to noun or vice versa, from transitive verb to intransitive verb (§5.3) or mass noun to count noun (§7.6).

There are two systems of arrangement based on the domain of attachment: a) a non-concatenative system of prefixes and infixes, which has its origins in Mon-Khmer; and b) a concatenative system of prefixes, suffixes and a circumfix acquired through contact with Austronesian, notably Malay. Both systems of attachment are prosodically driven.

Affixes may be syllabic, or they may be non-syllabic consonantal units defined in terms of syllable position, e.g. $\langle r \rangle$ 'CAUS' is a syllable coda infix $r]_S$. Syllabic affixes are fully prespecified, e.g. [par]_S 'CAUS(ative)'.

Where affixes exhibit allomorphy, this is conditioned primarily by the syllabic structure of the root and the domain of attachment: monosyllabic roots select heavy syllable prefixes, whilst disyllabic roots select either infixes or light syllable prefixes, dependent on the individual affix. These points are illustrated in the following discussion which focuses on affixes from the non-concatenative system that take the prosodic head as the domain of attachment.

The two forms given above for the causative morpheme, tar- 'MCAUS' and $\langle r \rangle$ 'CAUS', are licit for monosyllabic and disyllabic roots respectively:

- a) *Monosyllabic root: hɔp* 'to put food in (one's) mouth'
Affix tar- 'MCAUS':
tar-hɔp 'to put food in (s.o.'s) mouth'
- b) *Disyllabic root: jtɛk* [jtɛk] 'to sleep'
Infix $\langle r \rangle$ 'CAUS':
 $j\langle r \rangle tɛk$ [jərtɛk] 'to cause to sleep'

The most intriguing morphological process is the reduplicative process, which I term 'Copy'. This is a form of internal reduplication, whereby a phonologically underspecified morpheme template $[CC]_\mu$ equivalent to the maximal canonical syllable, is affixed in penultimate syllable position to the final syllable of the root. The underspecified segment, which may be a syllable onset or a coda, derives its phonemic content by copying that of the corresponding segment of the root. The underspecified vowel slot is filled by epenthesis.

With monosyllabic roots, the satisfaction of the prefixed template requires full reduplication of the root:

Monosyllabic root: sɔc 'to whistle'

- a) Prefix the template $[CC]_\mu$:
 $[CC]_\mu-sɔc$

² Morpheme boundaries of prefixes and suffixes are represented by a hyphen '-', infixes and the circumfix by parentheses ' $\langle \rangle$ ', underspecified affixes are enclosed by '+' and clitic boundaries are shown by '='.

- b) Onset and coda copy reduplicate the phonemic consonantal content of the root into the underspecified positions in the prefix:
 $sc-sɔc$
- c) Given that only consonants may be reduplicated, the underspecified nucleus receives phonetic content from the application of vowel realisation rules:
 $sc-sɔc [sɪcsɔc]$ 'to be whistling'

With disyllabic roots, the phonemic content of the penultimate syllable of the root is associated with the template. Coda copy takes place into the underspecified coda position:

Disyllabic root: catɛk 'to detach'

Coda copy:

- a) Prefix the template $[C]_\mu$ to the prosodic head and associate the phonemic content of the root:
 $ca\langle C \rangle_\mu-tɛk$
- b) Copy the phonemic content of the root coda into the underspecified coda position C:
 $ca\langle k \rangle tɛk$ 'to be detaching'

Coda copy may co-occur with the affixation of a partially prespecified affix like nominalising onset n- which has a morphemic template $[n C]_\mu$:

Monosyllabic root: sɔc 'to whistle'

- a) Prefix the template $[n C]_\mu$ and associate the phonemic content of the root:
 $[n C]_\mu-sɔc \rightarrow n C-sɔc$
- b) Coda copy reduplicates the phonemic consonantal content of the root into the underspecified position in the prefix:
 $nc-sɔc$
- c) The underspecified nucleus receives phonetic content from the application of vowel realisation rules:
 $nc-sɔc [nɪcsɔc]$ 'act of whistling'

Borrowed Malay lexemes are fully incorporated into the Semelai system. There are no restrictions on the combinatorial possibilities of indigenous and borrowed terms.

- a) Coda copy
Prefixation of the template $[CC]_\mu$ to the prosodic head of verbal root *tupuk* 'to point out, indicate' (< Malay *tunjuk*) derives:
tu<k>puk 'to be pointing'
- b) Nominalisation:
Infixation of the nominalisation template $[n C]_\mu$ into the root *sudu?* 'spoon' (< Malay *sudu*) derives a measure noun:
s<n>u<r>du? [sənu?du?] 'spoonful'

Roots can feed multiple affixation, combining affixes from both systems of attachment. In the following example, taking the verb *krep* 'to blink', the causative infix, $\langle r \rangle$ 'CAUS', is a morpheme of the non-concatenative type, and the prefixed happenstance morpheme, *t-* 'HAPP', is from the concatenative type:

- a) Infixation of <r> 'CAUS':
k<r>jep (blink<CAUS>) [kərjep] 'to make blink'
- b) Prefixation of t- 'HAPP':
t-k<r>jep (HAPP-blink<CAUS>) [təkərjep] 'to happen to make blink'

C. Word classes Semelai has clearly distinguished word classes, both syntactically and morphologically. The open classes are nominal, verb and expressive.

The nominal superclass includes nouns, as well as the following closed classes: numerals, pronouns, demonstratives and ignorative. Nominals function as the heads of NPs, predicates in non-verbal clauses and nominal modifiers in associative constructions.

Verbs function as predicates. The majority of verbs are clearly transitive or intransitive; there is also a small number of ambitransitive verbs. The distinction between stative and active verbs cuts across the transitive/intransitive division. Intransitive verbs also distinguish an adjective class.

Expressives³ are iconic utterances, which function to simultaneously provide information about both the predicate and its arguments, in the form of a single lexical item. They function as clausal adjuncts, or stand alone as minor clauses.

Members of this class express sensate imagery – aural, visual, oral, tactile – e.g. tɔŋbɔc 'short and fat (of people)'. Expressives frequently combine a cluster of properties, e.g. the lexical item br?ol '(s.th.) large, dark and motionless lying submerged in the water' combines dimension, colour and position. In addition, they exhibit irregular phonology, irregular reduplication patterns and vowel alternation.

The closed classes are preposition, adverb, auxiliary, existential and ascriptive predators, negator, connective and interjection (§4).

Lexemes must undergo derivational procedures, either morphological or syntactic, in order to function in a different word class.

D. Constituent order Semelai exhibits the characteristic head-dependent word order of the languages of mainland Southeast Asia: the attributive follows the noun (1) (§7.4); the possessor follows the possessee (2)–(3) (§7.5); the preposition precedes the noun (3)–(4) (§8), and relative clauses follow the head (5) (§11.1). At the phrase level, constituent order is fixed.

- (1) wøy ʔuh
knife be.sharp
sharp knife
- (2) wøy kmpen
knife wife
the wife's knife
- (3) wøy do kmpen
knife OF wife
the wife's knife
- (4) rom wøy
WITH knife
with a knife

³ 'Expressive' is the term used by scholars of Austroasiatic; the more widely used term is ideophone.

- (5) wøy mə=ʔuh
knife REL=be.sharp
the knife that is sharp

Constituent order at the clause level is fluid, with variation driven by pragmatic factors (§9). In the transitive clause the ordering is usually verb initial (6), although any argument may be pre-verbal, in which case it loses its role marking (7), see E below:

- (6) ki=tikam la=knlək hn=podəŋ rom lmeŋ
3A=stab A=husband O=tiger WITH spear
The husband stabbed the tiger with the spear.
- (7) podəŋ ki=tikam la=knlək rom lmeŋ
tiger 3A=stab A=husband WITH spear
The tiger, the husband stabbed with the spear.

In contrast, in the morphologically simplest, but 'statistically' marked transitive clause type – the 'universal' clause – the pattern is SVO, and there is no coding of grammatical relations:

- (8) podəŋ ea sma?
tiger eat person
Tigers eat people.

In the intransitive clause, the subject may also either precede or follow the verb.

- (9) kəhn swak
3S go
He went.⁴

- (10) swak kəhn
go 3S
He went.

E. Morphosyntax The syntactic typology of Semelai displays an intricate mix of features based on a basic transitive/perfective-intransitive/imperfective contrast.

Semelai exhibits both head and dependent marking, e.g. the head-marking pronominal proclitics on the transitive verb and the dependent-marking proclitics on post-verbal arguments which are introduced below.

Core grammatical relations are marked by a system of clitic cross-referencing in the transitive clause. The A is cross-referenced by a pronominal proclitic on the verb (6), and proclitic la= 'A' on the optional post-verbal NP. A system of differential object marking optionally encodes the object NP with proclitic hn= 'O'. The O is not cross-referenced on the verb, see example (6).

In the intransitive clause, the subject NP does not exhibit role marking, with the exception of the third person pronominal forms which have a fused enclitic hn 'S': kəhn '3S' (<kəh=hn 3=S), dehn '3plS' (<deh=hn 3pl=S). As with O, S is not cross-referenced on the verb, see example (9) above.

⁴ Third person pronouns are not distinguished for gender. Rather than using the clumsy '(s)he' for third person singular forms, they are translated throughout according to the context of the original text.

Indirect objects and obliques are coded by prepositions as in (6), with the exception of the benefactive which is not formally coded.

A semantically motivated split gives rise to two marking possibilities for intransitive verbs which express motion or emotion. The basis of the split is to mark a type of involuntary compulsion resulting from external causation as in (11). Curiously, the subject of the intransitive clause is encoded by the ergative pronominal proclitic like a transitive subject, although the clause is still monovalent, and the subject displays low agentivity (§5.5.1.2). Compare the following example with (9) above:

- (11) *The following comment was made when a person walked off from the group because someone offended him.*

ki=^cwak
3A=go
He went off.

Semelai exhibits odd transitivity behaviour, where lower semantic transitivity is not reflected by a change in morpho-syntactic marking. The iterative suffix -i? 'ITER', expressing the repetition of an activity (§5.3.6), and the happenstance prefix tr- 'HAPP', expressing lack of intention or volition (§5.3.7) both signal reduced transitivity, yet the clause retains the argument coding of a prototypical transitive clause.

F. Pronouns The pronominal system expresses categories of first, second and third person, with an inclusive/exclusive distinction for first person, a deference distinction for first and second persons singular, and a category of unidentified agent in the third person (§6.1). Interestingly, the personal pronouns have a greater range of number possibilities in the third person, where there are independent singular and plural forms in both the pronominal proclitics and free forms, but in first and second persons this distinction is not made for the proclitic forms, and is only optionally expressed in the free forms by the enclitic =en 'PL(ural)'.

G. Locative prepositions and directionals Semelai has a topologically based system of spatial deixis, manifested in locative prepositions and directionals, which conflate the locative relation with three degrees of height. These are arranged in a 'person' or 'speaker' oriented system mapped in relation to the topography of the local environment (§8).

The locative system consists of one proximate term, constituting the 'zero-point', and three distal terms (§8.5):

ha?	'AT'	immediate location of Speaker
he?	'AT:above'	space above Speaker
te?	'AT:across'	space on the same level, but away from Speaker
co?	'AT:below'	space below Speaker

There are four terms in the directional system, two neutral terms distinguished for specificity of direction, and two deictically specified terms (§8.6):

tet	'TO:spec'
te	'TO:unspec'

leg ~ leg	'TO:up'	upward, upstream, uphill
te?en	'TO:down'	downward, downstream, downhill

An example of a locative, co? 'AT:below' and a directional leg ~ leg 'TO:up' are shown in the following example:

- (12) ki=c^bok hn=hayam tom co? 2ate leg dol
3A=throw O=chicken SRC AT:below ground TO:up house
He threw the chicken from down on the ground up to the house.

Even when used to encode a core recipient relation like the indirect object, the locative preposition retains this semantic distinction, as illustrated below:

- (13) ki=c^bok hn=hayam co? ??e?
3A=throw O=chicken AT:below elder.sibling
He threw the chicken down to his elder sibling.

H. Avoidance speech style Semelai has an avoidance speech style, cokop b-sener (speak MID-tease.by.allusion), a system of word substitution, utilising the normal phonology and grammar of Semelai to avoid the utterance of tabooed words, see examples (15) and (16) below. Uttering tabooed words would result in the violation of the prohibition ma=pnon 'prohibitions relating to the consumption of food, and the uttering of tabooed words'.

Neglecting or being unable to fulfil one's desire is also ma=pnon, as a speaker explains below:

- (14) mahn da? da? jalu, hayam, da? sot ma=cəl.
if NEG EXIST pig chicken NEG permit IRR=utter
ma=pnon, yw?=hn ?psol da?, da? da? hal. kna?
IRR=taboo but=CONN when EXIST NEG EXIST problem incur
sko?. ppakit, ?nih do ha? ke, kna? b-blyan
malevolent.spirit illness be.ill OF AT there incur shamanic.ritual
If there is no pork or chicken, one shouldn't utter (the name). It is taboo. However, when there is, it doesn't matter. (When we don't have any, and yet we utter the name) (we) will be afflicted by malevolent spirits. The illness, being sick from them, is cured by a shamanic healing ritual.

When the Semelai enter the jungle to hunt, collect forest products, or prepare a swidden, it is imperative to employ this speech style. The jungle is seen as fraught with peril, so in order to avoid the danger of attack from sko?, 'malevolent spirits', this taboo is used. Failure to do so can result in a range of afflictions including soul-loss. Other consequences are falling victim to a tiger, crocodile, snake or centipede.

The avoidance style also exhibits variation connected to particular locations. In locations of peril, especially those connected to significant events in the creation of Tasek Bera, terms specific to the area are used. In the area of the Tembangau River, one must use cokop t^bmajə (speak Tembangau), where twar 'fish trap' is rhnəh, the fish species tumək is called jalay and ?us 'fire' is pnraŋ, not sma? gdo (person be.old).

Some speakers also extend the use of the avoidance speech style to the home, to prevent the intrusion of snakes or centipedes.

Although avoidance speech styles of this type are common amongst Aslian groups (see Diffloth 1980, and Benjamin in prep.), speech styles of this type are not exclusive to Aslian languages. The Austronesian-speaking collectors of the camphorwood found in Sumatra and Borneo and the southern part of the Malay Peninsula also used styles of this type. Skeat and Blagden (1906) dedicated a whole chapter to this particular linguistic feature.

Skeat and Blagden noted the significance of such speech styles: "The root-idea in all these languages is simple enough: it is merely the avoidance, in an indeterminate number of cases of the ordinary everyday word, and the substitution of something different and out of the common. The primary motive is ... a respectful fear of the superior powers, human, natural or supernatural, as the case may be ..." (1906: 417).

While many Aslian languages have avoidance-type speech styles (Semelai (Diffloth 1980); Cheq Wong and Jah Hut (Kruspe fieldnotes)), they are usually confined to the lexical replacement of nouns as in (15). Semelai is unusual in extending this to include verbs (G. Benjamin, N. Burenhult, G. Diffloth, p.c.), see (16).

- (15) The item replaced is *ʔus* 'fire':

hs=pte?	ʔen	[sma?]	gde]
1&2A=put	LOC	[person]	be.old]

We put (it) in the fire.

- (16) The item replaced is *k^bas* 'to die':

sma?	cløy
person	die

Someone has died.

The semantic domains of nouns covered by this taboo, along with some examples, are shown in Table 1.1. Some examples of verbs, organised into semantic domains, are given in Table 1.2.

TABLE 1.1 THE SEMANTIC DOMAINS OF NOUNS IN THE AVOIDANCE SPEECH STYLE

SEMANTIC DOMAIN	REGULAR TERM	AVOIDANCE TERM
Animals	cnj 'Lesser Mouse Deer'	mə=kurus ḥog (REL=COMP.be.thin foot) 'the one who is fine of hoof'
	tip 'snake'	mə=rus kba? (REL=drag body) 'the one who drags himself'
Forest produce	dre 'rattan'	?urat bri (sinew jungle) 'jungle sinews'
Consumables	ʔiby 'millet'	pos pədəŋ (tail tiger) 'tiger's tail'
	kupit 'turmeric'	jare? ka?ip (digit centipede) 'centipede's legs'
Artefacts	praho? 'dugout canoe'	db? 'large cooking vessel'
Biosphere	ʔare? 'rain'	mə=ramay (REL=be.many) 'the one who is many'

TABLE 1.2 THE SEMANTIC DOMAINS OF VERBS IN THE AVOIDANCE SPEECH STYLE

SEMANTIC DOMAIN	REGULAR TERM	AVOIDANCE TERM
Bodily activities or processes	c?ec 'to shit'	wunwen
	ca 'to eat'	chɔr 'peck at'
Food preparation	c?ɔŋ 'to roast'	kwe?
Cultural activities	yor leŋ d?oh (ascend TO:up swidden) 'to go to (one's) swidden'	yor leŋ lubuk (ascend TO:up pool) 'to go up to (one's) pool'
	c?or 'fire (a swidden)'	k< >wel (roast.(av.)<IMPERF>)
Dimension	t?ay 'be big'	k< >be? (body<HAVE>)
Colour	hitam 'be dark, black'	bihūt

One lexeme can have multiple avoidance lexemes, e.g. jaiu 'pig' is dil muh (be.blunt nose) 'the nose is blunt', mə=ʔəm (REL=pound) 'the one who pounds (the earth)' and at Bukit Rok is called mə=ptidah (meaning unknown). There is also some variation in the definition of terms, e.g. grphop is defined by some speakers as 'hot food', others as 'to eat'.

Lexical items are drawn from the following sources:

- a lexical item which is given a new sense: db? 'cooking vessels, crockery' for praho? 'dugout canoe';
- words exclusive to the register. These lexemes may have cognates in other Aslian languages: hūm 'to bathe' is mhmeħ which is possibly from Semoq Beri hmeh 'to bathe'; k^bas 'to die' is cløy, which is possibly related to Cheq Wong cløy 'to obstruct a path' (Kruspe fieldnotes);
- a lexical item with aberrant morphology: g<r>p>hop (be.hot, plus infixes <r> and <p>) 'cooked food';
- Malay loans with Semelai morphology: kmreŋ (be.dry plus infix <m> 'NMZ') 'water' ← *kering* 'dry';
- a lexicalised associative phrase: kapoh smut (egg ant) 'uncooked rice';
- a noun and an adjective: dloŋ ra-ket (wood COMP.be.small) 'gaharu wood';
- a verb phrase: reŋ dloŋ (seek wood) 'to seek gaharu wood', ko=glaŋ ?luc (3UA=swallow pass) 'to bathe'; gaŋ da? br-gaŋ (bite NEG MID-bite) 'chilli';
- a lexicalised relative clause: mə=kales planton (REL=wipe.anus chopping board) 'the one who wiped his anus (on the) chopping board' is k? 'Pig-tailed macaque (*Macaca nemestrina*)'.

Some terms are semantically transparent, e.g. býuku? 'Giant Malayan Terrapin (*Orlitia borneensis*)' is mə=ʔaj dol (REL=carry.on.back house) 'the one who carries his house on his back', whilst others require recourse to cultural knowledge. The term for the tawo 'White-handed gibbon (*Hylobates lar*)', whose flesh is the highest-prized of land animals, is called jlawat darat (fish sp. land) 'the jlawat of the land', the jlawat being the highest-prized of fish species.

Avoidance terms are generally based on the following:

- physical appearance: ?urat bri (sinew jungle) for dre 'rattan';
- characteristic behaviour: mə=rus kba? (REL=drag body) 'the one who drags himself' for snake;

- association with a traditional narrative, e.g. a lizard species, the *tke?*, gets his name *kapur kloc ditj* (lime inside bamboo), because while hiding inside bamboo he ate lime and betel nut without betel leaf and thus got his coughing-like call.

Many terms are semantically opaque to the speakers. The avoidance term for *hubi?* 'cassava' (from Malay *ubi* 'edible tuber') is *zi=cɔm* (Mr=dig up) 'Mr dig (him) up'. The Semelai lexical item for 'dig up' is *bɔy*, however in Mah Meri, a related Southern language, there is a cognate *cɔp* 'to dig', and in the Central language Jah Hut it is *cwɔm*.

It is possible that the presence of a system of lexeme replacement has resulted in the replacement of indigenous words by Malay loans, as with the word for cassava, above. In some instances, it would appear that avoidance terms have replaced indigenous terms in the ordinary lexicon, the lexical item for python, *dagen jlear* (meat be.long) has the form of an avoidance term.

1.2 The Aslian languages

Aslian languages are a sub-group of the Mon-Khmer language family. Along with Munda, Mon-Khmer is one of the two divisions that forms the Austroasiatic phylum. Austroasiatic languages are spread throughout mainland Southeast Asia, Northeast India and the Nicobar Islands.

The Austroasiatic languages have a long presence in this area, predating the Sino-Tibetan, Tai-Kadai, Miao-Yao, Austronesian, Dravidian and Indo-European families with which they coexist today. With the exception of three Mon-Khmer languages, Vietnamese, Khmer and Khasi (Assam, India), the languages are spoken by ethnic minorities.

The Mon-Khmer language family extends from Vietnam in the east, to Northeast India in the west, and from Southern China in the north, to the Aslian languages in the south, Semelai and Mah Meri (Besisi) being the southernmost members of the family. Mon-Khmer is divided into twelve branches (Diffloth and Zide 1992), although the precise relationship between the twelve is yet to be ascertained. As stated above, the Aslian branch lies firmly within the Mon-Khmer division, and is not a distinct division within Austroasiatic (Ruhlen 1986: 89-91). Aslian belongs in a Southern Branch along with the Monic languages, and possibly also those of the Nicobar islands (Diffloth p.c.).

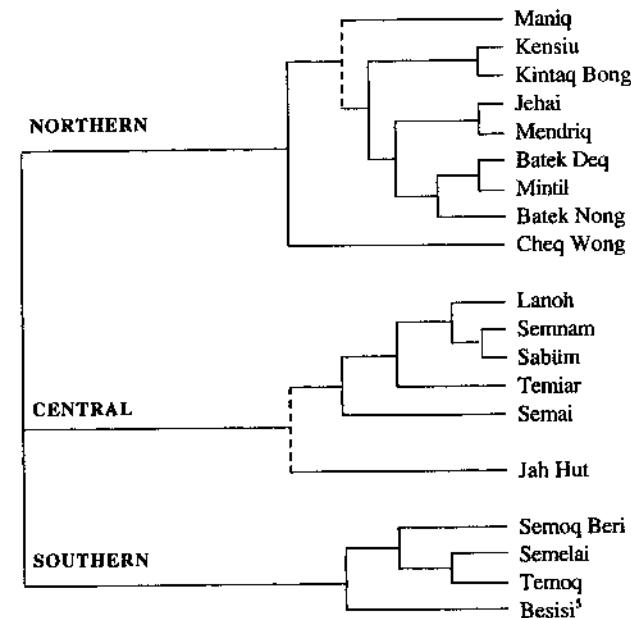
The question of internal subgroupings of Aslian is discussed in the following section.

1.2.1 Internal subgroupings of the Aslian languages

The Aslian languages are recognised as falling into three distinct subgroups: Northern, Central and Southern Aslian (Benjamin 1976a), or Jehaic, Senoic and Semelaic (Diffloth 1976b), see Map 2. Note that both groupings are the same, the difference is only in the labels. The first set is based on predominant geographic location within the Malay Peninsula, the second on either the major language of that group (Jehai, Semelai), or 'racial' group (Senoic). The actual number of Aslian languages is yet to be determined. In Figure 1.1 nineteen languages are represented following Benjamin

(in press). The dotted lines indicate a tentative relationship. For population figures see Table 1.5.

FIGURE 1.1 GENETIC AFFILIATION OF THE ASLIAN LANGUAGES (BENJAMIN 1999)



In an alternative analysis, Diffloth and Zide (1992) names sixteen languages. Subsequent to his earlier investigations Diffloth has since placed Jah Hut outside of the Central branch as an independent sub-branch, and included the Semnam and Sabum in the Central Group (Diffloth and Zide 1992: 139). Unfortunately, the reasons for the removal of Jah Hut from the Central Group were not given, although in his description of the language it was stated that other than lexicostatistical data, there was little evidence supporting the inclusion of this language in Senoic (Central) (1976c: 78).

1.2.2 Geographic distribution

Aslian languages are spoken in all states of the Malay Peninsula, with the exception of the north-west states of Perlis and Penang, see Map 2. Outside of Malaya there are communities of speakers in Southern Thailand: Kensiu in Yala, Trang, Satun and Pattalung Provinces, see Duangchand (1984), Nagata (1992), Bishop (1996a, b), and speakers of Jahai in Narathiwat Province, Benjamin (p.c.). A detailed linguistic map of Peninsular Malaysia and surrounding areas is available in Wurm and Hattori (1983). For a general discussion see §1.3.

⁵ The nomenclature of this group is confusing. The term Besisi is used here, as it is in Skeat (1896), and Benjamin (in press). They are also known as Mah Meri, the name used by Government administration, Betisek (Wazir-Jahan 1981) and Btsisi' (Nowak 1984). See also Kruspe, in prep. a.

1.2.3 The term Aslian

The term Aslian, used to identify the Austroasiatic languages spoken in the Malay Peninsula, was coined by Diffloth 1974 from '*Orang Asli*' ('original people'). Note that whilst 'Aslian' is a linguistic classification, the term *Orang Asli*, from which it is derived, is an 'ethnic' classification⁶ (see §1.4). The primary published adoption of the term was in Benjamin 1976a, and it has since become the standard term used by scholars involved in this area (see Asmah 1979: 4, Huffman 1986).

Prior to the use of Aslian to designate these languages, the suggested name was 'Malacca' languages (Pinnow 1963). The term Malakka or Malacca languages was used in the older literature by Schmidt (1903) and Schebesta (1926), based on the German term for the Malay Peninsula. In more recent times scholars not directly involved in this area have persisted with the term 'Malacca', e.g. Ruhlen (1986), Osada (1992), Peiros (1998). The term is however misrepresentative; Malacca is the name of a small state of Malaysia and the principle city thereof.

1.2.4 Aslian: an historical overview

The fact that the languages of the Aslian family fall into different subgroups was noted by the early investigators. Schmidt (1901) noted that there was divergence in everyday words in the Semang and Sakai dialects. He saw this as being suggestive of a non-Mon-Khmer strata in the Semang languages, possibly the remnants of a pre-Mon-Khmer language in the Peninsula. Schmidt drew evidence for this from the fact that the two groups did not share ethnological unity as discussed in §1.1.3.

As with the classification of Austroasiatic, the first internal classification of the Aslian languages was also put forward by Schmidt (1901):

SEMANG (northern)

SAKAI (southern): Sakei I
Sakei II

Using a test of distinctive vocabulary, Schmidt distinguished a north/south division, based on the presence of a non-Mon-Khmer stratum in the Semang languages, which was absent in the Sakai group where definite affinities with Mon-Khmer were evident. The second subgroup contained a further two divisions: Sakei I and II.

Sakei I is recognised as today's groupings of Temiar and Lanoh, and Sakei II as containing Semai and Besisi. Languages of the south such as Jakun were deemed not to be of the Austroasiatic stock. Besisi is now recognised as belonging to the Southern division (Benjamin 1976a: 38) and it has been discovered that there are languages in the south that are Austroasiatic, Semelai being one example.

Following the extensive work carried out in the linguistic survey of Skeat and Blagden (1906), Blagden reviewed his earlier classifications and came up with the following one which he maintained was both tentative and provisional. It was based not only on lexical data, but also took into consideration phonological criteria. Blagden's classification was based on the materials that comprised his 'Language and Comparative Vocabulary of Aboriginal Languages' (1906). This was put together

⁶ It is described in Government literature as 'a heterogenous category of numerically small and scattered indigenous "ethnic" groups in West Malaysia', *Jabatan Hal Ehwal Orang Asli* (1972). For problems with this classification see Hood (1990).

from word lists in a variety of orthographies which were collected by numerous authors untrained in linguistics, who often did not know the aboriginal language, and like the speaker they were dealing with, were not fluent in Malay. Blagden's groupings are given below in Figure 1.2. The bracketed language names on the far right follow Benjamin (1976a: 39) and denote the present day language names corresponding to Blagden's original groupings.

Blagden's classification of 1906 was the ground stone for the present day classification. Clarification took place as a result of further investigations involving the collection of more reliable materials and as clearer differentiation between individual languages was ascertained.

FIGURE 1.2 BLAGDEN'S ASLIAN LANGUAGE GROUPINGS (1906)

Semang:	1. Semang-Pangan 2. Low Country Semang	[Northern Aslian] [extinct]
Sakai:	1. Northern Sakai 2. Central Sakai 3. Southern Sakai: a) South Western Sakai b) South Eastern Sakai 4. Eastern Sakai: a) Inner group b) Outer Group	[Temiar, Lanoh] [Semai] [Besisi] [Semelai, Temoq] [Jah Hut] [Semaq Beri]

Blagden's classification of the northern languages was advanced upon by P. Schebesta (1926), as shown in Figure 1.3. However Schebesta's treatment of the Southern languages proved incorrect. Again the bracketed names on the right identify the languages as they are known today (Benjamin 1976a: 41). The names on the left indicate the major subgroup.

FIGURE 1.3 SCHEBESTA'S ASLIAN LANGUAGE CLASSIFICATION (1926)

1. North Semang (Meni')	a) Tonga Mos (Sth Thailand) b) Kensi'lu c) Kenta': (i) K. Nakii (ii) K. Bogn	[Kensiw, Maniq] [Kensiw] [Kentaq Bong]
2. East Semang (Menra')	a) Jahay b) Menri'	[Jehai] [Mendriq]
3. South Semang (Batek)	a) Batek Nogn b) Batek Hapen c) Batek Kleb d) Temo'	[Batek Nong] [Temoq] [Batek Kleb]

For the next half a century no further advances were made until Benjamin's classification of 1976. The Southern languages in particular had remained somewhat of an unknown entity up until this time. Williams-Hunt (1952: 23) had proposed a new classification, however, like so many other researchers he complicated facts by failing to understand that 'racial', cultural and linguistic criteria were not concordant. Classification of the 'Sakai' languages was plagued with inconsistency, partly due to this persistent oversight amongst investigators. Classification of the Semang group on

⁷ The Kensiw of South Thailand refer to themselves as Maniq. Bishop (1996b).

⁸ The Temoq are definitely misplaced in Schebesta's classification. They are in fact in the Southern group, not members of a southern branch of the Northern group as suggested here.

the other hand had always been relatively consistent, due no doubt to their distinct appearance and the fact that 'racial' and linguistic criteria are aligned.

Benjamin (1976a) preserves the Northern Semang group but finds evidence to class Schmidt's Sakai languages into two groups: Central and Southern. His paper presents a major advance in that the classification is based purely upon linguistic data. Benjamin collated word lists of twenty Aslian languages and dialects⁹ of the Malay Peninsula and carried out lexicostatistical and glottochronological surveys. The author describes the resulting classification (see Figure 1.1) as '... a heuristic classification of the Aslian languages ...' and suggests that future work may greatly alter the picture (1976a: 66). While Benjamin's work is to be commended, the fact remains, as he himself admitted, that it is largely based on a single word list for each language, and the presumption that all Aslian languages developed from one proto-language.

Unfortunately, to date no further work has been published in this area, other than Diffloth's slightly modified version which, however, doesn't provide any supporting argument (Diffloth and Zide 1992: 139).

Little information is available regarding the relationships of the Southern languages, other than what can be gleaned from available publications. The only linguistic information is word lists of Temoq (Collings 1949a) and Besisi (Skeat 1896, Wazir-Jahan 1981) and some preliminary remarks on Semoq Beri (Nik Safiah Karim and Ton binti Ibrahim 1979). Although there are only four languages recognised in this group (making it the smallest of the three) they represent a wide distribution, extending from the northern state of Kelantan down through Terengganu and into Pahang. In the west, the Besisi are the only coastal dwelling Aslian people, isolated from all other Aslian communities.

The Southern languages appear to have separated early, and developed into distinct languages. Semelai and Temoq appear to be the closest, but all four languages exhibit inter-language loan rates of less than 3 per cent and a low modal cognacy percentage of 38 per cent (Benjamin 1976a: 73).

1.2.5 Linguistic affiliation of Aslian with Mon-Khmer

The relationship of the languages of the Malay Peninsula and Mon-Khmer was first noted late in the nineteenth century, however there was reluctance amongst scholars to affirm a definite relationship. Kuhn (1899, cited in Schmidt 1903: 38) noted similarities in the mono-syllabic root stock of the Khasi-Mon-Khmer group with the languages of Nancowry (Nicobar) and the aboriginal dialects of the Malay Peninsula. He concluded, "We should not however be justified in deducing there from an ancestral connection with these partly polysyllabic languages."

Blagden refrained from claiming a definite relationship, arguing: "But even to assume that the aboriginal dialects are cognate languages which should be classified in the Mon-Amman family would be going further than our evidence justifies us in doing" (1894: 42).

⁹ The nineteen languages and dialects used by Benjamin (1976a): Kensiu, Kintuq Bong, Jehai, Mendriq, Batek Deq, Mintil, Batek Nong, Che' Wong, Semaq Beri, Temoq, Lanoh Yir, Lanoh Jengjeng, Semmam, Sabum, Temiar, two dialects of Semai, Jah Hut, Mah Meri, and Semelai. He personally collected all but two of the lists using Malay as the medium. The list for Batek Nong was collected by K. Endicott; Kintuq Bong data is from Asmah (1963).

Pater W. Schmidt is recognised as 'discovering' the Austroasiatic family. He grouped the Munda languages, Khasi, and the Nicobar, Palaung-Wa, Mon-Khmer, Malacca (Aslian) and Cham languages into an '*Austroasiatischer Sprachstamm*'. As a result of examining all published word lists and available texts, he presented what he considered to be the only possible conclusion, "... that there exists an inward and intimate condition between the Sakai and Semang languages and those of the Mon-Khmer" (Schmidt 1903: 43).

Schmidt was prepared to claim this despite a small number of words in the Aslian languages which did not have corresponding forms in Mon-Khmer. He felt the connection could not be discounted even though there existed a paucity of information regarding some of the out-of-the-way Mon-Khmer dialects.

Schmidt's attempt to classify the languages on the available data resulted in the following four Austroasiatic subfamilies:

1. Munda and Mon-Khmer, including a younger Malacca group Besisi and Jakun¹⁰ in the Mon-Khmer group;
2. Malacca, including Semang and Sakai;
3. Central Group including Khasi, Nicobar, Palaung-Wa;
4. 'S. E. Mixed Group' Chamic Languages.

In 1906, Blagden put forward a new classification based on his study of word structure, the morphological processes of prefixation and infixation, and vocabulary common to both Aslian and other members of the Mon-Khmer family. He believed there was a 'double relationship' between the Aslian languages and Mon-Khmer, reflecting the fact that the languages had first separated from Mon-Khmer, and then later had renewed direct contact with a Mon-Khmer population (1906: 453). Blagden claimed evidence for this in the diverse numeral systems exhibited by the Aslian languages, and on this basis established four language groups: Type I, the Northern Semang group, Types II and III the Sakai, now termed the Central group and Type IV, the languages recognised today as the Southern group. Languages of Types I-III had numeral systems limited to three, whilst those of Type IV were the only languages which possessed numeral systems extending beyond three cognate with Mon-Khmer. On the basis of this, his Type IV or outer languages were deemed to be directly related to Mon-Khmer proper. His Sakai Types II and III were "outlying members" like Palaungic, Khasi and Nicobarese, while his Type I, Semang, the Northern group, were not considered members of the family at all given that the numerals bore no resemblance either to Types II and III which were similar to each other, or Type IV (1906: 455-9).¹¹

Over the next half a century debate continued as to where the Aslian languages belonged in the Austroasiatic family. In the 1960s they were excluded from Austroasiatic altogether in Richard Salzner's *Sprachenatlas des Indopazifischen Raumes*, but reinstated by Pinnow, who considered the 'Malacca' languages, despite being gradually assimilated into Indonesian, as Austroasiatic because of the number of definitely Austroasiatic elements they contained (1963: 140-1).

¹⁰ Jakun is now recognised as an Austronesian language.

¹¹ For a recent analysis of Mon-Khmer numerals in Aslian languages see Diffloth (1976d), and for a qualification of the relationship of Senoiic and Mon-Khmer, cf. Diffloth (1977).

In a later classification of Mon-Khmer by Thomas and Headley, the Austroasiatic phylum is composed of at least four families: Munda, Mon-Khmer, Malacca, and Nicobarese (1970: 405). The 'Malacca' languages were tentatively considered a separate branch on the basis of low cognacy percentages.

It was only in the seventies with the work of Diffloth that the status of the Aslian languages within the Mon-Khmer and hence the Austroasiatic family was substantiated. Diffloth (1979) used comparative typological materials to establish the relationship, focussing on: phonological aspects, especially the presence of final palatals /c/, /ʃ/ and /p/ in Aslian words (final palatal stops, although present, are infrequent in Austronesian), and phonemic vowel length; the Mon-Khmer-type morphology associated with the nasal infix and processes of nominalisation; semantic similarities attested in the pronominal systems; and the existence of a word class of expressives.

Around this same period, as a result of his lexicostatistically-based investigation into the genetic affiliation of Aslian languages, Benjamin (1976a: 92-3) proposed that division within Proto-Aslian began earlier than within Mon-Khmer proper, and that Aslian may have been the first group to separate out. The conclusion was based on the fact that the lowest intra-Aslian cognacy rates (Northern-Southern Aslian) are in the same range as those of Temiar and Mon-Khmer, and intra-Aslian lexical evidence assumes a single proto-language.¹² He concluded, echoing Blagden's stance (1906), that following the first arrival of Austroasiatic speakers in the Peninsula, there must have been a second period of contact at a later date long after both groups began to separate (§1.2.6). The reader is referred to Benjamin (1976a).

In a complete turn around, Parkin (1991) who demonstrates little understanding of historical linguistics in his book, claims that there are four branches in Austroasiatic, and disregarding Diffloth and Benjamin, returns to Thomas and Headley's (1970) analysis placing Aslian outside of Mon-Khmer, giving the branches Munda, Nicobarese, Aslian and Mon-Khmer. He adopts the old argument of the low cognate percentages shared with Mon-Khmer languages, for instance only 11-16 per cent for Temiar, and also erroneously assesses Aslian as being closer to Munda than Mon-Khmer on the basis of structural similarity. An outline of the features of Austroasiatic, and the distinctive features of Mon-Khmer and Munda can be found in Diffloth and Zide (1992: 140-2).

Ultimately, until we have a more complete data base of all branches of Austroasiatic, such classifications are only provisional. For consensus on this issue see also Diffloth (1990: 125-6) and Peiros (1998: 104-14).

1.2.6 Prehistory of the Aslian languages

A full discussion of the prehistory of the Malay Peninsula is beyond the scope of the current work, however I will endeavour to outline the major issues here.

The Negritos of Peninsula Malaysia, the Andaman Islands and the Philippines are thought to be representative of the original peoples who populated South-east Asia. The Negrito presence in the Malay Peninsula, is estimated by Endicott and Bellwood at 10,000 years (Bellwood 1997: 133), and is believed to be either fully or partly ancestral to the present day Orang Asli. Bellwood's view, which is not shared by all,

¹² See Peiros (1996) for his provisional lexicostatistical classification of Austroasiatic languages.

is that the Negritos predate the arrival of the probably Austroasiatic-speaking Southern Mongoloid peoples from the north. However, the absence of any signs of change in the phenotypes in the sites excavated so far confuses the picture presented here, and supports an alternative view, that the Negritos are a recent offshoot of the Mongoloid peoples (Fix 1995 and 2000, Bulbeck 1996).

On the basis of archaeological evidence, the emergence of the Neolithic Ban Kao culture in the Peninsula marks the arrival of Austroasiatic speakers, possibly speaking a language related to Mon (Benjamin 1997b). The people associated with this culture travelled south from Thailand down the Isthmus and initially populated the western side of the Peninsula (Bellwood 1997). The Ban Kao culture in Thailand is well documented, and included rice cultivation and the exploitation of maritime resources. Artefacts linked to this period are dated from 3000 B.P. as far south as Central Malaysia, and represent a rapid migration from north to south, colonising the Peninsula (Bellwood 1997: 258). This migration is associated with the emergence of agriculture in the Peninsula, based on current archaeological records.

If the current stance that all the Aslian languages originated from a single proto-language is maintained, the three major splits that ensued spread the languages to the south and east across the main mountain range which runs parallel to the west coast (Benjamin 1997b). It is possibly as a result of this move that Austroasiatic languages replaced the original languages of the Semang populations, although the necessary lexical evidence to support such a claim is absent. In agreement with claims first made by Blagden (1906: 470), the three divergent language groups and their established life-styles suggest a long presence in the Peninsula. This is evident from the linguistic patterns of inter- and intra-Aslian relationships (Benjamin 1976a). This model is accommodated by supporting evidence from both biological and archaeological investigations (Fix 1995).¹³

Benjamin suggests that despite the early movement away from the rest of the mainland language family, some form of contact between Aslian languages and southern Mon-Khmer languages continued long after both groups began to diverge. This appears to be attested by the presence of two strata of Mon-Khmer in the Peninsula, first noted in Blagden (1906). One area where this is supported is the unusual feature of South Aslian languages having Mon-Khmer numerals up to five or seven. All other Aslian languages only have number terms up to two or three. These Southern communities are located along the main waterways of the interior which have been authenticated as ancient trade routes. These communities possibly borrowed the numeral terms from their Mon-Khmer speaking trading partners (Wheatley 1961).

The continued contact with Mon-Khmer speaking people is evident in historical records. Mon presence in the isthmus was only reduced by the arrival of Tai peoples in 1200 A.D. In the state of Pahang, it was only in 1455 that the Malaccan Malays took control and ended the long 'Siamese' influence there (Benjamin 1987, 1997c).

¹³ Preliminary investigations of the biological diversity of the Aslian populations in terms of mitochondrial DNA testing of a small sample of the Aslian population, and independent analyses of craniofacial and dental morphology have merely placed the Aslian population as a uniform group, descendants of a southern Mongoloid stock from the late Pleistocene, along with other Southeast Asian peoples, including Koreans and Malays (Fix 1995: 316). Independent studies of blood genetic markers (Fix 1995; Saha et al. 1995) have provided genetic evidence of a possible closer relationship between certain Aslian populations and the Khmer on the basis of a particular gene which is considered a Mon-Khmer gene, and has a low frequency in Malay populations (Fix 1995: 319).

This area has a long history of influence from Mon and Khmer culture.¹⁴ This is borne out by the obvious Mon-Khmer nature of Aslian languages, but the virtual lack of influence from Tai (Blagden 1906: 468).¹⁵

The arrival in the Peninsula of the Malay language is only recent. With the arrival of the Austronesians in the north of Borneo (present day Sabah) around 4000 B.P., the earliest possible date of arrival in the Malay Peninsula is 2,500 B.P. (Collins p.c.). Benjamin proposes 2000 B.P. in the south, and as little as 1000 B.P. in the north (Benjamin p.c.). It is assumed that prior to this the area was occupied entirely by the Austroasiatic speaking peoples mentioned above, possibly today's Aslian plus a Mon-Khmer presence, originating from mainland Asia to the north. Dunn (1975) maintains that even up until last century Malay communities were located primarily in the coastal regions, with internal settlements along the Pahang River which flows east, and in the hinterlands of Negeri Sembilan in the west which were settled by immigrants from Sumatra, notably the Minangkabau.

1.3 The Semelai and their language

Semelai is spoken predominantly in the state of Pahang, with smaller populations resident in the neighbouring states of Negeri Sembilan and Johore. There are four main areas where they reside today: Tasek Bera (Lake Bera), Sungai Bera (Bera River), Sungai Seriting (Seriting River) and Sungai Teriang (Triang River). The largest concentration of speakers is located at Tasek Bera.

In earlier times, the Semelai were reportedly more widespread, their territory extending south to the Endau and Palong River basins in Johore state (see §1.3.2) and north to at least the banks of the Pahang River. Today there are no known settlements in the former area and only isolated pockets in the latter.

Needham (1974) claims the Semelai are also located in Northern Pahang, but it seems more likely that he has confused them with the Semoq Beri, who are called Semelai by the Northern Aslian Batek Deq speakers (Kruspe fieldnotes).¹⁶

The Semelai have little contact with neighbouring ethnic groups. The Temuoq, who live to the west on the Jeram River, are the closest neighbours of the Semelai, both geographically and linguistically. They are referred to as *sma?* *bri* (people jungle), rather than Temuoq, reflecting their lifestyle as foragers. To the Semelai the term *tmo?* has pejorative connotations, and means a people less 'cultured' than themselves. According to Semelai narrative, the Temuoq were the original inhabitants of Tasek Bera (see §1.3.3.1).

The other neighbouring Orang Asli groups are speakers of Malay dialects: to the west are the Temuan, who the Semelai call *sma?* *nap* (person nap is the Temuan word for 'no'), and to the south and south east, the Jakun. For an analysis of Semelai ethnicity, interrelationships with other ethnic groups and their claims to shared origins with the Malays in Sumatra, see Gianno 1997.

¹⁴ The absence of Siamese influence until relatively modern times, 1000 A.D., indicates that it was more likely a Mon or Khmer presence.

¹⁵ For an interesting account of Aslian, and in turn Mon-Khmer influence on Malay toponyms and ethnic groups in the state of Pahang, see Benjamin (1997b).

¹⁶ The Northern Aslian Cheq Wong speakers have a traditional narrative featuring a tribe of amazonian women swidden-makers called the *bət̪ səməlay* (person Semelai). *səməlay* has no meaning in Cheq Wong.

The following Semelai creation myth establishes the relationship between themselves and the Malays. The world was created from the blood and flesh of a snake. The master created the prophet Adam from the clay and gave him his life and courage. Satan gave him his fear. A wife fell to earth and she bore forty-four children. The first forty were married to each other and became the peoples of the earth. Of the four remaining, two were chopped up and flung into the air. Some pieces fell to earth and became fish and meat, other pieces became vegetable foods. The two remaining brothers were knocked back to earth by the wind as they tried to climb to heaven. One fell to a sitting position and became the Malay Sultan. The other fell to an apparently subservient kneeling position as his quiver stuck in the ground. This brother was the first Semelai.

1.3.1 Ethnonym

The etymology of the ethnonym Semelai is uncertain. The Semelai themselves have no explanation, and the word *smlay*, does not occur elsewhere in the language. However, in the neighbouring Senoi language Jah Hut, there is a word *slay* – *hlay* 'swidden, ladang', from the Proto-Senoi **slay*, cognate with the Khmer *slay* 'fallow land'. There is an agentive infix <*m*> in Jah Hut, Semelai and other Aslian languages 'the one who/thing that does V': *k<m>bac* 'fishing pole' from *kbac* 'to fish with a pole'; *s<m>dər* 'a good rememberer' from *s<r>dər* imperfective of *sər* 'to remember'; *k<m>tkət* 'the class of stinging insects' from *kt-kət* imperfective of *kət* 'to sting, of insects'. It is possible that the ethnonym *s<m>lay* could be 'those who make swiddens'. For independent evidence for this etymology, see Benjamin (1997b: 111).

It is not known how long the Semelai have associated themselves with this name, but they do use it when referring to themselves in relation to non-Semelai society. Within their immediate environment they simply name themselves according to the location in which they dwell: the people of Tasek Bera refer to themselves as *sma? ha?* *tasik* (people at lake) 'people of the lake'; other people are known likewise by their geographical location: *sma? ɻayer hitam* (people water black) people from the kampong of Air Hitam (Seriting River).

1.3.2 Prior documentation of the Semelai

The recognition of the Semelai as a distinct group with a distinct language was only convincingly established in the 1940s. Prior to this, they remained as one of the 'mixed and doubtful' tribes of the relatively inaccessible southern region.

The earliest published mention of the inhabitants of the region of Tasek Bera is found in the posthumously published account by Gray (1852). Reputedly the first European man to penetrate the interior, Gray journeyed from Malacca to Pahang in 1827, and reported the presence of *Orang Hutan* (jungle people), as his native companions called them, frequenting the river to fish.

In 'Language and Comparative Vocabulary of Aboriginal Languages' (1906), Blagden mentions a manuscript containing words collected from a Mentera tribesman at Nyalias in Malacca Territory. The words are reportedly from the 'Semilai' tribe of an unknown locality. He noted that this was the same dialect as that which he had collected from the Seriting area in Pahang, however no connection was made between

the two. Examination of the Comparative Vocabulary reveals Semelai classified as South-eastern Sakai: "... cognate dialects (of Besisi) ... recorded in the Serting valley of the Negeri Sembilan, the Bera valley in South Pahang; and also on the Upper Pafong and Endau (or Indau) in Northern Johore, where, however, they are obsolescent, if not extinct" (Skeat and Blagden 1906: 397). This area conforms to that inhabited by the Semelai today. Skeat and Blagden also tentatively included the mixed Beduanda dialect of Chiong reportedly spoken at that time in Johol, Negeri Sembilan. Today, this area is inhabited by Malay speakers.

Specimens of the languages recorded in Blagden (1906) are given in Table 1.3. My data for Semelai has been added in the right-hand column.

TABLE 1.3 COMPARATIVE SOUTH-EASTERN SAKAI FORMS (BLAGDEN 1906)

	SERTING	BERA	ULU INDIAU	SEMELAI
BANANA	tiok		diok	[t̪iɔk]
BELLY	lepot	leput	lopot	[ləpət]
EYE	mot	muat	mot	[mət]
HOUSE	dol	dəl	dol	[dəl]
WATER	dak	dak	dak	[dak]
WIFE	kempun	kempung	kompotn	[kəmpə'n]

The confusion regarding the Semelai continued. Ivor H. N. Evans, assistant curator and ethnographical assistant for the Federated Malay States Museum, led a number of expeditions investigating the various aboriginal tribes of Pahang and Negeri Sembilan. In 1913, whilst on one of these expeditions, he encountered a Serting River person in the town of Bahau. Evans noted the language as being similar to that of the mixed peoples of the Pahang River. These 'Serting Jakun', as he referred to them, were undoubtedly Semelai. He reported that they were called *Orang Bukit* (people hill) by the Malays, or *Sakai Semlai* (or *Semleh*) (1915: 106). Evans understood the latter to refer to the language rather than the peoples themselves. However, he was unable to determine a reason for the term. He claimed that the people refer to themselves as *Berkтурk Chong*, which he erroneously translated as the equivalent of *Orang Bukit* 'Hill People'. 'Berkтурk' (obviously /bktək/) is recorded as the entry for 'person' in the accompanying word list. It in fact means 'weeds or undergrowth'. bktək 'undergrowth' is possibly a pun on the Malay word *semak* [səma?] 'undergrowth, scrub', a homonym with *sma?* [səma?] the Semelai word for 'person'.

In addition to an extensive word list of over 250 entries, of which 'person' is the only mistake, Evans also includes some of the first ethnographic information on Semelai habitat, livelihood, religion and superstitions, burials, marriage and circumcision.

In 1920, Evans published a second report, this time relating to a trip made in Pahang in 1917. Mention is made again of the '*Semilai*', this time as a 'wild tribe' occupying country some distance from the Pahang River, according to the Rompin Jakun from whom he learnt this information. In a footnote (Evans 1920: 27) Evans repeats his belief that the term *Semilai* is used by speakers of Malay to refer to speakers of Sakai. He believed them to be of 'mixed blood' as was the term for certain peoples at that time, i.e. proto-Malays speaking a Sakai language. Evans never reconciled the Serting and Bera peoples as one Semelai people, despite the excellent word lists he obtained.

Schebesta (1926) was the first person to associate the Jakun Semilai with the Bera River drainage basin, estimating a population of 2,000, the majority in the locality of Tasek Bera.

The first ethnographic account of the Semelai at Tasek Bera was recorded by Collings. Collings is remembered by some older members of the community who recall him as being conversant in their language. He visited the area in 1940 and then returned following the Second World War in 1947. He published two articles including a word list of Temoq in 1949. Ethnographic notes and the 'Myth of the Inundation', which he recorded at Tasek Bera, were published in Collings (1949b). Williams-Hunt (1952), like Evans before him (1915 and 1920), interchanged the terms Jakun and Semelai, failing to realise the two as different.

The first extensive ethnographic report was by Hoe Ban Seng (1964). In more recent times, investigation into Semelai ethnography and culture has been continued by Hood (1974, 1978) examining aspects of shamanism and healing rituals, and Gianno (1990) on culture and resin technology.

1.3.3 The Semelai at Tasek Bera

This section represents a brief summary of the area where fieldwork was carried out and the community that lives there. For detailed ethnographic studies of the Semelai see the references in the previous section.

1.3.3.1 Tasek Bera

Tasek Bera is less than 25m above sea level, located between the main east and west ranges in the southern lowlands of the Malay Peninsula. It is located 3° 5' N, 102° 38' E in Southwest Pahang State, close to the border of Negeri Sembilan. The nearest town is Bahau in Negeri Sembilan.

Tasek Bera is a tropical lowland freshwater alluvial peat swamp, containing species of fauna and flora unique to this location. It consists of a main channel system of areas of open water, interspersed with wide areas of Lepironia reeds, clumps of pandanus and impenetrable swamp forest. The water is black, due to low depth and the decaying vegetation on the waterbed. Four rivers and many smaller streams feed the swamp. The area around the lake is typical lowland forest dominated by Dipterocarpaceae and Buseraceae known as *meranti-keruing* in Malay.

Tasek Bera is situated on a north/south axis and drains through a single point northward into the Bera River. The aforementioned Serting and Teriang Rivers are tributaries of the Bera River. The Bera River drains into the Pahang River, the largest river in the Peninsula.

Tasek Bera experiences two wet periods each year. From late September to January the north-east monsoon brings heavy rains and the flow of the lake is reversed as the rivers flood downstream. This effects a change in water level between one to five metres and creates flooding within the swamp. The secondary wet period is from late March to May and is brought by the southwest monsoon. It is often little more than a few showers of rain which result in the area experiencing a very wet season during the main monsoon followed by an extended dry season.

It is estimated that the lake system evolved around 4,500 B.P. Soil sampling conducted by Lim, Furtado and Morley (1982) indicated that Tasek Bera has been settled by people practising swidden agriculture since 600 B.P. ±75. It is calculated

that at this point swidden cultivators migrated into the area, or alternatively those present adapted their life-style from one of hunter-gatherer to shifting agriculture. These people may have been the Semelai or their predecessors.

According to the traditional narrative, 'The Story of the Lake' (Gianno 1985: 461), the original inhabitants of Tasek Bera were the Temoq. The Temoq, the hunter gatherers who occupy an area to the east of the lake, invited the Semelai into their area, explaining that it was a pleasant place to live and food was abundant. The Semelai accepted, and after numerous attempts succeeded in entering the lake system. According to the narrative, the Semelai originated downstream. It is stressed in the telling of the story that the Semelai only arrived at Tasek Bera in recent history, a fact which would appear to coincide with the history of the area. The Semelai may have been forced into this area as state-formation intensified along the Pahang River, where they had originally been resident. This in turn can be related to the commencement of swidden farming at Tasek Bera which occurred at around the same time (based on the date of 600 B.P. from core sampling reported by Lim, Furtado and Morley 1982).

These observations are also reflected in traditional Malay narratives. One such legend, 'The Original Settlement of Pahang' (Sircum 1920) was told by a person of Minangkabau descent, in an area of Pahang which was formerly part of Negeri Sembilan. It relates how the first Malay settlement in the area was at Kuala Bera. Kuala Bera, the confluence of the Bera River and Pahang Rivers, was a traditional location of the Semelai. The settlement resulted after a Malay travelling through the area was given a Sakai princess whom he married. The children born from the marriage became the forefathers of all the Malays of Pahang. In the legend it is stated that 'Pahang was then thickly populated with Sakai'¹⁷ (Sircum 1920: 150). The name of the narrator's village '*Temai*', claimed by him to be a Sakai word, is indeed a Semelai word meaning 'earlier or previous', e.g. *troy tmay* (trail previous) 'the previous trail'. Many Semelai narratives also tell similar stories of intermarriage between the Semelai and the Malays, see §14.4.4.

Despite inhabiting an area rich in resources, Semelai population is relatively low. This could be attributed to the fact that this was an area where extensive slave-raiding, carried out by the Minangkabau Malays, was still occurring last century (Endicott 1983). Slavery was officially abolished in 1884, but continued up until the early 1900s (Gomes 1990). Elder Semelai still recall this period: one woman recounts how her grandmother was named *pnaoh* 'escape' because she was born during the flight from a slave raid.

Residential Location Prior to the period of the Communist Insurgency (1948-1960), residential location changed each year as a new swidden was established. Settlement patterns were dispersed, centred towards the lower reaches of the lake. The upper areas around Tasek Damparan and Paya Bangkong at the southern end of the lake were not inhabited, see Map 3.

Today, the majority of the Semelai community at Tasek Bera live on the south-west shores in a cluster of settlements around Pos Iskandar (see Map 3). Pos (Fort) Iskandar was established during the Communist Emergency in the 1950s by the British as a military post. The establishment of the fort was considered necessary to prevent the Orang Asli from being recruited by the Communists for their knowledge

of the hostile environment of the jungle. This area was the site of the last stronghold of the Communists in the southern part of the Peninsula. The result of the Communist Emergency has been the relocation of the Semelai into Government settlements: Pos Iskandar, Bukit Rok, Kampong Sungai Lui, Ayer Hitam, Pancur, and Batu Lima in Johore. Initially security installations, these are now seen as a means of settling the people in order to facilitate the development and assimilation of the Semelai into the wider community. Today Pos Iskandar is a Government-run outpost consisting of the homes of employees, an elementary school, meeting hall, rest house and paramedic clinic.

Houses are built in a non-traditional style from sawn timber and iron roofing, materials brought in from outside. There are traditional Semelai houses built by those unable to afford purchased materials, or those who prefer this style. Most households have access to a motor bike, and a few people own cars. Each kampong has a shop which stocks minimal supplies: rice, tinned condiments, cigarettes, alcohol, kerosene and petrol; and act as agents for *taukes* 'Chinese middlemen' who buy rubber, dammar and rattan. A road provides access to Ladang Geddes, a settlement in a nearby estate where shops and a medical service are available.

Around the kampongs are small holdings planted with rubber trees and then jungle in secondary growth, a dense undergrowth of grasses, scrub, climbing bracken and saplings. Some families still prefer to live away from the settlements temporarily if not permanently in order to maintain a more traditional lifestyle tending their swiddens. The two most popular locations are downstream around lubuk krweng and on the eastern side of the lake.

1.3.3.2 Ethnographic background

Social organisation Semelai society is unranked, and individuals have a great deal of autonomy. Within the family unit a husband has no given authority over his wife and parents have limited authority over their offspring. Children are mildly rebuked, but otherwise they do as they please. There are few domains in daily life which are divided on the basis of gender.

Households have fluid membership; they are primarily nuclear, but may include parents, in-laws, a sibling's offspring and so forth. Within a settlement, houses of related kin, usually parent/child, tend to be clustered together.

There is no individual ownership of land or the waterways. Individuals may have rights to the crops of fruit trees planted by that person or their kin, and in the past people had inherited rights to tapping specific dammar trees.

There is no indigenous political structure, beyond the level of the village. Each of the four main Semelai areas has an elected *baten* 'headman'. These people have a representative political role with the *Jabatan Hal Ehwal Orang Asli* (Aboriginal Affairs Department). Otherwise leadership is only for the jurisdiction of crimes: theft, adultery, divorce and theoretically murder, although personal violence is virtually unknown. These processes are carried out in consultation with selected elders.

Kinship The Semelai kinship terminology of same-generation relatives is of the Hawaiian type: parallel and cross-cousins, as well as siblings, are referred to by the same terms. There are no exogamous unilineal kin groups. All kinship terms are bilaterally symmetrical. Kinship terms are not used sociocentrally.

¹⁷ Sakai refers to the indigenous peoples of the Peninsula §1.4.1.

The kin terms are based on relative birth order, loosely based on the Malay system. Kin terms are commonly used to address, or refer to people unrelated to the speaker.

The basic terms for ascending generations above grandparents, and descending generations below ego, do not differentiate gender. The first three terms in the table, and the last three, also include the siblings of the particular relation, e.g. grandmother/grandfather, great-aunt/uncle and so forth.

The terms child, sibling's child and all other descending generations are not used as address terms; instead one uses a name.

Some speakers make an interesting distinction between 'elder male sibling' ??e? 'EB', and elder male cousin ?i?e? 'EB'; for other speakers, the two terms are interchangeable. This distinction is not made for elder female sibling and elder female cousin, kaka? 'EZ'. These terms, used for cousins, are calculated not by the chronological age of the siblings, but through the ascending generation, the age of one's parent relative to their sibling.

TABLE 1.4 KINSHIP REFERENCE AND ADDRESS TERMS

FEMALE	MALE	
munet		'great-great-grandparent'
mu?pnq		'great-grandparent'
tu?wan	tu?aki	'grandparent'
?ma?	bapa?	'parent'
ma?ambon	wa?ambon	'first born'
ma?nah	pa?nah	'second born'
ma?ondo?	?ondo?	'third born'
ma?itam	pa?itam	'fourth born'
?uda?	pa?uda?	'fifth born'
ma?lan	pa?lan	'sixth born'
kaka? (ga?u?)		'elder sibling, cousin'
	?e?e?	'elder sibling'
	?i?e?	'elder cousin'
?adi?		'younger sibling, cousin'
kn?n		'offspring'
km?n		'sibling's offspring'
cu?		'grandchild, siblings' grandchild'
cet		'great-grandchild'
kump?t		'great-great-grandchild'

The term for elder female sibling, ga?u?, has been replaced by the Malay term kaka? 'EZ', although speakers are still aware of the original term. Evans (1920) recorded ga?u? 'EZ', and it was still in use in Hood (1974), but rarely attested as reported in Gianno (1985). This seems to be in line with a general replacement of the indigenous terms by terms from the Malay system.

Note that, anomalously, the birth order terms for third-born male, ?ondo?, is not preceded by pa? 'father, or parent's male sibling' and the term for fifth-born female, ?uda?, does not include the term ma? 'mother, or parent's female sibling'.¹⁸ The non-gender-based address terms for the eldest and youngest-born parent's sibling are

¹⁸ ?ondo? is possibly from Malay *andak*, a family name for the 'fourth or fifth child', if female. The Malay term for a male is *pandak* (=Malay *pendek* 'short'). ?uda?, possibly Malay *uda* is a name for the fourth- or fifth-born child, or a general name for a parent's younger sibling, either male or female (*uda* is from the Malay *muda* 'young') (Wilkinson 1927).

bo? and bu?u? respectively. Comparative reference terms are me=bomonj 'the one who is eldest' and me=bu?u? 'the one who is youngest'.

One's parents may be referred to collectively as gade? 'parents', and parents' siblings as gade? prtwan.

Distinctions may also be made between ?adi?bradi? 'siblings, including one's cousins' and specifically one's first cousins, ?adi?bradi? bdiri? gade? (siblings' standing parents) 'first cousin'. A corresponding distinction is drawn between ego's offspring, kn?n 'offspring', and those of ego's sibling's km?n 'sibling's offspring', with a further distinction made for one's first cousin's child k<nan>mon bdiri? gade? (sibling's child <NMZ> standing parent) and one's sibling's child k<nan>mon mham (sibling's child <NMZ> blood).

A kinship term or *pajilan* 'calling name' can replace the use of personal names, the usage of which is restricted by various taboos, and is often preferred over the use of a pronoun. Qualification is often required to avoid ambiguity, achieved by the use of an attributive, usually a locative phrase as in (17), or a teknonym like ?ma? ?ame 'Amelia's mother' or a nick-name like monen 'morning'.

- (17) pa?jah c? bnal
2nd.born AT:below [name]
Second-born uncle down at *bnal* island. (= currently residing at *bnal*)

Standard diminutives are used to address children; frequently the use of these names is extended beyond childhood. The address terms d?y and m?y are derived from krdror 'female human' and rlmol 'male human' respectively. The terms dara?, from the Malay *dara* 'maiden', and trunp?, from the Malay *teruna* 'unmarried youth', are used to address prepubescent females and uncircumcised males respectively. The forms are listed below. They share a common gloss, 'girl' and 'boy' respectively:

FEMALE	MALE
krdror	rlmol
d?y	m?y
ya?	nana
la?	
dara?	trunp?

Affinal relations The reference term for one's spouse is km?n 'wife', or kn?k 'husband'. Other affinal reference terms do not differentiate gender. They are: mntuh? 'spouse's parent' from Malay *mer tua*, mntuh? soji dawon (parent-in-law end leaf) 'spouse's parent's sibling'; ?ipar 'sibling's spouse' from Malay *ipar*; biras 'spouse's sibling's spouse' from Malay *biras*, and kn?npupi 'one's offspring's spouse' and kn?n tir? 'stepchild' from Malay 'step relation'.

An in-law address term taboo prohibits the use of personal names or second person pronoun forms kb '2f' or ji '2', by ego to a parent-in-law, using instead either the birth order name or the third person singular pronoun k?h '3'.

The degree of relationship is ascertained by the elders prior to marriage. Marriage should be between members of the same generation; to marry someone of an ascendant generation is considered a form of incest, as is marriage to a sibling's offspring, parent-in-law or an adopted child. Marriage between known relatives or neighbours is desirable, cross-cousin marriage being encouraged, although there are

no distinct cross-cousin terms. Parallel first and second cousin marriage is classified as incest, however a small feast and payment of a fine negate the risk of retribution.

The selection of spouse is up to the individual, although families may intervene. Marriage takes place for girls any time after reaching puberty, however, there is no stigma attached to later marriage. It is left to personal choice. Males must be circumcised before being permitted to marry. The Semelai are the only Orang Asli who practise circumcision. The ceremony takes place at a feast and the actual circumcising is observed by the community, both male and female.

Following marriage, residence is not necessarily amongst kin, although immediately after marriage it tends to be matrilocal. Divorce is common and does not attract any social stigma; the elders are merely consulted and a fine paid. Remarriage to an ex-spouse's sibling is classified as incest. Polygyny is allowed although not widely practised. Polyandry is unknown.

Technology The Semelai derive their food from hunting wildlife and fishing in the lake. Hunting, which is a male occupation, was traditionally performed with a dog and spear or blowpipe. The blowpipe of the Semelai was constructed from wood rather than bamboo (as elsewhere), which is not available in the area. Various devices such as snares, basket traps, spear traps, birdlime and so forth are used. All of these are constructed from materials available within the jungle. Previously, many Semelai had old guns obtained during the communist emergency,¹⁹ which were loaned to those who did not have them. Most of these guns have now been reclaimed, but some members of the community still possess them. When meat from a hunt is abundant some of it is smoked and dried, and the remainder shared amongst the people. Firstly, if a gun was used on loan, the owner receives a portion. The rest is then distributed amongst members of the village, related or otherwise, at the hunter's discretion. Favoured meats are deer and pig, but monkey, chevrotin, squirrel, civet, monitor lizard, tortoise, riverine turtle, water snake and various types of bird are also eaten. Frogs and snakes are not generally consumed. There is a total taboo on the hunting of tigers, elephants and crocodiles.

Line fishing is carried out by both sexes, while trap fishing is predominantly a male activity, although I have encountered some females adept at the construction of traps. The vast number of fish types present in the lake streams and swamps is equalled by the different technologies employed in their exploitation: traps, pots, lines with various types of lure, spears, ichthyocide, and these days, town-bought nets. The waterways also yield riverine prawns, aquatic snails and a variety of molluscs.

Attempts by the *Jabatan Hal Ehwal Orang Asli* (Department of Orang Asli Affairs) to encourage animal husbandry have been unsuccessful. The Semelai keep chickens but do not collect the eggs or slaughter the animals for personal consumption; the explanation offered is that they become attached to the animals they raise.

Swidden cultivation is less common now than in the past. This is due mainly to the shortage of suitable jungle and a reluctance to leave the settlements for long periods on account of rubber trees that require tapping, but also a growing dependence on store-bought commodities, and the importance of children attending school.

¹⁹ Guns were formerly obtained through membership of the Government organisation R.E.L.A., a community-based internal security organisation (Hood 1990: 143).

Swidden cultivation is labour intensive for the often small return gained. Nevertheless, it is still fairly common for people to decamp and head downstream to make a swidden despite the less than perfect circumstances. This is perhaps motivated in part by sentimentality and an affection for their disappearing lifestyle. Swiddening begins as a communal effort between kin or friends. The felling of trees and slashing of undergrowth starts April-July, and then burning of the site follows in August or September. This coincides with the dry season in the middle of the year. Then the following crops are planted: a number of varieties of hill-rice including glutinous types, manioc, gourds, cucumber, corn, bananas, sweet potatoes, eggplant and tobacco. Rice planting takes place a month after the swidden is fired. The harvest takes place the following January and is carried out by the women, while the men stay at home and care for the children. Produce is kept by the family who planted the swidden; however, women assisting in the harvest are entitled to a third. Following the harvest one's neighbours are invited to attend a feast at which some of the new rice is consumed with game. This marks the celebration of a new year. Rice is considered a luxury more than a staple, possibly due to the small harvests, and people complain bitterly about eating cassava when the rice runs out.

It is worth noting that the Orang Asli have many varieties of rice, some unknown to the Malays (Skeat and Blagden vol. I 1906: 340, Evans 1920: 18). The rituals which surround rice growing, for instance the guarding of the rice soul prior to planting and harvest, are shared by the Malays. It is difficult to determine the origin of this rite. The arrival of the Orang Asli population, or their predecessors, in the Peninsula predates the advent of rice growing. However the indigenous varieties and names of rice among the Semelai suggest that they must have participated in its arrival, rather than acquired it from the Malays.

The Semelai have long exploited jungle produce which they traded to outsiders. Excursions were typically carried out by groups of nuclear families following the rice harvest, the whole family taking part. The produce was traditionally traded downstream to Kuala Bera or Temerloh on the Pahang River. They would either ship the goods downstream or middlemen would come upstream to the lake. In earlier times the Sultan of Pahang would send emissaries to the area to exchange goods for forest produce. The Semelai traded in dammar, *gaharu*²⁰ and *sapar* woods, rattan and bamboo, in return for betel nut, salt, crockery, coconuts, cloth and metal goods, such as knives, spear heads, fishing hooks and cooking pots. As recently as the 1960s Malays from Kuala Bera still travelled upstream to purchase the dammar torches manufactured by the Semelai.

Prior to the building of the road into Pos Iskandar the Semelai tended to focus downstream to Kuala Bera and Temerloh in Pahang. However, following its construction, they now look to Ladang Geddes, an outpost in the rubber estates in Negeri Sembilan. Before the road was constructed, Ladang Geddes could only be reached by travelling down the Bera River and then upstream on the Serling River.

Trade in forest produce continues today though to a degree diminishing in accordance with the decrease in jungle area and the increase in the use of synthetic compounds. Resettlement has removed people from the jungle, and the collection of rattan and dammar, still practised in the early 1990s, has now virtually ceased. But

²⁰ A scented wood (*Aquilaria spp.*) used as incense. *Sapar* is another type of this wood.

seasonal produce such as the fruit of the *petai* (*Parkia biglandulosa*, *P. Roxburghii*), *kerdas* and *jering* (*Pithecolobium lobatum*) are still collected for sale.

Many of the implements utilised in the various aspects of daily life are manufactured by the Semelai from materials readily available in the swamp and surrounding jungle. Pandanus and lepiironia are woven into sleeping and eating mats. Baskets used for collecting firewood, harvesting rice and so forth and pouches for holding tobacco and betel nut, are woven from pandanus leaves. Dippers used to collect well-water are made from the spathe of a species of palm. From rattan and bamboo they manufacture implements for fishing and hunting, such as traps, creels and the like.

About 40 years ago, as part of a development programme, the government planted the land around the kampongs with rubber trees, designating a small-holding of five acres to each family. This programme has further developed with major replanting of the trees in the early 1990s, and now comprises the major source of income for the people resident in the Task Bera area. Tapping is carried out by men and women, and the money earned is kept by the individual. This provides a source of cash income to the Semelai, enabling them to purchase food, clothing and other commodities like electricity and water. In the late 1990s two generators were installed, providing electricity from dusk until dawn for the settlements around Pos Iskandar, and in 2001 these settlements were connected to a piped water supply.

1.3.4 Current linguistic situation

In 1999, the population figure given in the government census was 4,103. This figure would be representative of the number of Semelai speakers. Semelai of all ages exhibit a preference for speaking their own language, and the use of Malay is limited to dealings with non-Semelai. Malay, the national language and the language used in education, is the lingua franca used with all other peoples, including those from other aboriginal minorities in the Peninsula.

It would appear that the rapid loss of Southern Aslian languages as feared by Benjamin (1976a: 89) is not in evidence as far as the Semelai of Tasek Bera are concerned. Benjamin (in press) remarks on the resilience of Semelai following a visit to the area which accords with my own recent visits to the community: young children were still monolingual, and I was unaware of any young Semelai speaking only Malay to each other.

An examination of Evans' word list (Evans 1915: 108–13), collected in early 1914 and containing 240 basic items, reveals that only two words are different today, almost 90 years on. They are the kin terms *apet* 'father' and *ga?u?* 'elder female sibling'. Both lexemes are still known by elders, but have been replaced by the Malay terms as mentioned above.

This is not to say that language loss is not occurring, and will not occur in the future. A lack of knowledge of many lexical terms within certain specific genres, for example technology and flora and fauna, is exhibited by many young people, especially those schooled away from home.

It is inevitable that the use of Semelai will decline as the Orang Asli take part in the wider Malaysian society. Education is in the Malay language; access to radio and television brings the Malay language and culture into the home.

Dialect variation I noted little variation in the Semelai language amongst the communities I visited, although I did not have the opportunity to make a systematic study of dialectal variation due to research permit requirements, which restricted my research to Pos Iskandar. The Semelai themselves maintain that there is little variation between speakers at Tasek Bera and Bukit Rok compared to speakers from the Serting River area.

1.4 The Orang Asli

Orang Asli is a modern Malay language term adopted post-independence to replace the colonial English ethnic term 'aborigine'.²¹ *Orang Asli* applies only to the indigenous minorities of the Malay Peninsula in contrast to the Malay, Chinese, Indian and Eurasian populations (see §1.3). It does not apply to the indigenous peoples of the Malaysian states of Sabah and Sarawak on the island of Borneo.

Whilst Aslian is a linguistic term, *Orang Asli* is an ethnic classification. Certain groups of *Orang Asli* are speakers of Austronesian, and not Austroasiatic languages: the Jakun, Temuan, Orang Seletar and Orang Kanaq speak their own dialects of Malay, whilst the *Orang Kuala* speak an Austronesian language (Kähler 1946–9).

According to current figures shown in Table 1.5, from the website of the *Jabatan Hal Ehwal Orang Asli*, the Government Department responsible for *Orang Asli* affairs, there is a total of 92,529 *Orang Asli* in Malaysia. The *Orang Asli* represent less than 0.5 per cent of the population. Of this figure, 56,515 are speakers of Aslian languages. The remainder of the population are speakers of Malay and related Austronesian dialects. The reason for including them as *Orang Asli* is discussed below. The Government census is organised into 'racial' groupings: Semang, Senoi and Melayu Asli (§1.4.1). The names of the Austronesian groups are in italics.

TABLE 1.5 ORANG ASLI POPULATION FIGURES (MAY 1999)

SEMANG					
Kintaq Bong	Kensiul	Jahai	Mendriq	Batek	Lanoh
235	224	1049	145	960	359
SENOI					
Temiar	Semai	Jah Hut	Cheq Wong	Semoq Beni	Besisi
15,122	26,049	3,193	403	2,488	2,185
MELAYU ASLI ²²					
Semelai	Temuan	Jakun	Orang Kanaq	Orang Seletar	
4,103	16,020	16,637	64	801	

Population figures are not given for the Temoq, the Semelai's closest neighbours, although the 1969 Census recorded a population figure of 100 (Asmah 1979: 8). Figures for the Temoq are presumably included with either the Semelai or Jakun. The Mintil, Batek Nong, Semnam and Sabum are also not distinguished in the Government's ethnic-based classification. Population figures for the Mintil and Batek

²¹ Carey (1975) does not attribute the origin of this term to the presence of communist insurgents as erroneously claimed by Parkin (1991: 41).

²² With the exception of Semelai, the remaining Melayu Asli are Austronesian speaking peoples who inhabit the southern portion of the Peninsula in the states of Pahang, Selangor, Negeri Sembilan, Melacca and Johore.

Nong are included with the Batek Deq, and Semnam and Sabum are included with the Lanoh figures.

Exact population figures for the Kensiu in Southern Thailand are unavailable, although an estimate of around 400 is reported in Benjamin (in press).

For a comprehensive bibliography of the Orang Asli see Lye (2001).

1.4.1 Semang, Senoi and Melayu Asli

Just as the Aslian languages are recognised as falling into three groups, there are three different 'racial' groupings recognised amongst the Orang Asli. These are not concordant with the three linguistic groupings, as will be shown in the final paragraph of this section. The three 'racial' groupings, based on supposedly 'typical' physiognomic characteristics, are Negrito or Semang, Senoi, and Melayu Asli. The Semang are located predominantly in the northern states, the Senoi in the central Malaysian highlands, and the Orang Melayu Asli in the south. Counter-intuitively the phenotypic differences do not reflect biological diversity (Fix 1995: 316, and §1.2.6), but are best viewed as indicators of societal tradition and linguistic groupings.

The Semang are speakers of the Northern group of Aslian languages. Traditionally they are classified as foragers. In earlier literature they were referred to as the 'Semang' and 'Pangan' (Skeat and Blagden 1906: 21). The term Semang has been adopted as the term to refer to those Negrito who follow a particular societal tradition (Benjamin 1985a).

The term 'Senoi' is taken from the Temiar *senɔɔ:y* 'people' or 'mankind'. As seen in the population table, they are the most numerous of the Orang Asli, and are speakers of languages from the Central group. They are generally described as being of a southern Mongoloid type. Traditionally the Senoi were involved primarily with swidden cultivation. The Senoi were referred to in earlier literature by the Malay term *Sakei* or *Sakai* 'slave, dependent'. This term was confusing as it was often extended to mean any Orang Asli, not just the Senoi. It is still common to hear Malaysians refer to the Orang Asli as Sakai, despite the fact that it is discouraged due to pejorative connotations. In Thailand this term is used to refer to the Aslian population there, Duangchand (1984).

The third group in the Peninsula is the Orang Melayu Asli as they are now known. With the exception of Semelai and Temoq, the remaining groups within this category are not Aslian. The term *Melayu Asli* is translated as 'Original Malays', or as 'Proto Malays' in Malaysian literature. The current term is based not on ethnic origin, but implies people who follow Malay (Melayu) type social organisation but do not practise Islam. In earlier literature this group was known as Jakun (Williams-Hunt 1952: 15), Orang Benua, and Mantra, amongst many other names.

Highly diverse modes of existence are exhibited by this group. Some practise swidden cultivation, the collection and trading of forest products, fishing (as in the case of the Semelai and Besisi, the former located on a lake and the latter on the coast), or farm small holdings in mainstream society. In complete contrast, the Temoq are hunters and gatherers.

The following discrepancies are to be noted between these 'racial' divisions and the languages: the Lanoh and Semnam are classified as Negrito but speak Central languages; the Cheq Wong have the physical characteristics of the Senoi, but speak a Northern language; and the Semoq Beri and the Besisi have the appearance of the

Senoi, but are members of the Southern linguistic group. In addition to this, prolonged contact and inter-marriage highlights the limitations of a 'racially-based' classification.

2 Phonology and phonotactics

Semelai has thirty-two consonant phonemes and twenty vowel phonemes. Consonants are discussed in §2.1.1 and vowels in §2.1.2. §2.2 provides a description of the syllable and the processes of syllabification. The structure of the syllable is central to the organisation of the phonology, and in turn an understanding of this structure provides the groundwork for the description of the structure of morphological processes in §3. The main discussion of segmental structure is contained in §2.3. The segmental and distributional possibilities of Semelai consonants illustrate that language-specific sequence structure constraints are not only necessary, but in fact reflect other morphological outcomes and the nature of the language's phonotactic system. §2.4 provides some comments on Malay loanwords and patterns of borrowing into Semelai. The chapter concludes with remarks on the comparative phonology of Aslian languages in §2.5.

2.1 Inventory of phonemes

2.1.1 Consonants

The five places of articulation are labial, alveolar, palatal, velar and glottal. Oral stops are produced at all five places of articulation. The three manners of articulation of oral stops are voiced, voiceless and voiceless aspirated. There is a corresponding series of nasal stops at each place of articulation of the oral stops.

Two places of articulation are distinguished for fricative segments (palatal and glottal); two for glides (labio-velar and palatal); and one for the rhotic and the lateral (both alveolar).

The sonorants, with the exception of the labio-velar glide, have a contrasting series with a glottal onset. Three voiceless nasals at the bilabial, alveolar and velar places of articulation are also attested, although they are highly infrequent. The reasons for treating the members of these two series as unitary phonemes are given in §2.3.1.1.

For easier transcription /j/, /s/ and /y/ are used instead of the IPA symbols /j/, /s/, and /ʃ/ respectively, and apart from this chapter, the voiceless nasals /m̥/, /n̥/ and /ŋ̥/ are represented by the following: /hm/, /hn/ and /ŋn/.

TABLE 2.1 CONSONANT PHONEMES

	BILABIAL	ALVEOLAR	PALATAL	VELAR	GLOTTAL			
STOPS aspirated	p p ^h	b t ^h	t d	c c ^h	j	k k ^h	g	?
NASALS +glottal	m m̥	n n̥	n n̥		m m̥	ŋ ŋ̥	ŋ ŋ̥	
FRICATIVES				s				
LATERALS +glottal			l l̥					
RHOTICS +glottal			r r̥					
GLIDES +glottal				y y̥		w		

The following lists contrast the consonant phonemes in initial and final positions.

	INITIAL POSITION		FINAL POSITION	
Stops:	/p/	/pɔy/ 'to make'	/la:p/ 'far'	
	/p ^h /	/p ^h ɔy/ 'to fan'		
	/b/	/bɔy/ 'refrain'		
	/t/	/tɔh/ 'to unfasten'		
	/t ^h /	/t ^h ɔh/ 'to spit'		/?et/ 'to extinguish'
	/d/	/dɔ/ 'of'		
	/tʃ/	/tʃɔ/ 'dog'		/?əc/ 'faeces'
	/t ^h ʃ/	/t ^h ʃɔ/ 'hill'		
	/y/	/jɔ/ 'foot'		
	/k/	/kɔp/ [name]		/bek/ 'to fasten'
Nasals:	/k ^h /	/k ^h ɔp/ 'to lie face down'		
	/g/	/gɔp/ 'Malay'		
	/ʔ/	/ʔɔɔ/ 'to howl'		/kəɔ:/ 'torso'
	/m/	/mɔ/ 'want'		/kəm/ 'to bury'
	/m̥/	/m̥ɔ/ 'what'		
	/n/	/nɔ/ 'before'		/lən/ 'to desire'
	/n̥/	/n̥ɔ/ 'thereupon'		
	/p/	/pɛp/ 'to winnow'		/?əp/ 'if'
	/p ^h /	/p ^h ɛp/ 'tiger's voice'		/ləp/ 'TO:up'
	/m/	/mɔt/ 'to mount'		
	/m̥/	/m̥ɔt/ 'eye'		
	/n/	/nɔt/ 'this'		
	/n̥/	/n̥ɔt/ 'this'		
	/ŋ/	/ŋɔt/ 'to be ugly'		
	/ŋ̥/	/ŋ̥ɔt/ 'to sense'		

¹ I am yet to find a suitable contrasting pair.

	INITIAL POSITION		FINAL POSITION		
Fricatives:	/h/	/ghɔp/	'to be hot'	/rəh/	'to probe'
	/s/	/sɔp/	'morning'	/θpəs/	'to be high'
Laterals and rhotics:	/l/	/lbh/	'to move transversely'	/jel/	'to bark'
	/r/	/rɔh/	'tree branch'	/pər/	'to fly'
	/ɻ/	/ɻuc/	'to pass'		
	/ɻ/	/ɻec/	'to masturbate'		
	/ɻt/	/ɻt̪ey/	'to be many'		
	/ɻt/	/ɻey/	'companion'		
Glides:	/w/	/wɔh/	'to wake'	/spew/	'plant sp.'
	/y/	/yon/	'tortoise sp.'	/t̪rey/	'to be many'
	/jy/	/jyən/	'to hear'		
	/y/	/yun/	(meaning unknown)		

2.1.1.1 The phonetic realisation of consonant phonemes

Oral stops At the phonetic level the series of stops can be identified as having six places of articulation. The series described above with a phonemically alveolar place of articulation in Table 2.1, are articulated at two distinct places: the voiceless and voiceless aspirated stops are apico-dental [t] and [t^h], the voiced stop [d] and nasal stop [n] are apico-alveolar. There is no evidence to justify this distinction at the phonemic level.

The voiceless stops /p, t, c, k/ are unreleased [p̚, t̚, c̚, k̚] in syllable final position. In word final position, as opposed to syllable final, the stop may have a delayed release: /kɔ.bəc/ [kɔ:t̚.bəc] 'to be fishing with a pole'.

In syllable initial position the voiceless lamino-palatal stop /c/, and the voiced /ɟ/, are slightly affricated [c^f] and [ɟ^f]: /cim/ [c^ft̚m] 'bird', /ɟalu/ [ɟ^falu:] 'pig'.

/ʔ/ is the voiceless glottal stop [ʔ]. Syllable finally, it is a fortis closure, /hmpɛʔ/ [həm.pɛʔ] 'three'. In syllable initial position it is less fortis, /t̚ɔŋ/ 'to howl'.

Nasal stops The nasal stops have the same place of articulation as their voiced oral counterparts. In syllable initial position, and as the coda in non-final syllables, /m, n, ɲ/ are realised as [m, n, ɲ]: /mūn/ [mū:h] 'nose', /knłek/ [kənłak] 'husband'.

Prestopped nasals Word final nasals which follow an oral vowel are prestopped by a denasalised homorganic voiced stop, i.e. nasality is delayed until after closure, e.g. /kimpən/ 'wife' [kmpa:n]. The articulation of the oral release onset takes place in the following manner. As the articulators adopt the position for the nasal stop, the delayed opening of the velum results in a voiced stop being articulated prior to the onset of the articulation of the nasal: /m/ is realised as [m̚], /gum/ [gu:m] 'to winnow rice vertically'; /n/ as [n̚]; /ron/ [rɔ:n] 'to straddle'; /ɲ/ as [ɲ̚] /tap/ [ta:p] 'to weave' and /ŋ/ as [ŋ̚] /jŋan/ [jɔ:ŋa:n] 'bone'.

The articulation of the nasal has prominence over the stop which is often barely audible. Following a nasal vowel, the distinction is neutralised and the final consonant is realised as a simple nasal stop: /hɔn/ [hɔ:n] 'where', /lmɔŋ/ [ləmɔ:ŋ] 'tooth'.

Prestopped nasals are considered a phonetic unit and there is no evidence to justify them at the phonemic level. Justification of this analysis is based on phonological and morphological evidence: a) consonant clusters are not permitted in

syllable final position; b) morphologically, nasal consonants reduplicated as non-final syllable codas are realised as simple nasals, not prestopped nasals, e.g. /tap/ [ta:p̚] 'to weave' → /tip-tap/ [tip-ta:p̚] 'to be weaving'. There is evidence that historically prestopped nasals arise from simple nasal consonants. Note that Malay loanwords which end in a final nasal consonant usually realise the coda as a prestopped nasal, /than/ [taha:n̚] 'hole'.

The term 'prestopped nasal' is standard phonetic terminology, an inversion of the term 'prenasalised stop' (Catford 1988: 114, Ladefoged and Maddieson 1996: 128-9). It is adopted here in preference to some of the terms which have been used elsewhere for this phenomenon in Aslian, e.g. 'decomposed nasals' (Matisoff 1990: 546), 'predenasalised stop' (Benjamin 1985b: 14-16, Diffloth 1976c: 111), and 'preplodled nasals' (Blust 1997).

Prestopped nasals are not to be confused with postnasalised stops, the nasal release associated with final stops, e.g. in the Aslian language Temiar: /tɔ:p/ [tɔ:ɔ:p̚] 'to hug', /tɔ:b/ [tɔ:ɔ:b̚] 'tiger' (Benjamin 1976b: 133-4).

Voiceless nasals The voiceless nasals, /ŋ/, /ɳ/ and /ɸ/ have a voiceless onset followed by some voicing in the transition to the following vowel: [ɸ ɳ ɳ̚] 'sound of the tiger's voice'.

Fricatives /s/ is approximately the voiceless lamino-prepalatal fricative [s]: [sət̚] 'laryngitis', [lps] 'tongue'. This is not to be confused with the English voiceless palatal fricative [ʃ]. The Semelai fricative has a much clearer quality, more like alveolar [s]; in fact it is very much like this sound, but is pre-palatal,² being produced at the edge of the palatal region, (as in diagram (c) Figure 33, Catford 1988: 96). It is a difficult segment to describe, with no distinct IPA symbol. In word final position, and particularly amongst older female speakers, /s/ is articulated with greater friction, producing a 'muddier' sound.

The voiced fricative [z] occurs as a marginal phoneme in initial position. It is only attested in Islamic personal names of Arabic origin, which are themselves loans from Malay, e.g. the female name [zaynən] 'Zainon'.

The glottal fricative [h] ranges from barely audible aspiration to strong articulation syllable and word-finally: [m̚d̚] or [m̚d̚h] 'nose', [cah.n̚t̚] 'to be indolent'.

Lateral /l/ is the clear apico-alveolar lateral continuant [l] in all positions: /lol/ [lɔ:l] 'a disease of the feet', /blɛŋ/ [bə.lɛŋ] 'arm', /bilhüt/ [bəl.hüt] 'black, dark, (av.)'.

Rhotic /r/ is the voiced apico-alveolar trill [r]: /rəc/ [rΔc] 'small bird sp.'. In final position it may be pronounced as a tap [ɾ]: /kr.dor/ [kər.dor ~ kər.ɾor] 'woman', /pər/ [pəɾ] 'to fly'.

/r/ has the phonetic realisation [ɻ] when preceded by the alveolar nasal /n/. The voiced homorganic stop transition is conditioned by a constraint disallowing sequences of homorganic sonorants: /gabn̚reŋ/ 'place name' [gabən̚r̚eŋ]. The morphological process of the second example is discussed in §3.2.1.3: /creh/ 'fish' → /c<n>r̚eh/, [c<n>r̚eŋ] 'a type of fish'.

² The term 'pre-palatal' is chosen over 'post-alveolar' in order to distinguish the palatal characteristics of this phoneme.

Glides Glides are always treated phonemically as consonantal. Sequences of glide plus vowel and vice versa are consonant vowel or vowel consonant sequences and not diphthongs.

/w/ is the labio-velar glide [w],³ corresponding to the high back rounded vowel [u]. In final position, the duration of [w] is reduced to an offglide: /?w.?aw/ [?uw.?aʷ] 'to be standing up', /khew/ [keheʷ] 'Crested Wood Partridge'.

/y/ is the lamino-palatal glide [y] corresponding to the high unrounded front vowel [i]. In final position it is an offglide: /bɔy/ [bɔʷ] 'NEG:IMP'.

Following a nasal vowel, the glides are nasalised: /tyāw/ [tyāʷ] 'to look down', /twāy/ [twāʷ] 'reaping-knife'.

Glottalised sonorants Glottalised sonorants are restricted to onset position. The onset of the sonorant is a barely audible glottal release represented phonetically as a superscript [ʔ]: /ʔəp/ [ʔəp] 'to be far'.

2.1.2 Vowels

Semelai has a rich vowel system, with twenty vowel phonemes in all: ten oral vowels, each with a nasal counterpart. Length is not phonemically contrastive. All vowel segments are monophthongs. Sequences of vowel plus glide are not considered diphthongs as a treatment of them as unit phonemes would conflict with the syllabic structure of the language (§2.2.1).

The greatest number of height distinctions, namely four, is made for the back vowels. The back vowels are: high /u/, mid-high /o/, mid-low /ɔ/, and low /ɑ/. The front and central vowels exhibit three height distinctions, front: high /i/, mid-high /e/, mid-low /ɛ/; central: high /a/, mid-central /ə/ and the low vowel /ɑ/.

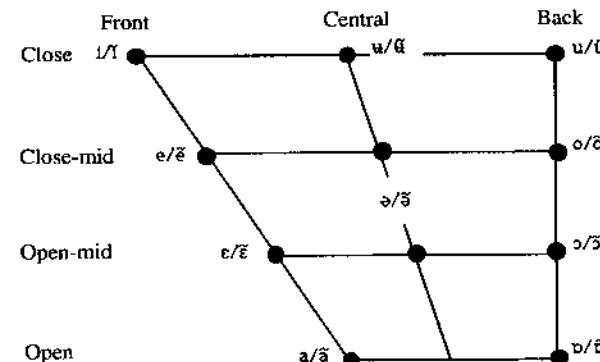
Front vowels /i/, /e/, /ɛ/, and the mid-central and low vowels /ə/ and /ɑ/ are unrounded, while all the back vowels /u/, /o/, /ɔ/, /ɑ/ and the high central vowel /a/ are rounded.

As stated above, all ten vowels have a phonemically nasal counterpart: /i/, /ɛ/, /ɛ/, /ɑ/, /ɔ/, /ɔ/, /ɑ/, /ɑ/, /ɑ/ and /ɑ/.

Figure 2.1 illustrates the oral and nasal vowel segments. These representations are not indicative of actual vowel spacing: the high and mid-high vowels, both front and back, are closer than the low and mid-low vowels. The mid-vowels are perceptually very close and difficult to distinguish. Phonetic realisation of the segments is given in §2.1.2.1.

³ The semi-vowel /w/ is a labio-velar approximant: a coordinate coarticulated sound, requiring not only lip-rounding, but also raising of the back of the tongue. In terms of articulatory setting it is the same as the vowel [u], the only difference being the duration of articulation: the semi-vowel is short whereas the vowel is sustained (Catford 1988: 83-4). Given the fact that the feature [+back] is integral to both /w/ and /u/ plus the distributional similarities between the two segments, it is important to note this. Further, in terms of the glottalised sonorants, the absence of a glottalised segment here is attributed to the fact that a velar sonorant does not have a glottalised counterpart *?ŋ.

FIGURE 2.1 PHONEMIC ORAL AND NASAL VOWELS



Contrasting positions of the oral vowels in near minimal pairs are given below:

FRONT	CENTRAL	BACK
/dɪŋ/ 'bamboo'	/rʌŋ/ 'bed bug'	/ruŋ/ 'termite'
/dɛŋ/ 'manner'		/coŋ/ 'trap pit'
/nɛŋ/ 'earlier'	/zəŋ/ 'if'	/jŋ/ 'foot'
	/caŋ/ 'pouch'	/caŋ/ 'mousedeer'

Minimal pairs differentiated by nasality alone are infrequent. Near minimal pairs illustrating the contrast are provided below:

/pit/ 'to shut one's eyes'	/ggaúk/ 'vine sp.'
/kit/ 'to be very small, tiny'	/sksɛk/ 'to cry until breathless'
/sleč/ 'bloodless'	/tut/ 'to blow'
/cěč/ 'to spray'	/bihlt/ 'dark, black (av.)'
/kɛt/ 'to sting'	/boh/ 'company'
/kɛt/ 'to be small'	/bkrɔbɔč/ 'the sound of a large animal in water'
/taŋ/ 'to weave'	/cɔč/ 'to gut (large fish)'
/?aŋ/ 'to carry (on the back)'	/cɔč/ [name]
/ləŋ/ 'to want'	/cʰɔr/ 'get urinated on'
/bkarən/ 'tousled'	/jɔr/ 'to urinate'

2.1.2.1 The phonetic realisation of vowel phonemes

As stated earlier, there is no phonemic length distinction in Semelai, but phonetically, vowels appear to be longer when they occur in open final syllables: /cɔč/ [cɔč] 'to gut (large fish)' compared to /cɔč/ [cɔč] 'dog', and when followed

by the glottal fricative /h/ [hu:^b] 'nose'. In contrast, when followed by a glottal stop [ʔ], they are perceptually shorter: see ceʔ [cɛq^ʔ] 'to disappear'.

All vowels are pronounced slightly higher when followed by the velar consonants /k/ and /tʃ/: /pk/ [pɒk] 'to chop (at)', /metʃ/ [mɛtʃ] 'cheek', and both higher and shorter when followed by the glottal stop /ʔ/: /ceʔ/ [cɛq^ʔ] 'to disappear'.

All vowels are equivalent or close to the cardinal vowels:

- /i/ is close to the high vowel [i] in open syllables and slightly lower [ɪ] in closed syllables: /ci/ [cɪ:] 'louse'; /cim/ [cɪ'm] 'bird';
- /e/ is the mid-high vowel [e]: /ple/ [pəle:] 'fruit', /khew/ [ke.hew] 'Crested Wood Partridge';
- /ɛ/ is the mid-low vowel [ɛ]: /ʔate/ [ʔa.tɛ:] 'earth', /ŋep/ [ŋɛp] 'to winnow cassava vertically'.

The back vowels have the acoustic quality of rounded vowels, although they are produced with minimal lip-rounding:

- /u/ is the high back vowel close to cardinal [u]: /blu/ [bə.lu:] 'thigh, lap', /ʔus/ [ʔus] 'fire';
- /o/ is the mid-high vowel [o]: /dol/ [dol] 'knife handle';
- /ɔ/ is the mid-low vowel [ɔ]: /ɔ:/ [cɔ:] 'dog', /dɔ/ [dɔl] 'house';
- /ɒ/ is a low retracted, almost pharyngeal vowel [ɒ]: /kp/ [kp] '2f', /wɒy/ [wɒy] 'knife'.

The set of central vowels contains both rounded and unrounded phonemes:

- /ə/ is the high central rounded vowel [ə]: /kum/ [ku^bm] 'to hold in the fist', /bəm/ [bəm] 'mountain';
- /ə/ is the non-high unrounded central vowel [ə]: /ptəm/ [pə'təm] 'to plant seeds'. In the stressed final syllable it is slightly back and lowered, approaching the articulatory setting of the unrounded back vowel [ʌ]: /jəl/ [jʌl] 'to bark', /bktək/ [bək'tək] 'undergrowth, weed';
- /ʌ/ is the low central unrounded vowel [ʌ]: /la/ [la:] 'A; BCS', /lat/ [lat:] 'dam barrier'.

Non-phonemic nasality All vowels may become non-significantly nasalised in certain environments. Nasality is always unidirectional, and is characterised as 'onset driven', spreading from left to right. There are two environments which condition nasalisation.

a) When an onset is a nasal stop, the following vowel is nasalised, neutralising the contrast between oral and nasal phonemes:

/mot/	[mɔt̪:]	'eye'
/kanɛʔ/	[kanɛʔ:]	'rodent, lesser gymnu're'
/majər/	[mä:jər:]	'Black Panther'
/napam/	[nä:päm]	[toponym]

Given that nasality is onset driven, a nasal coda, or a nasal onset of the following syllable cannot cause the nasalisation of the preceding vowel. In the following near minimal pair, the vowel in the first is phonemically nasal, the second is oral: /hɔn/ [hɔn] 'where', /hɔŋ/ [hɔŋ] 'hornet'.

b) Glottal phonemes are transparent to nasal spread. When the penultimate syllable onset is a nasal stop and the onset of the final syllable is a glottal segment, /ʔ/ or /h/, the nasality of the onset spreads to the vowels of both syllables: /mhəm/ [mähəm] 'blood', /mɛʔes/ [mɛʔəs] 'to sweat'.

Evidence for the unidirectionality of nasal spread is also found in examples like the following where nasality, located in the final syllable, has no influence on the vowel of the penult: /sɪʔlt/ [sɪ.'ʔlt] 'to smell rotten'.

2.2 The syllable

The phonotactics of Semelai are best described in terms of the syllable. In order to do this, a description of the syllable and syllabic structure will precede the discussion of the distribution of phonemes in §2.3. This description of syllable structure and the processes of syllabification is influenced by Clements and Keyser (1983) and Blevins (1995).

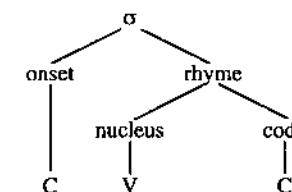
2.2.1 Syllable structure

The syllable template for Semelai is [CVC]_σ, consisting of a simple onset, nucleus and coda. Three constraints define the possible representations of the syllable in Semelai:

- a) Onsets are obligatory, hence there is a constraint against vowels in onset position: *_σ[V].
- b) Syllables may have a coda C_σ giving rise to the maximal syllable [CVC]_σ. It is universally acknowledged that if a language has C-final syllables, it also has V-final syllables (Clements and Keyser 1983), hence the minimal syllable in Semelai is [CV]_σ.
- c) Sequences of more than one consonant in the same syllable are inadmissible: *_σ[CC, or *CC]_σ, thus sequences of consonants are always heterosyllabic (§2.3) and never tautosyllabic.

The syllable is diagrammatically represented as a binary branching construct in Figure 2.2.

FIGURE 2.2 THE INTERNAL STRUCTURE OF THE SYLLABLE



This analysis is chosen over a flat structure which has no subconstituents, as there is evidence of the rhyme, (R), forming a constituent of which the nucleus (Nu) and the coda (Co) are daughter sub-constituents. The onset and rhyme are regarded as independent constituents, which for the most part combine freely, capturing the generalisation that onsets rarely influence the nucleus. Restrictions on the

combination of the nucleus and the coda are more frequently attested than restrictions on the onset and coda, although some are encountered. Evidence for positing a rhyme is also relevant at the morphological level where the coda conditions the vowel of the penult in reduplicated syllables (§2.3).

2.2.2 Stress

Stress is a syllable-level phenomenon. In Semelai the domain of word stress is the final syllable and there is no secondary stress. Only phonological words bear stress. In the case of words bearing suffixes, stress shifts from the root to the suffix: *pdr* 'to follow' when suffixed with the morpheme *-i?*, *p.'dr + -i?* is resyllabified and stress is reassigned to the final syllable giving *p.dr.i?* 'to follow (s.th.)' (§3.2.1.3).

2.2.3 Syllabification

It is generally held that syllable structure is not defined in underlying representations (UR) (Blevins 1995: 221). The evidence for this in Semelai is: a) there are no minimal pairs based on syllacticity, and b) individual morphemes, in the case of affixes and clitics, fail to conform to possible syllable types, e.g. the vowel initial suffix */-i?*.

Syllabification proceeds by attempting to map the minimal syllable [CVC]σ onto the UR, driven by the principles of syllable formation as set out above in §2.2.1. Note that there is no requirement that syllable structure be maximalised. Given that the right edge is the most stable portion of the word, and morphological processes are mapped from the right, syllabification takes place from right to left.

The first step is to locate the final vowel and associate it with a syllable node. The onset, which is obligatory, is then established, producing a minimal syllable [CV]σ in accordance with the requirement that minimum syllable wellformedness be met. Any unassociated consonant to the right of the vowel is then syllabified as a coda. The next vowel to the left is then located, and if there is one, the syllabification process is repeated. This process continues until the UR is exhaustively syllabified as illustrated in Derivation 1 and Figure 2.3 below.

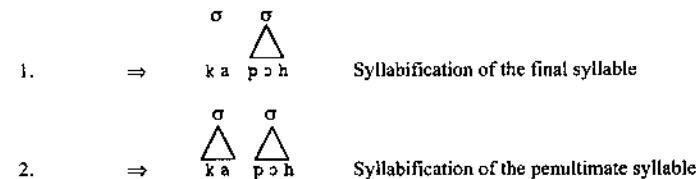
Derivation 1:

//kapoh// 'egg'

1. ka.poh Syllabification of final syllable
2. ka.poh Syllabification of penultimate syllable – with a phonemic vowel

FIGURE 2.3 DERIVATION 1

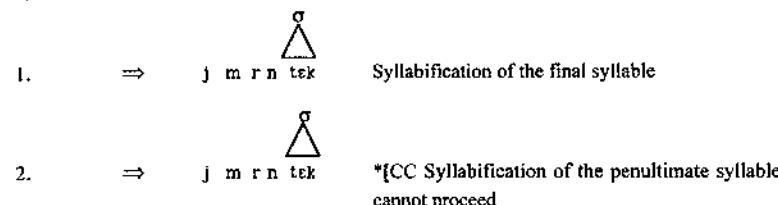
//kapoh// 'egg'



A problem is encountered with URs which contain sequences of unsyllabified consonants at the left syllable edge of the final or penultimate syllable, in violation of the minimal syllable [CVC]σ, e.g. //rmpa?// 'vine sp.' or //jmrntek// 'first basket of the rice harvest', where syllabification is arrested due to the string of consonants, see Figure 2.4.

FIGURE 2.4 DERIVATION 2

//jmrntek// 'first basket of the rice harvest'



If vowel epenthesis was invoked in the underlying form //rVmV.na?// at this stage of the syllabification process, syllabification would proceed with the correct output: [rə.mə.na?]. However, if vowel epenthesis was invoked in the underlying form //jmrntek// to derive //jVmVnVtek//, the incorrect output would result *[jə.mə.rə.nə.tek]. The correct syllabification would be two open syllables followed by a closed syllable: [CV]σ[CV]σ[CVC]σ[CVC]σ [jə.mə.rə.n.tek].

This indicates that vowel epenthesis can only occur after the consonant C₃, in a string of consonants [CCC₃.CVC], is identified as either a coda or an onset. If C₃ is designated as a coda, then vowel epenthesis will occur to the left of it, and the next consonant to the left will be deemed the obligatory onset [CC.CC₃.CVC]. If C₃ is not identified as a coda, it will be determined as the onset and the epenthetic vowel will occur to the right of it [CCC₃.CVC] in accordance with the requirement that minimal syllable wellformedness must be met. This was the case above with [rə.mə.na?] (although this form has not yet been justified).

The observation drawn from this data is that, in an unsyllabified sequence, some consonants have the potential to be syllabified as onsets and some as codas. It will be shown that the problem in predicting correct syllabification arises due to language specific constraints that determine which segments are licit codas in unsyllabified strings of prefinal consonants. By examining the segmental and distributional possibilities of Semelai consonants, it will further be shown that language-specific sequence structure constraints are not only necessary, but in fact reflect other morphological outcomes and the nature of the language's phonotactic system.

2.2.3.1 Codas and coda constraints in non-final syllables

In Semelai, the inventory of coda phonemes of final syllables, as set out in §2.1 and §2.3.2, is smaller than the inventory of onset phonemes. In non-final syllables potential codas are further constrained by the fact that they are in contact with adjacent segments. Only two types of coda are possible in non-final syllables, and these are further restricted by the position of the syllable within the word, penultimate or prepenultimate. Further to this, only one heterosyllabic cluster is licensed per word, and the preferred position is at the boundary of the final syllable.

The two types of possible coda in non-final syllables are: i) a consonant from the following subset of sonorants /m, n, ɳ, r, l/,⁴ and ii) a pre-established coda, i.e. a consonant which is identical to the coda of the final syllable.⁵ These are examined in detail below. Type i), the sonorant coda, may occur in any non-final syllable; type ii), the pre-established coda, is only licensed as a penultimate syllable coda.

i. Sonorant Codas Given that there are no vowel initial syllables, the non-final coda is always adjacent to the onset of the following syllable. These are examined below where it becomes evident that in certain cases the coda is linked to the following segment in terms of homorganicity; others are not homorganic. Generalisations on distribution can be explained in terms of the sonority of the adjacent segments.

The most common sequence of consonants across the syllable boundary is a sequence of homorganic segments: a) nasal + obstruent; b) nasal + liquid, or c) liquid + oral stop. The most frequently occurring is the homorganic nasal + stop pattern; in prepenultimate syllables this is the only possibility:

a) nasal + homorganic stop:

- /km.pən/ 'wife'
- /s.m.rŋ.kɛk/ 'the sound of people working together'
- /lm.bl.han/ 'refuse area under house'
- /pɳ.g.ra.wan/ 'a tree sp.'

In Semelai there is an absence of homorganic palatal sequences. Palatal stops always take the non-homorganic alveolar nasal /n/ in this position, never the palatal nasal /ɲ/: /kn.sɔp/ 'to throb' and /nɪn.ca/ 'food'.

b) nasal + liquid:

- /kn.lɛk/ 'husband'
- /cn.ɾɔŋ/ 'elephants' bathing place'

c) liquid + oral stop:

- /pl.to?/ 'explosion'
- /kr.te?/ 'personality'

In summary, the previous sets of data suggest that any sonorant /m, n, ɳ, r, l/ can occur as a penultimate syllable coda provided that it agrees in place of articulation with the onset of the final syllable. However, the following examples containing non-homorganic sequences are in violation of the claim made here:

d) liquid + obstruent:

- /cr.pek/ 'pointed, oval shaped'
- /kr.coʔ/ 'a sedge'

⁴ It should be noted that there is a phonological-morphological correlation between the set of nasal and liquid phonemes, and the set of infixing morphemes in Semelai: nominalising (<m>, <n>, <ɳ>), and causativising <r>. Hence, these infixes are identifiable as potential codas on both phonological and morphological grounds.

⁵ A similar pattern is found in the distribution of codas in penultimate syllables ('minor syllables') in Kammu, where permissible codas are ʂ, or a sonorant /m, n, ɳ, l, r/ (other than the palatal nasal /ɲ/), or a consonant identical to the coda of the final syllable (Svantesson 1983: 29).

/gr.hɛŋ/ 'to yell at (s.o.)'

/kl.cɛt/ (meaning unknown) from kray kicɛt 'glutinous rice sp.'

A sequence of sonorant – nasal, lateral or rhotic – plus an obstruent, must be syllabified as heterosyllabic, with the sonorant syllabified as the coda of the syllable to the left, and the obstruent as the onset of the preceding syllable on the right. The important point here is that the sonorant will be determined as a coda, not as an onset, because of its relationship with the following consonant, although the onset status of the following consonant would have been previously determined, for instance where it is the onset of the final syllable. Closer inspection of the syllabification outputs in the language reveals the following syllable contact constraints.

Syllable Contact Constraints:

In a sequence C₂C₁C₁V, C₂ is syllabified as syllable coda C₂VC₁C₁V, when:

- a) C₂ > C₁ in sonority; or
- b) C₂ and C₁ are homorganic.

OTHERWISE: C₂ is an onset.

CONDITION: Only one heterosyllabic cluster per word is allowed in the grammar.

The condition arises from word structure wellformedness requirements. Note that the preferred position for heterosyllabic clusters is at the boundary of the final syllable, as in //j.m.rn.ts̩k//, although they are also attested in prepenultimate position as well, e.g. /pr.j.g.ra.wan/ 'a tree sp.'

The syllable contact constraints predict the correct syllabification of forms such as //rmpaʔ// 'vine sp.' and //jmrntek// 'first basket of rice harvest' without the need to first invoke vowel epenthesis. The three derivations which follow illustrate the syllabification process. The second is a revised version of derivation 2.

Derivation 3:

//rmpaʔ// 'vine sp.'

- | | |
|--------------|---|
| 1. rm.pa? | Syllabification of the final syllable |
| 2. rmV.pa? | C ₂ = C ₁ in sonority. C ₂ is an onset |
| 3. rV.mV.pa? | Obligatory onset is required *σJV |

Derivation 4:

//jmrntek// 'first basket of the rice harvest'

- | | |
|-------------------|--|
| 1. jmr.n.tek | Syllabification of final syllable |
| 2. jmr.Vn.tek | C ₂ and C ₁ are homorganic. C ₂ is a coda |
| 3. jV.mV.r.Vn.tek | Condition: one heterosyllabic cluster per word. All remaining Cs are syllabified as onsets |

Another example is the form //gmrcɛŋ// 'flickering like flames' which is syllabified as [gə.mɛr.çɛŋ] not as *[gə.ma.ɾɛ.çɛŋ].

Derivation 5:

//gmrc^{bəŋ}// 'flickering like flames'

1. gmr.c^{bəŋ} Syllabification of final syllable
2. gV.mVr.c^{bəŋ} C₂ > C₁ in sonority. C₂ is a coda

It is interesting to note in relation to the set of consonantal infixes that they are all sonorous segments, and that prefixes like the causative have allomorphs ending in either an open syllable, /r/, or a conditioned nasal segment /N-/ (§3.2.2.3). The exception is the affix +raʔ+ 'COMP' (§3.2.2.1).

ii. Pre-established Codas The following constraint is restricted to the penultimate syllable, and reflects the distribution of a second set of segments – pre-established codas. The presence of these exceptional codas is driven by morphological processes: in the penultimate syllable any phoneme can be a coda provided that it is a copy of the coda of the final syllable. In effect these are segments which are licensed as coda segments in the final syllable, a privilege which allows them to be codas in the penultimate syllable. This morphologically conditioned form is associated with the reduplicative process referred to as coda copy, a morphological process described in §3.2.2.2. The examples below are synchronically monomorphemic, but the forms are identifiable as being complex, derived by morphological processes similar to those described in §3.2.2.2 and §3.2.4.1.

/t?r̩.r̩?/	[ta?r̩.r̩?]	'uncivilised person'
/b.ln.lan/	[bələniu?n]	'appearance of many lice crawling through hair'
/saŋ.lɛŋ/	[saŋle?ŋ]	'to be thirsting, craving (for s.th.)'
/k.pe?r̩.r̩?/	[kepe?r̩.r̩?]	'to call out in the forest to keep track of hunting partner'

A further parameter can be added to the syllable contact constraints described above. In a sequence //CC₃CVC₁// as in //kly?yu2// 'tree sp.', if C₃ is identical to C₁, C₃ is a coda and the output is [ke.lə.ya?yu2].

Derivation 6:

//kly?yu2// 'tree sp.'

1. kly?yu2 Syllabification of the final syllable
2. klyV₂.yu2 C₃ = C₁. C₃ is a coda
3. kV₁V₂.yV₂.yu2 Condition: one heterosyllabic cluster per word. All remaining Cs are syllabified as onsets

An UR such as //kcboŋ// 'an insect sp.' will automatically be syllabified as [ke.ca.boŋ] given the lack of identity between the successive codas.

Derivation 7:

//kcboŋ// 'insect sp.'

1. kc.boŋ Syllabification of the final syllable
2. kV₁cV₂.boŋ C₃ ≠ C₁ so C₃ is not a coda. C₃ is not homorganic with C₂. C₃ is an onset

Having determined the nature of possible syllable contacts and positing the linked coda constraints, it is possible to correctly predict non-final codas, which in turn paves the way for the realisation of epenthetic vowels.

2.2.3.2 Epenthesis

The epenthesis of vowels is best analysed as a response to violations of skeletal structure. It is used to build acceptable syllables from degenerate or unsyllabified sequences of consonants, which exist in the underlying representation, in accordance with the minimality condition of syllable wellformedness. In particular, epenthesis results from the need to meet wellformedness conditions and to build the syllable.

In order for the site of the epenthetic vowel to be correctly predicted, constraints on the phonotactic and prosodic possibilities of the language are taken into account (see §2.2.1 and §2.2.3). Realisation rules are not considered as separate rules, but as an intrinsic component of a general syllabification process that holds in the language. The domain of these rules is therefore the same as the domain of syllabification: at the underlying level of word formation.

Sequences of consonants requiring epenthesis only occur at the left edge of a word, where both codas and onsets can be syllabified within the sequence, demonstrating that epenthesis is independent of the direction of syllabification.

Epenthetic vowels are defined as those which are entirely predictable and therefore are of phonetic, not phonemic status, whereas phonemic vowels are not predictable, and therefore must always be specified. The default vowel is [ə], however, in Semelai all oral vowels have equal access and equity to either status – phonemic or phonetic (§2.3.3). A number of rules ensure the correct realisation of the epenthetic vowels. These are stated in terms of the adjacent right-hand segment: the non-final open syllable vowel is determined by the onset of the adjacent syllable, and the vowel of the closed syllable is determined by the coda of the rhyme.

i. The realisation of epenthetic vowels There are four possible realisations of the vowel as stated below.

a) The default epenthetic vowel is realised as the mid central vowel [ə]:

Ø → [ə] / _ . elsewhere

/d.r̩/	[dər̩]	'rattan'
/j.tek/	[jətek̩]	'to sleep'
/g.da.do/	[gədədo]	'huge, grand'
/b.s.jet/	[bəsəjet̩]	'to be dark'
/b.c.pon/	[bəcəpon]	[toponym]
/k.l.d ² .do ² /	[keləda?do ²]	'termite, in winged stage'

b) When the onset of the final syllable is a palatal sonorant, either the glide /y/ or the nasal /ɲ/, the epenthetic vowel of the preceding syllable is realised as [i]:

[ə] → [i] / _ . sonorant [+palatal]

/t.yek/	[tiyek̩]	'banana'
/sŋ.k.yon/	[səŋkiyo?ŋ]	'bamboo flute of windmill'
/s.jek/	[sijek̩]	'to be sweet'

The remaining non-sonorant palatal phonemes – the stops /ʃ/ and /ç/ and the fricative /s/ – do not determine the vowel in this environment and it remains as the 'elsewhere' vowel schwa:

/k.jah/ [kəjah] 'to be heavy'

/k.c.bɔŋ/	[kəcəbɔŋ]	'an insect sp.'
/r.səʔ/	[rəsəʔ]	'a fish sp.'

c) When the onset of the preceding syllable is the labio-velar glide /w/ the vowel of the preceding syllable is realised as [u]:

[ə] → [u] / _ . w

/j.wər/	[juwər]	'to sew'
/k.r.wan.ceŋ/	[keruwancəŋ]	'coral snake'

d) In words with an open penultimate syllable where the onset of the final syllable is a glottal segment /ʔ/ or /h/, the epenthetic vowel of the penultimate syllable will be the same as that of the final syllable (§2.3.3.1). Recall that nasality is not transported leftward across the syllable boundary (§2.1.2.1): /s.ʔit/ [sɪʔɪt] 'to smell of decay'. The rule is stated as:

[CV₂]σ '([h, ?]V₁C)σ, where V₂[+/-nasal] = V₁[+/-nasal]

Some examples are listed below:

/d.ʔoh/	[doʔoh]	'swidden site'
/b.ʔen/	[beʔən]	'NOT'
/p.hem/	[pehəm]	'to keep quiet about s.th.'
/g.hop/	[gohop]	'to be hot'
/ʔ.ʔeʔ/	[teʔeʔ]	'EB'

This rule never applies to prepenultimate syllables: /ʔ.ə.reʔ/ [ʔəʔareʔ] 'lower abdomen' and not *[ʔaʔareʔ].

ii. *The realisation of epenthetic vowels of morphophonemic origin* The realisation of these vowels results from the reduplication or copying of the final coda into the penultimate coda position. These rules hold whether the form is archaic, as in the case of the pre-established codas discussed in §2.2.3.1, or the product of a synchronic morphological process (§3.2.1.3).

There are four possible realisations of the vowel, forming a close parallel with the realisation of vowels of phonemic origin. The only differences are that in this group all members of the palatal series, both sonorants and obstruents, may determine the vowel, and the vowel and the determining phoneme always belong to the same syllable, except where the root has an open final syllable in which case the onset of the final syllable creates the conditioning environment: n-čo [ni.čo] 'kind of dog'. The realisation rules are as follows:

a) The epenthetic vowel is realised as [ə]:

ø → [ə] / _ C[-palatal, -glottal, -labio-velar] +

/rep/	'to thresh'	/rp-rep/	[rap̩.rəp̩]	'to be threshing'
/t̩em/	'to pound'	/tm-t̩em/	[təm.t̩əm]	'to be pounding'
/b.s.ŋet/	'to be dark'	/bs<ŋ>et/	[ba.sət̩.ŋet̩]	'darkness'
/c.man/	'to retell a myth'	/c<>n.man/	[cen.mən]	'to be retelling a myth'
/r.mol/	'to be male'	/r<>mol/	[rəl.məl]	'man'
/k.dor/	'to be female'	/k<>dor/	[ker.dər]	'woman'

/bek/	'to tie'	/bk-bek/	[bək̩.bək̩]	'to be tying'
/gɔŋ/	'to carry'	/gŋ-gɔŋ/	[gəŋ.goŋ]	'to be carrying'

b) Where the coda of the final syllable is a palatal segment, and is reduplicated as the coda of the penultimate syllable, the vowel of the penultimate syllable is realised as [i]:

[ə] → [i] / _ C[palatal] +

/hɔy/	'to yawn'	/hy-hɔy/	[hiy.hɔy]	'to be yawning'
/s.map/	'to ask for (s.th.)'	/s<>n.map/	[sɪn.map]	'to be asking'
/tac/	'to faint'	/tc-tac/	[tic̩-tac̩]	'to be fainting'
/c.les/	'to slice'	/c<>s.les/	[cis.les]	'to be slicing'

c) Where the coda of the final syllable is the labio-velar glide /w/, and is reduplicated as the coda of the penultimate syllable, the vowel of the penultimate syllable is realised as [u]:

[ə] → [u] / _ C[labio-velar] +

/ʔaw/	'to stand'	/ʔw-ʔaw/	[ʔuaw.ʔaw]	'to be standing up'
/c.ŋew/	'to look down'	/c<>w.ŋew/	[cuw.ŋew]	'to be looking down'

d) Where the coda of the final syllable is a glottal segment /ʔ/ or /h/, and is reduplicated as the coda of the penultimate syllable, the vowel of the penultimate syllable is realised as [a]:

[ə] → [a] / _ C[glottal] +

/rɔʔ/	'basket'	/nʔ-rɔʔ/	[naʔ.rɔʔ]	'basketful'
/p.doʔ/	'to sprout'	/p<>doʔ/	[paʔ.doʔ]	'to be sprouting'
/ʔɔh/	'to blowpipe'	/ʔh-ʔɔh/	[2ah.2ɔh]	'to be blowpiping'
/c.roh/	'to meet'	/c<>h.roh/	[cah.roh]	'to be meeting'

In many instances the penultimate syllable coda is presumed to be an archaic affix in what have become synchronically monomorphemic roots. On the basis of predictability they are treated in the same manner as the synchronically analysable forms. The processes are discussed in §3.2.4. In the following examples the postulated non-lexemic root (*v) is given in the last column:

/s.rp.tep/	'a plant'	[sərep̩.t̩ep̩]	*vstep
/sm.sm/	'bone marrow'	[sams̩.sm̩]	*vsm
/gt.gut/	'bird sp.'	[gət̩.gut̩]	*vgut
/kn.kn/	'child, not ego's'	[kənk̩.kn̩]	*vkən
/s.lc.goc/	'channel'	[salic̩.goc̩]	*vsgoc
/ty.toy/	'bird sp.'	[tiyt̩.tɔy]	*vtoy
/k.njŋ.puŋ/	'child-in-law'	[kaŋlŋpuŋ]	*vkpuŋ
/sk.sk/	'plaque'	[sək̩.sək̩]	*vsk
/cŋ.cŋ/	'elbow'	[cəŋ.cuŋ]	*vcŋ
/ts.ts/	'fire ant'	[tis.ts]	*vts
/p2.re2/	'ghost'	[paʔ.re2]	*vre2
/kh.kuh/	'termite mound'	[kahkuh]	*vkuh

The scope of these rules is expanded upon in §3.2.1.3.

2.3 The distribution of phonemes

The distribution of phonemes is directly related to the position they occupy in the syllable, and in turn the position of the syllable within the morpheme. The inventory of onset consonants includes all consonants in the language. The inventory of coda consonants is restricted. This is shown in Table 2.2, where the coda phonemes are shown as the enclosed set.

TABLE 2.2 ONSET AND CODA PHONEMES

p	t	c	k	?
p ^h	t ^h	c ^h	k ^h	
b	d	j	g	
m	n	ɲ	ŋ	
hm	hn		hŋ	
		s		h
		l		
		r		
		y	w	
?m	?n	?ɲ		
?l				
?r				
?	y			

Word structure in Semelai is $(\sigma_R)^\eta(\sigma)^\sigma$ where $\eta \leq 2$. There is a direct correlation between the distribution of phonemes and the syllable type in terms of its position: ultimate σ , penultimate σ , or prepenultimate σ_R (§2.2.1). As syllables are concatenated from right to left, the inventory of potential segments in the onset, coda and nucleus reduces. This is a common feature not only of Austronesian languages (Diffloth and Zide 1992, Benjamin 1985a), but of Mon-Khmer languages in general (see Svantesson 1983 for Kammu; Smith 1979 for Sedang).

2.3.1 Onset phonemes

Final syllable All thirty-two consonant phonemes can occur as the onset of the final syllable in monosyllabic words. There does not appear to be any distinction between open and closed syllables with regard to permissible onset consonants. The list of minimal pairs contrasting in onset position is found in §2.1.

In disyllabic and polysyllabic words, the voiceless aspirated stops never occur as the onset of the final syllable, with the exception of /t^h/ in /bt^həŋ/ 'to be afraid' and /pt^hac/ 'the sound of a string snapping as it breaks' and /k^h/ in the Malay loan /pek^her/ 'to think' and in /dik^hes/ ⁶ 'to be near'. The following consonants are also attested as occurring marginally, that is in less than five instances in this position: the palatals /c, j, ɲ, s/, and the voiced velar stop /g/.

Penultimate syllable Twenty-three phonemes are available as the onset of the penultimate syllable, however the following consonants all have highly limited distribution: the voiceless aspirated alveolar stop /t^h/, the palatal and velar nasals /ɲ/ and /ŋ/ and the glides /w/ and /y/. The glottalised sonorants and voiceless

nasals are never attested as syllable onsets in the penult. The labio-velar glide /w/, aspirated dental stop /t^h/ and velar nasal /ŋ/ each appear in only one word respectively: /wunwen/ 'to shit (av.)', /t^halɔc/ 'appearance or feeling of coming undone', /pren/ 'to be angry'. The series of voiced and voiceless stops and the fricative /s/ are attested most frequently as the onset in a penultimate syllable.

The following list of words exemplifies the onset of the penultimate syllable:

/p/	/p ^h ət/	'to stay'
/p ^h /	/p ^h əŋər/	'white hair'
/b/	/btɔc/	'to be lonesome'
/m/	/mrəŋ/	'to be itchy'
/t/	/t̪əl/	'to cut'
/t ^h /	/t ^h areŋ/	'canine tooth; fang'
/d/	/d̪əs/	'to be short'
/n/	/nr̪is/	'a plant sp.'
/c/	/cher/	'hunger pains'
/c ^h /	/c ^h ŋjir/	'the smell of ammonia'
/ʃ/	/ʃkɔs/	'Malayan porcupine'
/ɲ/	/ɲapah/	'blood (av.)'
/k/	/kbaʔ/	'torso'
/k ^h /	/k ^h bas/	'to die'
/g/	/gdo/	'to be elderly'
/ŋ/	/ŋren/	'to be angry'
/ʔ/	/ʔiʔət/	'EB' address term
/h/	/hayir/	'the smell of burning flesh'
/s/	/sdeč/	'to be cold'
/l/	/ipes/	'tongue'
/r/	/rbɔl/	'taro'
/y/	/yuyuy/	'gnat'
/w/	/wunwen/	'to shit (av.)'

Prepenultimate syllable In prepenultimate syllables σ_R , the inventory of onset phonemes is further reduced, permitting only voiceless and voiced stops, the fricative /s/ and the following sonorant segments: /m, n, ɲ, l, r/:

/s.l.new/	'the look of a rooster's sparsely feathered neck'
/r.s.tɔŋ/	'a condition associated with yaws'
/k,g,do/	'elephant (av.)'
/k,sn,du,duk/	' <i>Melastoma</i> sp. bush'
/j.m.rn.tek/	'the basket containing the first rice harvested from a swidden'

2.3.1.1 Complex onset phonemes

There are two series of marginal coarticulated segments – the glottalised sonorants /ʔm/, /ʔn/, /ʔl/, /ʔr/, /ʔp/, /ʔy/⁷ and the voiceless nasals /tŋ/, /hŋ/ and /h/.

These phonemes are only found in the onset of monosyllabic word forms, exhibiting a pattern of distribution like the series of voiceless aspirated stops.

⁶This form is possibly influenced by the Malay *dekat* 'close'.

⁷The coarticulation of a glottal stop with a velar nasal does not occur.

However, unlike the aspirated stops, they are highly infrequent and there is a dearth of minimal pairs. Some examples are:

/?mot/	'to mount'	/?nom/	'to be ripe'
/?pih/	'to be ill, in pain'	/?let/	'to extinguish, v.i.'
/?ris/	'to be alive'	/?ye/	'to see'

The voiceless nasals are even less frequent:

/m̥h/	'what'
/ŋe/	'shortly'
/ʃɔŋ/	'tiger's voice'

The question arises as to whether these complex segments are in fact a unitary phoneme, i.e. a single coarticulated segment, or if they constitute a cluster of two segments. As a unitary phoneme the glottalised sonorant is analysed as a sonorant consonant, the articulation of which is initiated by glottal pressure: [‘m, ‘n, ‘ŋ, ‘r, ‘l, ‘y]. Analysis as a unitary phoneme would liken the glottalised sonorants to aspirated voiceless stops, differentiated from the corresponding simple phoneme by manner of articulation.

Glottalised segments are generally rare in Asean languages. Some central dialects of Semai exhibit such stops, e.g. /?nar/ 'two' (Diffloth p.c.), but they are extremely rare even in these dialects. Similar phonemes are attested in other Mon-Khmer languages, e.g. the 'laryngealised glides' of Kammu, a Mon-Khmer language spoken in Laos (Svantesson 1983), and 'laryngealised' nasals and resonants in Sedang of Vietnam (Smith 1979). Kammu has a range of potential consonant clusters in the onset, but this sequence is not considered one. The following minimal pairs are provided by Svantesson as justification of these as unitary phonemes: /?wiat/ 'tired', /wiat/ 'to wilt' and /?yáaj/ 'rubber', /yáaj/ 'female'. Unfortunately it is unclear as to what criteria he has invoked as the basis of the distinction.

Minimal pairs, or near minimal pairs are once again difficult to locate in Semelai. Some minimal pairs are /?rey/ 'to be many', /rey/ 'companion', and /m̥h/ 'what', /m̥h/ 'to want' and the near minimal pairs: /?ris/ 'to be alive' /res/ 'root', /?let/ 'to extinguish' and /lat/ 'dam barrier'. Minimal or near-minimal pairs are yet to be established for /?n/, /?ŋ/ or /ŋ/. This does not constitute compelling evidence for positing a unitary phoneme as opposed to a cluster, however, the following arguments can be invoked in support of a unitary analysis:

- Distribution: The distribution is highly constrained. The sequence only occurs in syllable onset position in monosyllabic morphemes, never in the onset of any other syllable. In this respect, the glottal clusters are similar to aspirated voiceless consonants which also share a restriction to onset position, although the latter may occur as a penultimate onset in a few lexemes.
- The 'cluster' appears to be invisible to syllabification rules, there is no evidence of the presence of an epenthetic vowel, and the consonant sequence manifests as a single unit, the nature of the glottal segment lending itself to coarticulation with the following consonant. The word containing the cluster has only one sonority peak and the whole sequence receives stress, the

pattern attested for monosyllabic morphemes. Disyllabic morphemes have an unstressed initial syllable.

- Morphological evidence can be found in derivational processes (§3.2.1) where allomorphy of the affix is determined by the number of syllables in a root. Affixes are selected by calculating the number of syllables in the root: monosyllables select prefixes, disyllables select infixes. Roots containing glottalised sonorants are parsed as monosyllabic, selecting the prefixes associated with monosyllabic morphemes. In effect morphological processes fail to see the complex initial segment. For example, the causative is formed by a prefix *par-*, or *tar-* if monosyllabic, or a coda infix <r> if disyllabic (§3.2.3). Roots with a complex onset, pattern with monosyllabic word types and select a prefix. Note also the loss of the feature [+glottal] in the derivative (§3.2.1.3): ?*yot* 'to go back' → *tar-yot* 'to take sth. back', not *?<r>*yot*.

The failure of the process 'to see' the glottal feature provides further justification for a unitary analysis. The loss of this feature allows a comparison between these phonemes and the series of voiceless aspirated stops: when aspirated stops are reduplicated, the reduplication process fails to copy the feature [+aspiration]: *pʰoy* 'to fan' → *py-pʰoy* 'to be fanning' (§3.2.1.3).

2.3.2 Coda phonemes

Final syllable The inventory of consonants occurring in the coda of final syllables, σ, is smaller than that of onsets. The reduction is due to oral stops having only one manner of articulation in coda position, voiceless, as opposed to the three available in syllable onset position: voiceless, voiceless aspirated and voiced. The inventory of phonemes occupying coda position is given in §2.3, Table 2.2. The phonemes contrasted in near minimal pairs for the syllable coda are exemplified in §2.1.

There are also some restrictions on the co-occurrence of the glides, labio-velar /w/ and palatal /y/, and the preceding vowel in the rhyme. The labio-velar glide /w/ does not occur after the high front or high central vowel or any of the back vowels, ruling out the following sequences: */i:w, ɯ:w, uw, o:w, ɔ:w, ɒ:w/. Permissible sequences are exemplified below:

/ew/	/slew/	'to be slippery'
/ew/	/cnew/	'to look down upon (s.th.)'
/aw/	/traw/	'Long-tailed macaque'
/əw/	/pitəw/	'appearance of bulging eyes'

The palatal glide /y/ has a wider distribution than the labio-velar glide, occurring after all vowels with the exception of the high front vowel /i/, hence */iy/, and the mid-low vowel /ɛ/, hence */ey/.

/ey/	/sey/	'to be thin'
/ay/	/pay/	'to set aside part of a meal'
/əy/	/rey/	'friend'
/uy/	/kluy/	'millipede'

/uy/	/muy/	'one'
/oy/	/cloy/	'monkeys and lesser apes'
/ɔy/	/roy/	'fly'
/øy/	/cløy/	'to die (av.)'

Phonemically nasal vowels closed by a glide are less frequent. The following sequences are attested:

/əw/	kldəw	'out of alignment'
/əw/	brsəw	'to wipe off slime'
/əw/	mäw/	'be bland (fruit)'
/ɔw/	/tɔw/	'call of the Malay peacock-pheasant'
/ɔy/	/?iɔy/	'millet'
/øy/	/kanduwäy/	'eight (counting rhyme)'

Possible penultimate and prepenultimate syllable codas and related constraints were introduced in §2.2.3.1. They are summarised briefly below.

Penultimate syllable In simple monomorphemic disyllabic words the penult is either open with an empty coda, or in the case of a heavy penult, the coda can only be filled by a sonorant phoneme: /m, n, ŋ, l, r/. The palatal nasal /ɲ/ does not occur as a coda segment in this context:

/ka.ʈip/	'centipede'	/kal.tɔŋ/	'knee'
/bi.χəʔ/	'to ignore'	/kr.βɔʔ/	'crocodile'
/s.ɪeh/	'to be hungry'	/km.øŋ/	'wife'
/l.pcs/	'tongue'	/kn.ʈek/	'husband'
/ta.wɔ/	'White-handed gibbon'	/ʈɔŋ.kəʔ/	'Sambur deer (av.)'

In the case of the penultimate syllable, it is only in derived 'bimorphemic' words that phonemes licensed to occur in the final syllable coda can fill the coda slot of the penultimate syllable with an identical segment. This is the result of coda copy, as either, a) a synchronic process, or b) a fossilised form: a) /ʃl.jəl/ 'to be barking' ← /jəl/ 'to bark', /st.ɳɔt/ 'to be whimpering' ← /s.ɳɔt/ 'to whimper'; b) /ʃl.jiʔ/ 'to be dirty' ← *ʃl2 and /p2.reʔ/ 'ghost' ← *p2.re2.

Prepenultimate syllable The only potential codas in prepenultimate syllables are sequences of the following nasal segments: /m, n, ŋ/, followed by a homorganic stop: /kɳ.kɔ.ʈh/ 'nightjar', /sɳ.k.yoŋ/ 'the bamboo flute of the windmill'.

2.3.3 The nucleus

The distribution of vowel phonemes in the nucleus is described in relation to the position of the syllable in the word – final 'σ', penultimate σ, or prepenultimate σ_R⁸ – and whether the syllable is closed or open. The final and penultimate syllables are described together in §2.3.3.1, and the prepenultimate syllable in §2.3.3.2. In general, there are no restrictions on the distribution of vowels in the closed stressed syllable 'σ', where all vowels, both oral and nasal may occur. The possibilities

⁸ Prepenultimate refers to any syllable to the left of the penultimate syllable. The term antepenultimate is employed to refer specifically to the syllable immediately to the left of the penult.

become more restricted in the syllables concatenated to the left, reflecting the same generalisation made for the distribution of the consonants in §§2.3.1-2.

TABLE 2.3 DISTRIBUTION OF VOWEL PHONEMES

SYLLABLE	σ _R	σ	'σ
CLOSED		v v̄	v v̄
OPEN	(a, ɔ, ə)	v̄ v̄	i, u e, o ɛ, ɔ a, ɒ ə

Distribution of vowels in open final syllables is restricted, the inventory of oral vowels is smaller, and nasal vowels do not occur at all. /ə/ is only attested in bound morphemes. Phonemically nasal vowels are rare in syllables other than the closed final syllable 'σ, /k̩ət/ 'to be small'. Exceptions occur in the class of expressive words, /krā.'sāk/ 'the sound of fish biting, taking bait', plant names, /y᷑ŋk̩ɔs/ 'tree sp.', and personal names. Furthermore, phonemically nasal vowels do not occur in the rhyme of syllables where the syllable onset is a voiced palatal stop /j/, or the voiced velar stop /g/.⁹ Non-phonemic nasality is discussed in §2.1.2.1.

Syllables which contain a phonetic vowel are restricted to the left edge of the word, and never occur to the right of a syllable containing a phonemic vowel.

2.3.3.1 The nucleus in the final and penultimate syllables

Final syllable As stated above, any vowel can occur in the closed final syllable with the exception that the phonemically nasal vowels do not occur following the voiced palatal or velar stops, /j/ or /g/.

All oral vowels except the central vowels /u/¹⁰ and /ə/ occur in open final syllables in monosyllabic roots; /ə/ is attested in bound morphemes as in me= 'REL'.

/i/	/t̩i/	'hand'	/u/	/blu/	'thigh'
/e/	/t̩ye/	'to see'	/o/	/po/	'dream'
/ɛ/	/t̩yɛ/	'I'	/ɔ/	/co/	'dog'
/a/	/ca/	'to eat'	/ɑ/	/kɑ/	'2F'

In the final open syllable of disyllabic words the inventory of possible vowels is further reduced by the non-occurrence of the low back vowel /ɑ/, in addition to the two central vowels /u/ and /ə/. Disyllabic words with an open final syllable are always monomorphemic. The following are attested:

/i/	/ba.ʈi/	'rhizome'
/e/	/p̩a.re/	'monitor lizard'

⁹ A similar restriction applies in Jahai (Burenhult 2002), Jah-Hut (Diffloth 1976c: 103) and Kensiu (Duangchand 1984), where nasal vowels do not occur with an initial voiced velar stop.

¹⁰ /u/ is the least frequent vowel phoneme in Semelai. Diffloth reports that in Jah-Hut /u/ is also the least frequent vowel phoneme and amongst certain speakers is not distinguished from /ə/. /jəl/ 'to bark' (Diffloth 1976c: 103). In northeastern Semai it is found only in a small number of words, and in other dialects of Semai it merges with /ə/ (Diffloth 1977: 479).

/ɛ/	/tə.te/	'earth'
/a/	/ba.ba/	'unhusked rice'
/u/	/ja.lu/	'wild boar'
/ɔ/	/tɪ.jɔ/	'snake'

In trisyllabic words, only the mid-high vowel /o/, mid-low vowels /ɛ/ and /ɔ/ and the low vowel /a/ are attested in the final syllable, e.g. /kgdo/ 'elephant, (av.)', /gadad/ 'huge', /kmmbe/ 'five'(counting rhyme).

Penultimate syllable Any oral vowel may fill the nucleus of the penultimate syllable. The vowel with the highest frequency of occurrence in this context is /a/. Some examples of vowels in the penult are illustrated here:

/ɪ/	/tɪ.ʃy/	'millet'
/e/	/ləw,new/	'be unable to finish'
/ɛ/	/ce.çɛ/	'to hate'
/a/	/ka.pən/	'egg'
/u/	/dəh.yuh/	'to be shivering cold'
/ə/	/tə:en/	'TO:down'
/ɪ/	/tɪ:pɪŋ/	'smouldering log'
/o/	/c.nog.kɪŋ/	'hip'
/ɔ/	/tɔ:sok/	'placenta'
/ɒ/	/tɔ:n.lɒŋ/	'catwalk'

In many instances the vowel of the penult is the same as that of the final syllable. Motivation for this pattern of apparent vowel harmony can be explained only in the following case, as all other examples are synchronically unmotivated.

As explained previously in §2.2.3.2, in words with an open penultimate syllable, where the onset of the final syllable is a glottal segment /ʔ/ or /h/, the epenthetic vowel of the penultimate syllable will be the same as that of the final syllable, e.g. /c.ɔŋ/ [cɔŋ] 'to roast'.

This data suggests that the vowel in the former examples must be specified underlyingly, whilst those in the second set may be left unspecified.

In closed penultimate syllables the epenthetic vowel may be realised as [a, i, or u] determined by a reduplicated coda phoneme, either glottal /ʔ/, h/, a palatal segment /c, j, s, y/ or the labio-velar glide /w/, respectively. Elsewhere it is realised as [ə] (§2.2.3.2).

As stated previously, nasal vowels are only found in the penultimate syllable of words belonging to the class of expressives: /rɔ:p.rɔ:p/ 's.th. large walking over twigs', /rū:būp/ 'unkempt facial hair', plant names and in personal names /ga.tɔ:zɛ/ 'a woman's name'.

2.3.3.2 The nucleus in the prepenultimate syllable

In prepenultimate syllables σ_R, the only vowel generally available is phonetic [ə] or [u] as in a) below. The rare exceptions are b): the names of some animals, female names which may contain an initial syllable /ga/, and expressives.

a)	/b.k.pol/	[bəkəpol]	'cigarette smoke'
	/k.lk.bpk/	[kələkbpk]	'butterfly'
	/k.r.wan.cen/	[karuwancen]	'coral snake sp.'

b)	/ga.je.nah/	[woman's name]
	/ga.wen.wēt/	'moving in and out of view'

Note that in the last example in a) the vowel adjacent to the onset /w/ is realised as [u], but in the last example in b) under the same conditions, the vowel is phonemic [ə], not [ə].

2.3.3.3 A note on the vowel schwa

It could be suggested that the occurrence of schwa outside of the main syllable is almost entirely predictable, and therefore it is a non-phoneme. This is the stance traditionally adopted in the area of Mon-Khmer linguistics where it is argued that the presence of phonetic schwa [ə] is due only to the absence of any other vowel. However, with the exception of the Katuic branch which has disyllables, Mon-Khmer languages are monosyllabic, or at best 'sesquisyllabic'.¹¹ Aslian, as shown here, and also discussed in Diffloth (1976a), has true disyllabic forms containing non-predictable vowels including schwa in the non-final syllable.

2.4 Malay loanword phonology

Loanwords, which are extensive in Semelai, usually come either from, or via, Malay. There are also a few examples from English, from the period of the Communist Emergency when British troops were installed in the area (§2.4.5). These are noted at the end of the section. Generally there is little change to the form of the borrowed word, and it is treated as a Semelai word, available to the same morphological processes.

It is difficult to ascertain whether words which vary from the Standard Malay are Semelai innovations or simply a reflection of the form of Malay from which it was borrowed, either a regional variant, or an archaic form. It may be that the source of Malay borrowings in Semelai is the Aboriginal Malay dialects which surround the Semelai area, in particular the Jakun. The Jakun are one of the Orang Asli groups who speak varieties of Malay (§1.4). There is little data available on these dialects, although see Collins (1985) for Jakun spoken at Kampong Landai, Pahang; Blagden (1906: 433); Evans (1915) for Sakai-Jakun aboriginal dialects spoken in Ulu Langat and Ulu Kenaboi; and Noone (1939: 152-56) on the Benua Jakun language of Johore.

The situation is further compounded by the fact that certain Malay dialects preserve archaic forms. I have been unable to find any extensive information on this issue, and what has been located has been inconclusive, but some of these variations can be recognised from dialects of other regions.

Rather than attribute all derivations from Malay to Semelai adaptations, I take the stance that the form of the borrowing is largely reflective of the form as it was borrowed. For this reason, I use '=' here to indicate 'equivalence' rather than '←' indicating 'derived from'.

¹¹ A sesquisyllable is a syllable between a monosyllable and a disyllable, i.e. one and a half syllables. The half syllable consists of a consonant but no vowel, or a consonant and a reduced vowel (Matisoff 1990; Svartesson 1983).

Stress Placement The penultimate stress of Malay is replaced by the final syllable stress of Semelai: the Malay form *'ku.cing* 'cat' in Semelai becomes */ku.'cɪŋ/*.

Syllable Structure The following adjustments to syllable structure are noted:

- a) In Malay vowel initial words are acceptable, in Semelai this is corrected by the epenthesis of a glottal stop or glottal fricative. Although the forms may have been borrowed with the glottal onset present, there is also the distinct possibility that this may be a Semelai innovation in order to offset the inadmissibility of $\sigma[V:]^{12}$

/ʔubɒt/ 'medicine' = *ubat*
/ʔadiʔ/ 'YS' = *adik*

The presence of the /h/ is most likely a retention in the regional variety of Malay from which the forms were borrowed:

/hayam/ 'chicken' = *ayam*
/hubiʔ/ 'yam' = *ubi*
/habuʔ/ 'hearth' = *abu(k)*
/hatap/ 'thatch' = *atap*

- b) Semelai retains intervocalic /h/, no longer present in standard Malay, but present in some Malay dialects, e.g. Jakun: */tihɒŋ/* 'house pole' = *tiang (tihang)*; */prahōʔ/* 'boat' = *prau (prahu)*.
 c) Despite the fact that open final syllables $V]_σ$ are permitted in Semelai (§2.2.3.1), in Malay loans the coda is filled by a glottal segment, either /ʔ/ or /h/: */pasuh/* 'basin' = *pasu*, */hubiʔ/* 'cassava' = *ubi*. On the basis of the admissibility in Semelai, the 'innovation' probably belongs to the source. The choice of phoneme is not predictable.
 d) Intervocalic clusters of homorganic nasal plus voiced stop are reduced to the nasal, despite this sequence being possible in Semelai. Again, as in the previous point, this leads to the assumption that these forms have been borrowed in this form:

/ka.naw/ 'lemur' = *kandau*
/t.ma.kaw/ 'tobacco' = *tembakau*

Not all borrowings are subject to reduction, e.g. */mbuʔ/* 'bovine' retains the nasal + stop sequence.

- e) The sequence $lV.r(V)$, where the onset of the penultimate syllable is /l/, and the onset of the final syllable is /r/ is not allowed in Semelai, hence the metathesis to the acceptable sequence $rV.l$:

/ru.ius/ 'straight' = *turus*
/rɔ:l/ 'lorry' = *lori*
/p.robɛs/ 'Flores Island' = *Flores*

Variations of this type in Malay are recorded in Wilkinson (1927), for example *peruluh* vs. standard Malay *peluru* 'bullet'.

Metathesis also takes place between adjacent syllables where the offending segments are located in the onset of the penult and the coda of the final syllable,

$lV.w(V)r \rightarrow rV.w(V)l$: */s.r.wɒl/* 'trousers' = *seluar*, */k.r.wɒl/* 'to come out' = *keluar*.

Phonemes The following changes have been observed:

- a) Malay words ending in /f/, orthographic *k*, are realised as the voiceless velar stop /k/: */tasek/* 'lake' = *tasik* /taseʔ/.
 b) The voiceless labio-dental fricative phoneme /f/ is realised as the voiceless bilabial stop /p/, its closest equivalent in Semelai: */pekrər/* 'to think' = *fikir*, */protes/* 'Flores Island' = *Flores*. It may be that the phoneme is borrowed as /p/ from the local Malay dialect.

The voiced velar fricative /ɣ/ of many non-standard varieties of Malay is realised as /r/ everywhere: */rban/* 'pen', */critaʔ/* 'story'.

Note, for some speakers who have had contact with Malay /ɣ/ is a marginal phoneme, however other speakers who have not had direct contact, perceive it to be a Semelai phoneme, e.g. the personal name */layɪʔ/* = *lari* 'flee'.

- c) The final vowel, orthographic *a*, realised as schwa /ə/ in Malay, is a non-final vowel /a/ in Semelai, where it is always followed by a glottal stop: */cinaʔ/* 'Chinese' = *cina* /cɪnə/.

d) There is a change in the quality noted for some vowels. In some instances the realisation is systematic, in others it is sporadic, e.g. mid-high vowels realised as high vowels. Change tends only to occur in the final syllable. Again it is assumed that the words were borrowed in this form. See Collins (1985) for similar patterns in the Jakun of Kampung Landai, Pahang.

<i>/a/ → /ɔ:/</i>	<i>/smajɒt/</i>	'spirit of life'	= <i>semangat</i>
	<i>/sunɔ:t/</i>	'to circumcise'	= <i>sunat</i>
<i>/a/ → /ɛ:/</i>	<i>/mudɛh/</i>	'to be easy'	= <i>mudah</i>

There is one instance noted where the change applies to both syllables: */rentəŋ/* 'a creel' = *rantang*.

English Loanwords Semelai contains lexical items which are English in origin, e.g. */sigəret/* 'cigarette', but these are more likely to have been acquired as English loans in Malay: */stek/* 'slingshot' = *stek*, */rɔ:li/* 'lorry' = *lori*, */bɒm/* 'bomb' = *bom*.

During the era of the Communist Emergency, Pos Iskandar served as a base for British troops. Some English loans have been acquired directly from English, e.g. */resen/* 'rations'. Some of the older men who were boys at the time, claimed to have formerly spoken English. Two elderly men are known by names derived from English: */mɔ:nɪŋ/* 'morning', as a child he would always go around greeting the soldiers with 'Good Morning', and */pa:y pa:y/* 'five five', who is known by his English approximation of the price of a bottle of drink costing 55c.

Initial consonant clusters are reanalysed as an open penultimate syllable with the vowel [ə], */s.tate/* [sə.tə.tə] 'starter motor'. Other examples are: */s.ten/* 'Sten gun', */b.ren/* 'Bren machine gun' and */t.ray/* 'to try, attempt'.

The phoneme /ʃ/ is replaced by /p/ as in */s.tap nəs/* 'staff nurse' and the pronunciation of 'five' as */pa:y/*.

¹² Refer for instance to the variant -*2iʔ* which is suffixed to a vowel final item (§3.2.1.3).

2.5 Notes on comparative phonology of Aslian languages

The phonemic inventory below is largely adapted from a series of articles by Benjamin (1985b and 1986) for the *Orang Asli Studies Newsletter*. It provides the only comprehensive account of the phonemes of Aslian.

2.5.1 Consonants

There are five places of articulation, or six if alveolar and dental are distinguished. Some languages prefer alveolar to dental articulation, but the two are never contrasted (Benjamin 1985b: 8). All languages are reported to contain all of the consonants in the table, with the exception of those in italics which are explained below.

TABLE 2.4 PHONEMIC INVENTORY OF ASLIAN LANGUAGES

	BILABIAL	ALVEOLAR	PALATAL	VELAR	GLOTTAL				
STOPS (aspirated)	p (p ^h)	b	t (t ^h)	d	c (c ^h)	j	k (k ^h)	g	?
NASALS +glottal	(m)	m	(n)	n	(ŋ)	(?ŋ)	(ŋ)	ŋ	
FRICATIVES	(ɸ)			s	(z)				h
LATERALS +glottal				l (l ^h)					
RHOTICS +glottal				r (r ^h)					
GLIDES +glottal				w			y (y ^h)		

Only the Southern Aslian languages have a series of voiceless aspirated stops, an apparent innovation in this branch of the family.

The Northern Aslian group has two fricatives, bilabial /ɸ/ and lamino-palatal /z/: /nəɸhaʊf/ 'breath' (Mintil, Benjamin 1985b), /təd/ 'to scrape (blowpipe poison into a container)' (Cheq Wong, Kruspe fieldnotes) and /zəʔ/ 'new' (Bateg Deq, Kruspe fieldnotes). Bishop reports a phoneme /ɸ/ for Kensiu which only occurs syllable finally, e.g. /ʔeɸ/ 'to blow through one's fist' (Bishop 1996a: 234). It is not found in the Kensiu dialect described in Duangchand (1984).

The Kensiu dialect described in Duangchand lacks contrastive voicing for the velar stop (1984:183).

Final stops, are articulated in some languages as pre-stopped nasals, e.g. Semelai [ʔədʒɪ] 'I' (§2.1.1.1). Prestopped nasals are attested so far in Semai, Temiar, Jahai, Cheq Wong and Kensiu. A typological parallel is found outside Aslian in the related Mon-Khmer language Mon: [kɔd̩n] 'child, generic' and [kəta'b̩m] 'crab' (Bauer 1992b), and in Austronesian, primarily in the Land Dayak languages of Borneo, some of the languages of Sumatra, and the Chamic languages, but not elsewhere

(Blust 1997: 170).¹³ They are also attested in varieties of Malay spoken by some non-Austroasiatic Orang Asli.

2.5.2 Vowels

All Aslian languages recognise four degrees of height and three degrees of frontness, plus a phonemic contrast of nasalisation. The Central languages have a phonemic contrast in vowel length which is not found in either the Northern or Southern language groups, or the Central language Jah Hut.¹⁴

TABLE 2.5 INVENTORY OF STANDARD ASLIAN VOWEL PHONEMES

oral			nasal		
i	ɛ	ə	ɿ	ɛ	ə
e	ø	ə	ɔ	ɔ	ə
ɛ	a	ɔ	ɑ	ɔ	ə

Semelai has an additional low back rounded vowel /ɒ/ (§2.1.2); the Northern language Cheq Wong has an additional low front vowel /æ/. There are reportedly five degrees of vowel height for front and central vowels and four degrees for back vowels in Kensiu (Bishop 1996a: 228), while Duangchand reports only four in another dialect of Kensiu, with the high central vowel being realised as unrounded /i/ (1984: 102).

Semelai has a large inventory of nasal vowel phonemes (§2.1.2). The actual nasal phonemes may comprise only a subset of the oral vowels: Kensiu excludes the mid-high vowels /e, o/ and schwa /ə/ (Duangchand 1984); Temiar likewise does not have nasal counterparts of /e, o/ or /ə/ (Benjamin 1976b). In contrast Jah Hut has phonemically nasal counterparts for all oral vowels.

The nasal phonemes are of low frequency, Diffloth reports a frequency level of 6 per cent in the lexicon of Jah Hut (1976c: 103), and Duangchand records less than 100 out of a lexicon of 2200 items, the most frequent vowel being /ə/ (1984: 125).

Jah Hut (Diffloth 1976c), Kensiu (Duangchand 1984), Jahai (Burenhult 2002) and Semelai all record an absence of nasal vowels following the voiced velar stop onset /g/.

The Aslian languages are generally recognised as having neither contrastive register nor contour tone, although register distinctions are made in Mah Meri. For an account of the complex phonology of this language see Kruspe (in prep. a and b).

2.5.3 Syllable structure

Aslian languages have strong monosyllabic tendencies, although disyllabic and polysyllabic words are attested. The final syllable, often referred to as the 'major' syllable, exhibits fixed final stress, and is the most stable syllable in the word. To the

¹³ Adelaar (1995) questions whether they represent a residual areal feature of Borneo and the Peninsula, which is interesting given possible 'Aslian' evidence attested across Borneo, the Peninsula and Sumatra.

¹⁴ Benjamin (1976a: 67) reports that the Southern languages have phonemically long and short vowels. This is not born out by my research with the Semelai, Mah Meri or Semoq Beri (Kruspe fieldnotes).

left of this are the non-final or 'minor' syllables, non-stressed and generally exhibiting a smaller inventory of phonemes in terms of both consonants and vowels. This gives rise to what has become known in mainland South-east Asian typology as the 'sesquisyllable', half way between a monosyllable and a disyllable.¹⁵ Aslian languages also exhibit true disyllabic words:

- /garu/ 'termite' (Temiar)
- /kalto/ 'knee' (Semelai)
- /cu?ec/ 'to be inquisitive about something' (Semai)

All syllables must have an initial consonant. Any consonant phoneme may fill this position. The inventory of final segments is smaller than that of the initials, with voiceless aspirates and voiced consonants disallowed. Temiar, and some dialects of Semai represent a new innovation in the Aslian languages, with a set of final voiced consonants Benjamin (1976b): /gəb/ 'Malay' compared to Semelai /gəp/ 'Malay'. Final oral stops are always unreleased.

Southern languages have final open syllables as opposed to the closed syllables of other groups: e.g. Semelai /cə/ 'dog', /tʰi/ 'hand' compared to Jah Hut /cuə?/ 'dog' and /tʰi/ 'hand' and Mah Meri /ke/ 'to fall (of fruit)' compared to Cheq Wong /kar/ of the same meaning (Kruspe fieldnotes).

3 Morphology

In this chapter, the morphemic composition of words is examined, and an overview of the morphology is given in terms of the processes of word formation. The discussion of the various morphological processes will be organised according to the formal structures of process. A brief description of the function will be given, but the main discussion will be found in subsequent chapters.

The chapter begins with a description of theoretical preliminaries in §3.1 and an overview of the forms and processes. Affixation is organised into two sections according to the sub-systems which operate at this level: the indigenous system will be discussed in §3.2.2, and the loan system in §3.2.3. In §3.2.4 cases of archaic non-concatenative affixation are examined. Cliticisation is discussed in §3.3.

In the following discussion of morphology no particular theoretical model is adopted, although my treatment of the non-concatenative nature of the morphology processes draws broadly on the theories of prosodic morphology and template morphology. Descriptive tools have been adapted from consideration of the works of McCarthy (1981), Marantz (1982), Broselow and McCarthy (1983), Ter Mors (1983), and McCarthy and Prince (1990) to account for the features of Semelai morphology.¹

The term non-concatenative is adopted here *pace* McCarthy (1981) in its extended application used to account for reduplication processes (see also Spencer 1991, Gafos 1994).

3.1 Preliminaries

3.1.1 Terminology

Before proceeding to a discussion of the morphological processes, it will be useful to identify and define the structural units of the language; namely lexeme, grammatical word, clitic, root, affix and base. The following definitions are drawn from consideration of Lyons (1968), Spencer (1991) and Katamba (1993).

Lexeme A lexeme is the minimal free form, the smallest basic unit that exhibits lexical integrity and may exist independently. This is the form that would be listed in a dictionary as an independent entry, e.g. *dol* 'house', *2lap* 'be far'.

¹This account should not be considered in terms of this theory. The choice of some earlier sources is based on the objective of presenting an insightful and elegant account of the complex morphological structures encountered in Semelai, rather than relating the data to theoretical issues.

¹⁵In current phonological theory the 'sesquisyllable' is not recognised as a legitimate member of the prosodic hierarchy.

Grammatical Word The representation of a word associated with a certain morphosyntactic category is a grammatical word. It is necessary to define such a level with respect to clitic morphemes. Examples of grammatical words are: *ki=ca* (3A=eat) 'she ate', *co=hñ* (dog=3POSS) 'his dog'.

Clinic Clitics are unstressable phonologically bound morphemes, which attach to a host under certain syntactic environments. The 'host' is usually a phrasal category. In isolation, clitics lack meaning, but when attached to the 'host' their meanings are always predictable and devoid of the idiosyncrasy often associated with affixed forms.

Clitics may be prosodically deficient, failing to meet the necessary criterion of syllabic structure, e.g. the absolutive clitic *hn* 'ABS'. They are phonologically neutral, generally failing to undergo morphophonological processes. Enclitics never influence stress placement, a factor which distinguishes enclitics from suffixes.

A distinction is made between 'true' clitics, described in §3.3, and words which have the ability to cliticise, i.e. bound words. The former are unstressed as stated above, while the latter can be stressed in certain environments, e.g. the discourse clitics *=ce?* 'EM' and *=pa* 'CL' (§13.3). Clitic boundaries are indicated by '='.

Root Roots constitute the 'core' lexical items of the language, the majority of which exist as free lexical morphemes. Roots are monomorphemic units at the synchronic level, distinguished from affixes by their ability to occur freely, and the fact they exhibit lexical integrity. The very limited allomorphy and morphophonemic variation means that roots are always easily identifiable. There is no allomorphy of roots in relation to morphemic processes; allomorphy is only found in relation to the form of the affix, and there it is only morphophonemic.

Not all roots are free morphemes, certain roots existing only as bound forms, the postulated meaning bracketed: *√dəm* (lie down) → *dəm-dəm* 'to be lying down', *pa-dəm* 'to lay (s.th.) down'.

A further distinction can be made with regard to bound morphemes which are also non-occurring roots, most likely fossilised forms which do not enter into any other derivation. These 'non-lexemic' roots are identified thus, **√*, with the postulated gloss and morphemic structure bracketed (§3.2.6): **√kən* (offspring) → *knkn* (kn-kn) 'child'; *knən* (k<n>n) 'offspring'; *hankn* (nan-kn) 'childhood, youth'.

Affix Affixes are phonologically bound morphemes which operate in the domain of word formation. The class of affixes includes prefixes, infixes, suffixes, and a circumfix.² The form of a particular affix is frequently dependent on the prosodic structure of the root. The morphemic realisation of affixes vary. Some constitute well formed syllables, *tar-* 'MCAUS'; whilst others exist as subconstituents of syllabic templates: the causative morpheme *<r>* is a syllable coda in a morphemic template $[Cr]_\mu$, and the affix does not constitute a syllable itself (§2.2.1). The notation $[]_\mu$ indicates a morpheme template.

Affixes also vary in terms of their phonological specification. They may be either phonologically prespecified in terms of segmental and prosodic structure, as for the causative forms *par-/tar-* 'CAUS' (§3.2.2.1); or segmentally underspecified but prosodically specified, as in the case of coda copy, where a template of phonologically

empty units, $[CC]_\mu$, is affixed, and the phonological information is copied from the root (§3.2.2.2).

A third type exists as a combination of the previous two: a syllabic template consisting of a phonologically prespecified morpheme, and an underspecified one. The nominalisation process has a morphemic template $[nC]_\mu$, where *+n+* is the prespecified onset morpheme, and *+C+* represents the phonologically empty coda morpheme which sources its phonemic content from the root through the process of coda copy (§3.2.2.3). The empty vowel position is filled by epenthesis (§2.2.3.2).

Given the presence of complex morphemic units, it should be noted that not all constituents of these morphemic units are identifiable as 'meaningful' units (§§3.2.2.2–3).

In some instances, the derived root can undergo further derivation, making it necessary to identify a form 'base'.

In this grammar affix boundaries are marked by a dash '-' in the case of prefixes and suffixes, and by '< >' for infixes and the circumfix. Affixes which may surface as either a prefix or infix, are represented in the text enclosed by addition symbols '+ +', e.g. *+ra2+ 'COMP'*, but by the above conventions when affixed, e.g. *ra2-sey* (COMP-be.thin) 'to be thinner'. The addition symbol is also used to indicate general morpheme boundaries.

Base The base is any unit to which a morpheme can be added. The base may be either a root, or an affixed root which feeds further affixation. Thus *ca* 'to eat' combines with the causative prefix *pan-* 'CAUS', to derive *pan-ca* 'to feed' (§3.2.2.1). This base form can be further combined with the happenstance prefix *t-* 'HAPP' to create a new base *t-pan-ca* 'to happen to feed'.

Contenders for affixed bases are the imperfective (§3.2.2.2), the causative (§3.2.2.1 and §3.2.2.3) and partially reduplicated forms (§3.2.2.4). Nominalised forms (§3.2.2.3) are never bases at the level of derivational morphology.

3.1.2 Affixation and cliticisation

Two strata of attachment are recognised, affixation and cliticisation. Affixes occupy the inner level of derivational morphology, whilst clitics, which are syntactically driven, are located at the periphery, where they are arranged concatenatively. Phonotactically-based arguments support the distinction between affixes and clitics (§3.3).

Affixation is associated with the derivation of new morphemes by:

- a) placing the new lexeme in a different grammatical category, and potentially changing the meaning, e.g. the derivation of nominals from verbal roots: *glɔŋ* 'to swallow' → *g<nŋ>lɔŋ* (swallow<NMZ>) 'throat' (§3.2.2.3);
- b) placing the new lexeme in a different sub-class of the same grammatical category, e.g. the derivation of measure nouns from nouns: *rɔ?* 'basket' → *n2-rɔ?* (NMZ-basket) 'basketful' (§7.6.1.3), and likewise the formation of various verbal forms like detransitivising coda copy: *reŋ* 'to seek' → *rŋ-reŋ* (IMPERF-seek) 'to be seeking' (§3.2.2.2);

² For justification of the term circumfix see §3.2.3.3.

- c) modifying the meaning of the new lexeme without changing the grammatical category: certain types of verb to verb derivation: *jtek* 'to sleep' → *j<k>tek* (to sleep<IMPERF>) 'to camp out', and the derivation of the comparative of adjectives expressing dimension: *je?* 'to be short' → *j<r>a?>je?* (to be.short<COMP>) 'to be shorter' (§5.3).

Clitics modify a word in relation to phrasal or configurational properties by adding information relevant to its syntactic slot, e.g. the clitic *hn=* 'O' indicates that the following noun is a direct object, see (§3.3.4).

3.2 Affixation

3.2.1 Overview of forms and processes

There are two systems of affixation operating in Semelai: the non-concatenative system takes the left edge of the prosodic head as the domain of attachment, and is particular to the indigenous morphology; the concatenative system takes either the left or right edge of the base as the domain of attachment, and is particular to affixes borrowed from Malay.³ Light syllable reduplication (§3.2.2.4), which is considered a borrowed strategy here, is grouped with the non-concatenative affixes on the basis of structural observations. The one point of similarity between the two systems is the dependence on prosody, i.e. syllabic structure, to determine the realisation of the affix.

Affixes borrowed from Malay assign allomorphs on the basis of the prosody of the root and not segmental features as in Malay (§3.2.6). No such distinction is required for roots borrowed from Malay, which feed both non-concatenative and concatenative types of affixation.

3.2.1.1 Non-concatenative morphology

In non-concatenative processes, attachment is a left-edge phenomena resulting in two forms of affixation being identified: prefixation and infixation.

A specification of prosodic templates is an integral part of Semelai morphology. It is only through an understanding of the prosody of the language that morphological processes can be fully understood. To appreciate this, the issue of syllable structure is revisited.

In Semelai, words fall into three types based on their syllabic structure: monosyllabic, disyllabic or polysyllabic (§2.2.1). Monosyllabic roots and disyllabic roots with an open penultimate syllable feed both types of morphological processes; in general, disyllabic roots with a closed penultimate syllable feed only concatenative processes, and polysyllabic roots feed neither. The result of morphological processes

³ This distinction is problematic. The linguistic situation is extremely complex, and beyond the scope of the current description. Austroasiatic and Austronesian exhibit striking parallels in derivational morphology in terms of both form and function. Some reconstructed proto-forms illustrate this point, e.g. Causative: Austroasiatic *pa-/ap- and *ka- and Austronesian *pa-, *ka- and *paka-; Austroasiatic agentives *-um-, *ma-/am-, and Austronesian *mu-/um and *maRa- (Reid 1994). I have made a basic distinction between non-concatenative and concatenative processes which is roughly parallel to a division between the Austroasiatic morphemes, and those recognisable as borrowings from Malay.

means monosyllabic forms become disyllabic, and disyllabic forms either increase the weight of the penult to a heavy syllable, or derive trisyllabic forms.

There is a general though not absolute requirement that the penultimate syllable of a morphologically complex form be maximally expanded to a heavy syllable [CVC]₀. The affix may be of the form [CC]_μ, or it may only be the onset or coda of the penultimate syllable, the remaining consonants being constituents of the root. The underspecified C-slots of the morphemic template are filled by material copied from the final syllable of the base. Different morphological processes for the same morphological category are governed by the weight of the penultimate syllable of the base.

Affixes take one of three forms:

- a) A phonologically prespecified heavy syllable morpheme [C(V)C]_μ. The vowel may be prespecified as in +ra?+ 'COMP', or unspecified as in +pn+ 'NMZ'.
- b) A phonologically underspecified syllable, where the morphemic template [CC]_μ is phonologically empty. The phonetic information is copied from the final syllable of the root and prefixed or infixated back into it. This is a form of internal reduplication: *gos* 'to peel' → *gs-gos* 'to be peeling'; *kjep* 'to blink' → *k<p>jep* 'to be blinking'.
- c) The template is a combination of types a) and b): a phonologically specified constituent and an underspecified constituent. The phonologically specified material is a morpheme which is either a sonorant segment /n/, /m/ or /r/, or the voiceless bilabial stop /p/. The underspecified constituent is filled by the reduplication of the phonetic information of the root like that of type b) above: [nC]_μ 'NMZ' → *gos* 'to peel' → *ns-gos* 'act of peeling'; *kjep* 'to blink' → *k<n>jep* 'act of blinking'.

The stressed final syllable, or prosodic head, forms the base to which morphemic templates attach. Attachment only takes place at the left edge of this prosodic unit, hence, the domain of affixation is the prosodic head.

The morphological processes discussed in §3.2.2 all take the left edge of this unit as the domain of affixation regardless of the number of syllables in the root. Such is the strength of this requirement that where there is no penultimate syllable, as in the case of monosyllabic roots, a penultimate syllable is constructed to host the affix, i.e. the template cannot be infixated into the prosodic head. In the case of disyllabic roots, the penultimate syllable is reorganised in order to accommodate the affix in this position (§3.2.2.2). The original onset of the root is displaced into pre-penultimate position, whilst the vowel remains in the penult.

For the purpose of the exposition of morphological processes in Semelai, it is insightful to take the level of analysis one step further and recognise that the affixation of monosyllabic and disyllabic forms in fact constitutes the same process and should be understood as an identical operation, the prefixation of a morphemic template to the

prosodic head.⁴ In adopting this stance, there is the potential to minimise the need to appeal to syllabic types, or consonantal structure of the roots (Broeselow and McCarthy 1983; Gafos 1994). The importance of the foot as a prosodic constituent is further supported by the fact that a) it is only the final syllable which can bear stress, b) it is only this portion of the melody which is licensed to reduplicate, and c) the preferred site of heterosyllabic clusters is at the left edge of the prosodic head.

One morphological process stands apart from the others – light syllable reduplication. The domain of affixation is the left edge of the base. The reduplicant is the whole base, excluding the coda of the final syllable. Both consonants and vowels are copied, unlike other forms of reduplication where only consonants are licit. Reduplication of this type is a common feature in Malay dialects where it is applicable to both nominal and verbal forms. The process is described in §3.2.2.4.

In this section, the interaction between phonology and non-concatenative morphology has been introduced. Before turning to an overview of concatenative morphology, the underlying principles of the form of the template and its association to the base are summarised.

Principles relating to the form of the morphemic template – The Prosodic Morphology Hypothesis; The Template Satisfaction Condition and The Prosodic Circumspection of Domains – as well as the association of the template to the base, follow McCarthy and Prince (1990: 209–210).

The three Mapping Principles observed are: a) prefix the template to the final syllable; b) associate any prespecified material in the template to the base; c) copy the phonemic information of the base into any underspecified consonant (C) positions in the template from right to left. Any empty vowel positions are filled by means of vowel realisation rules (§3.2.1.3).

3.2.1.2 Concatenative morphology

A distinction is required in relation to the morphological processes borrowed from Malay which are arranged in a linear manner. Concatenative morphology in Semelai takes both the left edge and right edge of the base as the domain of affixation, manifesting as processes of prefixation, suffixation and circumfixation.

In this system the affix is always syllabic, and consonants are fully specified. Empty nuclei are filled by schwa epenthesis (§3.2.1.3). Roots never have allomorphs, only affixes, a feature shared with the non-concatenative system. Prosody still comes into play with regard to the choice of allomorph, as the syllabic structure of the affix is dependent on the number of syllables in the base: monosyllabic roots generally take heavy syllable allomorphs [CC]_μ, disyllabic roots, or bases of two or more syllables affix light syllable allomorphs [C]_μ.

3.2.1.3 General morphophonemic processes

The following morphophonemic processes are valid for both systems of affixation. They occur in response to inadmissible sequences of segments in terms of the syllable contact constraints (§2.2.3.1), e.g. /r/ dissimilation, violations of syllable structure as in coda resyllabification and violations of the phonotactics of the syllable (§2.3).

⁴ This observation was also made by Radhakrishnan in his account of causative affixation in the Nicobar language Nancowry where the causative morpheme ‘prefixes’ to the stressed final syllable (Radhakrishnan 1970).

Morphophonemic processes can be stated in terms of i) prefixing and infixing operations and ii) suffixing operations.

i. **PREFIXATION AND INFIXATION** There is only one morphophonemic process, glottal deletion, which affects the structure of the root. All other processes alter the realisation of the affix.

Glottal deletion Roots with an initial glottalised sonorant /ʔC/ lose the glottal feature when prefixed: ?yəŋ ‘to hear, listen’ → par-yəŋ (CAUS-hear) ‘to inform’.

/r/-dissimilation Geminate /r/ is not tolerated in Semelai (§2.2.3.1), furthermore, there is an absence of words with an /r/ in consecutive codas.⁵ Violations of these constraints can arise as a result of morphological operations.

The rules as set out in a) below pertain only to the non-concatenative system (§3.2.3). The rules for the concatenative system are motivated by the same conditioning factors, but manifest differently as shown in b).

a) If the onset of the penultimate syllable is /r/, the causative infix <r> is realised as <n>. The sequence *n]σ·σ[r is in violation of the syllable contact constraints (§2.2.3.1) triggering the onset of the following syllable to be realised as the allophone [r] (§2.1.1.1): carek ‘to tear’ → *ca<r>rek → ca<n>.rek [can.⁴rek] ‘to tear (s.th.)’; ?ris ‘to be alive’ → pan-ris [pan.⁴ris] ‘to raise (s.o.)’. This also occurs when the nominalising infix <n> or affix pn- is adjacent to an onset /r/: /reŋ/ ‘to seek’ → /pn.reŋ/, realised as [pn.⁴reŋ] ‘thing sought’.

Certain Malay loan words take an allophone /ŋ/: sanar ‘to lean against’ → sa<p>nar ‘to lean (s.th.) against (s.th.)’; rulus ‘be straight’ → ru<q>lus ‘to straighten’. Compare the following which has the regular allomorph <n>: kurang ‘be less’ → ku<n>raŋ ‘to decrease’.

Note that dissimilation does not occur where <r> occurs in consecutive codas as a result of coda copy: bər ‘to fasten’ → br-bər ‘to be fastening; jwər ‘to sew’ → j<nr>wər ‘act of sewing’.

b) /r/-dissimilation has the following realisation for the concatenative affixes. Where the onset of the root is the phoneme /r/, dissimilation is not attested, in its place the light prefix form b- ‘MID’ or t- ‘HAPP’ is selected: roc ‘to uproot’ → b-roc (MID-uproot), not *br-roc [bən⁴.roc]. Where the coda of the root is /r/, the prefix is the light syllable b- ‘MID’ or t- ‘HAPP’, in compliance with the constraint against sequences of the phoneme /r/ in consecutive codas: yvr ‘to ascend’ → t-yvr (HAPP-ascend) ‘to happen to ascend’, not *tr-yvr.

The causative prefixes have a number of alternant forms. In some cases these are clearly phonologically conditioned, as above, while in others the motivating factor is not readily apparent. The alternant forms are discussed in §3.2.2.1.

⁵ The onset and coda of the monosyllable cannot both be filled by r: */r Vr/. This is a common feature in Mon-Khmer languages, attested also in the Aslian languages Jah Hut (Diffloth 1976c) and Jahai (Burenhult 2002), and in Kammu of Laos (Svantesson 1983) and Sre of Vietnam (Manley 1972).

Onset deaspiration Onset deaspiration is particular to onset copy. Aspirated onsets lose the feature of aspiration when copied into the onset slot of the penultimate syllable: kʰɛ? 'to know' → k?·kʰɛ? 'to be knowing'; pʰɔy 'to fan' → py·pʰɔy 'to be fanning'.

The realisation of epenthetic vowels The epenthetic vowel is [ə]. Conditioned variants are determined by the coda of the affix. When the conditioning segment is a palatal, labio-velar or glottal segment, the epenthetic vowel is realised as [i, u or a] respectively, see §2.2.3.2.

- a) [ə] → [i] / _C[palatal] +
 t_{mep} 'to smile' → t_nmep [tip.məp] 'to be smiling'

b) [ə] → [u] / _C[labio-velar] +
 ?aw 'to stand' → ?w-?aw [?uw.?aw] 'to be standing'

c) [ə] → [a] / _C[glottal] +
 bə? 'to carry (child in sling)' → b? -bə? [ba? .bə?] 'to be carrying'

The application of morphophonemic processes overrides phonetic ones. In disyllabic roots which have a glottal onset in the final syllable, the vowel of the penultimate syllable may be realised as the vowel of the final syllable (§2.2.3.2): *c?ɔŋ* [c?ɔŋ] 'to roast'. However, when a morphemic template is infixated, the environment no longer permits the copy of the nucleus to proceed as it is blocked by the intervening morpheme: *c<ŋ>?ɔŋ* [c<ŋ>?ɔŋ] (roast<|IMPERF|>) 'to be roasting'.

ii. SUFFIXATION Suffixes are not a typological feature of Mon-Khmer languages. Semelai, unlike other Aslian languages, has incorporated two suffixes, -i? 'APPL/ITER' and -an 'NMZ', and a circumfix b<...>ah 'TOG', all items borrowed from Malay. The rules discussed here apply to the suffixes and the suffixing portion of the circumfix. In order to accommodate the suffixes in terms of syllable structure and stress placement, two morphophonemic strategies are required: a) Resyllabification and b) Stress reassignment.

a) Resyllabification takes place to avoid a vowel in the onset of the final syllable (§2.2.1). Roots which do not have a coda, select an allomorph which has a glottal stop in the onset, correcting the syllable structure to prevent violation of the constraint $*_o[V$ (§2.2.2): -?i? 'APPL/ITER' and b<...>-?an 'TOG'. An alternative analysis, would be to argue for a single suffix -i? 'TR', and a rule of glottal stop insertion to correct the violation in syllabic structure. The choice here, of allomorphy over epenthesis, is based on the persistent feature of Semelai morphology, where allomorphy is conditioned by the syllabic structure of the root.

b) Stress reassignment shifts stress from the penultimate to the final syllable (§2.2.2). The following example illustrates the above strategies: 'ca 'to eat' → 'ca-?i? (eat-ITER) → ca '7-1? 'to gobble up'.

3.2.2 Non-concatenative morphology

This section is organised according to the form of various affixes, rather than the function. Prespecified syllabic affixes are discussed in §3.2.2.1, underspecified non-syllabic affixes in §3.2.2.2 and prespecified non-syllabic affixes in §3.2.2.3.

3.2.2.1 Prespecified syllabic morphemes

A prespecified syllabic morpheme takes the form of a phonologically specified maximal syllable, where both the onset and coda positions are prespecified. The vowel position may be empty, or prespecified. There are three morphemes in this group, one with an empty vowel position, and two with fully specified vowels. The morphemes are: comparative *+ra?*+ 'COMP', nominalising *+pn+* 'NMZ' and causative *par-* and *tar-* 'CAUS'.

The morpheme is prefixed to the prosodic head and if the vowel position is unfilled, vowel epenthesis rules are invoked.

i. Comparative +ra2+ 'COMP' Monosyllabic roots are derived by prefixing the morpheme ra2+ to the prosodic head, producing a disyllabic form:

$$[\text{CVCl}]_0 \Rightarrow [\text{ra?}]_0 \cdot [\text{CVCl}]_0$$

This is illustrated with *sey* 'to be thin' → [ra?sey] vs. *sey* 'to be thinner' → [ra?sey].



In the case of disyllabic roots, the affix is realised as an infix

$$[\text{Cl}]_S [\text{CYC}]_S \Rightarrow [\text{Cl}]_{\leq S} [\text{ra3}]_{\leq S} [\text{CYC}]_S$$

The infix morpheme displaces the onset of the penultimate syllable of the root into the pre-penultimate position, producing a trisyllabic derivative: *jʌrəŋ* 'to be long' → *j<crə>jʌrəŋ* [jə.ra?.ʌrəŋ] 'to be longer'. The vowel position of the new pre-penultimate syllable is filled by enethesis (§3.2.1.3).



Some further examples are: *kēt* 'to be small' → *ra-* *kēt* 'to be smaller'; *dpeš* 'to be low, short' → *d* <*ra?*> *pēš* 'to be lower, shorter'.

This derivation, which derives the comparative of dimension adjectives, is discussed in (§§ 3.8).

ii. Nominalisation +pn+ 'NMZ' This morpheme shares the same process of attachment to the root as *+ra2+* 'COMP' above:

$$[CVC]_\sigma \Rightarrow [pn]_\mu - [CVC]_\sigma$$

and

$$[C]_\sigma [CVC]_\sigma \Rightarrow [C]_\sigma < [pn]_\mu > [CVC]_\sigma^6$$

The empty vowel position is filled by the 'elsewhere' realisation rule: $\emptyset \rightarrow [\emptyset]$. The monosyllabic $?oh$ 'to shoot with a blowpipe' derives $pn\text{-}?oh$ [pən.?oh] 'thing blowpiped'. The disyllabic form $cūm$ 'to wrap, bandage' gives rise to trisyllabic $c<pn>\text{-}ūm$ [cə<pən>?ūm] 'a package wrapped in a traditional leaf wrapper'. Some further examples are: $?ye$ 'to see' → $pn\text{-}ye$ 'sight', $swak$ to walk' → $s<pn>wak$ 'route'. See also §7.6.2.

iii. Causative par-/tar- 'CAUS' The morphemes used to derive causative verb forms are fully phonologically specified. In monosyllabic roots, one of the syllabic morphemes, selected on semantic grounds, is prefixed. In disyllabic roots where the penultimate syllable onset and nucleus are already filled, the choice of affix is syllable coda <r>, see (§3.2.2.3). The causative affixes differ from the previous affixes discussed above, because rather than infixing the whole prefix into a disyllabic root, there is just the non-syllabic consonantal allomorph <r>. Certain roots utilise a third process, prefixation of p-, discussed in §3.2.3.1.

The prefix for the monosyllabic form is a fully phonologically specified morpheme conforming to the maximal canonical syllable [CVC] $_\sigma$. The unmarked forms of the prefix are par- and tar- 'CAUS', identified on the basis of having the widest distribution. These morphemes are unusual in the number of allomorphs attested for them.

The prefixes par- and tar- and the allomorph for disyllabic forms, infix <r>, all feature a segment /r/ in the coda position, which possibly conveys the causation. Distribution of the monosyllabic prefixes⁷ is motivated semantically, and not phonotactically, see (§5.5.3.2).⁸

The association of the prefix takes the same form as the prefixing forms above: ce? 'to be lost' → par-ce? 'to cause (s.o./s.th.) to be lost'; wes 'to peel' → tar-wes 'to cause (s.o.) to peel (s.th.)'. Note that as a result of prefixation, the usual morphophonemic constraint applies where glottalised sonorant phonemes lose the glottal feature, e.g. ?pat 'to be ugly, malformed' → par-pat 'to make (s.th.) ugly' (§3.2.1.3).

The causative base feeds concatenative nominalisation, happenstance, middle voice and imperfective derivations (§3.2.3).

⁶ This affix is generally problematic. I suspect that the onset morpheme is the nominalising onset <p> and the coda is <n> (§3.2.2.3). Unfortunately, there are too few examples to draw either structural or semantic motivations for the two forms. The affix also resembles nominalising pen- a Malay affix, however the non-concatenative nature of attachment would appear to exclude this analysis, unless it is a very old borrowing.

⁷ One possible explanation for this form is that the prefix is causative <r>, like the disyllabic form, and the p-/t- defines the type of causation.

⁸ This is in contrast to the situation in Temiar where the monosyllabic root has a choice of two allomorphs, ber-, ter- which are phonologically rather than semantically conditioned. The conditioning is triggered by the onset of the root: ber- in the environment of an onset /t/ or /c/, and ter- elsewhere. Disyllabic roots have an allomorph <r> (Benjamin 1976b: 169).

Causative Allomorphs The causative prefixes have a number of alternate forms. In some cases, these are clearly phonologically conditioned; in others the motivating factor is not readily apparent. The following allomorphs are attested for par- 'CAUS' and tar- 'MCAUS': for the former pan-, paŋ-, pa- and p^har-; and for the latter tar-. Note that the morpheme *ta- is never attested as a light syllable like pa-.

- a) The prefix par- is realised as pan- when the onset of the root is /r/: ?ris 'to be alive' → pan-ris [pan.'ris] 'to raise (s.o.)'. This is the /r/-dissimilation process as described in §2.1.1.1 and §3.2.1.3.

It has been attested earlier (§2.2.3.2) that geminate /r/ is not tolerated in Semelai, and furthermore, there is an absence of words with an /r/ in both the onset and coda of the monosyllable. In b) below, the allomorph appears to be conditioned not by an identical coda/onset sequence, but by a sequence of identical codas /r/.

- b) This case looks like another instance of /r/ dissimilation although here it is the coda which triggers the allomorphy. It is not clear why the nasal is realised as a velar segment rather than alveolar, other than a possible constraint against the onset and coda of the affix having the same place of articulation in this construction.

The coda of the causative prefix tar- is realised as a velar nasal stop tar-, where the coda of the root is /r/: jor 'to urinate' → tar-jor 'to make (s.o.) urinate'; cɔr 'to fire (s.th.)' → tar-cɔr 'to make (s.o.) fire (s.th.)'.

The alternative analysis would be that /ŋ/ is the coda when the onset of the root is a coronal segment. However motivations on this basis are not attested elsewhere, unlike /r/ dissimilation.

- c) The prefix paŋ- is only attested as an alternant for par- in the following example: paŋ-woh 'to make (s.th.) upright' ← √woh (get up from lying). An alternative form is pan-woh.

d) The allomorph pa- also appears to be determined by the coda of the root. The prefix is realised as pa- where the coda of the root is bilabial /m/, or /p/, or labio-velar /w/: hūm 'to bathe' → pa-hūm 'to bathe (s.o.)'; ?lep 'to be far' → pa-lep 'to move (s.th.) apart' and √law (standing up) → pa-?aw 'to stand (s.th.) up, make upright'. An alternative form for √aw is /pan-?aw/ 'to stand (s.th.) up'.

The tar- prefix is insensitive to bilabial codas as shown in the following example: hop 'to put food in (one's) mouth' → tar-hop 'to put food in (s.o.'s) mouth'.

- e) par- is realised as pa- when prefixed to monosyllabic nouns: dɔl 'house' → pa-dol 'to house (s.o.)', dak 'water' → pa-dak 'to wet (s.th.)'.

- f) The prefixes par- and tar- are realised as p^har-, p^her- or t^har- when prefixed to a monosyllabic root with an onset phoneme /l/ and the coda consonant is a dental or palatal stop. It is not clear why the onset of the prefix is aspirated. It may be that the loss of the glottal feature manifests as the feature aspiration in the onset of the prefix: ?let 'to go out' → p^har-let 'to extinguish a flame, turn off', ?uc 'to pass, become free' → p^her-luc/t^har-luc 'to set (o.self) free/'to free (s.o.)'. Note the counter example where the coda of the root is /m/: ?em 'to be content' → pa-ləm 'to make (s.o.) content'.

g) The allomorph tam- appears to be taken one step further and the coda of the prefix is conditioned by the onset of the root with which it has assimilated homophonically: tam-p^hal 'to place (s.th.) down' ← p^hal 'to come down'. The problem is that this assumption requires that the prefix coda is /n/ in order for assimilation to take place, however the conditioning factors are not found in the data here.

With the exception of roots with /r/ in the onset, discussed in a) above, it is not possible to account for the alternant pan-, although it too could be considered a case of assimilation, recalling that /n/ is the realisation of nasal assimilation with the palatal series (§2.2.3.1): cin 'cooked' → pan-cin 'to cook', also par-cin; ca 'to eat' → pan-ca 'to feed'.⁹

3.2.2.2 Underspecified non-syllabic morphemes

Underspecified non-syllabic morphemes are representative of a form of internal reduplication which has been the focus of studies in prosodic and templatic morphology (McCarthy 1982, Broselow and McCarthy 1983, Sloan 1988), and in Optimality Theory (Gafos 1994). There are two non-concatenative processes of reduplication in Semelai: coda copy and onset copy.

Copy is a form of reduplication where a phonologically underspecified morpheme template is prefixed to the final syllable of the root. The phonemic content of the underspecified morpheme is copied from the prosodic head. Only this portion of the root is licensed to reduplicate. It is not just the segmental features that are copied, but also the prosodic identity of the segment: an onset is always copied as an onset, and a coda is always copied as a coda. Only consonants are copied, never the nucleus. An empty nucleus slot in the template is filled either by the root vowel of the penultimate syllable in disyllabic forms, or by the process of epenthesis if the vowel position is empty (§2.2.3).

Direction of copying is only ever from right to left, following the direction of syllabification, from the final syllable into the penultimate syllable position, and never beyond into the pre-penultimate syllable.

Coda copy is representative of infixing reduplication,¹⁰ where a copy of a segment of the root is infixated into the root. Coda copy is also attested as part of the template in nominalisation processes discussed in §3.2.2.3.

Onset copy is both phonologically and semantically underspecified. In essence, the morphemic unit resulting from the operation of onset copy is an empty morph, a formative required to satisfy the template satisfaction condition (§3.2.1.1). In subsequent nominalisation operations, coda copy also presents in the role of what appears to be a structurally determined formative, empty of any semantic content (§3.2.2.3). The lack of phonological specification of coda-copied morphemes is paralleled by a lack of semantic specification, the function being determined by the word class of the root to which it is affixed. These are briefly surveyed at the end of §3.2.3.

⁹ Further evidence that /r/ does not dissimilate with the palatal stop /c/ is found in the following example where the excessive agent prefix par- 'XS' is affixed to the root ca 'to eat' giving rise to the form par-ca (XS-eat) 'glutton'.

¹⁰ Matisoff (1989: 543) has coined a term 'incopyfixation' to label this process in Aslian and Lamet. This phenomenon is also present in certain Austronesian languages, Levantine Arabic, and some Amerindian languages of the Salishan family.

We now pass to individual treatments of coda, and onset, copy.

i. *Coda copy* The template for this process is the structure [CVC]_μ, equivalent to the maximal canonical syllable. It is prefixed to the final syllable, expanding the penultimate syllable position from either a light syllable [CV]_σ to a heavy syllable [CVC]_σ (3), or creating a penultimate syllable where there was previously none (4). This can be formulated:

$$(3) \quad [CV]_{\sigma} [CVC]_{\sigma} \Rightarrow [CV < C_{\mu} >]_{\sigma} [CVC]_{\sigma}$$

$$(4) \quad [CVC]_{\sigma} \Rightarrow [CVC]_{\mu} - [CVC]_{\sigma}$$

In (3) the morpheme is a single segment C_μ infixated as the penultimate syllable coda. In (4) it is also the single segment C_μ infixated as the penultimate syllable coda, but in a constructed penultimate syllable.

Disyllabic Roots Essentially the morphemic template consists of an empty or phonologically underspecified coda position which is prefixed to the final syllable. Although the morpheme does not constitute a syllable, it requires a host syllable. For disyllabic roots which contain a phonemic vowel in the penultimate syllable, an onset and nucleus are already prespecified (3), and the phonemic information is simply linked to the template. Disyllabic roots with an empty vowel position associate the onset of the root to the template and, following affixation, invoke realisation rules to fill in the phonetic vowel of the penultimate syllable (4).

We will begin by examining the process of coda copy as it is applied to disyllabic roots.¹¹ The root used to exemplify this is rabac 'to scratch' → ra <> bac 'to be scratching'.

(5a) Prefix the morphemic template to the final syllable and associate any prespecified phonemic information from the penultimate syllable with the template:

$$\begin{array}{cc} [CV]_{\sigma} & [CVC]_{\sigma} \\ | & | \\ r & a \\ | & | \\ b & a c \end{array} \Rightarrow \begin{array}{cc} [CV < C_{\mu} >]_{\sigma} & [CVC]_{\sigma} \\ | & | \\ r & a \\ | & | \\ b & a c \end{array}$$

(5b) Maximally expand the template by copying the phonemic content of the coda of the final syllable into the empty C position:

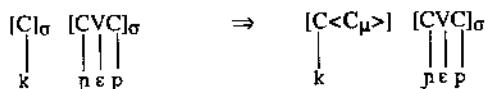
$$\Rightarrow \begin{array}{cc} & c \\ & | \\ & r a \\ & | \\ & b a c \end{array} \quad [CV < C_{\mu} >]_{\sigma} \quad [CVC]_{\sigma}$$

The output is the disyllabic structure ra <> bac [CV < C>]_σ[CVC]_σ.

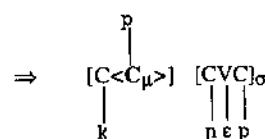
¹¹ The order of the presentation of the data does not set out to claim the disyllabic root and its process of affixation as primary, and the monosyllabic template as a conditioned variant, or vice versa. The aim here is to provide an economic and insightful account of the data at hand.

The derivation of roots which lack a vowel in the penultimate syllable is parallel to the derivation of roots with a prespecified vowel, apart from the realisation of the epenthetic vowel. The resulting derivation is simply parsed with the realisation rules in order to satisfy the template. An example is set out below. The root used to exemplify the derivation is *kjep* 'to blink' giving rise to *k<P>jep* 'to be blinking'. The initial syllabic structure of *kjep* is disyllabic [CVC]_σ[CVC]_σ.

(6a) Prefix the morphemic template and associate any prespecified phonemic content from the penultimate syllable with the template:



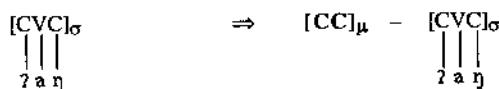
(6b) Copy the coda of the root into the empty coda position of the template:



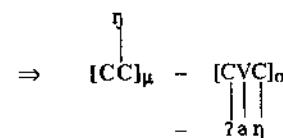
(6c) Apply the morphophonemically motivated epenthesis rules to supply the vowel. *k<P>jep* is realised as [kəpəpəp] in accordance with the 'elsewhere' rule *ŋ* → [ə] / _C +. The syllabic structure of the output *k<P>jep* is [CVC]_σ[CVC]_σ.

Monosyllabic roots Monosyllabic roots initially present as problematic. If we simply affix the template and invoke copying of the root coda onto the template, the output will present with an unacceptable structure *C_σ in the penultimate syllable, as illustrated with the root *?an* 'to open (mouth)':

(7a) The template is prefixed to the root. There is no penultimate syllable, so no information is available to be associated with the template.



(7b) Invoke coda copy:

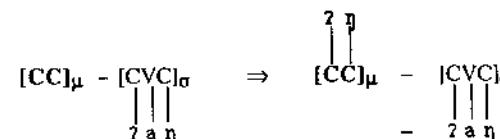


The operation now stalls as syllabic structure is violated in the form *n.?an. The response to this impasse is to invoke an operation that will fill the onset position of the template in order to satisfy the obligatory onset requirement. A strategy already available in the language is utilised; namely copy. Just as the coda is available to copy, so too is the onset, resulting in the process of 'onset copy'.

ii. *Onset copy* Monosyllabic roots require expansion to maximally satisfy the morphemic template [CVC]_μ, recalling that syllable structure requires obligatory onsets, *_σV (§2.2). The strategy for filling the empty onset position is to copy the phonemic content of the root onset into the empty template slot, giving rise to the formative σ[C_μ. Onset copy is restricted to occurring with monosyllabic roots involving coda copy; the onset of the final syllable of disyllabic roots is never licensed to be copied.

(8a) The template is prefixed to the root as shown in (7) above.

(8b) Onset copy is implemented subsequent to coda copy in accordance with the direction of mapping which operates from right to left:



(8c) Vowels are not reduplicated in Semelai, so the empty vowel of the template is filled according to the vowel realisation rules set out in §3.2.1.3: ?ŋ-?aŋ 'to be opening' is realised as [?əŋ-?əŋ].

Although the above analysis has been chosen, alternatively, in order to illustrate the parallel structure of the process in relation to disyllabic roots, association could take place first in relation to the onset, given that onsets are obligatory. Infixation of the coda would be into a structure resembling the disyllabic root, i.e. a word with a light penultimate syllable. In fact, the ordering would not be of any significance, and placing coda copy subsequent to onset copy would produce the same result.

iii. *Function* The lack of phonological specification of coda copy is reflected in the lack of semantic specification associated with the affix. The functions of this process are varied, ranging from the derivation of verbs to the derivation of nouns, dependent on the word class of the root. These are outlined below with cross-references to the relevant sections:

- Transitive verb to intransitive verb derivation (§5.3.1): *cək* 'to plant' → *ck-čək* 'to be planting'.
- Some verbs have a citation form which is the result of coda copy to a non-occurring free root (§5.2.1.2): *√ku?* (vomit) → *k2-k^hu?* 'to be vomiting'; **√kol* (rest head) → *klkol* 'to lay o's head down, be lying with head resting (on s.th.)'.
- Noun to stative verb derivation (§5.4.1.1): *tuh* 'breast' → *th-tuh* 'to have breasts'.

- d) Numeral to stative verb derivation (§5.4.1.2): *msoŋ* 'five' → *sŋ-sŋi* 'to be in a group of five'.¹²
- e) Adjective to noun derivation (§7.6.2.1): *jłəŋ* 'to be long' → *j<ŋ>ləŋ* 'length'.

3.2.2.3 Prespecified non-syllabic morphemes

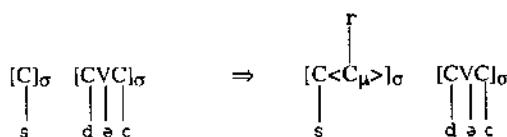
Prespecified non-syllabic morphemes are constituents of morphemic templates which are prespecified for only one position, either the onset or the coda. Those which prespecify a coda are only applicable to disyllabic roots, e.g. causative <r>. Prespecified onset morphemes fill either the empty coda position by means of coda copy where applicable, or where coda copy is not applicable, the allomorph is a syllable coda, e.g. nominalising <n>.

i. *Causative <r>* The formation of the causative of disyllabic roots is a phonologically prespecified syllable coda <r>. The morphemic template is represented as [CVr]μ.¹³

The morpheme <r> has allomorphs <n> and <p> which arise from the general condition of /r/-dissimilation discussed above in §3.2.1.3. When the onset of the final syllable is /r/, the causative morpheme is realised usually as <n>: *carek* 'to tear' → ca<r>rek → ca<n>rek 'to tear (s.th.)', but some roots take <p>: *tiris* 'to leak' → ti<p>ris 'to make (s.th.) leak'.

The causative morpheme is parsed into the empty coda slot of the penultimate syllable. The onset and nucleus of the penultimate syllable satisfy the syllabic requirements for the morpheme to be a constituent of a heavy syllable template [Cr]σ. This process is shown below where the causative form s<r>dəc 'to cool (s.th.)' is derived from *sdec* 'to be cool'. The initial syllabic structure of *sdec* is [C]σ[CVC]σ.

(9a) Affix the morphemic template into the penultimate syllable position and associate the phonemic information:



(9b) Determine the epenthetic vowel. The output is *s<r>dəc* [sər.dəc].

Some further examples are: *ma<r>suk* 'to put (s.th.) in, make (s.th.) enter' ← *masuk* 'to enter'; *p<r>cah* 'to smash (s.th.)' ← *peah* 'to be smashed'.

ii. *Nominalisation +n+ 'NMZ'* The primary method of nominalisation is the affixation of a morpheme +n+ 'NMZ'. There are no segmentally conditioned allomorphs of this morpheme, only prosodic ones. In the unmarked case the

¹² With certain numerals, the initial syllable deletes and the final syllable is reduplicated as a monosyllable (cf. § 5.4.1.2).

¹³ Cross-linguistically the use of an infix to express causation is the least common means (Song 1996). Song records this feature for only two languages, Nancowry, a Nicobar language (Radhakrishnan 1970: 53), and Kammu, a Mon-Khmer language (Svantesson 1983: 103), which have allomorphs for the causative, one of which is an infix. In Nancowry the infix is /um/, in Kammu it is /m/. Like Semelai, the Aslian languages Mah Meri, Temiar, Semai and Jahai, have a causativising infix, see also §5.3.5.4.

morpheme is an onset affix σ[n]; in the marked case the allomorph is the coda infix n]σ.

Nominalisation involves a phonologically prespecified onset morpheme, which is contained in a morphemic template: [nC]μ. Coda copy is invoked to satisfy syllabic requirements. This is always required for monosyllabic roots and for the majority of disyllabic roots. A subset of disyllabic roots infixes the morpheme as a coda in a template [Cn]μ.

The nominalisation process for monosyllabic and disyllabic roots consists of three steps:

- prefix the morpheme template [nC]μ to the prosodic head, and associate the prespecified phonemic content;
- in order to satisfy the syllabic requirements, invoke coda copy to copy the coda of the root into the empty coda position of the template (§3.2.3.1);
- if the vowel of the penultimate syllable is not specified, implement the morphophonemically motivated realisation rules as required (§3.1.2.3).

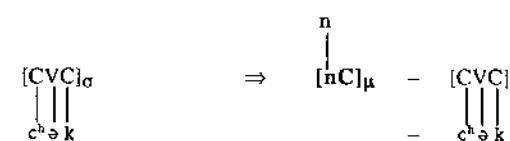
In the case of disyllabic forms, step a) results in the splitting of the penultimate syllable, so that the onset of the penult is displaced as an extrasyllabic segment into the antepenultimate syllable position, while the nucleus remains in the penultimate syllable. This is discussed below.

The disyllabic sub-type which infixes a nominal coda requires only step a) from above and is also discussed below. Nominalised forms do not feed further derivations.

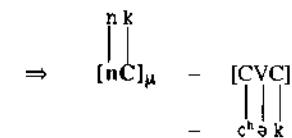
A. *Onset morpheme +n+* The discussion begins with the monosyllabic root. All roots, whether verbal or nominal, undergo the same process. Some notes on monosyllabic forms are given at the end of the section.

Monosyllabic Roots The verb *cək* 'to throw' derives the form *nk-cək* 'projectile, act of throwing'.

(10a) The affix is prefixed to the prosodic head:



(10b) Coda copy is invoked to maximally satisfy the template:



(10c) Realisation rules associate a vowel with the empty V position in the penultimate syllable: *nk-cək* is realised as [nək-cək].

Some further examples are: *gos* 'to peel' → *ns-gos* 'act, manner of peeling rattan'; *pot* 'latex cup' → *nt-pot* 'a cupful (measure noun)'.

Roots which have an open final syllable prefix an onset morpheme in a light syllable: *jalu* 'pig' → *j<n>alu* 'type of pig'. The vowel of the penultimate syllable is determined by the onset of the root: *co* 'dog' → *ni-co* 'type of dog'.

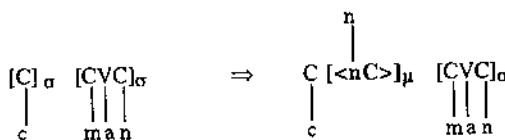
Disyllabic Roots The prefixation of the syllable template into a disyllabic root ignores syllabic structure and causes the displacement of the penultimate syllable onset of the root into the onset position of the antepenultimate syllable. The nucleus of the penultimate syllable of the root does not change position. Evidence of this is attested in roots which contain a prespecified vowel, e.g. *catek* 'to detach' → *c<n>a<k>tak* 'the endpoint'. This pattern of infixation contrasts with that of the prespecified morphemes which displace the whole penultimate syllable into antepenultimate position, see §3.2.2.1 and §3.2.2.2.

The extrasyllabic phoneme, a prespecified onset, presents a problem given that the minimal permissible syllable is [CV]ᵣ. There can be no empty vowel slot given that the original nucleus has been seconded by the penultimate syllable.

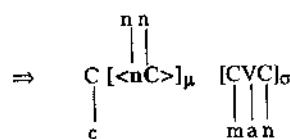
Given that the onset of the following syllable is always the morpheme *n*, the vowel will always be realised as [ə] in this construction. This is precisely the result which would be determined by the vowel realisation rules for phonetic vowels in the syllabification process set out in §2.2.3.2.

The derivation shown is of the root *cman* 'to retell a traditional story' which derives the form *c<nn>man* 'the retelling of a traditional story, a traditional story'. The first three steps are identical to those described for the monosyllabic root in (10) above.

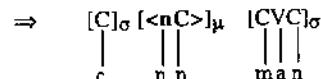
(11a) The morphemic template is prefixed to the prosodic head, displacing the original penultimate syllable onset into antepenultimate position. The prespecified information is associated:



(11b) The template must be heavy, so coda copy is invoked:



(11c) Invoke vowel realisation rules to fill the empty vowel slots in the non-final syllables. The output is the trisyllabic form: *c<nn>man* → [ce.ə.nən.man].



Another example is *jɔh* 'to drink' → *j<nh>?ɔh* 'act of drinking'.

Nominalisation is idiosyncratic in terms of both form and function. Verbs which are morphologically complex and do not have lexemic roots cannot undergo this process, *jɔjɔ?* 'to be raw' ← **√jɔ?*, does not derive **n?jɔ?*. This is in contrast to active intransitive verbs which have independent lexemic root forms, and do feed nominalisation: *dm-dəm* 'to lie down, be lying' ← *√dəm* → *nm-dəm* 'location of lying'; *k<c>bəc* 'to be fishing with a pole' → *√kbəc* → *k<nc>bəc* 'location being fished'.

The process of noun to noun derivation (§7.6.2), displays a degree of idiosyncratic variation. Nouns may affix either an onset or a coda morpheme. These are not predictable, and some disyllabic nouns have a variant form, often co-existing with a derived form with infixated syllable coda ⟨n.⟩: *sma?* 'person, generic' → *s<n>ma?* ~ *s<n?>ma?* 'type of person'; *sudu?* 'spoon' → *s<n>udu?* ~ *s<n?>du?* 'a spoonful'.

Noun roots with a closed penultimate syllable which is either a copy of the root coda, or a nasal segment, prefix the onset morpheme into the penult: *p?lb?* 'animal' → *p.<n>la?* 'a type of creature'.

Some roots utilise the onset morpheme, but do not implement coda copy to close the penultimate syllable. The reason is unclear, although all the roots have a velar nasal in the coda: *malar* 'side' → *m<n>alaj* 'area of a side'; *perej* 'plate' → *p<n>erej* 'a plateful'.

b. Coda morpheme +n.+ In the examples in the preceding section, the process of nominalisation occurred in conjunction with coda copy. Usually this process is only possible where the root permits coda copy, although certain nouns also follow this pattern. With transitive verb roots, coda copy forms an intransitive verb (§5.3.1). Stative verbs, which by virtue of being intransitive do not derive intransitive forms, utilise the allomorph coda ⟨n.⟩ in the underspecified coda slot, giving a template [Ch]ᵢᵣ. The structure of the process takes the same form as coda copy (§3.2.2.2) or causative formation in disyllabic roots, (9) above. This affix is restricted to disyllabic roots.

An example of coda ⟨n.⟩ derivation is illustrated with *jtek* 'to sleep', deriving the disyllabic form *j<n>tek* 'location of sleeping, sleeping area'.

(12a) Affix the coda morpheme to the penultimate syllable and associate all prespecified information:



(12b) Implement the vowel realisation rules: *j<n>tek* is realised as [jən̩tek].

Further examples are: *c?əc* 'to defecate' (**c<c>?əc*) → *c<n>?əc* 'place for/of defecating'; *p?ɔt* 'to remain' (**p<t>?ɔt*) → *p<n>?ɔt* 'location occupied'.

iii. Process nominalisation +p+ 'NMZ' The onset morpheme +p+ is prefixed to the stressed syllable as a template [pC]ᵢᵣ, and coda copy is invoked to satisfy syllabic

requirements: *sot* 'to become, permit' → *pt-sot* 'source, origin'. This derivation is discussed in §7.6.2.

Interestingly with this affix, although there is evidence of what appears to be an imperfective base, created through coda copy, there are some stative verbs which do not have imperfective forms, yet utilise this process: *?pih* 'to be ill, in pain' → (**jnh-jnh*) → *ph-jnh* 'pain, lesion'. This departs from the pattern established for the +n+ onset morpheme. Recall that with +n+ onset nominalisation, ability to form the imperfective was a pre-requisite for the selection of the syllable onset morpheme.

Monosyllabic roots are derived in the following manner, which parallels the nominalisation process for monosyllables discussed above (10):

$$[\text{CVC}]_\sigma \Rightarrow [\langle \text{pC} \rangle]_\mu \cdot [\text{CVC}]_\sigma$$

go? 'to fell (tree)' → *p? - go?* [pa?.go?] 'tree-felling'

The derivation of the disyllabic root is also the same:

$$[\text{C}]_\sigma [\text{CVC}]_\sigma \Rightarrow [\text{C}]_\sigma [\langle \text{pC} \rangle]_\mu [\text{CVC}]_\sigma$$

rmo? 'to reap' → *r <p?>mō?* [rə.pa?.mɔ?] 'reaping'

iv. *Agent nominalisation* <m(C)> 'NMZ' The morpheme <m> is a syllable onset morpheme: {m(C)}_σ. It is infixated into the penultimate syllable of disyllabic roots, or an imperfective or causative base derived from a monosyllabic root. This morpheme does not exhibit the productivity of nominalising +n+.

The structural process is identical to that of onset nominalisation with the template [nCV]_σ, as in (11) above.

For disyllabic roots the process is schematised as:

$$[\text{C(V)}]_\sigma [\text{CVC}]_\sigma \Rightarrow [\text{C}]_\sigma [\langle \text{m(V)C} \rangle]_\mu [\text{CVC}]_\sigma$$

sawel 'left-hand' → *s <m.al>wel* [sə.mal.wel] 'user of the left hand'

Other examples include: *satəm* 'right-hand(side)' → *s <mam>təm* 'user of the right-hand'; *sdər* 'to remember, recall' → *s <mr>dər* 'a good rememberer'. In an isolated case, <m> fills the empty penultimate syllable coda position, like coda nominalisation [C<n_μ>]: *vbəc* 'fish with a pole' → *k <m>bəc* 'fishing pole'.

For monosyllabic roots, the derivation of the imperfective base or a causative base must precede the affixation of the nominalising morpheme given that the onset affix is infixated inside the imperfective or causative prefix:

$$[\text{CVC}]_\mu \cdot [\text{CVC}]_\sigma \Rightarrow [\text{CV}]_\sigma [\langle \text{m}_\mu \text{VC} \rangle]_\mu [\text{CVC}]_\sigma$$

The following examples are derived from an imperfective base and from a causative base respectively. In both instances, the onset of the imperfective and causative morphemes is displaced: *k^bet* 'to sting, of insects' → *kt-k^bet* (IMPERF-sting) 'to be stinging' → *k <m>t-k^bet* [kə.*<m>et*.k^bet] (IMPERF <NMZ>-sting) 'stinging insects, generic'; *lət* 'to go out, of a flame' → *p^bar-lət* (CAUS-to go out) 'to extinguish a flame', → *p <m>ar-lət* [pə.*<m>ar*.lət] (CAUS<NMZ>-to go out) 'apparatus to extinguish a flame'.

An onset morpheme <m>, is also used to derive nominals from verbs for the avoidance speech style. It differs from the affix described above in that it is infixated into an open syllable:

$$[\text{CV}]_\sigma [\text{CVC}]_\sigma \Rightarrow [\text{C}]_\sigma [\langle \text{m} \rangle]_\mu [\text{CVC}]_\sigma$$

Interestingly, in this case all these roots are borrowed from Malay. The failure of the penult to be realised as a heavy syllable has a parallel in the affixation of nominalising +n+ into Malay loanwords, described above: *krej* 'to be dry' → *k <m>rej* 'alcoholic beverage (av.)'; *pdas* 'to be hot, spicy' → *p <m>das* 'chilli (av.)'.

3.2.2.4 Light syllable reduplication

This form of reduplication rightfully belongs in the following section on concatenative morphology given that it is most probably a Malay loan, however it is included here on the basis of its non-concatenative nature, i.e. a reduplicative process. Light syllable reduplication differs in a number of important ways from the non-concatenative processes discussed in the previous sections. These differences are discussed below.

This process of reduplication is licensed to copy both consonant and vowel segments of the base up to, but not inclusive of, the final syllable coda. In other words, the base, including all but the coda is licit with regard to reduplication. The syllabic structure of the root is maintained in the reduplicated portion, but it does not bear stress. The morphemic template for monosyllabic roots is [[CV]_σ]_μ, for a disyllabic base it is [[CV]_σ[CV]_σ]_μ, and so forth. The phonemic information of both consonants and vowels is then associated with the template.

Morphophonemic processes (§3.2.1.3) do not apply to this operation, e.g. the morphophonemic rule concerning the loss of the glottal feature of glottalised sonorant phonemes is not applicable: *?le?* 'to be long, of time' → *?le-?le?* 'eventually'.

This type of reduplication is common to Malay dialects, where both nominal and verbal forms are reduplicated. In Semelai it tends to be only verb roots which are reduplicated. The copied sequence is prefixed to the root as in (13), or the base as in (14):

- | | | |
|------|---|---|
| (13) | $[\text{CVC}]_\sigma$ | $\rightarrow [\text{CV}]_\sigma \cdot [\text{CVC}]_\sigma$ |
| | <i>?yot</i> 'to return' | $\rightarrow ?yo\text{-}?yot 'to return'$ |
| | $[\text{CV}]_\sigma [\text{CVC}]_\sigma$ | $\rightarrow [\text{CV}]_\sigma [\text{CV}]_\sigma \cdot [\text{CV}]_\sigma \cdot [\text{CVC}]_\sigma$ |
| | <i>dk^bes</i> 'to be near' | $\rightarrow dk^b\text{-}dk^b\text{-}es 'to be really near'$ |
| | $[\text{CVC}]_\sigma \cdot [\text{CVC}]_\sigma$ | $\rightarrow [\text{CVC}]_\sigma \cdot [\text{CVC}]_\sigma \cdot [\text{CVC}]_\sigma \cdot [\text{CVC}]_\sigma$ |
| | <i>mandeh</i> 'what' | $\rightarrow mande\text{-}mandeh 'whatever'$ |
| | $[\text{C}]_\sigma [\text{CVC}]_\sigma \cdot [\text{CVC}]_\sigma$ | $\rightarrow [\text{C}]_\sigma [\text{CVC}]_\sigma \cdot [\text{CVC}]_\sigma \cdot [\text{C}]_\sigma [\text{CVC}]_\sigma \cdot [\text{CVC}]_\sigma$ |
| | <i>krabat</i> 'royal court official' | $\rightarrow kraba\text{-}krabat 'all levels of royal court officials'$ |
| (14) | $[\text{C}]_\sigma [\text{CVCl}]_\sigma \cdot [\text{CVCl}]_\sigma$ | $\rightarrow [\text{C}]_\sigma [\text{CVCl}]_\sigma \cdot [\text{CVCl}]_\sigma \cdot [\text{C}]_\sigma [\text{CVCl}]_\sigma \cdot [\text{CVCl}]_\sigma$ |
| | <i>d <rə?>pəs</i> 'to be lower' | $\rightarrow d\text{-}rə\text{-}pə\text{-}d\text{-}rə\text{-}pəs 'to be really lower'$ |

Bases derived by light syllable reduplication, feed the formation of the reciprocal in conjunction with the affixation of br- 'MID' and the uncontrolled incessant action in conjunction with the affix tr- 'HAPP' (§3.2.3). The reduplication of verb roots, the 'intensive', is discussed in §5.3.9, and nominal roots in §7.6.1.4.

3.2.3 Concatenative morphology

The following affixes, which exhibit a concatenative system of arrangement, are borrowed from Malay. The crucial difference between this system and the indigenous one described in §3.2.1, is the way in which the root is parsed: the affix is not sensitive to the prosodic head of the root, only to the number of syllables it contains. Consequently, borrowed morphology is characterised by: a) its concatenative nature of arrangement, and b) the fixed nature of the morphemes: either an open or a closed monosyllable, prefixed or suffixed to the edge of the word.

The weight of the prefix is determined by the number of syllables in the root, a light syllable prefix is attested in pre-penultimate position, i.e. prefixed to disyllables, and a heavy prefix is attested in penultimate syllable position, prefixed to monosyllables. The allomorph of the excessive agent/performer prefix par- 'XS' is realised not as a light syllable in pre-penultimate position, but by the allomorph pr- [pər]; [ə] is the preferred vowel in antepenultimate syllable position. These factors are in accordance with Semelai constraints on the concatenation of syllables (§2.2).

The same allomorphy, light versus heavy syllables, is attested in Malay, the language from which the prefixes were borrowed. However the motivating factor in Malay is not prosodic, but segmental.¹⁴

The affixes borrowed from Malay are: middle voice br- 'MID'; happenstance tr- 'HAPP'; causative p- 'CAUS'; excessive agent/performer prefix par- 'XS', imperfective m(N)- 'PERFM', applicative -iʔ 'APPL/ITER', nominalising -an 'NMZ' and collective b- ... -an 'TOG'.

The same morphophonemic processes and vowel realisation rules presented in §3.2.1.3 are relevant to this section and are not repeated here.

The prefixes are discussed first in §3.2.3.1, the suffixes in §3.2.3.2 and the circumfix in §3.2.3.3.

3.2.3.1 Prefixes

i. *Middle voice b(r)- 'MID'* The prefix br-, used exclusively with monosyllabic roots as in (15), has an allomorph b- which is determined prosodically by the constraint disallowing a heavy syllable from occurring in pre-penultimate syllable position (16):

- (15) b-[CVC]_σ:
 - bɔy 'to dig' → br-bɔy 'be dug up'
 - dɔl 'house' → br-dɔl 'to have a house'

- (16) b-[CV]_σ [CVC]_σ:
 - tanre? 'to show (s.th.)' → b-tanre? 'be shown'
 - prec 'wing' → b-prec 'to have wings'

¹⁴ In Malay, the heavy prefix is used for vowel initial roots, the light in all other cases. Orthographically it is always written as the heavy syllable, e.g. ber- (Mintz 1994: 134).

- b-[CV]_σ [CV]_σ [CVC]_σ:
 - kreta? 'car' → b-kreta? 'to have a car'

Partially reduplicated bases (§3.2.2.4), feed prefixation of b- as shown below. The bases are either disyllabic (17), or quadrisyllabic (18):

- (17) b-[CV]_σ [CVC]_σ
 - tuy 'to dig (s.o.) in the ribs' → b-tu-tuy 'to dig each other in the ribs'

- (18) b-[CV]_σ [CV]_σ [CV]_σ [CVC]_σ
 - tmup 'to smile to (s.o.)' → b-tmua-tmup 'to smile to one another in acknowledgment on meeting'

ii. *Happenstance t(r)- 'HAPP'* The prefix tr- 'HAPP', (§5.3.7) used exclusively with monosyllabic roots (19), has an allomorph t- 'HAPP' which is determined prosodically by the constraint disallowing a heavy syllable to occur in pre-penultimate syllable position (20):

- (19) tr-[CVC]_σ:
 - ca 'to eat' → tr-ca 'to happen to eat'
 - yok 'to take' → tr-yok 'to happen to take'

- (20) t-[CV]_σ [CVC]_σ:
 - jtek 'to sleep' → t-jtek 'to happen to fall asleep'

The following bases feed the prefixation of t- 'HAPP': a) a causative base form (§3.2.2.1 and §3.2.2.3);

- t-pa-dəm (HAPP-CAUS-to lie down) 'to happen to fall to a lying position' ← pa-dəm (CAUS-lie down) 'to lay down' ← dəm 'to lie down';
 - t-c<n>roh (HAPP-CAUS-encounter) 'to happen to meet (s.o.)' ← c<n>roh (CAUS-encounter) 'to meet (s.o.)' ← croh 'to meet';
- or b) a base formed from the partial reduplication of the root (§3.2.2.4):
- t-ca-ca (HAPP-eat-eat) 'to happen to gobble up, with no concern for others' ← ca 'to eat'.

iii. *Causative p- 'CAUS'* Certain roots take the alternative causative prefix p- 'CAUS'.¹⁵ This is the least frequent of the causative morphemes. It is difficult to state the motivation for this morpheme. It is sometimes due to the heavy penultimate syllable which blocks normal processes of affixation (§3.2.1): jʔjiʔ 'dirty' → p-jʔjiʔ 'to dirty (s.th.)'. In other cases it is not possible to motivate a reason: dloŋ 'stick' → p-dloŋ 'to provide (s.o.) with a stick'; kasut 'shoes' → p-kasut 'to put shoes on (s.o.)'. Some roots exhibit alternant forms, either this prefix or a causative form from the indigenous system: jreh 'tired, from working' → p-jreh or j<n>reh 'to tire (one self) out' (§3.2.2.3).

¹⁵ The assigning of this morpheme to the class of borrowed affixes is done in full knowledge of the fact that it has a cognate in other Austronesian and Mon-Khmer languages. It is classed here primarily on account of its structural features, and because it cannot be motivated as an allomorph of par- 'CAUS' on the data in my corpus.

iv. *Excessive performer par-/pr- 'XS'* This form has two allomorphs *par-* 'XS', and *pr-* 'XS', which affix to the root of monosyllables and disyllables respectively (§7.6.2). The choice of allomorph is dependent on the syllable position of the prefix. In penultimate syllable position the allomorph is *par-* 'XS' as in a); in the antepenultimate syllable position, the allomorph is *pr-* 'XS', realised as [pərl], given that [ə] is the preferred vowel in this position as shown in b):

- a) ca 'to eat' → *par-ca* 'one who eats incessantly, a glutton'
- b) *ʃeəs* 'to fart' → *par-ʃəs* 'one who always farts'
- c) *pre?* 'to tease' → *pr-pre?* 'one who always teases'
- d) *jtek* 'to sleep' → *pr-jtek* 'one who always sleeps'

Context differentiates the monosyllabic prefix *par-* from the homophonous monosyllabic causative prefix *par-* 'CAUS' (§3.2.2.3)

v. *Perform m(N)- 'PERFM'* The process of attachment is the same as for Malay. When the onset of the root or base is a voiceless stop or fricative, the nasal mutates to a segment homorganic with the onset: *tar-beh* 'to deny' → *m-nar-beh* 'be denying'; for voiced stops, the nasal is homorganic with the onset of the root: *glar* 'name' → *mg-glar* 'to bestow a name'. When the onset of the root is a sonorant the prefix is realised as *m-*, e.g. *m-moto* 'ride (a motorbike)'.

The major function of this prefix is to derive an intransitive verb from a noun, 'to perform with an N'. It is also used as an alternative means of deriving the imperfective when a heavy penultimate syllable prevents coda copy: *pan-cin* 'to cook' → *m-man-cin* [mə-mən-cin] 'be cooking' (§5.4.2).

3.2.3.2 Suffixes

There are only two suffixes in Semelai, -(?)i? 'APPL/ITER' (§5.3.6), and -an 'NMZ' (§7.6.2). They are also the only vowel initial morphemes. Two morphophonemic rules are implemented to correct improper syllabification and improper stress assignment. They are resyllabification and stress reassignment, mentioned in §3.2.1.3.

a) Resyllabification takes place to avoid a vowel in the onset of the final syllable (§3.2.1.3). This is achieved by syllabifying the coda of the penultimate syllable as the onset of the suffix: *g.lok* 'to smile, laugh' → *glo.k-i?* 'to laugh unkindly at (s.o.)' (Suffixation) → *g.lo.k-i?* (Resyllabification). If the root has an open syllable a glottal stop is inserted as the onset of the suffix; alternatively it could be argued that there is an allomorph /-ʔi?/ which is selected by a vowel final base: *ca* 'to eat' → *ca-ʔi?* 'to keep on eating'.

b) Stress is reassigned to the final syllable: 'go? 'to fell' → *go.-i?* 'to keep on felling'.

Note that -an is a relatively infrequent suffix used primarily where a heavy penultimate syllable blocks the normal nominalisation process: *pan-cin* 'to cook' → *pan-cin-an* 'cookery'; *j2ji?* 'be dirty' → *j2ji?-an* 'dirtiness, filth'. In the case of Malay loans, where the penult is not heavy, it is unclear whether the lexemes were borrowed already affixed, or whether it is a productive process: *pajelan* 'address term' ← *panggilan* 'address term'; *tarek* 'to pull' → *tarek-an* 'handle' ← *tarik* 'to pull' (see §7.6.2).

3.2.3.3 The circumfix

The collective activity affix *b-...-(?)an* 'TOG' is analysed on language-internal grounds as a circumfix, rather than as an example of 'parasynthesis', the combination of a prefix and a suffix (Spencer 1991: 12-13). The 'suffix' occurs independently in Semelai, but does not appear to be related semantically. The prefixed portion, which is always in pre-penultimate syllable position is realised as *b-*. The usual morphophonemic processes associated with the suffix, as described above, are required to ensure correct syllabification and stress assignment: *paloh* 'to flee' → *b<pa'loh>an* 'many people fleeing together' (Affixation) → *b.pa.lo.han* (Resyllabification) → *b.pa.lo.han* (Stress reassignment).

The derivation is also fed by the imperfective base (§3.2.2.2): *?j-?ɔŋ* 'to be hiding' → *b<j-?ɔŋ>an* 'many people hiding together'. There is one instance where the affix is applied to a lexicalised verb phrase: *sar h̄m* (descend bathe) 'to go down to bathe' → *b<sar h̄m>an* 'to go down to bathe together'. See also §5.3.4.

3.2.4 Non-productive lexicalisation patterns

Many lexical items in Semelai display the morphological complexity of bimorphemic structure, but synchronically are monomorphemic roots.

This may be due to: a) processes identical to those of affixation (§3.2), where productive morphology is identified in fused forms; b) aberrant morphological processes which utilise the same type of affixation processes found in present day Semelai, as in the case of lexical items from the avoidance register.

There is a correlation between the structure of some morphemes and specific domains, in particular the names related to fauna and flora (§§3.2.4.1-2).

3.2.4.1 Coda copy

Coda copy, the process discussed in §3.2.2.2, is a marginal morphological process for deriving nouns in Semelai. However, there are many nouns in the lexicon that have the structure of a reduplicated form, but for which there is no lexemic root form. The realisation rules for vowels of morphophonemic origin apply to these forms (§2.2.3.2). Some examples are listed here according to semantic domains:

- *Life forms (generic)*: *knkon* 'offspring, not ego's' (*kn-kon*) ← **ʃkɔn* (offspring); *p?re?* 'ghost of one who has died a bloody death' (*p<?>re?*) ← **ʃpre?*; *p?la?* 'animal' (*p<?>la?*) ← **ʃpla?*.
- *Animals (lower order)*: *krk̄er* 'brush-tailed porcupine' (*kr-k̄er*) ← **ʃk̄er*; *?k-?ɔk* 'crow' (*?k-?ɔk*) ← **ʃ?ɔk*; *tytoy* 'small bird sp.' (*ty-toy*) ← **ʃtoy*; *?ŋ-?aŋ* 'termite, large variety' (*?ŋ-?aŋ*) ← **ʃ?aŋ*. The following example has the structure of a nominalised monosyllable (§3.2.2.3): *n?we?* 'python (av.)' (*n?we?*) ← **ʃwe?*.
- *Natural phenomena*: *khkuh* 'termite mound' (*kh-kuh*) ← **ʃkuh*; *?l?ol* 'mist seen in the morning or after rainfall' (*?l-?ol*) ← **ʃ?ol*; *djɔdɔŋ* 'log' (*dj-ɔdɔŋ*) ← **ʃdɔŋ*, note *dlɔŋ* 'tree'.
- *Flora (lower order)*: *prhgoh* 'a tree sp.' (*pr<h>goh*) ← **ʃprgoh*; *dij spletj* 'a bamboo sp.' (*s<p>pletj*) ← **ʃpletj*.

- *Body parts*: cncun 'elbow' (cn-cun) ← *v_{CON}; smsm 'bone marrow' (sm-sm) ← *v_{CON}; sksek 'plaque' (sk-sek) ← *v_{CON}. The last example is a nominalised form, a combination of the morpheme <n> plus coda copy (§3.2.2.3): cnɔŋkñi 'hip' (c<nɔŋ>kñi) ← *v_{CON}kñi.
- *Stative intransitive verbs expressing physical or mental conditions*: jʔji? 'to be dirty' (jʔ-ji?) ← *v_{CON}?; jʔjp? 'to be raw, of meat, rice' (jʔ-jp?) ← *v_{CON}?; lʔlb? 'to be of unsound mind' (lʔ-lb?) ← *v_{CON}?; shpeh 'to be lazy, disinclined' (s<h>peh) ← *v_{CON}h; sʔr? 'to be longing' (s<ʔ>r?) ← *v_{CON}r?; saŋjeŋ 'to be craving' (sa<ŋ>jeŋ) ← *v_{CON}eŋ.

3.2.4.2 Infixation

Various forms of infixation are set out in the following sections.

i. *Root infixation* Infixation into the root itself is not attested in Semelai, although it is a productive process in Temiar (Benjamin 1976b).

The non-occurring free root v_{KON} (child) has related forms: k<n>on 'offspring, ego's offspring'; k<m>on 'sibling's offspring'; k<n>mon 'related as sibling's offspring'.

ii. <JC> The following processes display recognised morphological processes, but the prespecified formatives are synchronically unattested.

There are a number of nouns which have a penultimate syllable onset morpheme <i>, which occurs in conjunction with coda copy: [IC]_μ. There are no actual root forms attested except possibly for the following: klmtam 'scorpion' (k<l>m>tam) ← ktam 'crab'; klpdop 'the appearance of baby hair, soft and fine and sticking up at all angles'. This is possibly derived from kdop 'fontanelle'.

Nominals following the former pattern denote various insect types¹⁶ and expressives: klkbok 'butterfly' (k<l>bok) ← *v_{KON}bok; klkap 'spider' (k<l>p>kap) ← *v_{KON}kap; kljnɔŋ 'dragonfly' (k<l>jnɔŋ) ← *v_{KON}jnɔŋ; kltpup 'cockroach' (k<l>p>tup) ← *v_{KON}tup; kltput 'the appearance of s.th. bobbing up and down, in and out of view' (k<l>tput) ← *v_{KON}tput.

The next two nouns listed, are possibly derived by prefixing k-1-: pre-penultimate onset k- and penultimate onset 1- without coda copy to a disyllabic base in an adaptation of the above process: klbeber 'a type of edible frilled fungus, found on decaying logs'; (k-1-beber) ← beber 'lips'; kld?do? 'termites in their winged life stage' (k-1-d?do?) ← *d?do? ← *v_{DO}?

Affixation of this type is also attested in Jah Hut (Diffloth 1976c: 100). He suggests that the Jah Hut forms are expressive type nominals, the sound symbolism conveying an underlying idea of 'erratic movement'. Some examples are /klkbak/ 'butterfly' and /kljeh/ 'a small bird sp.'

iii. <rC> Certain words exhibit morphological complexity, a penultimate syllable onset morpheme /r/ in conjunction with coda copy [rC]_μ: grhp>p 'food, to eat (av.)' (g<r><p>hp) ← ghop 'to be hot'; krmnom 'bladder' (k<r>m>nom) ← *v_{KON}nom. The following form sets a record containing three infixes: onset /m/, onset /r/ as

above, and coda /n/: jmrntek 'first basket of the rice harvest which is left to "sleep" until the harvest is over' j<m>n>r><n>tek ← jtek 'to sleep'.

3.2.4.3 Causative affixes

In addition to the causative forms discussed above in §3.2.2.1 and §3.2.2.3, certain lexemes have what appears to be a causative prefix kar-, k- and ha-. These three prefixes have cognates in other Mon-Khmer languages,¹⁷ including Aslian. Prefixes of this form are noted in Temiar: kar-, ha- (Benjamin 1976b: 139), and in Semai ka- is a malevolent causative prefix. The Nicobar language Nancowry has a productive causative prefix ha- and a non-productive causative stative prefix ka- (Radhakrishnan 1970: 46).

None of the Semelai items listed here have a recognised root, although possible sources are given for the first form: krlec (kr-lec) 'to cause to emerge'. This has the same meaning as p<r>iec 'to cause to emerge' ← plec 'to emerge'. Other examples are: karcēs (kar-cēs) 'to strike a light'; karsōc (kar-sōc) 'to wet (s.th.)'; kroy ~ haroy (k-roj ~ ha-roj) 'to sprinkle (s.th.)'.

3.3 Cliticisation

Clitics are bound morphemes which are primarily integrated at phrase rather than word level, where they function to fit words into syntactic structures.

Clitics are difficult to analyse: they are not independent words, but nor are they clearly affixes. Some have narrow distribution, attaching only to certain word classes (e.g. the pronominal clitics in §3.3.6), others have a broader range (e.g. hn 'ABS' (§3.3.4)); some have only a syntactic function, whilst others appear to substitute for an independent word (e.g. də= ~ ddə= [dədə] 'OF' (§3.3.6.2)), replacing də 'OF' (§8.9). Clitics are devoid of the morphological, syntactic and semantic idiosyncrasies associated with affixes. It is on the basis of these features that they are defined as clitics.

The individual clitics are described according to the following features from Klavans (1985): a) the phrase type which constitutes the domain of attachment of the clitic, b) the location of the clitic within that domain, c) the location of the clitic with respect to its host, proclitic or enclitic and d) the syntactic host and the phonological host are the same or different.

The domain of attachment and location within that domain is peculiar to the individual clitic, although in general the domain of attachment is a phrase, and the position within that domain is the initial constituent.

In Semelai the clitics are primarily proclitics, with the exception of absolute (=)hn(=) 'ABS', which may be either proclitic or enclitic to its host, dependent on its function (§3.3.4), and the pronominal enclitics =2en AUG and =h3? 'ATTN' (§3.3.6).

The following discussion effectively differentiates clitics from affixes.

a) *Syllabic structure* Clitics maintain phonological independence. They are typically of the form of the minimal syllable [CV]_μ, e.g. ma= 'IRR' 'imrealis', ko=

¹⁷ A prefix *ka- is reconstructed for Proto-Austroasiatic.

¹⁶ Note that many insect names contain an onset /k/: kaʔip 'centipede', khay 'millipede', kkibn 'slater bug', k'mur 'caterpillar', kmus 'type of fly'.

'3UA' 'unidentified agent', and do not require a heavy syllable template like affixes. The clitic is always fully phonologically prespecified. The open syllable of the clitic reduces the likelihood of phonotactic or morphophonemic processes motivated by the juncture of two consonants.

b) *Domain of attachment* The domain of attachment of clitics is the phrase. Clitics are located at the periphery of the base, either the left or right edge, peripheral to affixes. The form of the morpheme is insensitive to the prosodic structure of the base to which it cliticises. Compare *ma=yɔk* (IRR=take) 'would fetch' and *ma=yk-yɔk* (IRR=IMPERF=take) 'would be fetching'.

c) *Allomorphy* Affixes display allomorphy which is dependent on the prosodic structure of the root, while clitics are insensitive to this. For example in the formation of the causative, if the root is monosyllabic it takes a fully specified syllabic morpheme *par-*, but if the root is disyllabic the affix is realised as a syllable coda morpheme: <*r>* (§3.2.2.3). The following examples provide evidence for the insensitivity of the clitic to the syllabic structure of the root or base: monosyllabic *tʰɔh* 'to spit' → *ki=tʰɔh* (3A=spit) 'he spat'; disyllabic *jɔh* 'to drink' → *ki=jɔh* (3A=drink) 'he drank'.

3.3.1 Pronominal proclitics

The pronominal proclitic forms an integral part of the transitive verb phrase where its primary function is to cross-reference the optional external post-verbal NP_A (§5.5.1 and §9.1.3.1): *Pro_i=V (A=NP_i) (O)*.

- (21) *ki=tikam ia=knlək*
3A=stab A=husband
The husband stabbed (it).

Phonologically the proclitic is attached to the verb, but syntactically it coreferences the external NP.

Pronominal proclitics are distinguished from free pronouns by their inability to bear stress, and their position within the VP. The free form cannot appear inside the verbal unit, e.g. between the negator *da?* 'NEG' and the verb, nor can the free pronominal substitute for the bound morpheme, or vice versa. Compare the position of the bound pronominal in the transitive clause (22) with that of the free form in the resulting ungrammatical clause in (23).

- (22) *kəh, da? ki=ca*
3 NEG 3A=eat
Him, he didn't eat (it).

- (23) **da? kəh ca*
NEG 3 cat
*He didn't eat

A table of the bound pronominal forms is given in §5.5.1. For a full pronominal table, and a discussion of the semantics, see §6.1.1.

The domain of attachment is not the VP, but the initial verbal word within the VP, as in (24), and given that under elision of the verb it attaches to the directional preposition as in (25), this is evidence of cliticisation as distinct from affixation.

In a serial verb construction as in (24), the pronominal attaches to the initial verb in the sequence, even though this may be a supplemental verb. In this case, the second verb, not the first, determines the form of the pronominal morpheme and is therefore the main verb; the verb *sec* 'to steal' is functioning as a modifying supplemental verb 'surreptitiously' (§11.3.1).

- (24) *ki=sec yɔk*
3A=surreptitiously take
He surreptitiously took (it).

The clitic is only ever proclitic to verbs, as in the preceding examples, or to directionals when the verb is not represented as in (25), where the directional preposition *ta?en* 'TO:down' hosts the proclitic.

- (25) *ki=ta?en bri.kmuc*
3A=TO:down forest.ghost
He (went) down to the underworld.

First Person Allomorphs The first person singular familiar pronoun *2əŋ=* '1A=' has an allomorph /N=/: *2əŋ=* is reduced to a nasal (*m=*, *n=*, *ŋ=*) homorganic with the oral stop or palatal glide onset of the verb: *2əŋ=lən* → *n=lən* (1fA=want) 'I wanted (it)'. A familiar constraint is again attested: where the onset is a palatal obstruent /c, t^h, j, s/, the allomorph is realised as the alveolar nasal /n=/ as in *2əŋ=ca* → *n=ca* (1fA=eat) 'I ate (it)', (see §2.2.3). This rule does not affect the palatal glide /y/: *2əŋ=2ye* → *n=ye* (1fA=see) 'I saw (it)'.

3.3.2 Irrealis *ma* = 'IRR'

The proclitic *ma* = 'IRR' occurs primarily in conjunction with transitive verbs to mark the irrealis in a range of clause types (§10.2). It is mutually exclusive with the pronominal proclitic, the presence of *ma* = 'IRR' being associated with the suppression of the A in the clause. Unlike the proclitic, it does not coreference an external NP: *ma=V ((hn)=O)*.

ma = has the same range of distribution as the pronominal proclitics, cliticising to the initial verbal constituent of the verbal word (§5.5.2) as in the negative imperative clause in (26).

- (26) *boŋ ma=yɔk!*
NEG:IMP IRR=fetch
Don't fetch (it)!

3.3.3 Agentive/causal *ia* = 'A'/'BCS'

The morpheme *ia* = 'A', 'BCS' is proclitic either to the word it directly marks, e.g. a nominal or verb, or to the first constituent of a larger syntactic unit, such as an NP or

a VP. Examples of distribution are given below. On the basis of distributional grounds *la*=‘A’, ‘BCS’ as a marker of the ergative argument in the transitive clause (§5.5.1), and a marker of causal arguments (§11.5.3.1), is considered another example of a clitic morpheme. In all instances *la*=‘A’, ‘BCS’ has a constant form. There are no allomorphs and it is not subject to any morphophonological processes.

Distribution a) *la*=‘A’ marks the post-verbal NP_A in a transitive clause. It is cross-referenced by the pronominal proclitic: *Pro_i*=V *la*=A_i (O).

- (27) *ki*=?arjkot *la*=kmpen *hn*=dak
3A=fetch A=wife O=water
The wife fetched the water.

b) Cliticised to a phrase or clause, *la*=‘BCS’ has a causal or reason function. In (28) the initial constituent is already host to the pronominal clitic *ki*=‘3A’.

- (28) luka? nehneh, *la*=ki=ckit kmpen=hn
wound just BCS=3A=pinch.off.pieces wife=3POSS
(His ears were) just (covered in) wounds, because his wife had pinched bits off (them).

3.3.4 The clitic *hn*

The most striking difference between *hn* and the other proclitics is its phonological and morphological independence. *hn* can be realised as either a proclitic or an enclitic dependent on the category which it is marking, and the position of that category in the clause. The position of the clitic in relation to the host is determined syntactically. Further, it can appear in a clause in spite of the elision of the constituent it marks. In this respect it is closer to a word as compared to the other clitics discussed so far.

hn conforms to the typical non-final syllable type [CC]_S, where the vowel is a non-phonemic schwa and the final nasal has the phonotactics of a non-final syllable, [n] not [^tn] (the latter is the usual realisation of a final nasal following an oral vowel (§2.1.1.1)). In speech *hn* tends to be realised as a voiceless nasal [ɳ], as reflected in the transcriptions.

hn maintains phonological independence except when attached to third person pronouns, as discussed below. This is exemplified in the following account with regard to stress placement. As an enclitic third person possessive =*hn* ‘3POSS’ does not affect the assignment of stress, nor does it syllabify. This is in contrast to suffixes which cause resyllabification and the reassignment of stress (§3.2.1.3). Compare the following, where stress remains on the final syllable: *ba'pa?* ‘father’ → *ba'pa?*=*hn* (father=3POSS) ‘his father’, with a base suffixed with -i? ‘ITER’ which causes resyllabification of the root coda as suffix initial, and the reassignment of stress to the ‘new’ final syllable: *yɔk* ‘to take’ → *yɔ'.k-i?* ‘to borrow s.th. without returning (it)’ (§3.2.3.2).

Distribution The distribution of the clitic is directly related to its function. It is defined here in terms of distribution.

Proclitic. a) *absolutive hn*= marks the pos-tverbal core argument O¹⁸ (§9.3.1.2):

- (29) *ki=yɔk la=knlək hn=bantal*
3A=take A=husband O=pillow
The husband fetched a pillow.

In the event that an O has zero representation in the clause, =*hn* may appear enclitic to the verb:

- (30) *dɔs he?* dol, *ki=pan-cin=hn*
reach AT:above house 3A=CAUS-cook=O
(When he) arrived home, he cooked (it).

b) In the case of prepositional phrases, the clitic attaches not to the initial constituent of the phrase, but directly to the NP. Therefore, it can be argued that the domain of attachment is the NP within the phrase:

- (31) *huc ki=pan-ca ha? hn=kəh*
rice 3A=CAUS-eat AT ABS=3
Rice, she fed to her.

- (32) *pɔ:t rɔm hn=?əŋ?*
stay WITH ABS=1
Stay with me!

Enclitic. a) =*hn* ‘3POSS’ is enclitic to the NP when it signals third person possessor (§7.5.1). It attaches to NPs in both pre- and post-verbal position:

- (33) *co=hn ki=g<i>lɔr-i?* ɔi=nac̥t
dog=O 3A=name-APPL<NMZ> mr=[name]
His dog, he named Mr. nac̥t.

b) In the formation of clausal adjuncts, =*hn* ‘CONN’ is enclitic to the connective (§11.5.2): *la*lu?*=hn* (then=CONN) ‘then’.

3.3.5 The proclitic *me*=

a) *me*=‘REL’ is a proclitic which attaches to the initial constituent of a VP or PP to signal a relative clause (§11.1). In (34), *me*=‘REL’ is hosted by an adjective, and in (35) by the negative marker *da?* ‘NEG’.

- (34) *knon me=k dor*
offspring REL=be.female
(my) child who is female

- (35) *me=da? k^te?* km-k^tom
REL=NEG know IMPERF-sit

¹⁸ As a marker of S, it is only used for pronominal third person forms where it is fused in enclitic position to the pronominal. Stress is assigned in the following manner *kəhn* ‘3S’.

the one who doesn't know (how to) sit
monitor lizard (av.)

In (36) the proclitic is hosted by a PP.

- (36) mə=ʔen karom dol
REL=LOC under house
the one under the house
dog (av.)

b) The numeral *muy* 'one' has an homophonous proclitic form *mə* = 'one' which attaches to classifier phrases, measure nouns and numeral units (§7.3.1):

- (37) mə=pŋəl dlŋ
one=side wood
one end of the stick

3.3.6 Clitics to free pronouns

3.3.6.1 Augmented =ʔen 'AUG'

To express plurality in the first and second person free pronouns, =ʔen 'AUG' is encliticised to the non-familiar form of the pronoun: ye=ʔen '1=AUG' ← ye '1', he=ʔen '1&2=AUG' ← he '1&2' and je=ʔen '2=AUG' ← ji '2'. The distinction between familiar and non-familiar is neutralised in the plural form.

Evidence that this morpheme should be treated as a clitic rather than an affix is based on the failure of the morpheme to alter stress-placement. The morpheme fails to attract stress, like the enclitic functions of hn 'ABS', (§3.3.4) above, and unlike the affixes discussed in §3.2.3.

The second person plural form lowers the vowel /i/ to that of the final syllable /e/. The vowel /e/, attested in the first person forms, is not affected by the enclitic.

3.3.6.2 Possessive proclitic də= ~ ddə= 'OF'

Possessive pronouns are formed by the cliticisation of də= or ddə= [dədə] 'OF' to the free pronominal form (§6.1.4). də= ~ ddə= is an allomorph or clitic form of the possessive preposition də 'OF' (§8.9). də is used with nouns and demonstratives, whilst the reduced forms də= and ddə= are limited to co-occurring with the pronouns; də= can attach to either monosyllabic or disyllabic forms, ddə= can only attach to the former: də-kəh ~ ddə=kəh 'his, hers, its'; də=je=ʔen 'yours'.

3.3.6.3 Focus proclitics

The following two proclitics provide focus. Both forms may also host the attention clitic =h3? 'ATTN' as in (38). =h3? encliticises to personal pronouns, but only ever in conjunction with the proclitics described below.

i. *Possessor focus pə= 'PFOC'* The possessor focus proclitic pə= 'PFOC' can only be hosted by second and third person free form possessor pronouns:

- (38) A: məndehməh pə=kəh=h3? B: pinan
WHAT PFOC=3=ATTN betel.nut
A: What is his (one)? B: Betel nut

ii. *Third person absolutive focus ?a= 'DET'* The morpheme ?a= 'DET' is proclitic to the free third person pronouns, kəh '3', and deh '3pl', and the absolutive forms kəhn '3S', dehn '3pls' (§6.1.3):

- (39) sampay ?leʔleʔ=hn, sar hūm ?a=deh
until eventually=CONN descend bathe DET=3pl
So eventually, them, they went down to bathe.

3.3.7 Name ?i= 'NM'

The proclitic ?i= 'NM' has two domains of distribution:

a) Title ?i= 'NM', proclitic to a nominal or verb, forms a primarily male reference term, although the following term was used with reference to a female child: ?i-par-gən (NM-XS-bite) 'Miss Biter'. It is most commonly found in traditional narratives, ?i=miskin 'Mr. Poorman' ← miskin 'poor', and in the formation of some avoidance words,¹⁹ e.g. ?i=soc 'a type of babbler' ← soc 'to whistle', ?i=cɔm 'Mr. Cassava' avoidance word for 'cassava'.²⁰ The following is possibly derived in this manner: ?i=ʔe? 'elder male cousin' ← ?e?e? 'elder male sibling'.

b) A dysphemistic expression is derived when ?i= is cliticised to a noun and used in a negative existential clause (see §7.6.1.2) as in the following clause, where a husband querulously refutes the wife's claim that stones are used:

- (40) A: b-jɔy rom batu? B: da? da? ?i=batu?
A: MID-make WITH stone B: NEG EXIST NM=stone
A: (It) is made from stone. B: There aren't any damned stones!

This clitic is also used with nominalised nouns, denoting a lower order type or individuated members (§7.6.1.1), rather than the class as a whole: ?i=p<n>?la? 'damned animal' → p<n>?la? (animal<INDIV>) 'type of animal' → p?la? 'animal, generic'.

¹⁹ Diffloth (1976c: 100) notes a similar use of /i-/ as a reference marker for Jah Hut. It is used with kin terms and to derive pronouns from possessives, personal prefixes and deictics. He also notes a morpheme of the same form which functions as a definite article and third person possessive in Semai; in Temiar there is a Subject marker ?i- (Benjamin 1976b: 164).

²⁰ There is no word cɔm in Semelai, however in Jah Hut there is cwɔm 'to dig tubers' (Diffloth 1976c: 86), and in Mah Meri cɔp 'to dig (bait)' (Kruspe, fieldnotes).

4 Word classes

The following word classes are distinguished for Semelai: nominal (§4.2.1); verb (§4.2.2); expressive (§4.2.3); preposition (§4.3.1); adverb (§4.3.2); auxiliary (§4.3.3); existential and ascriptive predicates (§4.3.4); negator (§4.3.5); connective (§4.3.6) and interjection (§4.3.7).

Noun, a sub-class of the nominal superclass, verb and expressive are all open classes, the remainder are closed classes.

There are also various types of bound morphemes, including verbal (§5.3) and nominal (§7.6) derivational affixes, various verbal (§5.5) and nominal proclitics (§9.3) and discourse clitics (§13.3). These are discussed in the relevant sections as indicated.

4.1 Introductory remarks

The majority of lexemes in Semelai can be assigned unproblematically to a particular word class. Lexemes must undergo derivational procedures in order to function in a different word class. Verbs generally require 'deverbalising' before being able to function as a syntactic noun, either by morphological nominalisation: *yɔk* → *nk-yɔk* (NMZ-take) 'taking', or syntactic nominalisation *hitam*: 'to be black' → *ma=hitam* (REL-be.black) 'the black one'.

Nouns undergo verbalising processes in order to function as verbs, e.g. *bnih* 'seed' → *b<h>nih* (seed<IMPERF>) 'to sow rice seed'; *baju?* 'clothes' → *p-baju?* (EQUIP-clothes) 'to dress (s.o.)'. In other words, any non-prototypical function of a lexeme will be either morphologically or syntactically marked (see Table 4.1). In some cases there is more than one strategy available.

TABLE 4.1 LEXICAL, SEMANTIC AND PRAGMATIC CATEGORIES

	REFERENT	ATTRIBUTE	PREDICATE
NOUN (object)	<i>dɔl</i> 'house'	<i>məɔɔm dɔl</i> 'house-like'	<i>br-dɔl</i> 'to be housed' <i>par-dɔl</i> 'to house (s.o.)'
VERB (property)	<i>r<h>mɔl</i> 'male, man' <i>ma=rɔl</i> 'the male one'		<i>rmɔl</i> 'to be male'
(action)	<i>n?-gɔ?</i> 'felling'	<i>br-gɔ?</i> 'felled'	<i>gɔ?</i> 'to fell a tree'

There is a small set of bound roots which exhibit idiosyncratic behaviour (§5.2.2). They are associated primarily with verbal derivatives, e.g. *√yup* (sleep covered) *yp-yup* 'to sleep with a cover'. However, the root may also be used in conjunction with a noun to form a nominal compound, without first being nominalised. This is atypical for an associative phrase of this type (§7.5). Compare *kayen yup* (cloth (sleep covered)) 'sleeping cover' with the compound for 'drinking water' *dak j<h>?ɔh* (water drink<NMZ>). Until further investigation is undertaken, it is

difficult to ascertain whether these are indeed verbal roots, or if they have some sort of precategorial status.

For the most part, the lexemes which present problems to the system in terms of multifunctionality are loanwords from Malay, e.g. *turköt* 'stick', *turköt* 'to prop up', where there appears to be a case of zero conversion operating between the word classes of noun and verb. The treatment I give to such lexemes is to posit a case of heterosemy.¹ Not all loanwords behave in this manner, and many, like *krja?* 'to work', need to undergo overt derivational change, e.g. *krja?* 'to work' → *k<h>rja?* 'work, task'. In the domain of indigenous Semelai lexemes, zero conversion usually only takes place within subclasses of a main class, e.g. transitivity shifts such as the labile verb *p^hul* 'to come down, intr./to come down (s.th.), tr.', or shifts from main verbs to supplemental verbs or adverbs, e.g. *?ayon* 'to be slow' and adverbial 'slowly'.

We will now examine the general features of class membership in Semelai. The major contrast is between open and closed word classes. Within this division, it may be possible to further organise the membership into minor classes identifiable on the basis of combinatorial possibilities. Not all distinctions at this level of minor classification necessarily signal different word class membership.

An open class is distinguished by the following criteria:

- a) It will have a potentially unlimited number of members, and thus can be freely added to.
- b) It may have the potential to feed derivational processes which function productively to augment the membership of the class to which the form belongs, e.g. causative derivation of the verb *ca* 'to eat (s.th.)' → *pan-ca* 'to feed (s.th.) to (s.o.)' (§5.3.6).
- c) It may contain subclasses which are themselves closed, e.g. the verb class contains a subclass of eight dimension adjectives (§4.2.2). Classes which exhibit finite membership are grouped within the open class on the basis of primary criteria, e.g. the potential to function as the head of a verb phrase.
- d) The members of an open class are generally free lexemes.

A closed class can be distinguished by the following characteristics:

- a) It has a finite number of members and thus cannot generally be added to.
- b) It does not contain items which result from the application of derivational processes to members of the open word classes. In general, the members of the class do not have the potential to feed derivational processes (for exceptions see second person pronouns (§5.4.3.3), and the locative prepositions (§6.3.1.2)).

¹ This is in fact a widespread problem in Austronesian, which has received a range of treatments – see Broschart (1997) for Tongan, Schachter (1985) for Tagalog, and Gil (1994) for Riau Indonesian. The same problem has also been claimed in Austroasiatic, for instance Hoffmann (1903) for Munda and more recently for Khmer (Schiller 1992). However, for a reanalysis of the Munda problem, see Evans (2000), and for word classes in Korku, see Nagaraja (1999: 30). The situation in Semelai corresponds well with the description of word classes in the Central Asian language Semai (which Diffloth notes with surprise given the situation in other Austroasiatic languages, and Austronesian), where morphological and syntactic criteria clearly define the two classes of noun and verb, with only minor evidence of overlap (Diffloth 1976a: 249-51).

- c) It may constitute a subclass of an open class (see point c) above).
- d) Members of the closed class may be either free lexemes or bound lexemes.
- e) Closed classes tend to exhibit distinctive and paradigmatically organised patterns of semantic opposition, e.g. the categories of person, number and deference in the pronominal system (§6.1.1), the spatial parameter of proximate and distal in demonstratives (§6.3.1), the antonymic oppositions in the adjectives of dimension (§5.2.1), and the spatial distinctions in the deictic locative and directional prepositions (§§8.4–8.6).

In the analysis given here, the word classes are distinguished by structural or formal properties, in terms of the relations they form with other categories. Essentially, membership in classes is unique, although there are some areas where a lexical item may exhibit features of two classes, suggesting an overlap. This apparent overlap, where certain items appear to exhibit membership of a different class is treated as a case of heterosemy, e.g. locative nominals which may enter into a construction which is equivalent to a locative preposition (§7.5.3).

Membership of a class does not automatically attribute the same status to all members of that class (Croft 1990: 190). At the core are the prototypical members which display all attributes; less prototypical members which may display only a subset of the characteristics of the class or exhibit their own unique features, and members which sit on the boundary, exhibiting features of intersecting classes, e.g. deverbal nouns, where the derivative does not enjoy the same membership as prototypical members of the class. This point is taken up in §4.2.1.

Derivational morphology creates the means of transfer from one class to another, e.g. nominal to verbal and vice versa, but for the most part, movement between classes is unidirectional. Derivational morphology is at its most productive for verb to verb derivations, more constrained for verb to noun derivations, restricted for noun to noun derivation and highly restricted for noun to verb derivation.

Intraclass derivations can be characterised as imparting information about a specific facet of an event or entity, e.g. verbal morphology imparts information about aspect or changes in valence structure; nominal morphology derives nouns which refer to a subset of referents of a generic noun. It should be pointed out that where affixes share the same phonological form, e.g. deverbal nominalising *+n+* 'NMZ' and the nominal unit deriving affix *+n+* 'UNIT' (§3.2.4), they are considered as a case of affixal homophony and it does not suggest a case of convergence between lexical categories.

Derivation is semantically driven, and often idiosyncratic, hence it is not reliable as the sole means of establishing distinctions either between or within particular classes. If anything, derivation suggests the necessity of defining specific semantic field categories. For example, the application of verb-deriving morphology to nominals is restricted to classes of words which have body parts, clothes, instruments/implements, kin or rituals as their referents, deriving verbs with the meaning 'to have/use/provide with N', e.g. *tpuŋ* 'flour' → *t-<ŋ>-puŋ* 'to perform the *tpuŋ* tawar "flour" healing ritual'.

The following discussion of the various word classes is divided into two main sections, the first deals with open classes (§4.2), including the superordinate category of nominal (§4.2.1), and the second section deals with the numerous closed classes (§4.3). Nouns are distinguished on the basis of syntactic rather than

morphological criteria, whereas verbs also invoke morphological criteria, although this functions to set up various minor classes rather than provide conclusive evidence of a class of verb.

Two specific exceptions to the general complementary distribution of noun and verb are the ability of restrictive adverbs to modify both verbs and nouns in terms of scope, *ca ?en* (eat just) 'just eat', *dak ?en* (water just) 'plain water' (§10.3.1); and the ability of both nominal and verbal phrases to be marked with *la=* 'A; BCS' with similar semantic consequences (§11.5).

Two categories, expressive (§4.2.3) and interjection (§4.3.7), function independently without syntactic connection to another clause, but maintain a pragmatic connection within the discourse context. The discourse function of connectives (§4.3.6) also shares this feature.

4.2 Open classes

4.2.1 Nominals

The following word classes are subsumed under a 'nominal' superclass: noun, numeral, pronoun, demonstrative and ignorative.

Nominals are free lexical roots which have the ability to head a noun phrase (NP). The primary member of the group is the open class of nouns which function referentially to indicate an entity. The major syntactic function is as arguments of predicates.

Numerals represent an open class (§4.2.1), although within this class there is a small closed class of indigenous terms.

Pronouns, demonstratives and ignoratives (§4.2.1) are all closed classes. The first two are distinguished by their ability to function as determiners, while all three classes are distinguished in terms of their ability to head an NP.

4.2.1.1 Nouns

The main function of the noun is as an argument of a verb. They occur as phrasal constituents at the head of an NP (§7.2). Typically, nouns appear with a referentialising phrase (§7.3), an attributive phrase (§7.4) or an associative phrase (§7.5). They may also mark relationships between nouns, e.g. in an associative phrase, or a possessive phrase, e.g. *pjam dɔ ?əŋ* (NMZ-feel OF 1f) 'my opinion' (§8.9). As the arguments of verbs, they may host role marking proclitics (§9.1).

A marginal function of nouns is the predicative function (§10.1). Nominal predicates are distinguished from verbal predicates by virtue of the fact that they do not subcategorise, and with the exception of the imminent aspect clitic *ga=* 'IMM', do not co-occur with aspectual adverbs.

Nouns can be identified as either proper or common, which affects the manner in which they are modified. Proper nouns have unique reference and therefore are neither modified by attributives nor co-occur with quantifiers or demonstratives, although they may include a title, e.g. *puyŋŋ jo?an* 'shaman jo?an'. Kin terms are used to both refer to and address people, and may be modified by either a demonstrative or a locative phrase or another kin term or pronoun: *?ude?* 'parent's third-born female sibling'; *pa?nah he?* *jlawat* 'parent's second-born male sibling'

up at *jlawat'*, *kmon ye* 'my sibling's offspring'. Common nouns may refer to a single entity: *malej* 'sky'; a set: *?ma?*.*bapa?* 'parents', or a mass: *gar?m* 'salt'. This restricts the range of constituents which can co-occur with the noun. Mass nouns, for example *gar?m* 'salt', cannot be individuated other than as a measured quantity: *gar?m m?=?c<pn>?um* (salt one=wrap<NMZ>) 'one package of salt' (§7.3). Generic or superordinate nouns, e.g. *cim* 'bird', have derived forms which specify a lower order set or individual referent, e.g. *nm-cim* (INDIV-bird) 'kind of bird' (§7.6.1). Some examples of nouns are given below.

Proper nouns include: place names, *t?mayer* 'Tembangau River'; individual person's names (these are rarely used), *lec* 'a woman's name', *lajkog* 'a man's name', *r?mgasi* 'the name of an ogre'; nicknames, *paypay* 'a man's nickname'.

Kin terms include: *?i?e?* 'elder male cousin', *wa?*.*mboj* 'parent's eldest sibling (male)', *busu?* 'youngest-born sibling', *?ipar* 'sibling-in-law' (§1.3.3.2).

Common nouns may have concrete objects as their referents, e.g. humans (*sma?* 'person') including *kin* (?*ma?*.*bapa?* 'parents') and other relationships (*b?i?e?* 'unrelated friend'); afterlife forms (*kmuc* 'ghost'); artefacts (*luk* 'quiver for blowpipe darts', *lat* 'dam barrier'); body parts (*t?n* 'ear, place where handle is attached to a basket', *suk* 'hair, feather, fur'); rituals (*tpuj* *tawar* (flour exchange) "flour" healing ritual); environmental features (*bri* 'jungle', *siegoc* 'channel'); flora (*bkaw* 'flower', *sak?* 'pandanus'); fauna (*co* 'dog', *yuyuy* 'gnat'); celestial bodies (*bulan* 'moon'); meteorological phenomena (*musim* 'season', *tijo rbol* (snake aroid) 'rainbow'); times (*som* 'morning'); and quantities and measures (*k?n>opi* 'tinful' ← *kopi* 'tin').

Nouns with abstract referents are usually derived from verbs, e.g. abstracts of dimension like *nsp?es* 'height' ← *?p?es* 'to be high' or other properties like *p?<t>ret* 'heat' ← *pret* 'to be hot', and concepts like *s<n>u<n>m?n* 'act of incest' ← *sum?n* 'to commit incest'. In general such abstract concepts, in particular actions and feelings, cannot be expressed by nouns, but require expression as full clausal statements.

Deverbal nouns (§7.6.2) do not generally function as arguments of verbs, but may function either as a subject (1), or a predicate of non-verbal clauses (2), or in possessive interrogatives (§10.5).

- (1) *s<n>wak k?h da? btol*
walk<NMZ> 3 NEG be.correct
His manner of walking is not correct.

- (2) *tmpot [ga=nl-p?ul k?h]pred. k?b?ej*
place [IMM=NMZ-come.down 3]Pred QUOTE
"(This) place is going to (be) (the one of) his descent," (they) said.

4.2.1.2 Pronouns

The free personal pronouns are a small closed class (Table 6.1). There is also a corresponding set of proclitic forms (§5.5.1). Pronouns are restricted to having animate or potent inanimate entities as their referents.

Pronouns are distinguished for number in all three persons, unlike nouns which do not have a number distinction. There is an inclusive/exclusive distinction in the first person, a distinction in deference in the first and second person minimal and an absolute distinction in third person singular and plural forms (§6.1).

Other than this, pronouns are not distinct from the class of nouns, and may co-occur with demonstratives and other determinatives, role-marking clitics (§6.1.2) and the possessive proclitic (§6.1.4).

4.2.1.3 Numerals

The indigenous numerals which form a small closed set of seven members do not form compounds, either amongst their own subclass or outside of it, and have limited distribution with other categories (§7.3.1). The set of borrowed numerals do not exhibit such restrictions.

Numerals function either alone as cardinal numerals, or in classifier expressions when used attributively or pronominally (§7.3). They may also be used pronominally without a classifier.

Classifiers, both mensural and sortal, are a small closed functional class (usually derived from nouns), which are restricted to co-occurring with numerals *m?=?ikur* (o (one=CLF dog) 'one dog' (§7.3.2).

The indigenous numerals *hmpe?* 'three' and *himp?on* 'four' take the forms *pe?* 'three' and *pon* 'four' when used other than as cardinal numerals.

The class of numerals is homogeneous in relation to the expression 'to be in a group of X number of persons', however the two types of numeral employ different morphological strategies in deriving the verbal form: *hmpe?* 'three' → *p?_pe?* (BE-three) 'to be in a group of three' (§5.4.1), and *dwa?* 'two' → *b-dwa?* (HAVE-two) 'to be in a group of two' (§5.4.3).

4.2.1.4 Demonstratives

Demonstratives are a closed class of two primary members, encoding the distinction of two degrees of distance: proximate *?n?o?* 'this' and distal *ke* 'that'. Reduplication augments the number of members in the class by adding *n?o?n?o?* 'this here' and *kke* 'that there'.

Demonstratives function as deictic determiners in the NP. They have three functions: a) they point out a referent in the situational context, either things or places; b) they refer to an entity in the discourse or linguistic context, and c) they point out an intended referent, something which is in neither a) nor b), but with which the speaker judges the hearer to be familiar. These functions are discussed at length in §6.3.3.

When the demonstratives occur as the complement of a locative preposition the referent is a place rather than an object: *ha?* *n?o?* (AT this) 'here'.

There is also a set of derived demonstratives which are based on locative prepositions (§6.3.1), which may function as demonstrative pronouns *na?*.*ha?* (DEM-AT) 'this here'.

4.2.1.5 Ignoratives

Ignoratives (§6.2) are a closed class of words which express a lack of knowledge, e.g. wh- words, and indefinite pronouns. The primary function is in the formation of interrogative clauses (§10.5), where the following ontological categories are expressed: person, thing, reason, manner, place, quantity and time.

Less central notions, like 'whose', 'whither' and 'whence' are expressed by phrasal lexemes, e.g. with the possessive preposition: *do* *kadeh* (of who) 'whose', or with the directional *te h?n* (TO:unspec where) 'to where'.

Forms like *mande* 'why' and *hɔn* 'where', also function as indefinite pronouns (§6.2.1), either alone, or in compounds, e.g. *bɔrnj-hɔn* (thing-place) 'somewhere'. Other categories are expressed by generic nouns, e.g. the generic noun for 'person' *sma?* is 'anyone', and in negative existential clauses 'no-one'.

4.2.2 Verbs

Verbs are an open class. They occur as the head of a verb phrase (VP) where they determine the argument structure of the clause. They express actions (*pər* 'to fly', *tkəl* 'to cut', *rabac* 'to scratch'), events (*cəh* 'to be born', *lsəm* 'to rain') and human states (*trɔ?* 'to have fever', *?pih* 'to be ill, in pain'). There is no distinct adjectival class in Semelai, however verbs which express the following properties do form an identifiable adjectival subclass: dimension, *k̄t* 'to be small'; colour, *hijaw* 'to be green, light-blue', and physical property; *m̄en* 'to be flexible' (Kruspe 2004). Verbs also function as attributives. As an attributive they may modify either other verbs (3), or nouns (4):

- | | |
|---------------------|--------------------|
| (3) swak j̄es | (4) dak j̄es |
| walk be.swift | water be.swift |
| walk swiftly | swift water |

Verbs must first submit to a process of nominalisation in order to be eligible to function as the head of an NP, e.g. *p̄tem* 'to plant' → *p<n>t̄em* (NMZ-plant) 'planting' as in the following NP:

- | |
|--|
| (5) b- <i>j̄oy</i> tikar pandan p<n> <i>t̄em</i> solog tahan |
| MID-make mat pandanus plant<NMZ> first year |
| A mat was made from the new year's planting (of) pandanus. |

Nominalisation processes derive deverbal nouns from verbs. There are numerous forms which derive forms with specific types of referent, e.g. action nouns like *cəl* 'to pronounce → *n̄l-cəl* (NMZ-pronounce) 'pronunciation', agent/instrument nouns like *k̄t* 'to sting' → *k<m>t-k̄t* 'class of stinging insects', excessive agent nouns like *pr̄?* 'to tease' → *pr-pr̄?* (xs-tease) 'an habitual tease' (§7.6.2). Deverbal nouns retain arguments, but not verbal status, e.g. in (6) the aspectual is hosted by the locative prepositional phrase and not the deverbal noun.

- | |
|--|
| (6) ga=t̄e hɔn ji s<n>wak? |
| IMM=TO:unspec where 2 walk<NMZ> |
| Where are you going to (with) (that) walking? |

The primary subclass of verbs is based on inherent lexical transitivity: transitive verbs subcategorise for two or more core arguments, intransitive verbs subcategorise for only one (§5.2). This basic division manifests itself in a number of differences. The various morphological categories which may be marked on the verb are determined by the transitivity of the root. Transitive verbs host a pronominal proclitic, cross-referencing the subject; intransitive verbs do not generally cross-reference the subject. Transitive verbs can host the unrealis proclitic, whereas intransitive verbs do not mark the unrealis category in this manner. The semantic class of intransitive verbs which express human activities and emotions can encode

the subject with transitive pronominal proclitics (§5.3.1). Stative verbs within this particular intransitive subclass and the subclass of adjectives also feed derivational processes, which as we will see below, is atypical of stative intransitive verbs.

Transitive roots may be marked for aspect (iterative), middle voice and valency increasing and decreasing derivations, such as the imperfective and causative respectively. Intransitive verbs are more restricted in terms of availability for derivational procedures.

The two categories of transitivity type allow for further divisions. Transitive verbs can be classified according to the number of arguments, i.e. divalent or trivalent. Trivalent verbs, apart from subcategorising for three arguments, also exhibit restrictions in availability for morphological processes, e.g. detransitivising, and nominalisation processes (§5.2.1).

Both classes of verbs distinguish active and stative verbs. The distinction is important to the class of intransitive verbs. Stative verbs, including adjectives, are restricted morphologically. They cannot generally derive causative forms, or feed nominalisation processes (§5.2.2) as freely as active verbs.

Within the class of adjectives are two minor classes of particular interest: a) adjectives of dimension and b) adjectives of colour. Both classes exhibit closed membership.

a) The class of eight dimension adjectives contrasts four antonymic pairs:

size:	<i>they</i>	'be big'	<i>k̄t</i>	'be small'
length:	<i>j̄eq</i>	'be long'	<i>j̄ɛ?</i>	'be short'
height:	<i>?j̄es</i>	'be high'	<i>d̄p̄es</i>	'be low, short'
thickness:	<i>s̄e?</i>	'be thick'	<i>sey</i>	'be thin'

The term which expresses the positive dimension can be nominalised: *t̄ey* 'to be big' → *ny-t̄ey* 'size'. The term *j̄eq* 'to be long' is derived by means of coda copy, *j̄<p>l̄eq* 'length'. All eight members of the set can feed the comparative derivation *j̄ɛ?* 'to be short' → *j̄<r̄a?>ɛ?* 'to be shorter' (§5.3.8), and can also form the intensive *k̄t* 'to be small' → *k̄t-k̄t* 'to be really small' (§5.3.9).

b) The colour terms are also a finite set. Six members of this class are borrowed from Malay, the remaining two are indigenous terms.

<i>putih</i>	'to be white, silver, metallic grey' from Malay <i>putih</i> 'white'
<i>hitam</i>	'to be black' from Malay <i>hitam</i> 'black'
<i>mirah</i>	'to be red, orange, red-brown' from Malay <i>merah</i> 'red'
<i>kunig</i>	'to be yellow, gold' from Malay <i>kuning</i> 'yellow'
<i>hijaw</i>	'to be green/light blue' from Malay <i>hijau</i> 'green'
<i>biru?</i>	'to be darker blue' from Malay <i>biru</i> 'blue'
<i>jhor</i>	'to be russet'
<i>jhor</i>	'to be lighter russet'

Other indigenous colour terms belong to the class of expressives (§12).

Colour terms do not feed any derivational processes other than the collective activity *b<...>an* 'TOG' (§5.3.4). They can be nominalised syntactically with a referential meaning: *ma=mirah* (REL=be.red) 'the red one'. Lighter and darker

shades may be indicated by qualifying a colour term with either *putih* 'light', or *hitam* 'dark', e.g. *mirah putih* 'to be light-red'.

4.2.3 Expressives

Expressives are words which may stand alone as a minor clause (§12). They are iconic utterances which express sensate imagery, e.g. the aural expressive *crālāp* 'the sound of (s.o/s.th) entering the undergrowth'. They function to simultaneously provide information about both the predicate and its arguments, providing a summary in the form of a single lexical item.

Expressives are distinguished by their aberrant phonology and apparent morphological complexity, although this is not clearly understood.

The nature of this word class would suggest that there is the potential to coin new lexemes. On the basis of this they constitute an open class, see also Diffloth (1976b: 249).

4.3 Closed classes

4.3.1 Prepositions

Prepositions function as the head of a PP. They form a phrasal constituent with nominals which may be in one of a range of functions (other than a core argument of the verb), e.g. locative prepositions encode the indirect object (§9.3.1). Prepositional phrases also express subcategorised non-core complements (7), as well as the non-subcategorised circumstantial roles of location (8), direction, source, comitative/instrument and possession, (see §8.3).

- (7) ga=k1=tukar-i2 rōm hayam
IMM=3A=exchange-APPL WITH chicken
(She) is going to exchange (her) for some chickens.

- (8) dōs tōm c>? bna1
arrive SRC AT:below [name]
(They) arrived from down at *bna1*.

The locative and directional prepositions are deictically specified for three degrees of height. These are arranged in a 'person' or 'speaker' oriented system mapped in relation to the topography of the local environment.

The locative system consists of one proximate term, constituting the 'zero-point', and three distal terms (§8.5). These forms are also unusual in deriving nominalised forms which function as demonstratives (§6.3).

There are four terms in the directional system, two neutral terms distinguished for specificity of direction, and two deictically specified terms (§8.6).

4.3.2 Adverbs

Lexical adverbs provide information which has the full clause in its scope. They contain information which may relate to aspectual or circumstantial information, e.g. aspect, manner, degree or frequency.

Lexical items from other parts of speech which may function adverbially, but which are not to be accorded membership of the adverb class, are nominal adjuncts expressing time or similitude (§10.3.2), numerals and verbal adjuncts (§10.3.3).

Certain aspectual adverbs (§10.3.1) may have developed from verbs, e.g. the aspectual adverb *?areh* 'recently' and the verb *?areh* 'to be new'. The aspectual adverb is identifiable by its predicated position, failure to be predicated and inability to be modified.

Adverbs of manner express the way in which an action is performed, e.g. repetition *?pacə?* 'likewise', or restriction *?en* 'just, nothing but'. Restrictive adverbs are unique in having the ability to restrict the reference of nouns, e.g. *dak ?en* 'plain water'. The expression of manner, in terms of human characteristics is achieved by using a nominal adjunct *c<nan>lon* 'backwards', a verbal adjunct, or a serial verb like *sec* 'surreptitiously'. Adverbs of quantification express frequency or degree, e.g. *jō?* 'very, really'.

4.3.3 Auxiliaries and supplemental verbs

When some verbs are not heads of their expressions, they may have an alternative interpretation, e.g. modal auxiliaries (*lən* 'want' ← *lən* 'to desire') (§5.6) and modifying serial verbs (*sec* 'surreptitiously' ← *sec* 'to steal') (§11.3). They are, however, semantically and syntactically distinct enough to be considered categories in their own right.

Auxiliaries are identifiable by their fixed predicated position, failure to be predicated, and inability to be modified other than by the negator *da?* 'NEG'.

Modifying, or supplemental serial verbs may either precede or follow the main verb, depending on the individual lexemes (Table 11.1). Like auxiliaries, they are not predicated, however unlike auxiliaries, the modifying serial verb may host pronominal proclitics.

4.3.4 Existential and ascriptive predators

There are two non-verbal predators: a) existential *da?* 'EXIST' (§10.1.2) and b) metamorphic *sot* 'to metamorphose into' (§10.1.2). They are distinguished by their inability to subcategorise, and their inability to be modified, either aspectually or adverbially.

4.3.5 Negators

There are four negators – the internal negator *da?* 'NEG', the external negators *bəh* 'NO' and *b2en* 'NOT' and the negative imperative *boj* 'NEG:IMP' (§10.4) – plus three complex negators, which combine the internal negator with a post-verbal aspectual marker (§10.4.1).

4.3.6 Connectives

Connectives operate at both the sentential and the discourse level linking clauses together into complex sentences, or larger tracts of text into a cohesive discourse (§11.5). Connectives express cause, condition and temporal sequence. Many of these are etymologically derived from verbs, but are distinguished from verbs by virtue of

the fact that like adverbs they are not predicated, cannot be modified, and are semantically distinct enough to be considered lexemes in their own right.

4.3.7 Interjections

Interjections are a closed class of words which may stand alone as complete utterances (§13.2). They are generally monosyllabic forms, and include expressions of negative attitude, delight, bewilderment or surprise, commands and vocatives.

5 The verb

In Chapter 5 the categories that are marked morphologically on the verb are examined. Prior to discussing the functions of the morphemes, the structure of the verbal word is examined in §5.1. We then pass to a semantic classification of verbs in §5.2.

In Chapter 3 the structural processes of morphology were examined. The focus in this chapter is on the function of the derivational morphemes (§5.3 and §5.4), and verbal clitics (§5.5).

In §5.6 we diverge from categories marked morphologically on the verb, to a description of modal auxiliaries, treated systematically here on the basis of semantic category.

5.1 The structure of the verbal word

The structure of the verbal word is determined by whether the argument structure of the verb is intransitive or transitive (§5.2). Intransitive verbs are generally more simple in structural terms than transitive verbs, as there are fewer categorial distinctions available. In its simplest form the verbal word exists in the root form; at its most complex it may be host to a number of affixes and/or clitics.

Derivational affixes predominantly take the form of prefixes and infixes, but also include a suffix and a circumfix. These are immediately adjacent, located either on the left or right edge, or in the case of infixes, internal to the root. Peripheral to derivational morphemes are the proclitics. The pronominal A= (§5.5.1) and the irrealis TRR= proclitics (§5.5.2), which are mutually exclusive, are primarily associated with transitive roots, derived transitive forms, and to a limited degree with certain intransitive roots (§5.5.1). Peripheral to these is the imminent aspect clitic ga= 'IMM' (§5.5.3) which may co-occur with bases hosting derivational affixes or clitics. Clitics are only able to cliticise to verbal roots or bases. Although the clitics are formally attached to the verb, semantically they function in relation to the whole proposition.

The first set of examples, (1–3), illustrate the combinatorial possibilities for a transitive root, while the second set (4–5) illustrate the possibilities for an intransitive root.

kəŋ 'to hit'

kəŋ-kəŋ (IMPERF-hit) 'to be hitting'
tar-kəŋ (MCAUS-hit) 'to get (s.o.) to hit (s.th.)'

(1)	ki=kɔŋ cɔ 3A=hit dog He hit the dog.	(2)	ma=kɔŋ cɔ IRR=hit dog He would hit the dog.	(3)	ga=ki=kɔŋ cɔ IMM=3A=hit dog He is going to hit the dog.
<i>?yam</i> 'to cry'					
	?	yam-i? (cry-APPL) 'to cry (for s.o.)'			
(4)	kəhn ?yam 3S cry She is crying.	(5)	kəhn ga=?yam 3S IMM=cry She's going to cry.		

Figure 5.1 provides a representation of a basic verbal structure. In more complex constructions which contain a sequence of verbs, e.g. serial verbs, the clitic attaches to the initial constituent, which may not necessarily be the main verb, see §11.3.

FIGURE 5.1 THE STRUCTURE OF THE VERBAL WORD

Clitics=	+Affixes+	Verb
ga= 'IMM'	PRO= 'A'	
ma= 'IRR'	(Derivation)"	Root

5.2 Verb classification

The primary categorisation of the verb is in terms of the number of arguments for which it subcategorises. Verbs fall into two basic classes: transitive verbs (§5.2.1.1), and intransitive verbs (§5.2.1.2).¹

Semelai is a language with fixed lexical transitivity, i.e. there is a clear distinction between verbs which have two or three participants, i.e. transitive verbs, and those which have one participant, i.e. intransitive verbs. Most verbs belong strictly to one category or the other, and any divergence from the prototypical conceptualisation of the event expressed by the verb requires overt marking (§5.3). This situation obtains for the majority of verbs in the lexicon, with only a small class of specific 'minority verbs' constituting an exception (§9.1).²

The morphosyntactic consequences of this distinction in terms of participant coding are discussed in full in §9. The prototypical one-participant frame encodes the single argument in an absolute frame, while in the prototypical two-participant event the pattern of encoding of the arguments is ergative. These two frames hold also for derived verbs, the derivative adopting the frame consistent with the number of participants for which it subcategorises.

The importance of the transitive/intransitive dichotomy in Semelai is not reflected simply in the differing number of arguments, but is also associated with a distinction in *Aktionsart*, or lexical aspect. Prototypical transitive verbs are inherently bounded

¹ The following classification does not represent a complete classification of verbs in Semelai. Rather it focuses on the distinctions which are necessary for an analysis of derivational morphology. Further refinement of the classification can be found in §9 in relation to verbal argument types (subcategorisation patterns) and grammatical relations and coding strategies.

² These verbs have only one form, which does not appear to have predetermined transitivity and instead may function either transitively or intransitively with zero modification of the root (§9.1), e.g. *ca* 'to eat, intr./tr.'; *yor* 'to ascend, intr./to climb, tr.'

or perfective, in terms of the event having a conceptual initiation and endpoint, whilst prototypical intransitive verbs express events which are unbounded or imperfective, there being no reference to the point of inception or termination.

Any change in the number of participants in an event, e.g. as a result of the application of a detransitivising or transitivising process, represents not only a change in the transitivity of the event, but is also correlated with a change in *Aktionsart*.

These distinctions are imperative to an understanding of the processes of derivational morphology. Essentially, derivational morphology functions to index an alternative conceptualisation, either in terms of greater or lesser elaboration of the event.

5.2.1 Verbal sub-classes

5.2.1.1 Transitive verbs

The prototypical transitive event is one where an action proceeds from a volitional being and is directed toward an external goal. This construction is used to express all two-participant events, including those which are low in terms of semantic transitivity, e.g. non-process events like verbs of perception, cognition, speech, etc. The unifying feature of all transitive verbs appears to be the presence of two participants, where the agent is necessarily agentive, i.e. controlling or volitional. These receive transitive morphosyntactic treatment. Divalent verbs take two arguments: an agent 'A' and a patient 'O' (§9.3.1):

(6)	ki=reŋ knɛk A=V O 3A=seek husband She sought (her) husband.
-----	--

A sub-class of transitive verbs are the trivalent transitive verbs. These verbs differ in two important ways from divalent verbs:

- a) They subcategorise for three arguments: an agent 'A', a patient 'O', and a recipient 'REC' expressed as a locative PP complement (§9.3.1):

(7)	rəkɔk ki=jon he? gdo cigarette 3A=give AT:above be.old She gave the cigarettes upward to the old (man).
-----	---

- b) Trivalent verbs exhibit a general inability to feed derivational processes. This feature seems to be dictated by the fact that: a) they are inherently punctual and cannot be extended, demonstrated by their inability to co-occur with imperfective or iterative aspectual marking; and b) they cannot feed valency increasing processes, apparently because they are subcategorised for the maximum number of possible arguments. Trivalent verbs also exhibit inconsistencies in relation to the ability to feed valency decreasing processes, although the reason for this is unclear. Some trivalent verbs are: *?ur* 'to instruct'; *smap* 'to request'; *jon* 'to give'; *?ajak* 'to invite'.

Transitive roots are morphologically simple – for example, *tap* 'to weave', *pam* 'to sense', *yok* 'to take', *reŋ* 'to seek', *kɔŋ* 'to strike, hit', *sec* 'to steal', *cʰek* 'to

throw', cman 'to retell a myth', rus 'to drag', sboñ 'to save', j?oy 'to make', tojoy 'to wait for' – with these exceptions:

tanre?	'to show (s.th. to s.o.)'	(? <i>tar</i> . ³ 'CAUS' + *√ <i>re</i> ?)
pagoh	'to put (s.th.) out of reach'	(? <i>pa-</i> 'CAUS' + *√ <i>goh</i>)
trsñk	'to overhear'	(? <i>tr-</i> 'HAPP' + *√ <i>sñk</i>)
h?gi?	'to give (s.th. to s.o.)'	(? <i>Malay bahagi</i> 'to allot')

5.2.1.2 Intransitive verbs

Intransitive verbs are those which subcategorise for a single argument, typically an agent or an experiencer (§9.1). The defining characteristic of intransitive verbs is that they involve one participant, and the action or state expressed by the verb affects the participant, i.e. in the case of action verbs, the action does not proceed beyond the participant, e.g. hñm 'to bathe (oneself)'.

All intransitive verbs whether active or stative, encode the core argument of a basic clause in the same manner (§9.3.2). In (8) the verb is active, in (9) stative:

- | | |
|---------------------|---------------------|
| (8) kəhn gr̩k | (9) kəhn k̩st |
| 3S fall | 3S be.pregnant |
| She fell. | She is pregnant. |

A subset of intransitive verbs is distinguished by having two possible morphosyntactic treatments of the subject, expressing a contrast between internal versus imputed external agency (§5.5.1).

Some examples of intransitive verbs are listed below:

pl̩c	'to emerge'	pret	'to be hot, sunny'
kñmñj	'to end, finish'	lrñr	'to oversleep'
k̩bñs	'to die'	k<c>bñc	'to fish with a pole'
?yam	'to cry'	shñch	'to be lazy'
?ris	'to live, be alive'	bñbñh	'to be faint, of sound'
k̩st	'to be pregnant'	wc-wñc	'to curl, be curled up'
?yot	'to return'	tc-tuc	'to faint'

The class of intransitive verbs distinguishes a number of sub-classes, the major division being made on the emic distinction of whether the event is active or stative. Active is defined simply as an event requiring an agentive participant who initiates the action, while a stative event has a non-agentive participant. Stative intransitive verbs, including the class of adjectives, feed very few derivational processes.

Active intransitive verbs include those classified as translational motion, e.g. swak 'to walk, go', posture kp-kñp 'to lie face down', bodily processes gñk 'to smile, laugh', hy-hñy 'to yawn' and physical processes/changes like coh 'to be born', guc 'to burn'. Stative intransitive verbs express conditions, e.g. l?ñb? 'to be mentally unstable' and physical properties, e.g. cren 'to be brittle'.

As the examples listed above reveal, many intransitive roots exhibit morphological complexity in the citation form. This is particularly true amongst two classes of verbs, those which express posture and bodily processes (which are inherently reflexive verbs), and those which express states.

³ Causative /r/ → /n/ where _+rVC, e.g. /tanre?/ [tan're?].

Such morphologically complex verb lexemes appear to have the structure of a derived form. For example there is apparent evidence of coda copy in the following lexemes: s<ñ>pñh 'to be disinclined' and j?jñ? 'to be raw'.

Some verbs have an identifiable root which forms a base for other derivations, indicated by morpheme boundaries, e.g. √dñm 'lie down'. Others have no identifiable formative, indicated by an absence of segmentation, e.g. j?jñ? 'to be raw', for which no root *√jñ? is attested. Given that there is no identifiable underlying root form, the latter 'non-reversible' type would appear to be unable to constitute a base for further derivations given that the maximality condition in relation to syllable weight is already fulfilled (§3.2.1). However, some exceptions do exist, e.g. p-j?jñ? (CAUS-be dirty) 'to dirty (s.th.)', pan-cin-an (CAUS-be cooked-NMZ) 'cooking', where the restriction is quashed by using an alternative borrowed prefixing or suffixing morpheme (§3.2.3).

This distinction between forms with and without underlying roots brings us back to the active/stative distinction. Active verbs have a simple root, as shown above, and feed other derivations; stative verbs are restricted in their availability to feed further derivations. The following example shows the citation form, followed by the root (\) then the derived nominalised form:

- | | |
|----------------|--|
| (10) ki=kñl | 'to lay, be laying one's head down' |
| ← *√kñl | (lay head down) |
| → pn=kñl | (NMZ-(lay head down)) 'head-rest; place to lay one's head' |

Such verbs which have a citation form which resembles the structure of a form derived by means of coda copy, are inherently intransitive. They are not derived from a free transitive root, but a bound root, which never occurs as a citation form, and which exhibits highly restricted combinatorial possibilities, like (10) above. Some bound roots can be used in a transitive construction with zero conversion. For example, the citation form jr-jñr 'to urinate' which has a bound root √jñr (urinate), may be used transitively as in the following example:

- | | |
|---|-------|
| (11) ki=jñr | srwñl |
| 3A=urinate | pants |
| He urinated (his) pants/(in his) pants. | |

If a bound root is causativised, it derives a two – and not a three – place predicate, e.g. √jñr (urinate) → tan-jñr 'to make (s.o.) urinate'; √kñom (sit) → pa-kñom 'to set' (s.th. s.wh.).

- | | | | |
|------------------------------|-------------|-----|------|
| (12) kñej | ki=pa-kñom | ha? | 7us |
| pot | 3A=CAUS-sit | AT | fire |
| She set the pot on the fire. | | | |

Not all bound roots have the ability to be used in transitive clauses with zero-conversion, form the causative and so forth. Given the idiosyncrasies of the individual roots, the exact properties of the bound root at this stage remain unclear.⁴

⁴ The situation is further complicated by the fact that some bound roots function in nominal compounds with zero-conversion in a position usually reserved for nouns, e.g. dak jñr (water urinate) 'urine', (§3.5).

Further complex intransitive roots are discussed in the section on middle marking in §5.3.3.

5.3 Derivational morphology

The morphological processes of prefixation, infixation and suffixation are used to derive verbs from either verbal or nominal roots. The structural processes of the derivational morphemes have been discussed in Chapter 3. Here we examine the semantic and syntactic consequences of the affixes and the roots with which they combine. The affixes discussed here are presented in Table 5.1, and cross-referenced with the relevant sections in this chapter.

TABLE 5.1 DERIVATIONAL MORPHEMES

FORM				FUNCTION	REFERENCE
PREFIX	INFIX	SUFFIX	CIRCUMFIX		
+C+				Imperfective	§5.3.1
mN-				Imperfective	§5.3.2
b(r)-				Middle voice	§5.3.3
		b....-(7)an		Collective	§5.3.4
par-				Causative	§5.3.5
tar-					
p-					
	<r>				
	-?(?)i?			Applicative	§5.3.6
t(r)-				Happenstance	§5.3.7
	+r?+			Comparative	§5.3.8
LSR				Intensive	§5.3.9

Roots may host combinations of up to two affixes, either two prefixes, or a prefix and the suffix. The combinatorial possibilities are discussed in the relevant sections.

The particular function of the morphemes is dependent on the root to which it is applied. For example, while the most frequently attested function of coda copy is to convey the imperfective of transitive roots, the actual function of the process is dependent on the word class of the root, whether a transitive verb, an active intransitive verb, a stative intransitive verb, a bound root or a nominal. This range of possible functions is shown in Table 5.2 below.

TABLE 5.2 THE RANGE OF FUNCTIONS OF CODA COPY

ROOT	FUNCTION	EXAMPLE	REFERENCE
Transitive verb sec 'to steal'	Imperfective	sc-sec 'to be stealing'	§5.3.1
Active intransitive verb poh 'to grow, of people'	New lexical item	ph-poh 'to struggle, thrash about'	§5.3.1
Adjective jʌŋ 'be long'	Nominalisation	j<r>ʌŋ 'length'	§7.6.2.1
Bound root vdem (lie)	Intransitive verb	dm-dəm 'to be lying down'	§5.3.1
Nominal pos 'tail'	Stative intransitive verb	ps-pos 'to have a tail'	§5.4.1

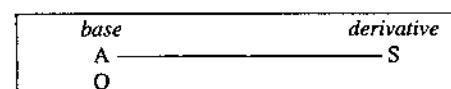
In each section below a formula is given to illustrate the comparative participant involvement of the root with that of the derived form. The root is shown on the left, labelled as the 'base', and on the right the derivative is provided. An association line indicates the relationship between the participants in the base and derived forms. An unassociated participant in the base form signals its absence in the derivative; an unassociated participant in the derived form signals its introduction. The formula for the imperfective derived from a transitive base is shown in Figure 5.2 below.

5.3.1 Imperfective +C+ 'IMPERF'

The reduplicative process of coda copy simultaneously indicates derived intransitivity and the imperfective aspect associated with intransitive events. It would be equally valid to label this a detransitivising derivation. The structural process is discussed in §3.2.3.

The application of the detransitivising process reduces a prototypical two-participant event to a derived one-participant event, in other words the transitive verb derives an intransitive equivalent as illustrated in Figure 5.2.

FIGURE 5.2 THE IMPERFECTIVE



In the imperfective, the A of the transitive verb is encoded as the S of the derived intransitive verb. Example (13) is transitive, (14) is intransitive.

- | | |
|--|--|
| (13) k ^l =go ² dlog | (14) kəhn g ² -g ² ? |
| 3A=fell tree | 3S IMPERF-fell |
| He felled the tree. | |
| He was felling. | |

As the examples reveal, there is a concomitant change in the aspectual category. Transitive verbs are associated with unitised perfective aspect (§5.2.1.1), whereas intransitive verbs have inherently unbounded imperfective aspect (§5.2.1.2). The primary function of the process therefore can be characterised as facilitating an imperfective interpretation. By deriving the imperfective, the inherent aspect of the transitive event is modified, and as a result the imperfective is always associated with an action which is in progress or incomplete. The association between lowered transitivity and imperfective aspect is noted extensively cross-linguistically, see for example Hopper and Thompson 1980. The removal of the focus from an externally directed transitive predicate results in a shift in focus to the agent's participation/involvement in the event. This in turn facilitates the use of this form to express reflexive events.

Some examples are shown here:

croh	'to find, meet'	→ c<h>roh	'to be meeting'
gɔŋ	'to carry'	→ gŋ-gɔŋ	'to be carrying'
?əm	'to lie'	→ ?m-?əm	'to be lying'
?ɔŋ	'to hide'	→ ?ŋ-?ɔŋ	'to be hiding'
rabac	'to scratch'	→ r<c>bac	'to be scratching'
k ^h e?	'to know'	→ k ² -k ^h e?	'to be knowing'

moc	'to dive into'	→	mc-moc	'to be diving'
smaŋ	'to ask for'	→	s<n>maj	'to be asking'
jŋok	'to look at'	→	ja<k>jŋok	'to be looking'

The imperfective form does not feed further verbal derivation, but does play a part in nominalisation processes (§7.6.1): gɔ? 'to fell' → g?·gɔ? 'to be felling' → n?·gɔ? 'the act of felling', tʰoh 'to spit' → th·tʰoh 'to be spitting' → nh·tʰoh 'spitoon'.

As illustrated in Table 5.2 above, bound roots derive an intransitive verb by means of coda copy, in particular verbs of posture and verbs of bodily processes: √dəm (lie) → dm-dəm (IMPERF-lie) 'to be lying down'; √kʰu? (vomit) → k?·kʰu? (IMPERF-vomit) 'to vomit'.

- (15) kəhn tc-tac
 3S IMPERF-faint
She faints.

Active intransitive verbs may feed coda copy, deriving not the imperfective form, but a verb with a new meaning: jtek 'to sleep' → j<k>tek (sleep<IMPERF>) 'to camp out, sleep away from home', e.g. while hunting, rattan collecting or preparing a new swidden.

- (16) jtek ye, jtek he? ke
camp.out I camp.out AT:above there
I camped out, camped out up there.

5.3.1.1 The function of the imperfective

The function of the imperfective is to provide aspectual information, in relation to the structure of the event. Transitive verbs were characterised above as describing an event which is viewed externally as a unitised punctual event (§5.2.1.1). In contrast, by using the derived imperfective form an event can be viewed internally as an ongoing event, in the same manner as an intransitive verb (§5.2.1.2). The imperfective describes an event in progress (17), i.e. one at which someone is occupied, as the following examples illustrate:

- (17) kmpan tp-taj
wife IMPERF-weave
The wife was weaving.

It may indicate a persistent event:

- (18) This is the title of a traditional story about a person felling trees for a new swidden.
He is cursed. No matter how long he chops, the particular tree refuses to fall.

c<nn>man sma? g?·gɔ?
retell.myth<NMZ> person IMPERF-fell.tree
The myth of the person (who was) tree-felling.

- (19) ?əŋi jm-jam lən b-knən
I IMPERF-feel want HAVE-offspring
I am feeling (I) want to give birth.

It may indicate an habitual activity:

- (20) Hamedi is a known thief.
hamedi sc-sec
[name] IMPERF-steal
Hamidi steals.

The focus on the agent's involvement is illustrated in the following examples. Speakers would characterise an activity in the imperfective as being in progress, which also highlights the focus on the agent, providing a paraphrase with tjah 'middle' as in the example in (22).

- (21) kəhn pk-pok
3S IMPERF-wash.clothes
She is washing clothes.

- (22) kəhn tjah pk-pok
3S midst IMPERF-wash.clothes
The woman is in the middle of washing clothes.

This pair was contrasted with an example of a clause containing the transitive verb form in (23) and a paraphrase with the perfective adverb loc 'already' a prior event, but also indicating the focus on the result where the verb is prefixed with the middle marker br- 'MID' (24):

- (23) ki=pok baju? knən (24) loc br-pok
3A=wash.clothes clothing child already MID-wash.clothes
She washed the child's clothing.
(The clothes are) already washed.

Some further examples of the imperfective follow:

- (25) A man and his dog are walking through the jungle when suddenly the dog stops to bark at something. The owner is unable to make him budge. He waits in anticipation for something to happen as the dog continues to bark.
jl-jəl jačə? ha? ke
IMPERF-bark likewise AT there
(The dog) was again and again barking there.

There is an implication in (26) that the action may take place in the future; the speaker just hasn't got around to starting yet, hence the presence of the negator.

- (26) ye da? j<y>pɔy
I NEG make<IMPERF>
I am not making (it yet).

In the following clause, there are two potential motivating factors for the use of the imperfective. It is used because the protagonists hide themselves, not some other object, and secondly, because the event is open-ended.

- (27) Two brothers lie in wait for the resident of the house to return.
2luc.ton de=ca, deh ?intə?, ?p-?ɔŋ karom pɔntɔŋ
after 3plA=eat 3pl exit.house IMPERF-hide under firewood
After they had eaten, they left the house, hiding (themselves) under the firewood.

Given that the imperfective typically frames other events, it is used in clauses which are background to some further consequence, e.g. to express the open-endedness of the event itself:

- (28) *Some people have finished sowing for the day, but the task itself being unfinished, the following day they will continue with the sowing.*

2luc.tom	p<m>təm	ke,	ga=ʔyot		
after	sow<IMPERF>	that	IMM=return		
yɔk	sarek,	dɔs	lagi?	ha?	d2oh
take	tomorrow	arrive	again	AT	swidden

p<m>təm
sow<IMPERF>

After sowing there, (they) are going to return (home). The next day, (they) will come again to the swidden (and continue) sowing.

Before turning to the function of the imperfective in complex constructions, the following should be noted. The imperfective is characterised by its inability to co-occur with the agentive pronominal proclitics, or with specified objects. In the event that an apparent 'object' NP occurs in the clause, there is coalescence between the verb and the O and as a result the NP lacks all features associated with a transitive object – for example, the NP cannot host proclitic hn= 'O', be fronted, or relativised (§9.3.1). In (29) the focus is on the activity, not on the effect or resultant state of any specific cassavas.

- (29) kəhn cə-cək hubi?
3S IMPERF-plant.cutting cassava
She's cassava planting.

The imperfective is primarily connected with the intransitive clause. However, in certain contexts the form may be used in a transitive clause to express the internal iterativity of a bounded event. Aspectually, the implication is that the action was repeated over a period of time, but has now been completed.

- (30) *A boy had to pull and pull at the stick before he succeeded in retrieving it from the pool.*

ki=tə<k>rek	tŋkol	dloŋ	
3A=pull<IMPERF>	length	tree	

He had been pulling (and pulling and pulling) the length of wood.

5.3.1.2 The imperfective in complex clauses

The imperfective is employed in a range of complex constructions – simultaneous actions, serial constructions and complement clauses – illustrated below. This pattern of distribution is also attested for imperfective mN- 'IMPERF', see §5.3.2.

i. *Simultaneous action* The imperfective is used to describe an ongoing background action against which another punctual action is performed. The structures in the following clauses are discussed in detail in §11.4.1 and §11.5 respectively.

- (31) kna? dɔs sma? gdo, dehn pon tŋ-təŋ
happen arrive person be.old 3pIS too IMPERF-serve.food
(It) happened (that) (when) the old man arrived, they were dishing up.

ii. *Serial constructions* The imperfective is used in purposive serial constructions and reflexive serial constructions. Purposive serial constructions with the verb swak 'to go' express an habitual cultural activity (§11.3.2): swak by-bɔy (go IMPERF-dig) 'to go digging wild tubers'; swak ʔh-ʔɔh (go IMPERF-shoot.blowpipe) 'to go hunting with a blow pipe'.

The imperfective is also used in reflexive serial constructions involving the verb goŋ 'to bring' (§11.3.2):

- (32) *A woman is fed up with her husband, so she takes herself off to the forest.*
kl=goŋ swak kloc bri
3A=take walk inside forest
She took (herself) walking in the forest.

iii. *Complement clauses* In complement clauses the imperfective form is used in complements of verbs of desire like lən 'to desire', məh 'to want', with kʰeʔ 'to know' and susah 'to be difficult'. The subject of the matrix and complement are the same in all types. The focus is on the activity, in terms of lack of knowledge (33), or difficulty in performing a task (34).

- (33) *The speaker asks the woman if she knows how to construct a dammar torch:*
ji=kʰeʔ py-pɔy?
2a=know IMPERF-make
Do you know (how) to make (them)?

- (34) susah ma=yk-yɔk
difficult IRR=IMPERF-fetch
(It) will be difficult to keep on retrieving (them).

5.3.1.3 Comparative notes

Reduplicative processes appear to be synonymous with imperfective, progressive or continuative verbal forms amongst the Aslian languages.

In Jah Hut coda copy signals 'an action in progress': ca? 'to eat' → c?ca? 'to be eating', kləŋ 'to speak' → k?ŋləŋ 'to be speaking'. These forms are characterised by a) the inability to co-occur with agentive subject prefixes, and b) their co-occurrence with non-specific objects (Diffloth 1976c: 96). These features also hold for Semelai.

In Temiar a reduplicated form involving initial and final consonant reduplication marks the continuative aspect: kɔw 'to call' → kewkɔw 'to be calling', səluh 'to shoot' → sehluh 'to be shooting' (Benjamin 1976b: 169, 171-2). The continuative form is only optionally transitive.

5.3.2 Imperfective mN- 'IMPERF'

Lexemes affixed with the Malay prefix *meN-* are attested in Semelai. The presence of the morpheme is associated with the derivation of an intransitive verb with a single argument S, which semantically is an agent (see Fig. 5.2). The derivative describes a process, focussing on the agent's participation in, or performance of, an event:

- (35) ʔyot tuhən=sən tom mn-(t)ugəl
return master=SC SRC IMPERF-dibble.rice
The master returned from dibbling rice.

The function of forms which exhibit this affix closely parallels, and in some cases overlaps with, those of coda-copy derivatives (§5.3.1), not only expressing the imperfective, but also occurring in the reflexive and purposive serial constructions (§11.3.2) and deriving intransitive verbs from nouns (§5.4.2).

The prefix *mN-* 'IMPERF' is not highly productive in Semelai. It occurs primarily on Malay loans which due to phonological restrictions cannot employ coda copy, e.g. *mm-(p)antaj* 'practise a taboo' ← *pantaj* 'taboo'. The standard means of expressing the imperfective category on roots loaned from Malay is to employ coda copy, *tujuk* 'to point' → *tu<k>juk* 'to be pointing' (§5.3.1). It is also fed by some Aslian bases which have a heavy final syllable. In most cases, the base is a causative derived from a monosyllabic root: *mm-(p)ar-sey* 'to be thinning' ← *par-sey* (CAUS-be thin) 'to thin (s.th.) out'. Causatives do not usually feed the imperfective.

The presence of *mN-* 'IMPERF' as a marginal prefix is in consonance with its reported distribution in colloquial varieties of Malay, where it is used in most varieties when the verb is active and refers to some definite process, or as a marker of durative, habitual or casual Aktionsart (Benjamin in press).

There are two types of prefixed words: those which exist only in the prefixed form, and those which have an identifiable root or base. I have reflected this in the gloss by avoiding a morpheme by morpheme gloss for those forms which do not have a root, and instead employing a colon, e.g. *mrajonj* (IMPERF:crawl) 'to crawl'. These are treated as monomorphemic in the lexicon.

Those forms, which are derived from verbs which exist independently, have been segmented and glossed as prefixed forms, even though the relationship between the root and the derivative was not always recognised by the speakers, in the same way for instance as a root and its coda-copied derivative. The morphological process is the same as that in Malay: the nasal segment assimilates to the place of articulation of the onset of the root. The onset is retained if it is a voiced segment, but is elided if voiceless. In the case of voiceless onsets the elided consonant is shown in brackets, e.g. *torəh* 'to tap rubber' → *mn-(t)orəh* 'to be rubber tapping' (§5.3.2). Affixed to sonorants, the morpheme is realised as *m-* 'IMPERF', e.g. *laŋkah* 'to step across (s.th.)' → *m-laŋkah* 'to step across'.

The class of verbs which exhibit marking with *mN-* 'IMPERF', both derived and fixed forms, express cultural events, in particular the preparation of swiddens, and physical activities, e.g. crawling, smoking, etc. Lists of some of these verbs are provided below:

<i>tbas</i>	'to slash undergrowth'	→ <i>mn-(t)abas</i> 'to be slashing' ⁵
<i>saløy</i>	'to prepare a swidden in working party'	→ <i>mp-(s)aløy</i> 'to be preparing a swidden in working party'
<i>jaga?</i>	'to guard (s.th.)'	→ <i>mn-jaga?</i> 'to be guarding'
<i>jamah</i>	'to taste'	→ <i>mn-jamah</i> 'to taste, associated with the activity of negating a desire' (< Malay <i>jamah</i> 'physical possession, particularly in relation to women')

⁵ The appearance of the /ə/ in this form can possibly be attributed to the form *tabas*, an alternative to *tebas* [təbas], particularly in areas under Minangkabau influence where schwa is realised as /ə/.

The following words do not have a recognisable root in Semelai. They belong primarily to the domain of cultural activities, related to swidden making and so forth: *mjalb?* 'to go digging for old cassava in an abandoned swidden', *mparaya?* 'to work in a party'. In Ulu Muar Malay, a non-standard variety in an area bordering the Semelai, there is a verb *mfrayo* 'gets a task done by inviting neighbours to a work party' (Hendon 1966: 75).

- (36) dehn ha? mn-(t)ugol. dom, kira?, mparaya?
3S AT IMPERF-dibble.rice AFF figure IMPERF:work.in.a.group
They thereabouts were dibbling rice, yes, I reckon preparing the swidden in a working party.

In (37) *mN-* is used in a subordinate clause.

- (37) ye ga=kmoŋ mn-(t)ulis
I IMM=finish IMPERF-write
I'm going to finish writing.

The function of this form, paralleling that of the imperfective expressed by means of coda copy (§5.3.1), has already been commented upon in the introduction. The two imperfective forms also share the feature of subject affectedness with some attestations of *b(r)-* 'MID' forms, e.g. *brlhuh* 'to run'. The correspondences between the three intransitive morphemes expressing 'subject involvement' are illustrated in the following example, where one verb hosts the prefix *b(r)-* 'MID', another the imperfective coda copy, and the third the imperfective *mN-*:

- (38) saløy, b-jamu?, g?-go?, mn-(t)abas
work.together MID-party IMPERF-fell IMPERF-slash
saløy (means) to (get together) to work in a party, fellng (and) slashing (the swidden site).

5.3.3 Middle voice *b(r)-* 'MID'

Overt marking of the middle voice is achieved by the affixation of the multifunctional prefix *b(r)-* 'MID'. The prefix *b(r)-* 'MID' belongs to the inventory of affixes borrowed from Malay.⁶ The structural specifics are discussed in §3.2.3.

Although the actual range of functions of *b(r)-* 'MID' is diverse, this variation is contained within a limited semantic range, typical of middle marking. Semantically, it is characterised here as expressing 'subject-affectedness'.⁷ It is used in the following middle situation types: passive-like decausatives, reciprocals, reflexives and intensive experience.

The lexical meaning of the verb is the primary determinant in establishing the meaning of the derived form: transitive verbs typically derive either decausative or reciprocal forms; both transitive and intransitive (non-stative) verbs derive reflexive meanings, and stative intransitive verbs typically derive an intensive experience

⁶ In Malay *ber-* has numerous functions which are equally wide ranging and elusive, but fall within the range of middle marking (Arbab 1984: 70; Liaw 1985: 118-22; Benjamin 1993: 371-85, 1997a). A brief comparative note is made at the end of each section in the ensuing discussion.

⁷ The middle voice is characterised by Lyons as 'the "action" or "state" affects the subject of the verb or his interests' (Lyons 1968: 373). See Geniusiene 1987 on reflexives and Kemmer 1993 on middle marking for recent accounts.

meaning. It is not always clear how a verb should be classed, and indeed there does appear to be an overlap in some instances.

Syntactically, middle marking is associated with intransitivity. For instance, in decausative constructions there is a single argument O, and the agent is disposed of (§5.3.3.1); whereas reciprocal constructions conflate the A and O into a single argument S (§5.3.3.2). This category differs from the 'imperfective' one-participant type event, in that the participant in the middle construction is understood as being 'affected' by the event, whereas the imperfective expresses the agent's 'performance' of the event. This can be seen most clearly in the 'passive' function of middle marking where the single participant is the O of the corresponding transitive verb:

- (39) *tʰl* *jpŋ* *br-bək*
hand foot MID-bind
(Her) hands (and) feet were bound.

Table 5.3 provides a summary of middle marking on verbal roots.

TABLE 5.3 MIDDLE MARKING

FUNCTION	VERB	SUBJECT	REFERENCE
Decausative	transitive verb	O	§5.3.3.1
Reciprocal	intensive base of active transitive verb	S	§5.3.3.2
Intensive experience	intensive base of stative intransitive verb	A	§5.3.3.3
Auto-causative reflexive	active verb	S	§5.3.3.4
Reflexive reciprocal	active verb	S	§5.3.3.5

The derivation of verbs with b(r)- from nominal roots is discussed in §5.4.3. This process derives a form expressing possession, glossed as 'HAVE', or to use 'USE' the referent of the nominal root. Often the meaning conflates the two senses, so that it means to both 'have and use'.

Finally, some apparently morphologically complex verbs in Semelai exhibit the prefix b(r)- 'MID', but lack an underived root form, indicated in the gloss by the absence of a morpheme boundary. These are termed 'non-reversible' reflexives by Geniusiene (1987: 299, 339). These verbs express inherent middle semantic meaning, e.g. non-translational motion or body movements (*bsanar* 'to lean (on s.th.)', *bnare?* 'to dance', *bsurək* 'to shout'); translational motion (e.g. *brlhuh* 'to run', *bjancket* 'to be spread (of disease)', *btuluk* 'to set off') and inherently reciprocal events (*bklahi?* 'to quarrel', *bkumpol* 'to assemble (a crowd)'). Some examples are shown here:

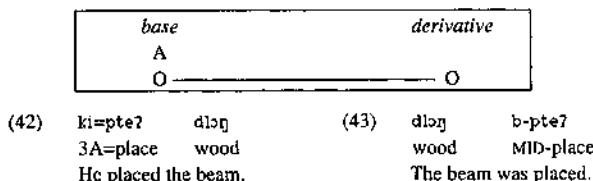
- (40) *bsanar* *?en* *tom* *dloŋ*
lean LOC bole tree
(She) leant (herself) against the bole of a tree.
- (41) *The speaker is explaining characteristics of different skin problems. This one isn't contagious.*
kul, *kudes* *da?* *bjancket*
[name] skin.condition NEG spread
kul (is a) skin condition that doesn't spread.

The following discussion begins with the most frequently occurring and transparent function, and concludes with the most difficult.

5.3.3.1 *b(r)- 'MID' and the decausative*

The affixation of b(r)- 'MID' to a transitive verb derives an intransitive verb. In Semelai it is the most frequently occurring use of b(r)- 'MID'. The O of the transitive clause (42), becomes the single argument of the intransitive clause (43). The agent is not represented in the clause. This is schematised in Figure 5.3.

FIGURE 5.3 THE DECAUSATIVE



With respect to role-marking clitics, the single argument continues to be coded as an object (§9.3.1.4):

- (44) *b-tajam* *hn=dloŋ*
MID-plane O=wood
The wood is planed.

This process is labelled the decausative construction, because the causative component is subtracted from the event schema (cf. Geniusiene 1987: 97ff.). The low elaboration of events results in a non-specific, or non-individuated agent. The event is understood to result from the action of an autonomous agent (i.e. it is a non-autonomous event) but the agent is considered unimportant in the conceptualisation of the event, and is never represented in this clause type. The derived form functions in effect as a resultative quasi- or medio-passive.

This decausative derivation occurs with transitive verbs, either transitive roots or derived transitive forms, which contain a causative component. Some examples are listed here:

<i>kʰə?</i>	'to know'	→ <i>br-kʰə?</i>	'to be known'
<i>ptəm</i>	'to sow, plant'	→ <i>b-ptəm</i>	'to be planted'
<i>jon</i>	'to give'	→ <i>br-jon</i>	'to be given'
<i>tanre?</i>	'to show'	→ <i>b-tanre?</i>	'to be shown'
<i>?yəŋ</i>	'to hear'	→ <i>br-yəŋ</i>	'to be heard'
<i>?ye</i>	'to see'	→ <i>br-ye</i>	'to be seen'
<i>campor</i>	'to mix, combine'	→ <i>b-campor</i>	'to be mixed'

The decausative is used in a range of contexts, where the verbal action is viewed in relation to the resultant state of the referent. (Non-resultative or potential passives are expressed by the irrealis proclitic *ma-* 'IRR' (§10.2).) It may indicate the outcome of a specific event as in (45) and (46), or be used in universal type statements to provide a definition (47), or describe a procedure or characteristic behaviour (48).

- (45) *roti?* *da?* *da?* *wə?,* *kməŋ* *br-ca*
bread NEG EXIST longer finish MID-eat
(There's) no more bread, (it's) all been eaten.

- (46) da⁷ br-yok, da² br-k^bm
NEG MID-fetch NEG MID-get
(You) don't fetch (it), (you) won't get (it).
- (47) cēp, cēp, cēp, swara? bras br-gum
winnow.rice.inexperienced! sound husked.rice MID-winnow
'cēp cēp cēp' (is) the sound (of) husked rice (being) winnowed (by someone inexperienced).
- (48) br-yəŋ t̪nay.doy
MID-hear previous.afternoon
(They) are heard (at) dusk.

b(r)- 'MID' is also used to convert transitive verbs into attributive modifiers:

- (49) t̪i jog br-bək, ki=təh
hand foot MID-bind 3A=untie
(Her) bound hands (and) feet, he untied.
- (50) ki=jyok ?ate br-bəy
3A=look.at earth MID-dig
He looked at the dug up earth/the earth which had been dug up.

This form is also commonly used to make polite enquiries:

- (51) br-lən, beh?
MID-want NO
(Do you) want (it), (or) not? (Lit. Is (it) wanted, (or) not?)
- (52) A: br-k^bm k<c>bəc?
MID-succeed fish.with.pole<IMPERF>
(Any) fish caught?
- B: br-k^bm
MID-succeed
(Fish) were caught.

It is important to note that this construction is concerned neither with the deliberate suppression of the agent's identity, nor lack of knowledge of the identity of the agent. In such cases a clause like that in (53) containing the unidentified agent proclitic ko= '3UA' would be used instead.

- (53) ko=bəy ?ate
3UA=dig earth
Someone has dug up the earth.

Interestingly, this particular function of the middle marker is considered one of the more peripheral of the middle domains in Kemmer's typological survey of this category (1993: Ch. 4), however Geniusiene (1987) in her typological survey of reflexives includes this as a well-attested category cross-linguistically. The prevalence of this construction may well be due to the fact that there is no actual passive or resultative construction in Semelai which can be used to focus on the endpoint of the event. (The 'passive' function of b(r)- 'MID' is attested in Malay, e.g. *tulis* 'to write → ber-tulis 'written', *jahit* 'to sew' → ber-jahit 'sewn' (Arbak 1984: 75).)

5.3.3.2 b- 'MID' and the reciprocal

Prefixed to the intensive form of a transitive verb, b- 'MID' conveys reciprocal action. The intensive (INTNS) expresses the intensity of the performance of an action (§5.3.9).

Reciprocal action is expressed as a one-participant event, the agent and experiencer represented as a single entity by a plural subject.

- tuy 'to dig (s.o.) in the ribs'
→ b-tu-tuy 'to dig each other in the ribs'
- tmap 'to smile (at s.o.) in acknowledgment on meeting'
→ b-tmu-tmap 'to smile to one another in acknowledgment on meeting'
- (54) "he b-reg-reg ci" k^bəŋ
1&2 MID-INTNS-seek lice QUOTE
"We'll look for each others' lice," (he) said.
- (55) Two men speak back and forth to each other.
b-cəkə-cəkəp rəm twanko? ke
MID-INTNS-converse WITH prince that
He conversed with the prince.

Reciprocal actions may also be derived by prefixing the middle marker to a reduplicated nominal (§5.4.3).

Comparative note The standard Malay formation of the reciprocal is *ber-RDP-V-an*, *ber-sambut-sambut-an* 'exchange visits' (§5.3.4). However, *be-V* without reduplication, is attested in some local varieties, reciprocity being implied through the semantics of the verb, e.g. *bertikam* (*ber* + *stab*) 'to stab (each other)'.

5.3.3.3 b(r)- 'MID' and intensive stative verbs

The partial reduplication of stative verbs is associated with intensity of the experience (§5.3.9). When b(r)- 'MID' is prefixed to an intensive stative base, the derivation highlights or emphasises the expressed state.

In the first example, the drummers at an exorcism have been beating their drums, taking turns for seven days. Consequently, the result, expressed by the middle form, is swollen fingers. This would also account for example (57) where the wife is buried by her husband while she is still very much alive.

- (56) sma? pok rbana? sampay b-?b-?bs j<2>re? t̪i
person beat drum until MID-INTNS-be.swollen digit<HAVE> hand
The people beat the drums until their fingers were swollen.
- (57) ki=kəm b-ri-?ris
3A=bury MID-INTNS-be.alive
He buried (her) alive.
- (58) Some children bury their deceased mother. They give her a proper funeral despite the fact that she hadn't been a normal human, only a head.
de=jyoy məcəm b-cuku-cukup t̪i, jəŋ, kba?
3plA=do like MID-INTNS-be.enough hand foot body
They did (a burial) as if (she had been) complete, hands, feet and body.

5.3.3.4 Auto-causative reflexive *b(r)- 'MID'*

The affixation of *b(r)- 'MID'* may derive an auto-causative, or self-causative reflexive (Geniusiene 1987). This construction always has a singular, animate subject who is simultaneously both the agent and experiencer. The semantic roles of A (agent/experiencer) and O (patient), realised as the intransitive subject S, are co-referential. Verbs which mark this relationship with *b r- 'MID'* are relatively infrequent, the auto-causative usually being expressed covertly in Semelai by imperfective morphology, rather than distinctive reflexive marking.⁸

The unaffixed verb in (59) focuses on the transitive event; in (60) it is an auto-causative reflexive.

- (59) ki=kr ja?
3A=work
She worked (at s.th.).

- (60) km̥ən da? ki=jon b-kr ja?
wife NEG 3A=allow MID-work
He didn't allow (his) wife to (experience) work (herself).

This form is also used with intransitive verbs as in the following example:

- (61) jtek b-swak
sleep MID-walk
(to) sleep walk

Verbs derived from nominals may also be used with an auto-causative meaning (§5.4.3):

- (62) k1=jɔy b-ksum, mɔɔdm ksum klkbk
3A=make HAVE-cooon like cocoon butterfly
He made (himself) cocooned, like a butterfly's cocoon.

5.3.3.5 Reflexive reciprocals

Middle marking is used with inherently reciprocal activities which are also auto-causative reflexives.

In (63)–(64) the verbs of meeting (e.g. *croh rom* 'to meet with (s.o.)'), and separation (e.g. *cray rom* 'to separate from (s.o.)'), are inherently reciprocal.

- (63) ?inat da? ga=b-croh rom knɛk
assume NEG IMM=MID-meet WITH husband
(He) assumed the new husband wasn't going to be encountered (by himself).

- (64) ?areh b-crav rom habu?
just MID-separate WITH kitchen
Just separated (myself) from the kitchen. (Lit. (I) just came from the kitchen.)

Example (65) cannot have a purely reciprocal interpretation given the subject is singular. Someone sleeps restlessly, tossing to and fro, indicated by the reduplication of the verb *guls?* 'to roll over'.

⁸ Inherently transitive verbs used as auto-causative reflexive body moves mark the co-reference by filling the O position with a 'desemanticised' body noun, either *kbo?* 'body', *səc* 'flesh', or less frequently *dri?* '1self' as illustrated here: *ki=rus kbo?* (3A=drag body) 'He dragged himself'.

- (65) dm-dəm kəh, b-gule-gule?
IMPERF-lie 3 MID-INTNS-roll.over
He lay, tossing himself to (and) fro.

5.3.4 Collective *br<>an 'TOG'*

The circumfix *br<>an 'TOG'* (see §3.2.3.3) expresses a simultaneous activity or state of two or more participants who are perceived as forming a group. The circumfix, used with both active and stative intransitive verbs, has a modifying function.

The affix is a loan from Malay, where it is considered to be a discontinuous combination of the prefix *ber-* and the nominalising suffix *-an* (Abdullah 1974: 44). It expresses mutuality or generality, or the reciprocal if the verb is reduplicated, e.g. *ber-tumbuk-tumbuk-an* (*ber-punch-punch-an*) 'to punch each other'. The reciprocal form is infrequent in modern Malay.

Kermmer reports the widespread cross-linguistic distribution of polysemy between the reciprocal and collective functions (1993: 98). This generally attested similarity between the collective and reciprocal forms would appear to be reflected in the relationship between the reciprocal in Malay and the collective in Semelai.

It is important to note that reciprocity is never implied in Semelai usage, and although the participants may be involved in the same event, it does not express cooperative actions, but the portrayal of a perceived group. The context referred to may simply be the result of pure coincidence as shown in the first two examples:

- (66) A bystander remarks on the children who are all lying asleep on the verandah:
knən b<jtek>an
child TOG<sleep>
Children all asleep together.

- (67) A group of people are described as being dark-skinned:
b<hitam>an
TOG<be.dark>
(people) all dark (skinned)

In contrast, in the following example, the circumcision ceremony is an organised event, and the participants will be companions in the ceremony:

- (68) sma? ga=b<sunt>an
person IMM=TOG<circumcise>
People are going to be circumcised together.

The following examples provide some context, the first with an active verb *zyam* 'to cry', the second stative *c2en* 'to be happy'.

- (69) A traumatised family return home:
knən pon dos ha? dəl b<zyam>an,
child THEN arrive AT house TOG<cry>
bapa?=hn zyam, knən pon zyam
father=3POSS cry child too cry

The children then arrived at the house, crying together (with their father). Their father was crying, the children were crying too.

- (70) kira?, b<c?en>an, sparoh bsuruk deh=son, rpok t^{hi}
 figure TOG<happy> some shout 3pl=SC clap hand
 (I guess you could say they) were all being happy, some were shouting, see, (and)
 clapping (their) hands.

5.3.5 The causative

There are four causativising affixes: par-, tar-, <r> and p-. The morphological specifics of the first two are discussed in §3.2.2.1, the third in §3.2.2.3 and the last in §3.2.3.1.

The causative prefix attaches to verbal roots and nominal roots. The derivation of nominal roots is discussed in §5.4.4. Causative bases derived from verbal roots feed derivations of the iterative -i? 'ITER' (§5.3.6) and happenstance t(r)- 'HAPP', e.g. pr-cēc 'to wet' → t-pr-cēc 'to happen to splash' (§5.3.7). Causatives derived from nominals do not feed further verbal derivations, e.g. *t-pa-dak (HAPP-EQUIP-water) 'to happen to wet' ← pa-dak (HAPP-EQUIP-water) 'to wet' is unacceptable (§5.4.4).

The canonical causative construction expresses unity of the relationship between two events, where a consequence results from a prior event. Causatives can be distinguished in terms of the degree of involvement of the participants, and the directness of the cause.

Before launching into the discussion of the causative construction, it should be noted that this is a highly restricted domain. Firstly, causers, i.e. agents, must be animate beings, usually humans, anthropomorphised animals or meteorological entities. The pronominal proclitics can only have such entities as their referents (§5.5 and §6.1.1). Therefore, in a causative construction only animate beings can be the causer.

- (71) p?re? par-pih sma?
 ghost.violent.death CAUS-be.ill person
 The ghosts (of those who have died a violent death) make people ill.

Indirect causation, where either the causing entity is inanimate, or there is a lack of temporal unity, is expressed peripherastically by an adjunct clause (see (73) below, and §11.5.3).

Secondly, the degree or force of causation must be understood against a society where personal autonomy is paramount at all levels of personal interaction. At most, causation merely represents the assistance in, or the facilitation of an event, where the causee is an animate being, based on the causee's willingness to participate. This excludes the derivation of such events as 'making someone be happy/work' etc. For instance, it is not possible to derive causative forms, 'to make s.o. V', from verbs expressing bodily activities: ?yam 'to cry', pam 'to feel', h̄doy 'yawn', ḡp 'to bite (of animals)', c̄rt 'to bite (of humans)', are all such verbs. Nor is it possible with verbs expressing states or conditions of human propensity: trɔ? 'to have fever/be feverish', ḡren 'to be angry', risaw 'to be sad', l̄lb? 'to be unsound of the mind'. In general, stative verbs, including adjectives, do not feed causativisation (see Kruspe 2004). An analytic, or syntactic construction is used to express causation of ineligible verbs, with the source of causation expressed externally in a causal clause introduced by la= 'BCS' (§11.5.3.1):

- (72) ki=c?en [la=bapa? 2yot]
 3A=content [BCS=father return]
 She was happy because her father returned.

The following list of roots fails to feed the causative derivation, verbs of bodily activity, verbs of displacement and transfer, including ditransitive verbs, and verbs of social interaction:

Body action verbs	Verbs of transfer	Social interaction
jal 'to bark'	k̄bm 'to get, obtain'	?ajak 'to invite'
?puk 'to suck'	sec 'to steal'	smaji 'to request'
?yam 'to cry'	pay 'to set food aside'	tolon 'to help'
pam 'to sense, feel'	jon 'to give'	
soc 'to whistle'	paloh 'to flee'	
glbñ 'to swallow'	pagoh 'to move (s.th.) aside'	
ḡp 'to bite'		

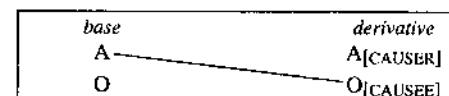
The syntactic consequence of causativisation is an increase in the valency of the predicate, which forms an important part of the following discussion. Interestingly, in a small number of cases application of the process apparently fails to increase valency, and the result is a reflexive causative where a one-participant event is coded as a prototypical two-participant event. This is discussed in §5.3.5.2. We will begin with a characterisation of the valency changing properties of the causative in §5.3.5.1, before moving to a discussion of the individual morphemes in §5.3.5.2.

5.3.5.1 Valency change

In the causative construction the causer is encoded as the subject, and in both intransitive and transitive clauses, this is the introduced (i.e., the new) argument in the construction. It is always the causee which must change its syntactic position (Comrie 1985: 336), as shown in Figures 5.4 and 5.5. The usurped agency of the causee is expressed by encoding the argument as the O in a derived two-participant event, or as the OBL in a derived three-participant event.

The argument frame of the derived two-participant event is schematised in Figure 5.4.

FIGURE 5.4 THE TWO PLACE CAUSATIVE CONSTRUCTION



The participants of the derivative are encoded as a prototypical two-participant event, where the introduced agent is cross-referenced by the agentive proclitic, see §9:

- | | |
|--|---|
| (73) knon h̄m
offspring bathe
The child bathed. | (74) knon ki=pa-h̄m ia=2ma?
offspring 3A=CAUS-bathe A=mother
The mother bathed (her) child. |
| (75) cin creh
be.cooked fish
The fish is cooked. | (76) ki=pan-cin creh
3A=CAUS-be.cooked fish
She cooked the fish. |

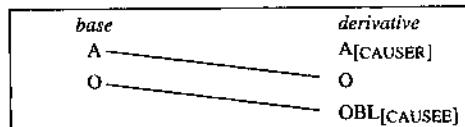
An anomalous pattern is found with the verb *susah* 'to be difficult', which encodes the subject as a locative complement; this remains as a locative complement in the corresponding causative and an A is introduced:

- (77) susah ʔen sma?
be.difficult LOC people
(It) is difficult for people.

- (78) ki=p-susah ʔen sma?
3A=CAUS-be.difficult LOC people
He causes difficulty for people.

The argument frame of the verb derived from a root expressing a two-participant event is a three-participant event which is schematised in Figure 5.5.

FIGURE 5.5 THE THREE PLACE CAUSATIVE CONSTRUCTION



The participants of the derivative are encoded as a prototypical three-participant event, where the introduced agent is cross-referenced by the agentive pronominal proclitic (§9). The oblique, or indirect object, is coded either by the neutral locative preposition ʔen 'LOC' (§8.4.2), or by one of the deictic locative prepositions (§8.5.2), adopting the argument frame of trivalent verbs (§9.3.1).

The following pairs of examples compare the underived with the derived form:

- (79) de=jɔh hn=təh
3plA=drink O=tea
They drank the tea.

- (80) ye=j<r>?ɔh hn=təh hə? hn=dəh
1A=drink<CAUS> O=tea AT ABS=3pl
I gave tea to them to drink.

- (81) ki=?yəŋ la=bapa?
3A=hear A=father
The father heard (it).

In example (82), the teacher reports to the father using *he?* 'AT:above', signifying that the father is located geographically higher than the speaker:

- (82) ki=par-yəŋ la=cɪ?gu he? bapa?
3A=CAUS-hear A=teacher AT:above father
The teacher informed the father (of it).

There is no causative derivation from three-participant events. This is possibly based on the fact that there is no typical four-participant event structure available in the language.

5.3.5.2 The semantics of the causative morphemes

The affixes *par-* 'CAUS', *tar-* 'MCAUS', and <*r*> 'CAUS' are highly productive, although *tar-* 'MCAUS' does not co-occur with nominal roots. *p-* 'CAUS' is the least frequent of the morphemes, and is particular to the class of roots least available to causativisation: stative intransitive verbs and disyllabic loans, both verbal and nominal.

The distinction between the two morphemes available to monosyllabic roots, *par-* 'CAUS' and *tar-* 'MCAUS', is of semantic consequence as indicated by the gloss. While both morphemes express direct causation, *tar-* 'MCAUS' suggests a more assistive 'hands on' involvement, and is glossed 'MCAUS' accordingly. This point will be expanded on in the following discussion in ii) below. The distinction in the type of causation as expressed with monosyllables is not available to disyllables, where only the one morpheme <*r*> or *p-* is available, the choice determined by structural factors. This is summarised in Table 5.4.

TABLE 5.4 DISTRIBUTION OF CAUSATIVE AFFIXES

ROOT		AFFIX			
		par-	tar-	< <i>r</i> >	<i>p-</i>
monosyllabic	verb	+	+		
	noun	+			
disyllabic	verb			+	+
	noun			+	+

i. *par-* 'CAUS' The function of this prefix can be characterised as 'facilitative'. Derivatives with *par-* 'CAUS' indicate a situation where the agent facilitates the event, but does not retain control over the outcome. The inability to manipulate the outcome to the intended result is in part due to the requirement of cooperation from the causee.

Roots which feed this derivation include transitive and intransitive roots, both active and stative. Recall that this is also the morpheme which is used to derive verbs from monosyllabic nominal roots, not *tar-* 'MCAUS' as we will see below. Overall there appears to be a correlation between low semantic transitivity of the root, and low causative force, or reduced agent causation.

The roots which feed derivations with *par-* 'CAUS' are comprehensively listed below based on the present corpus of data. The verbs primarily express body actions or states.

ca	'to eat, intr./tr.'	→ pan-ca	'to feed (s.o.)'
tut	'to blow, tr.'	→ par-tut	'to cause (s.th.) to blow'
cəl	'to pronounce, utter'	→ par-cəl	'to cause (s.o.) to utter'
?ləm	'to be pleasant'	→ pa-ləm	'to cheer up, settle (s.o.)'
?yəŋ	'to hear, listen'	→ par-yəŋ	'to inform (s.o.) of (s.th.)'
tɔh	'to spit'	→ par-tɔh	'to make (s.o.) spit'
(83)			
ki=pa-ləm	gres	knən	
3A=CAUS-be.pleasant	liver	offspring	
She settled (her) baby.			
?pih	'to be ill'	→ par-pih	'to make ill'
?ris	'to be alive'	→ pan-ris	'to raise (s.o.)'

poŋ?	'to suckle, intr.'	→ pa-ŋ?	'to suckle, tr.'
ʔlat	'to extinguish, intr.'	→ p'ar-lat	'to extinguish, tr.'
cəc	'to splash, intr.'	→ pr-cəc	'to splash, spray, tr.' ⁹
təŋ	'to dish out (food)'	→ par-təŋ	'to make (s.o.) dish out (food)'
hūm	'to bathe, intr.'	→ pa-hūm	'to bathe, tr.'
rus	'to drag, tr.'	→ pa-rus	'to make (s.o.) drag (sth.)'
ce?	'to be lost, intr.'	→ par-ce?	'to get (s.th.) lost'
(84)	ki=pr-cəc	seitok	
	3A=CAUS-splash	insecticide	
	He sprayed the insecticide.		

The following body action derivatives are based on bound roots, discussed in §5.2.1.2.

√ʔaw	(stand)	→ paŋ-ʔaw	'to stand (s.th.) up'
√woh	(rise)	→ paŋ-woh	'to make (s.th.) upright'
√kʰu?	(vomit)	→ par-kʰu?	'to cause (s.o.) to vomit'
√dəm	(to lay)	→ pa-dəm	'to lay (s.o./s.th.) down'
√kʰom	(to sit)	→ pa-kʰom	'to sit (s.o./s.th.) up/down'
(85)	kdeŋ	ki=pa-kʰom	ha?
	pot	3A=CAUS-sit	AT
	She set the pot on the fire.		

ii. tar- 'MCAUS' Derivatives with tar- 'MCAUS' can be characterised as mediated causation. tar- 'MCAUS' derivatives suggest a direct 'hands on' physical manipulation by the agent, in the sense of assisting or directing someone to achieve an outcome, but not exercising control over them.

Roots which select this form are typically transitive process verbs. There is a correlation between high transitivity of the predicate, and higher, or sustained causative force of the causer. Body action verbs which are derived with this morpheme indicate direct physical manipulation, e.g. pit 'to shut (one's eyes)' → tar-pit 'to shut (s.o.'s eyes)', like the eyes of a corpse, for instance, or √jor 'urinate' → tar-jor 'to make (s.o.) urinate', i.e. to get a child to urinate before bed by kneading the lower abdomen.

Verbs which typically feed derivations with this prefix are listed below. The list is not exhaustive.

koh	'to come undone'	→ tar-koh	'to take off'
c'ek	'to throw'	→ tar-c'ek	'to get (s.o.) to throw'
bb?	'to carry (child) in a sling'	→ tar-bb?	'to get (s.o.) to carry (child) in a sling'
tap	'to weave'	→ tar-tap	'to get (s.o.) to weave'
c'br	'to fire'	→ tar-c'br	'to get (s.o.) to fire (s.th.)'
geh	'to scratch'	→ tar-geh	'to get (s.o.) to scratch'
wes	'to peel'	→ tar-wes	'to get (s.o.) to peel (s.th.)'
cob	'to be born'	→ tar-cob	'to deliver (s.o.)'
wen	'to throw away'	→ tar-wen	'to make (s.o.) throw (s.th.) away'
per	'to fly'	→ tar-per	'to get (s.th.) to fly'

⁹ pr- 'CAUS' is an irregular form of the affix, and occurs only in this one form.

?pat	'to be ugly'	→ tar-pat	'to disfigure (s.th.)'
gos	'to peel'	→ tar-gos	'to make (s.o.) peel (s.th.)'
pok	'to wash clothes'	→ tar-pok	'to get (s.o.) to wash clothes'
hop	'to put food in the mouth'	→ tar-hop	'to put food in (s.o.'s) mouth'
ʔoh	'to blowpipe'	→ tar-ʔoh	'to get (s.o.) to shoot'
kɔŋ	'to hit'	→ tar-kɔŋ	'to get (s.o.) to hit'
kəm	'to bury'	→ tar-kəm	'to get (s.o.) to bury'
gəl	'to float'	→ tar-gəl	'to get (s.o.) to float'
?ur	'to instruct'	→ tar-?ur	'to get (s.o.) to instruct'
sar	'to descend'	→ tar-sar	'to place (s.th.) lower'
p'ul	'to come down'	→ tam-p'ul	'to place (s.th.) down'
ʔyot	'to return'	→ tar-ʔyot	'to send (s.th.) back'
dɔs	'to come, arrive'	→ tar-dɔs	'to make (s.th.) arrive'
moc	'to dive'	→ tar-moc	'to dunk (s.o.)'

With a verb of motion, the causer enables the movement of the causee, although the causee does not necessarily perform the action. In (86) the verb dɔs 'to come, arrive' is used in the underived form.

(86)	ʔanda?	dɔs	tom	co?	bna!
fourth.born.male arrive SRC AT:below [top]					

In the derivative in (87), the causer acts directly on the causee facilitating the event.

(87)	creh	ye	ki=tar-dɔs	la=ʔanda?	tsm	co?	bna!
fish 1 3A=CAUS-arrive A=fourth.born.male SRC AT:below [top]							

iii. Variable prefixation The following roots may take either form of the prefix par- 'CAUS' or tar- 'MCAUS'. These are the only roots attested so far:

ʔluc	'to pass'
p'ar-luc	'to release o.self'
tar-luc	'to instruct to release (s.o.)'
√yup	(sleep with cover)
par-yup	'to cover o.self'
tar-yup	'to instruct to cover (s.o.)'
√ʔaw	(stand upright)
paŋ-ʔaw	'to stand (s.th.) up, get (s.th.) to stand up o.self'
tar-ʔaw	'to instruct to raise, make upright (e.g. fibre when weaving)'
√sot	(wear sarong)
par-sot	'to dress in a sarong o.self'
tar-sot	'to dress (s.o.) in a sarong'

The following two examples contrast the distinction:

(88)	ki=p'ar-luc	(89)	"ʔec beh!"	k'laŋ,	"tar-luc	ʔəŋ!"
3A=CAUS-pass eh NO QUOTE CAUS-pass If						

She released (herself).

"Ah no!" (she) said, "Release me!"

The form derived with par- 'CAUS' is used in the context where one covers oneself, or covers someone else, but where the agent is self-motivated:

- (90) ki=par-yup knon
3A=CAUS-sleep.cover offspring
She covered (her) child.

The morpheme tar- 'MCAUS' is used to mediate causation, to instruct someone else to perform the event:

- (91) ki=?ur tar-yup knon
3A=instruct CAUS-sleep.cover offspring
She instructed (her) to cover (her) child.

iv. <r.> and p- 'CAUS' The remaining causative morphemes, <r.> and p- 'CAUS', are purely structurally conditioned. This is illustrated by the following alternants, which are phonologically rather than semantically conditioned.

ddes	'to be near'	→ par-des	'to move (s.th.) near'
dk ^h es	'to be near'	→ d<r>k ^h es	'to move (s.th.) near'

A similar situation is found with p- and <r.> in the following case with jreh 'to be weary': p-jreh 'to weary (o.s)' and j<r>jreh 'to weary (o.s)'.

A list of derived forms, followed by some examples, is given below:

- | | | | |
|---------|----------------------|------------|---------------------------------------|
| swak | 'to go' | → s<r>wak | 'to move (s.th.)' |
| ghop | 'to be hot' | → g<r>hop | 'to heat (s.th.)' |
| grək | 'to fall' | → g<n>rək | 'to make (s.th.) fall, drop' |
| tarek | 'to pull' | → ta<n>rek | 'to make pull' |
| clew | 'to bathe (av.)' | → c<r>lew | 'to wet (the throat), drink alcohol' |
| j?ji? | 'to be dirty' | → p-ja?ji? | 'to dirty (s.th.)' |
| susah | 'to be difficult' | → p-susah | 'to cause trouble (?en LOC 'fors.o.') |
| hamponj | 'to be light-weight' | → ha<r>pɔŋ | 'to lighten' |
- (92) de=b<r>tol praho?
3plA=be.correct<CAUS> boat
They righted the boat.
- (93) cɔŋ ki=ma<r>suk kloc kampit
mousedeer 3A=enter<CAUS> inside sack
He got the mousedeer to enter inside the sack.
- (94) cɔ ha?, ki=?ur bək k<r>mɔŋ
dog AT 3A=instruct tie finish<CAUS>
The dogs thereabouts, he instructed (them) to tie (them) all up.

Example (95) is a negative transitive imperative form:

- (95) boy ma=p-jreh kba?!
- NEG:IMP IRR=CAUS-be.weary body
Don't wear yourself out!

v. The cooperative auto-causative Certain verbs have a causative derivative, termed the 'cooperative auto-causative', which retains the same set of arguments as

the underived root. The auto-causative shares the same morphology as other causative derivations, either <r.> or p- 'CAUS' (all attested roots are disyllabic), but differs in two respects: a) the affixation fails to alter the valency of the clause, although it may alter the coding of arguments (Figure 5.6), and b) it has an added semantic component, the auto-causative, where the agent is simultaneously both causer and causee, and acts according to a cooperative social principle. The argument structure is examined briefly, before turning to the semantics of this construction.

The failure of the process to increase the number of participants in the expressed event is unlike usual causative constructions (cf. Figures 5.4 and 5.5 above), as illustrated in the following examples. One such verb is the intransitive verb siyah 'to call in, break a journey'. Examples (96) and (97) of the underived verb, provide a comparison for (98) where the derivative is used. In (96), the speaker makes an invitation.

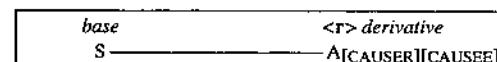
- (96) siyah ha? doi ye!
break.journey AT house 1
Break (your) journey at my house!

- (97) ?yot tom b-layar, siyah ha? doi ?ma?bapa?
return SRC MID-sail break.journey AT house parents
Returning from sailing, (he) broke (his) journey at (his) parents' house.

In (98), where the verb is in the causative form, there is a change in the coding structure of the arguments (Figure 5.6), accompanied by a change in meaning.

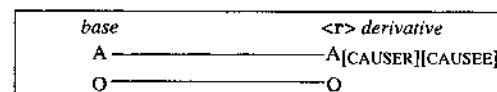
- (98) de=si<r>jah ha? ke
3plA=break.journey<CAUS> AT there
They thought to break (their) journey there.

FIGURE 5.6 THE ONE PARTICIPANT AUTO-CAUSATIVE



The majority of verbs which occur in this construction are transitive. The event schema for the derivation of a transitive root is illustrated in Figure 5.7, and in examples (99) and (100). Here there is no change in argument coding, but there is however, a change in meaning.

FIGURE 5.7 THE TWO PARTICIPANT AUTO-CAUSATIVE



Sentence (99) is an example of the underived root, (100) shows the derivative.

- (99) de=porok ha? ?amɔŋ
3plA=put AT basket
They put (the fruit) in the basket.

In (100), an imperative clause, typical of this clause type, the agent is not coded on the verb.

- (100) *The group are out collectively gathering fruit. One of them gives an instruction:*

"jl,	po<n>rök	ha?	no?"
2	put<CAUS>	AT	this
"You, (mind and) put (them) in here!"			

A disparate group of verbs is involved in this derivation:

slet	'to insert'	→ s<r>let	'to insert'
porok	'to put'	→ p<n>rök	'to put'
sban	'to save, tr.'	→ s<r>bən	'to cause (s.th.) to be saved'
kirim	'to send (s.th.) to (s.o.)'	→ ki<n>rim	'to make (o.s.) send (s.th.) to (s.o.)'
cmot	'to act to avoid transgression'	→ c<r>mot	'to act to negate desire (on behalf of s.o.)'
cpat	'to be quick'	→ c<r>pat	'to make (o.s.) quick'
sijah	'to call in'	→ si<r>jah	'to make (o.s.) call in'

The auto-causative construction is used in expressing cooperative social interaction, of what are necessarily externally oriented actions. This extension follows when the type of possible causative events in Semelai are taken into account. Recall that causatives are associated with a low level of control where both participants are animate, so that the prototypical causative event is a cooperative one between two individuals (see introduction to §5.3.5). I characterise the construction as expressing 'to take it upon oneself to do V for someone else's/cooperative benefit'.

A clear example can be seen in (101). The context in which this statement is made must be explained. cmot is a symbolic act of partaking in the consumption of food (e.g. a pot of food on a stove, or a meal someone is eating) or potential food source (e.g. a carcass brought in from a hunt), in order to negate one's own desire. In (102), the speaker's stepson, acting of his own accord, dips his finger in a pot of food and wipes it on the speaker's ankle. The causer performs the action on behalf of some one else; the causee or beneficiary is entirely passive in relation to the event. I have included the expression 'thought to' in the translations in an attempt to express the 'forethought' or 'mindfulness' of the participant in executing the action.

The overt expression of the beneficiary is optional. In (100) and (102) it is represented, in all other examples it is not.

- (101) ki=c<r>mot! ye
3A=act.to.negate.desire<CAUS>!
He thought to negate the desire (for) me. (Lit. He caused the negation of desire (for) me.)

In the first member of the pair of examples in (102) and (103), the narrator relates the stranger's actions on his arrival at the house.

- (102) ki=slet wəy=hn ʔen sən bṇul
3A=insert knife=3POSS AT end beam
He inserted his knife at the end of the beam.

Later on in the story, the knife has become the focal point in the narrative. In (103), after some confusion regarding the knife, the protagonist removes it from where the stranger had placed it earlier on entering the house. This time, in her recapitulation of the event, the narrator uses the derived form of the verb s<r>let insert <CAUS>, indicating that he has not come with malevolent intent.

- (103) dom ki=s<r>let la=tuhən ʔen denerj
AFF 3A=insert<CAUS> A=owner LOC wall
Yes, the owner (had) thought to insert (it) in the wall.

Some further examples and the contexts in which they were used are provided below.

- (104) *The act of giving must always be reciprocated. Rice is sent for this purpose.*

bras	ki=ki<n>rim	ha?	bapa?	ʔame
rice	3A=send<CAUS>	AT	father	Amelia

The rice, he had thought to send to Amelia's father.

- (105) *The speaker's late grandmother, pnāloh, put weights in the bottom of the basket in order to trick the sultan's emissaries into believing the basket to be full of money. This action prevented her people from being subject to any retribution.*

ki=tipoh	rantay	jala?	pō<n>rök	ʔen	karom
3A=fool	weight	casting.net	put<CAUS>	LOC	underneath

She fooled (them) (with) casting net weights she'd thought to put underneath.

As (100) above reveals, the form is also used in imperatives. Whether this is a separate function, or still the same as that above is unclear.

It was suggested by a speaker that this causative form is used to instruct someone to cooperate, ?ur ʔen smat? (instruct 1.OC person) 'instruct s.o.'. In the following examples, the action carried out by the person receiving the instruction will be of direct benefit to them (106), and to the addressee in (107). The underived form would not express this.

- (106) c<r>pat!
be.quick<CAUS>
Make (yourself) quick!
Move it!

In (107), the verb p-su<r>ku? is unusually complex. The underived root is a noun suku? 'a section, portion', which derives a verb by the prefixation of causative <r> 'EQUIP': s u<r>ku? 'to divide into portions' (§5.4.4). This stem feeds further causative derivation, prefixing the morpheme p- 'CAUS': p-su<r>ku? (CAUS-portion<EQUIP>). Note that the auto-causative form p-su<r>ku? is only used when addressing the woman.

- (107) *A squirrel instructs the woman how to prepare the betel nuts in order to woo her errant husband back.*

"ple	ca?-i?	kke,	ma=speck	p-su<r>ku?"
fruit	eat-TR	that	IRR=be.sweet	CAUS-portion<EQUIP>

ki=ʔur	su<r>ku?	mə=malan	ki=ʔur	su<r>ku?
3A=instruct	portion<EQUIP>	REL=be.toxic	3A=instruct	portion<EQUIP>
"The fruits (to be) eaten, (the ones which) should be sweet, be mindful to divide (those ones) into portions!" He instructed (her) to divide (them) into portions. The intoxicating ones he instructed (her) to divide up (too).				

5.3.5.3 Comparative notes

In the related Southern language Mah Meri, the causative is only marginally productive, but fossilised forms show monosyllabic roots prefixed with *pa-* (*lek* 'to get up' → *pa-lek* 'to wake (s.o.) up'), *ta-* (*miŋ* 'to be near' → *ta-miŋ* 'to move (s.th.) near', and *te(N)-cʰŋj* 'to fall off' → *te-n-cʰŋj* 'to remove (s.th.)') and an infixed <*N*> for disyllabic roots (*kəbəs* 'to die' → *kə<m>bəs* 'to kill'). *N* is a nasal which assimilates to the onset of the following syllable (Kruspe in prep.b).

In the Central Aslian languages Temiar and Semai, the causative is formed by the infixation of <*r*> in disyllabic roots, or by prefixation of an allomorph to monosyllabic roots.

In Temiar, there are two variants for monosyllabic roots *ter-* and *ber-*. The choice of prefix is conditioned by the initial of the root, an initial /c/ or /t/ requires the allomorph *ber-*; *cā?* 'to eat' → *ber-cā?* 'to feed', *səg* 'to get trapped' → *tersəg* 'to trap', *səŋl* 'to wake up' → *sergil* 'to wake s.o. up' (Benjamin 1976b: 168-9).

Semai forms the causative of CCVC roots with an -*r*- infix: *kdey* 'not to know' → *k<r>dtey* 'to cause not to know'. Monosyllabics use the prefix *br-*. There is also a malevolent causative, i.e. causation done with evil intent, expressed by the prefix *kr-*: *caa?* 'to eat' → *br-caa?* 'to feed' and *kr-caa?* 'to poison'.

In contrast, Jah Hut, also a Central language, has only causative prefixes, either *p-* or *pr-*: *cyek* 'to sleep' → *p-cyek* 'to put s.o. to sleep', *ca?* 'to eat' → *p-pea?* or *pr-ca?* 'to feed s.o.', *tias* 'to escape' → *pn-las* 'to release' (Diffloth 1976c: 96-7). He notes there are also forms /*tr-*/ and /*kr-*/, as in Temiar. There is a similar, but unproductive causative prefix *kar-* / *kr-* in Semelai (§3.2.4.3).

The Northern language Jahai exhibits a range of causative morphemes, both prefixing and infixing, exhibiting differing degrees of productivity (Burenhult 2002: 152-6). Monosyllabic roots may take a prefix *p-* plus a copy of the root coda (?*ek* 'to give' → *pk-?ek* 'to make someone give'), *pi-* (*muc* 'to eat' → *pi-muc* 'to feed') or *pr-* (*wek* 'to go back' → *pr-wek* 'to cause some one to go back'). Sesqui- or disyllabic roots take an infix <*r1*> (*bkit* 'to be hot' → *b<r1>kitt* 'to heat').

Amongst the other Northern Aslian languages for which data is available, only prefixes are recorded. In Kintak Bong and Kensiu, the causative prefix is *pi-*, e.g. Kintak Bong: *sa* 'descends' → *pi-sa* 'causes to descend', *teg* 'sleeps' → *pi-teg* 'causes to sleep' (Asmah 1976: 954); Kensiu [ʔutbm] 'to fall, collapse naturally' → [pi-ʔutbm] 'to lower (s.th.)'; [*wek*] 'to return' → [pi-*wek*] 'to return by carrying'; [*kəs*] 'to exist', [*pi-kəs*] 'to raise (animals)'; [*puh*] 'to wake by oneself' → [pi-*puh*] 'to wake (s.o.)' (Bauer 1992a).

Austronesian languages also have similar causative forms. Modern Malay has an archaic causative prefix *per-*, Karo-Batak of Sumatra exhibits a form *pə-*. The *pi-* morpheme of the Northern languages is attributed to Austronesian influence by Bauer (1992a: 536). However, a causative prefix of the form *pə-* is a shared feature of Mon-Khmer languages and also occurs in Munda.

5.3.6 Valency increasing -i? 'APPL', 'ITER'

-i? is the only productive suffix in Semelai. The morphological specifics of the suffix are discussed in §3.2.3.2.

-i? has two functions in Semelai, depending on whether it is affixed to an intransitive or a transitive verb:

- a) suffixed to an intransitive verb, -i? 'APPL' increases the valency by introducing an argument O (§5.3.6.1);
- b) suffixed to a transitive verb, -i? 'ITER' marks iterative aspect (§5.3.6.2).

Semantically, the affix can be characterised as indicating a greater affectedness of the object.

The suffix is probably a borrowing of the Malay suffix *-i*, where it indicates a locative relation toward the object. In other words, it directs the action to a location, or directs an action toward the source of an emotion (see §5.3.6.1 below). Note that in Malay, these clauses are not transitive. In colloquial varieties of Malay where the suffix *-i* does not otherwise exist, there are certain lexemes where the suffix is attested as being borrowed attached to a verb. In Semelai -i? is a fully productive affix.

5.3.6.1 Applicative -i? 'APPL' and intransitive verbs

The suffix -i? is attested with verbs of motion and verbs of emotion, where it functions as an applicative. The function of the morpheme is to bring into focus the goal argument, be it a location, or the source of the emotion in the case of emotion verbs. The action is focused toward the O, but does not imply physical impact upon it.

The presence of the suffix results in increased valency, and is not just a case of valency rearrangement. Evidence that the intransitive verb functions as a transitive verb is the cross-referencing of the postposed agent with the pronominal proclitics (109), the potential for the verb to be marked for the irrealis *ma*= 'IRR' and form the transitive negative imperative (111), both features that are usually associated with the transitive verb.

(108)	kəhn	gren	(109)	ki=gren-i?	kmpən
	3S	be.angry		3A=be.angry-APPL	wife
		He is angry.			He is angry (at/toward) (his) wife.

The occurrence of this form in a transitive negative imperative construction confirms the increased valence of the derived form. Compare the negative imperative of the underived form of the verb in (110) with the transitive negative imperative form of (111).

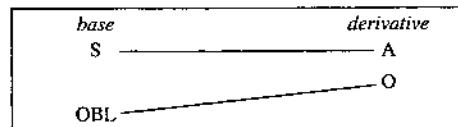
(110)	boy	yɔr!
	NEG:IMP	ascend
		Don't go up (there)!

(111) Some parents warn their children from approaching the troops:

"boy	ma=yɔr-i?"	kʰləŋ,	"he	hnc	x'bəs,	de=bunuh"
NEG:IMP	IRR=ascend-APPL	QUOTE	1&2	THEN	die	3plA=kill
						(us.)"

i. *Directional raising 'APPL'* The suffixation of -i? 'APPL' to intransitive verbs of motion which take a 'directional' complement (TO:NP) (see §8.6) raises the complement to O status. This process, which I term 'directional raising', is schematised in Figure 5.8. Cross-linguistically, motion towards a goal is reportedly one of the most common semantic roles of additional NPs resulting from an applicative process (Comrie 1985: 316).

FIGURE 5.8 DIRECTIONAL RAISING APPLICATIVES



The pair of examples in (112)–(113) illustrates the process. The goal is expressed as a prepositional phrase, shown in brackets in (112); it is expressed as the O in (113).

- (112) ga=ypr [lej ma?itam]_{PP}
IMM=ascend [TO:up fourth.born.aunt]
(I'm) going up to fourth-born aunt's (place).
- (113) ga=ypr-i?_?ma?itam
IMM=ascend-APPL fourth.born.aunt
(I'm) going up (to) fourth-born aunt.

The choice of the directional raising construction in (112) is due to conditions of discourse salience. The clause containing the applicative in (113) suggests a greater sense of purpose. The speaker is going up specifically to see her aunt, perhaps with something in mind, and not just casually wandering up to her aunt's place to visit whoever might be at home.

Some further examples are provided below. Example (114) compares an underived form, the prepositional phrase bracketed, with a derived one in (115).

- (114) tom som bspjet, mudik [lej k^boy dak]_{PP}
SRC morning be.dark go:upstream [TO:up head water]
From early in the morning, (he) went upstream up to the head waters.
- (115) A woman is seeking Salacca fruits which grow alongside the stream.
ki=mudik-i? paya?
3A=go:upstream-APPL stream
She went upstream (on the) stream.

The intransitive motion verb krwl 'to come out, emerge' exhibits a variable transitivity frame and can be used underived with an object krwl 'to put [NP] out', as it is in (116). The focus is on the location:

- (116) masuk k^beh kloc ?ate, ki=krwl-i? io? dyal, ?anu?, pasir
enter 3 inside earth 3A=come.out-APPL penis above [HES] sand
He got inside the hole, and put (his) penis out above the, umm, sand. (Lit. He entered inside the earth ...)

ii. *Directed emotion -i? 'APPL'* When suffixed to verbs of emotion, -i? 'APPL' introduces the external focus of emotion, raising an oblique argument OBL to O, see Figure 5.8 above. The complement which undergoes raising is only optionally expressed as either an oblique NP in the base form, or as the O in the derivative. The following two clauses illustrate the process.

- (117) glok kmpen
laugh wife
The wife was laughing.

In (118) the derivative has negative connotations, glok 'to laugh' → glok-i? 'to laugh mockingly (at s.o.)'.

- (118) ki=glok-i? ?en sma?
3A=laugh-APPL LOC person
She laughed mockingly at the people.

The absence of the proclitic in (119) is due to factors of discourse salience.

- (119) A woman's husband lay down to rest in the incision of a tree. The tree closed over him, and all that remained of him was a protruding lock of hair. The woman's crying was directed at this.
?yam-i? knl^bk n? ne?
cry-APPL husband this before
(She) cried (for) (her) aforementioned husband here.

- (120) ?pac k^bst, ?eh ki=risaw-i? la=bapa?
sight! be.pregnant eh 3A=be.grieve-APPL A=father
The bewildering sight (of the rounded belly), pregnant, ah, the father grieved (for her).

In the causative construction the focus is on the emotion as a result; in the applicative, on the emotion as a cause of further actions or emotional developments. bt^boj 'to be afraid' → bt^boj-i? 'to be afraid toward [NP]':

- (121) k^bhn bt^boj
3S be.afraid
He is afraid.

- (122) ki=bt^boj-i? ?ej
3A=be.afraid-APPL If
He is afraid (of) me.

The increased valency of the verb is confirmed in the sentence in (123) where the derived verb is in the transitive negative imperative construction (§10.6.3).

- (123) boy ma=bt^boj-i? ?ej!
NEG:IMP IRR=be.afraid-APPL If
Don't be afraid (of) me!

The raised argument in the applicative construction is never represented by a causal adjunct. This construction is not the same as the external causation construction, illustrated in (124)–(125) (§5.5.1.2), which implies a source of the emotion, but where the outward manifestation of the fear is not directed externally.

- (124) ki=bt^bŋŋ la=?ap
3A=be.afraid BCS=1f
He is frightened because of me.

In (125), the emotion verb is in the intransitive form, followed by an adjunct expressing the cause of his feelings. Here, the emotion is internally oriented, as opposed to the applicative clause where the orientation is external, directed at the source.

- (125) kəhn risaw [la=kaka? k^bbas]
3S be.sad [BCS=EZ be.dead]
He was sad because his elder sister was dead.

5.3.6.2 Iterativity -i? 'ITER' and transitive verbs

The function of -i? 'ITER' with transitive verbs, is not to increase valence, but to express the iterativity of the event. Transitive verbs are inherently bounded representing unitised events. The presence of an iterative marker on a transitive verb expresses either repeated instances of that event directed toward one object, giving an 'iterative' reading; or the repetition of that unitised event toward a number of individual objects, i.e. a 'distributive' reading.¹⁰ In both possible readings, the focus is on the repetition of the event and consequently the impact of the event on the objects involved. The morpheme is only found on active transitive roots, and always has a perfective interpretation. It is translated as 'kept on V-ing'.

There is no overt means of distinguishing the two readings, which are only discernable contextually, and in light of this the one gloss is utilised for both functions. The following example is plausible with either reading:

- (126) ki=go^b2-i?
3A=fell-ITER tree
Iterative: He kept on felling the tree (in one go without pausing).
Distributive: He felled a series of trees.

In (127), an iterative reading would indicate that the whole area under the verandah was dug up; the distributive would suggest a series of locations under the verandah were dug up. Note that the iterative derivation forms a stem for the middle voice.

- (127) A baby has slipped through the midwife's hands, down underneath the verandah, and disappeared.
br-bɔy-i?, beh da? br-ye
MID-dig-ITER NO NEG MID-see
(The earth) was dug up (all over), (but) no, (the baby) wasn't seen.

In (128) the affix is added to a causative stem.

- (128) An ill-omened child, rejected by his siblings, is surreptitiously nurtured throughout his infancy by his sisters-in-law.
de=sec pa=jip*i?* la=kaka?, ?anu?, la=?ipar
3PLA=steal CAUS-suckle-ITER A=EZ [HES] A=sibling,in.law
The elder sisters, umm, (sisters) in-law, surreptitiously kept on suckling (him).

¹⁰ Internal iterativity of a perfective event is expressed by the imperfective verb form in a transitive clause, e.g. ki=cə<k>rek (3A=tear<IMPERF>) 'She tore and tore (it)'.

Compare the following examples (129)–(130). In (129) the verb is in the unmarked form expressing a single unitised event with no indication as to its internal structure.

- (129) ki=go^b? d^bŋŋ
3A=fell tree
He felled the tree/trees.

A clause like (130a) indicates that he felled the tree, blow after blow, without pausing until the task was completed. A speaker's paraphrase to illustrate this point is given in (130b).

- | | |
|--|---|
| (130a) ki=go ^b -i?
3A=fell-ITER tree
He felled the tree (in one go
without pausing). | (130b) trus ki=go ^b ?
be.straight 3A=fell tree
Without stopping, he felled the tree(s).
(Distr: He felled a series of trees.) |
|--|---|

The temporal repetition of the action is shown here:

- (131) A father is addressing his son who has been missing. He would not have been able to find him, given that the son has been down to the underworld.
jnuh ?əp=reŋ-i?, beh da? j=ye
exhaustively 1FA=seek-ITER NO NEG 1FA=see
I kept on seeking (you) (until I'd had) enough, (and) no, I didn't see (you).

In the following examples the O is elided. Here, the focus would appear to be on the repetition of the action, in (132), possibly on a number of items, but the modifying clause indicates repeated action. In (133) it is with reference to only one object, the addressee's only pair of jeans, and in (134), the wife of the jealous husband.

- | | |
|--|--|
| (132) ki=tut-i?
3A=blow-ITER wind until fall AT:below ground
The wind kept blowing (the clothes) until (they) fell on the ground. | |
| (133) boy ma=pok-i?
NEG:IMP IRR=wash-ITER be.difficult IRR=be.dry BCS=rain
Don't keep washing (that pair of jeans all the time)! (It's) difficult to dry (them). | |
| (134) "ha? nɔ?", k ^b ləŋ, "de=?ye-i?
AT here QUOTE 3PLA=see-ITER person
"Here," (he) said, "people keep on seeing (you)". | |

The following examples indicate that an event is affecting a series, or a number of objects, clearly indicating the 'distributive' reading which is common to them.

- (135) This man compulsively collected rubber-like exudate everytime he found some as he made his way through the forest:
ki=yok-i?
3A=take-ITER exudate
He gathered exudate (everytime he found some).

- (136) *A man and his dog spend the day making a series of traps:*
 yok sarek, dehn swak j?sy-i? twar=c?2
 take tomorrow 3pLS walk make-ITER trap=EM
 The next day, they went (and) made (a series of) traps.
- (137) *The narrator describes how the Japanese slaughtered the different groups of people in the surrounding area during World War Two.*
 sma? putih, gop, cina? de=krja?i? la=jpon
 person be.white malay chinese 3A=work-ITER A=japanesc
 Europeans, Malays and Chinese, the Japanese slaughtered them all. (Lit. The Japanese worked them (all) over.)

Compare examples (138)–(139). The woman informs a series of people, and the distributive nature of the event is reflected in the verbal marking; in the negated clause in (139) mention is made of only one group, and the suffix is not used.

- (138) *The narrator is talking about her cousin who was a known informer during the Communist Emergency.*
 ki=par-yen-i? ?en sma? putih, ?en cina?
 3A=CAUS-hear-ITER LOC person be.white LOC chinese
 She informed (everywhere), to the Europeans, to the Chinese.
- (139) *The narrator contrasts her elder sister's behaviour with that of her informing cousin.*
 r?m sma? putih, da? ki=par-yen
 WITH person be.white NEG 3A=CAUS-hear
 With the Europeans, she didn't inform (them).

5.3.7 Happenstance t(r)- 'HAPP'

The prefix t(r)- 'HAPP' is a verbal affix borrowed from Malay *ter-*. Allomorphy and other morphological specifics are discussed in §3.2.3.1.

tr- 'HAPP' has various functions, all of which share a common theme of being resultative, having taken place prior to the moment of speech. The various usages also contain recurrent themes of non-volition, and inability.

The primary function of this morpheme is to express the non-intentional result of an action (§5.3.7.1). The focus is on the completed state. It is never used to indicate a non-completed state, and is thus not used in conjunction with aspectual markers, nor in conjunction with negation.

In §5.3.7.2 I discuss the use of the prefix in conjunction with the negator *da?* 'NEG' to express inability; prefiguration to a reduplicated stem which expresses intensive absorption is discussed in §5.3.7.3; the use of *tr-* to express involuntary acts or uncontrolled causation where *t-* is prefixed to a causative verb in §5.3.7.4, and the discussion closes with the comparative function of stative verbs expressing degrees of speed in §5.3.7.5.

In Malay *ter-* is used with verbs to express a final or completed state, and with adjectives to express the superlative (Liaw 1985: 112–13). It has two functions as a verbal prefix: a) to indicate a final or completed state, achieved actively or passively, where intention is irrelevant; and b) *ter-* to mark an unintentional or non-volitional action (Liaw 1985: 122–4). A rare form involving the negator *tidak* indicates inability.

tr- 'HAPP' functions exclusively as a verbal affix in Semelai, and while it is not used to express the superlative, it closely parallels the functions of *ter-* in Malay.

5.3.7.1 Happenstance 'HAPP'

This is an action which is performed without intent, resulting in an unexpected, spontaneous or accidental result. The use of the prefix *tr-* indicates not that it is the action which is unintended, but the result which is unexpected, or beyond the control of the agent. A suitable paraphrase would be 'happened to V', Frawley's 'happenstance' (1992: 205), and it is glossed here as 'HAPP' in the light of this.

- (140) *I was about to open a tin of milk, but the tin slipped out from under the blade, and the opener pierced my big toe.*
 ga=y=e=tbk susu?, y?2=hn t-tbk gad??.j?n
 IMM=1A=open.tin milk but=CONN HAPP-dig.hole parent.foot
 I was about to open (the tin of) milk, but (I) inadvertently dug a hole in my big toe.

The type of verb prefixed by *t(r)-* in this function is highly restricted, co-occurring exclusively with transitive verbs with external orientation. The verb is generally transitive, or an intransitive verb which first requires transitivising with the causative morpheme (§§5.3.7.4–5).

yok	'to take'	→	tr-yok	'to happen to take'
y?r	'to ascend, climb'	→	t-y?r	'to happen to ascend'
curah	'to tip out'	→	t-curah	'to happen to tip out'

The following verbs are excluded on the basis of the definition above, that they are not externally oriented:

- a) The verb cannot be an intransitive verb of posture, or a bodily activity, with the exception of *jtek* 'to go to sleep' → *t-jtek* 'to fall asleep':

swak	'to walk, go'	→	*t-swak
?yam	'to cry'	→	*tr-yam
k?k?u?	'to vomit'	→	*t-k?u?

- b) The verb cannot be a stative verb of emotion:

c?en	'to be content'	→	*t-c?en
------	-----------------	---	---------

- c) The agent must be animate. The derived transitive verbs retain the agent pronominal proclitic, but the subject is never attested post-verbally by an external NP marked with *la* = 'BCS'. In other words, neither the subject nor object show signs of reduced transitivity.

- (141) ki=tr-ye wdy kke
 3A=HAPP-see knife that
 She happened to see that knife.

The wife is actually seeking lice, but seeking does not presuppose finding them:

- (142) ki=ksep tr-ye ci, kapoh ci ki=ca
 3A=suck HAPP-sce lice egg lice 3A=cat
 She sucked the lice (she) happened to see, (and) the nits she ate.

The object can be marked by the proclitic *hn* = 'O':

- (143) t-biŋy hn=mata?are?
HAPP-look.upward O=sun
(He) inadvertently looked upwards (at) the sun.

In example (144) the irrealis *ma*=‘IRR’ is prefixed to the verb, replacing the agent proclitic. *tr-* is used here in the explanation because one does not go around intentionally annoying millipedes, as they are associated with the transgression of certain taboos.

- (144) ?pspl ma=t-gnen, wc-wsc deh
when IRR=HAPP-disturb IMPERF-curl.up 3pl
When (one) happens to disturb (centipedes), they curl up.

The absence of the proclitics in (145) is due to the generic nature of the statement.

- (145) *crwēt* is a visual-based expressive. This example provides a definition of the word.
The possible context given for this happening is whilst bathing.
crwēt! tr-ye r̩s b?lə?
catch.sight.of.female.genitals! HAPP-see genitals.female friend
crwēt (means) to happen to see a female's genitals.

Some examples now follow:

- (146) A woman and her child live with a tiger. One day whilst napping, the tiger slaps at something, not realising it to be the child. The woman reprimands him, realising he wouldn't normally have done such a thing.
kp=t-tampar hne knon
2fA=HAPP-slap just.now offspring
You inadvertently slapped (at) the child just now.

- (147) A possessed wife is feeling under the boat with the aim of grabbing her husband who is hiding underneath. Her hand comes into contact with his ear, but she was not feeling specifically for the ear.
ki=tr-jam dawon.təŋ knlek
3A=HAPP-feel leaf.ear husband
She happened to feel (her) husband's ear lobe.

- (148) The protagonist is on her way to fetch a knife when she sees one belonging to the visitor. She was not expecting the knife to be there.
ga=m-lanjah ha? bñul ke, ki=tr-ye wpy kke
IMM=IMPERF-step.across AT beam that 3A=HAPP-see knife that
(She) was going to step across the beam, (when) she happened to see that knife.

In the following clause the O is also the subject of the conjoined clause:

- (149) yok sarsk, tr-ye hn=b?lə? ga=k<?>rū?
take tomorrow HAPP-sec O=friend IMM=fish.poison<IMPERF>
The next day, (he) happened to see his friends going fish-poisoning.

In (150) the subject is an experiencer, no longer an agent, and as a consequence is not encoded by a pronominal proclitic.

- (150) A short-sighted man has been cured:
sta? t-bukp? mot, ki=k³nal j?oy k<n>rja? b̩rɔŋ ?a=deŋ
after HAPP-open eye 3A=know do work<NMZ> thing DET=like
After (his) eyes were opened, he knew how to do work (and) what-like.
After his blindness was cured...

i. Dreams and thoughts The notion of passivity, of being subjected to unintentional external action is found in the following examples where *t-* ‘HAPP’ is prefixed to a verb describing the coming to one of a dream, or something entering one’s thoughts. These can be neither expected nor controlled.

- (151) The younger sibling has an idea:
t-masuk ha? gres ?adi?
HAPP-enter AT liver YS
(It) happened to enter younger sibling's mind.

- (152) A father seeks a name for his newborn son. The name comes to their servant in a dream.
t-masuk po skmbanj ?ɔraŋ gaji deh
HAPP-enter dream servant person wage 3pl
(It) happened to enter the dream of their wage-earning servant.

- (153) ye ke, sampay da? moh ca, la=t-peker hn=kəhn
1 that until NEG want eat BCS=HAPP-think O=3S
Me there, (got to a state where I) didn't want to eat on account of thinking of him.

- (154) ?luc.tom masak, t-peker iŋ jtek war
after cook HAPP-think want sleep instead
After cooking (she) (had the) thought that (she) wanted to sleep instead.

5.3.7.2 Inability

A verb with the prefix *tr-* in a clause negated by *da?* ‘NEG’ expresses the inability to perform an action. The abilitative clause is resultative, and the agent is not expressed in the clause. It is usually mitigating circumstances which prevent the event from taking place, and not the physical ability of the protagonist.

- (155) da? tr-yot tet dol
NEG HAPP-return TO:spec house
He was unable to return to the house.

- (156) “da? dapat t-paloh” k³ley
NEG succeed HAPP-flee QUOTE
“(They) weren't able to flee,” (he) said.

In (157), it is possible to have either an inabilitative a), or a happenstance reading b):

- (157) A woman describes the power of evil shaman as being so potent that people die before signs of illness are able to manifest.
da? tr-pih, cəŋ k³bəs la=pham dehn
NEG HAPP-be.ill immediately die BCS=think 3plS
a) Unable to be ill, (one) immediately dies because of their (evil) thoughts.
b) Before signs of illness happen/are able to manifest, one dies straight away.

5.3.7.3 Intensive absorption

In conjunction with partially reduplicated 'intensive' forms of the root (§5.3.9), verbs are derived with the meaning 'to perform an action completely lacking in self-awareness', where involuntary response is involved.¹¹ As with the causative construction (§5.3.7.4), the root form of the verb is not otherwise available to t-prefixation, only the derived stem form. Note that both transitive and intransitive roots feed this derivation:

- | | | | |
|------|----------|--------------|--|
| ca | 'to eat' | → t-ca-ca | 'to gobble up, with no concern for others' |
| ?yam | 'to cry' | → t-?ya-?yam | 'to cry inconsolably' |

In the following example, the reduplication of the verb merely indicates intensity. The woman gazes upward, so self-absorbed that she doesn't realise people are watching her. Compare this with sentence (143) above where the verb is not reduplicated.

- (158) t-blo-bloy kəh, ki-jjɔk srlər
 HAPP-INTNS-look.upward 3 3A=watch leaves.moving.in.breeze
 Gazing upward self-absorbed, she watched the leaves moving in the breeze.

- (159) t-pyo-pyoh ?pih mət
 HAPP-INTNS-thrash.about be.sore eye
 He thrashed about self-absorbed, (because his) eyes were hurting.

The presence of the negator da? 'NEG' in the clause merely negates the verb, and is not related to the non-abilitative reading (§5.3.7.2).

- (160) A woman has been crying inconsolably over the death of her mother:

- | | | | | | | |
|-----|---|---------|------|-----|----------|------|
| tom | ha? | kke, | beih | da? | de=2yam | wɔ? |
| SRC | AT | there | NO | NEG | 3plA=cry | more |
| da? | t-?ya-?yam | loc | | | | |
| | HAPP-INTNS-cry | already | | | | |
| | From there on, no, they didn't cry any more. (They) didn't cry inconsolably (any longer). | | | | | |

The final example is set in context:

- (161) ʔəp loc jnuh ca crh. da? t-ca-ca wɔ?
 If already be.surfeited eat fish NEG HAPP-INTNS-eat more
 I've had enough of fish-eating. (I'm) not eating (it) anymore.

5.3.7.4 tr- and the causative base

In §5.3.7.1 in point (a), it was noted that intransitive verbs of posture, bodily activity and motion don't feed the prefixation of t- 'HAPP'. However, the prefix can be added to causative bases of these verbs (§5.3.5) to indicate an involuntary result.

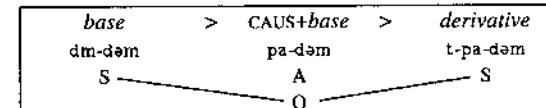
- | | | | |
|------|------------|------------|--|
| √dəm | (lie down) | → pa-dəm | 'to lie (s.th.) down' |
| | | → t-pa-dəm | 'to be involuntarily caused to fall to a lying position' |

¹¹ As far as I am aware this construction is not attested in Malay.

- | | | | |
|--------------------|------------|--------------------------|--|
| √k ^b om | (sit down) | → pa-k ^b om | 'to set (s.th.) down' |
| | | → t-pa-k ^b om | 'to be involuntarily caused to fall to a sitting position' |

The argument frame of the derived forms is given in Figure 5.9.

FIGURE 5.9 ARGUMENT FRAME OF THE HAPPENSTANCE CAUSATIVE



In an interesting twist, the patient, the object of the causative form, is coded as the subject, and the agent is not specified (162). The presence of the prefix appears to detransitivise the verb:

- (162) kəh pon t-pa-k^bom, cəj t-pa-dəm ʔa=kəh
 3S THEN HAPP-CAUS-sit immediately HAPP-CAUS-lie DET=3
 She was caused to fall to a sitting position (and then) immediately she was caused to fall down (flat).

In the following examples (163)–(164), there is a sense of fate dictating the event as the causative force:

- (163) ʔare? muy, t-p<r>loc ʔen dol sma?
 day one HAPP-emerge<CAUS> LOC house person
 One day, he happened to come out (of the undergrowth) at someone's house.

This construction is also used where the causative force is obviously inanimate and therefore unable to direct the action at the undergoer. In the following examples the causative force is his intoxication (164), the sun (165)–(167), and in (168) again something akin to fate.

Where the root is an intransitive verb the subject takes the usual pronominal form. The causative force can be expressed in a causal la= 'BCS' clause in (164) and (166)–(167).

- (164) da? sðər wɔ?, t-pa-dəm ha? kke, la=malan
 NEG remember longer HAPP-CAUS-lie AT that BCS=bc.intoxicated
 (He) didn't remember any longer, (and so he) was made to fall to a lying position there, because of the intoxication.

In the following example mataʔare? 'the sun' is the object of t-bloy 'to involuntarily look upward', and the causative force of k<r>ŋep 'to make blink<CAUS>':

- (165) t-bloy bn=mataʔare?, t-k<r>ŋep ye
 HAPP-look.upward O=sun HAPP-blink<CAUS> 1
 Inadvertently looking upwards at the sun, happened to make me blink.

Only forms collected by direct elicitation have pronominal agent cross-referencing, as attested in (166)–(167).

- (166) ki=t-k<r>ne_p la=?ame, la=mata?are?
 3A=HAPP-blink<CAUS> A=Amelia BCS=sun
 Amelia was made to blink, because of the sun.
 Amelia happened to blink, because of the sun.

The following example with a first person inclusive subject proclitic illustrates that the proclitic is coding the undergoer and not the causative force:

- (167) he=t-k<r>ne_p la=mata?are?
 1&2A=HAPP-blink<CAUS> BCS=sun
 We were made to blink because of the sun.

This construction is also attested with the verb croh 'to meet' → c-n-roh 'to meet (s.o.)' → t-c<n>roh 'to have a chance meeting/happen to meet up (with s.o.)'. The clause in (168) expresses spontaneity or lack of input by the protagonist, who just happens to meet up with various creatures as he travels through the forest.

- (168) t-c<n>roh rɔm brwaj, podəŋ, gajah
 HAPP-meet<CAUS> WITH bear tiger elephant
 (He) happened to meet up with a bear, a tiger and an elephant.

5.3.7.5 Comparative *t-* and verbs of speed

A similar usage to that in the previous section is attested where *t-* 'HAPP' occurs with the causativised derivatives of adjectives of speed. This function conveys the comparative. It would appear to be related to a similar situation in Malay.

The derivation is only available to the following three roots:

cpat	'to be quick'		
→ c<r>pat (be quick <CAUS>)	'to make (s.th.) move fast'		
→ t-c<r>pat (HAPP-be.quick <CAUS>)	'to make (s.th.) move faster'		
jres	'to be fast, swift'		
→ j<n>res (be.fast <CAUS>)	'to make (s.th.) move fast'		
→ t-j<n>res (HAPP-be.fast <CAUS>)	'to make (s.th.) move faster'		
?ayon	'to be slow'		
→ ?a<r>yon (be.slow <CAUS>)	'to make (s.th.) move slow'		
→ t-a<r>yon (HAPP-be.slow <CAUS>)	'to make (s.th.) move slower'		

This form is only used in relation to an animate agent, usually as an imperative 'to walk, or move faster or slower':

- (169) t-c<r>pat!
 HAPP-fast<CAUS>
 Move it! (Lit. Make yourself move faster!)

5.3.8 Comparative +ra?+ 'COMP'

The morpheme +ra?+ 'COMP' indicates a comparative relationship between two or more entities.

This affix has the most restricted distribution of the derivational morphemes, occurring exclusively with the class of eight dimension adjectives listed below.

t ^b ay	'to be big'	→ ra?t ^b ay	'to be bigger'
k ^b et	'to be small'	→ ra?k ^b et	'to be smaller'
?jəs	'to be high'	→ ra?jəs	'to be higher'
d ^b əs	'to be low, short'	→ d ^b ra?d ^b əs	'to be lower, shorter'
ləŋj	'to be long'	→ j<r>a?ləŋj	'to be longer'
ləɛ?	'to be short'	→ j<r>a?ləɛ?	'to be shorter'
sɛ?	'to be thick'	→ ra?sɛ?	'to be thicker'
sey	'to be thin'	→ ra?sey	'to be thinner'

The morphological specifics of the process are presented in §3.2.2.1.

The derived form may be used predicatively (170) or attributively (171):

- (170) mə=mirah ha? ra?k^bet, ke=sən, ra?t^bay
 REL=be.red AT COMP-be.small that=SC COMP-be.big
 The red ones there are smaller, these see, are bigger.

- (171) boy ga=p?la? ra?t^bay, p?la? ra?k^bet, ki=?ye
 NEG:IMP IMM=animal COMP-be.big animal COMP-be.small 3A=see
 Don't (think it was) going to be bigger animals, the smaller ones, he saw (as well).

The function of this morpheme is to make a relative comparison of dimension in terms of one of the four antonymic pairs of this particular class of stative verbs: size (big/small), height (high/low), length (long/short) and thickness (thick/thin).

In (172), note that the evaluative adjectives ?ilok 'to be good' and buruk 'to be broken, worthless' are not marked for the comparative, despite the same comparative relationship implicit to the pair. The relationship can only be expressed by juxtaposition for those terms which do not derive the comparative.

- (172) *The narrator compares the size and condition of the dwellings in relation to the distribution of money from the British forces.*

dəl ra?t^bay, ra?k^bet ?ilok, buruk,
 house COMP-be.big COMP-be.small be.good be.worthless

mə=ratus=jə de=jən
 one=hundred=CL 3plA=give

(Whether) the house was big (or) small, in good (condition) (or) worthless, they gave (us) one hundred dollars, see.

When any of the eight adjectives of dimension are used without the morpheme they have a purely descriptive reading (173), rather than a comparative one as in (174).

- (173) 2le=le? de=kde? dəl t^bay, dəl k^bet
 so 3plA=dwell.in house be.big house be.small
 So, they lived in big houses (and) small houses.

- (174) *On returning home the size of three pigs, spotted earlier running through the plantation, is recounted.*

yε=?en, yε=?ye jalu, dwa? ?ikur ra?t^bay
 1=AUG 1A=sec pig two CLF COMP-be.big

rɔm mə=?ikur ra?k^bet
 WITH one=CLF COMP-be.small
 Us, we saw (some) pigs, two bigger and one smaller.

The object of comparison can be expressed in a prepositional phrase with *tom* 'SRC' (§8.7):

- (175) d<ra?>p̥es tom ye
be.short<COMP> SRC 1
(She's) shorter than me.

In drawing the comparison, it is not necessary to have both dimensions explicitly stated. In (176), the hill is compared with those previously traversed, as the car makes its way along a hilly track. The speaker cannot account for any hills that may lie ahead, but she passes judgement:

- (176) c̥ɔŋ ra?̥-n̥es
hill COMP-be.high
A high(er) hill.

In most examples there is no obvious object of comparison other than some kind of prototype entity, as is demonstrated in the following examples.

- (177) *The expressive kikmuk is used to describe someone who has bigger cheeks than what is considered to be the norm.*
kikmuk, sma? ra?̥-t̥hey metj
big.cheeks! person COMP-be.big cheek
kikmuk (means) someone with bigger cheeks.

In (178) it is not the cow which is the object of comparison, but the class of animals in general.

- (178) kmus p?la? ra?̥-k̥et, g̥n-g̥n, sukp? dk̥es lmbu?
[midge] animal COMP-be.small IMPERF-bite like be.near bovine
kmus are smaller animals (which) bite. (They) like being near cattle.

- (179) kayen ra?̥-sey mirah, m̥bom kayen p̥asot
cloth COMP-be.thin be.red like cloth parachute
(It) is thin red cloth, like parachute cloth.

5.3.8.1 Comparative notes

Morphemes of the shape *ra?*, *-r-* and *<1>* are associated with pluralising and replicative functions in other Aslian languages. Some of these are noted below.

In the related Southern language Mah Meri, the affix *<1>* or *<1a>* indicates multiple referents or replication of subjects of positional verbs: *t̥omp̥et* 'be lying on the ground (one item)' → *t<1>omp̥et* 'many (things) lying on the ground' (Kruspe in prep. b).

In Semai there is a distributive infix *-ra-* indicating 'several referents', or a 'simultaneous plural'. It occurs with expressives: *dyɔ̥i* 'an object floating downstream' → *drayɔ̥i* 'several objects floating downstream together' (Diffloth 1972: 440, 1976b: 252-3).

A morpheme *ra?*, glossed as the superlative, is noted by Diffloth in the Aslian language Jah Hut (Diffloth 1976c: 97). He provides the following examples: *num* 'to be ripe' → *ra?-num* 'very ripe'; *hluk* 'heavy' → *s<ra?>luk* 'very heavy'. The gloss however does not conform to the label 'superlative'.

In some Northern Aslian languages there is an infix *-r-* which is used to indicate the plural in human nouns, e.g. Batek *k̥dah* 'young woman' → *k̥eradah* 'young women'.

Benjamin notes a Mon form meaning 'like, manner' indicating a relationship 'in a series', which he relates to a similar meaning in Temiar. The Mon forms which he cites are Old Mon *row* and spoken Mon *r̥ea* (Benjamin ms.). In the wider context, Diffloth (1984: 311) notes an Austroasiatic infix *-ra- indicating 'simultaneous'.

5.3.9 Intensive 'INTNS'

The intensive is a marginal derivational process which takes the form of light syllable reduplication (§3.2.5). It denotes intensity, either in the degree of a state: *?p̥e-?p̥es* 'to be really high', ← *?p̥es* 'to be high'; or intensity in the performance of an action: *tare-tarek* 'really pulling' ← *tarek* 'to pull'.

The primary function of the derivative is as a verbal adjunct with an adverbial reading of intensification (§10.3.4).

- (180) ki=br̥kas ?ilɔ-?ilɔk
3A=tic.together INTNS-be.good
She tied (them) together really well.

The derivative may also function as a verb:

- | | |
|----------------------------|-------------------------------|
| (181) k̥ehn pon dk̥e-dk̥es | (182) r̥po-r̥pok t̥hi |
| 3S THEN INTNS-be.close | INTNS-clap hand |
| She was really close. | really clapping (their) hands |
- (183) k̥na? ki=c̥je-c̥jew, ki=j̥p̥k sma?
happen 3A=INTNS-look.down 3A=observe person
It happened (that) (when) he looked down, he saw someone.

The intensive stem feeds derivations with the middle morpheme *b(r)-* 'MID', where it has two separate functions, and the happenstance morpheme *t(r)-* 'HAPP'. An example of each function is provided below.

- a) The intensive of verbs prefixed with *b-* 'MID' forms the reciprocal, e.g. *b-c̥kpa-c̥kpa* (MID-INTNS-speak) 'to converse (with each other)' (§5.3.3.2):

- (184) p̥ot, b-c̥kpa-c̥kpa r̥om twanko? ke
stay MID-INTNS-speak WITH sultan that
(He) stayed, conversing with the sultan.

- b) The intensive of stative verbs prefixed with *b-* 'MID' (§5.3.3.3), indicates the experience of an intensive state:

- (185) *The whole kampong is involved in the dibbling of the rice.*
k̥ehn dl̥h mn-(t)ug̥l b-r̥aina-ramay
3S go.across IMPERF-dibble.rice MID-INTNS-be.many
He went across to dibble rice (as one of) many, many (people).

- c) In combination with the happenstance prefix *t-* 'HAPP', the derivative indicates self-absorption in an activity (§5.3.7.3):

- (186) *A girl is so absorbed in the leaves that she fails to notice that she is being stalked.*
 t-blo-bloy kəh, ki=jpɔk sr̥lər
 HAPP-INTNS-look.upward 3 3A=watch leaves.moving.in.breeze
 Gazing upward self-absorbed, she watched the leaves moving in the breeze.

5.4 Derivation: nominal to verb

The derivation of verbs from other parts of speech is achieved by the same morphological processes as those used for verb to verb derivation: coda copy 'USE', prefixation of mN- PERFORM 'PERFM' ('IMPERF(ective)' on verbs), the prefixation of b(r)- glossed here as 'HAVE', or 'USE' ('MID(icle marking)' on verbs), and affixation par-, <r> or p- 'EQUIP' ('CAUS(ative)' on verbs). The structural aspects of these processes are discussed in §3.2. The basic function of the processes is the same as for verb to verb derivations, with some limitations, but semantically, they are more constrained. Coda copy and causative derivations only occur with quite specific classes of nominals rendering them relatively unproductive processes.

Coda copy, mN- 'PERFM' and b(r)- 'HAVE', 'USE' derive intransitive verbs, while the causative 'EQUIP' derives transitive verbs. Aspects of the syntax are noted below. Details relating to clausal syntax and argument coding are discussed in detail in §9.

The processes which are discussed here are summarised in Table 5.5.

TABLE 5.5 SUMMARY OF VERBAL DERIVATION STRATEGIES

MEANING	MORPHEME USED			
	+C+ 'IMPERF'	b(r)- 'MID'	m(N)- 'IMPERF'	par-/<r>/p- 'CAUS'
'to have N' ('HAVE')	+	+		
'to utilise N' ('USE')	+	+	+	
'to equip with N' ('EQUIP')				+

5.4.1 Coda copy

Coda copy is used to derive verbs from nouns expressing body parts, implements and clothing, and to derive verbs from the set of Mon-Khmer numerals (§7.3.1). These derivatives are used in stative predicates. The process is not productive; the comparable process for nouns which do not feed reduplication, and for numerals borrowed from Malay, is to prefix b(r)- 'HAVE' (§5.4.2). The latter process, prefixation of b(r)- 'HAVE', is the more productive of the two processes. This pattern, where the prefix compensates for the limited productiveness of coda copy is also witnessed for verb to verb derivation (§5.3.3). See also (§5.4.2) below.

It is not understood why one form is preferred over the other, since there is nothing in the morphological process, or the syllabic structure of the roots which would disallow it, and both Semelai words and Malay loanwords feed coda copy. In general, there is no constraint on Malay roots feeding Aslian processes, nor vice versa, and roots may exhibit a preference for one form or allow either as illustrated: p̥əc 'wing, fin' → b-p̥əc/p<c>rəc 'to have wings, fins'.

5.4.1.1 Noun to verb

Certain nominals expressing body parts and clothing can derive a verb: i) 'to HAVE N[body part]', or ii) 'to be utilising N, USE'. The full extent of productivity for this process has not been determined.

i. 'to HAVE N[body part]' The following forms are the only ones attested for this process:

pes	'tail'	→ ps-pes	'to have a tail'
tuh	'breast'	→ th-tuh	'to have breasts'
suk	'hair, fur, feathers'	→ sk-suk	'to have hair, fur, feathers'
jare?	'digit'	→ j<?>re?	'to have digits'

- (187) ?rey ja<?>re? mirah
 many digit<HAVE> be.red
 (That millipede) has many red feet.

The three derivatives which follow express an inherent property of the subject, 'to emit N':

k ^b lem	'odour'	→ k ^b <m>lem	'to emit an odour'
jlək	'smoke'	→ j<?>lk	'to emit smoke, (of fires)'
?ahom	'breath'	→ ?a<m>hom	'to breathe'

ii. 'to utilise N', 'USE' Some examples of this process are:

bnih	'seed'	→ b<h>nih	'to sow rice, of women'
cru?	'tray used to winnow rice'	→ c<?>ru?	'to winnow rice with a tray'
syuy	'loincloth, nappy'	→ sy-suy	'to wear a loincloth, nappy'
satom	'right hand'	→ sa<m>tom	'to use the right hand'
sawel	'left hand'	→ sa<?>wel	'to use the left hand'
tpunj	'flour'	→ t<?>punj	'to perform a flour healing ritual'

- (188) Uttered to excuse unavoidable use of the left hand:

sa<?>wel ye huc
 left.hand<USE> 1 rice
 Left-handed I (pass) the rice.

The following two derivatives are based on the bound roots √yup and √sot respectively: yup-yup 'to use a sleeping cover' ← kayen.yup 'sleeping cloth'; st-sot 'to wear a sarong at the waist' ← kayen.sot 'sarong worn at waist'.¹²

- (189) ki=k^bom yup-yup, kəh, ki=k^bom st-sot
 3A=get USE=sleeping.cloth 3 3A=get USE=sarong
 (He) got to use a sleeping cloth, (and) her, she got to wear a sarong.

5.4.1.2 Numeral to verb

The Semelai ordinal numerals, three to seven, feed coda copy deriving a verb 'to be (in) a group of X members'. The members of the group must be animate, either people, ghosts or animals.

¹² The point occurring between the two lexical items in these compounds is to indicate that neither yup nor sot are independent roots, unlike suy 'loincloth, nappy'.

Numerals three to five delete the initial syllable, prior to coda copy. Prior to implementing coda copy the numerals six and seven drop the initial syllable and prefix the morpheme /t-/,¹³ the meaning of which is unknown. The derivations are exemplified below:

hmpe?	'three'	→ p?-pe?	'to be a group of three'
hmpoñ	'four'	→ pn-pon	'to be a group of four'
msoñ	'five'	→ sñ-sñg	'to be a group of five'
pru?	'six'	→ t<ñ>ru?	'to be a group of six'
tmpoh	'seven'	→ t<ñ>poh	'to be a group of seven'

In (190) the derivative is used predicatively, and in (191) it is a modifier.

- (190) dehn t2-ru?
3plS BE-seven
They were in a group of seven.
- (191) dos p?-pe?
arrive BE-three
(They) arrived as a group of three.

Only the Semelai Mon-Khmer numerals feed the process described here. The Malay numerals feed the prefixation of b(r)- 'BE': dwa? 'two' → b-dwa? 'to be two' (§5.4.3), and 'to be/do alone' is expressed as ki=ba?i 'he alone', de=ba?i 'they alone'.

5.4.2 Performative mN- 'PERFM'

The prefix mN- 'PERFM' derives intransitive verbs, loosely defined as 'to perform with an N', paralleling the function of coda copy described in the previous section. The single argument is an agent and the clause is imperfective. Elsewhere, it is used to derive the imperfective form of verbs (§5.3.2). The prefix is borrowed from Malay, and is only applicable to nouns which are also borrowed from Malay, where the corresponding forms exist, e.g. *tuba* 'a plant with sap used as fish-poison' → *menuba* 'to fish using this poison'.¹⁴

sayor	'vegetable'	→ mñ-(s)ayor	'to stew meat with vegetables'
gube?	'mortar and pestle for pounding betel nuts'	→ mñ-gube?	'to pound betel'
stek	'slingshot'	→ mñ-(s)tek	'to shoot with a slingshot'
rökök	'cigarette' ¹⁵	→ mñ-rökök	'to smoke tailor-made cigarettes'
jala?	'a casting net'	→ mn-jala?	'to fish with a casting net'

¹³ The form of the prefix is /t-/ rather than /ta-/. Prefixation of /t-/ precedes left coda copy. The infixation of the final syllable in numerals six and seven, /t/ and /h/ respectively, results in the realisation of [ə] ← [ə] (§3.2.1.3).

¹⁴ The noun *sayur* 'vegetable' ← Malay *sayur* 'vegetable' has the verbal form *məsayor* 'to cook vegetables' in Semelai, while the Malay form is *sayurkan* 'to cook vegetables'. The suffix -kan is not attested in Semelai. Formerly the Malay form was *menyayor*, according to Winstedt (1957).

¹⁵ To smoke traditional rolled cigarettes is expressed as ?puk dawon (suck leaf).

The following form deviates from the rules given in §5.3.2: lsoñ hindí? 'rice-mortar whose pestle is worked with the foot' → mñhindí? 'to pound rice in a mortar using a pestle worked with the foot'.

It is not always easy to determine whether the root should be analysed as a noun or verb, for example in the case of mnjala? 'to fish with a casting net' ← jala? 'casting net', given that this can also function as a transitive verb. In this case either IMPERF or PERFM is an appropriate gloss:

- (192) ki=jala?=hn
3A=fish.with.casting.net=O
He caught her in the casting net.

In (193) the derivative is used in a serial construction.

- (193) ki=gøj mn-jala? la=kñlæk neñ
3A=take PERFM-casting.net A=husband before
The husband took (himself) fishing with a casting net.

A few examples do not fit the definition given above. One expresses destination: darat 'land' → mn-darat 'to land'. A second is glør 'name' → mñ-glør 'to bestow a name'. A final example involves the verb mñpot 'to weed a swidden'. It comes from the noun rompot 'non-cultivar', from the Malay *rumput* 'weed'. The <m> appears to have been analysed as a nominalising agentive infix (§7.6.1), and dropped in order to derive the verbal form.

- (194) kñhn mñpot
3S PERFM:weed
She is weeding.

5.4.3 Middle voice b(r)- 'to HAVE/USE'

The prefixation of b(r)- to a noun derives a stative predicate. As with coda copy (§5.4.1), the derived form either expresses (a) possession, 'to HAVE N', or (b) utilisation of an object, glossed here as 'to USE N'. The subject is the possessor of the object or attribute (§9).

The use of middle or 'reflexive' marking to derive a denominal form is a peripheral, but well-attested morphological process cross-linguistically (Geniusiene 1987: 340).

5.4.3.1 b(r)- + N possession 'to HAVE N'

The possessed entity may be a body part, 'to HAVE N body-part', an attribute, 'to HAVE N attribute', or an object.

This is used to express the possession of a body part (195)–(196), an object (197), or to express possession of physical attributes: aesthetic properties, physical, or geographical features (198)–(199).

- (195) ?oh ?ma? da? br-køy
oh mother NEG HAVE-head
Oh, the mother (was) headless!

- (196) smə? b-misay, b-jarot, b-huban
person HAVE-moustache HAVE-beard HAVE-grey.hair
a bearded, moustached, grey-haired person
- (197) da? br-wpəy kəh
NEG HAVE-knife 3
He didn't have a knife.
- (198) da? b-bri wə?
NEG HAVE-forest more
(It does) not have forest anymore.
- (199) da? b-gaya? dol wpəy
NEG HAVE-appearance handle knife
(It) didn't have the appearance of a knife handle.

It is also used to express the following affinal relations:

- b-knək 'to have a husband, be married, (from a female perspective)'
b-kmpən 'to have a wife, be married, (from a male perspective)'

- (200) ye da? lən b-kmpən
I NEG desire HAVE-wife
I don't want to have a wife/be married.

b-knən (HAVE-offspring) 'to give birth; have children' refers primarily to the process of giving birth rather than the state of having children. In the following example the woman is experiencing labour pains:

- (201) ?əp jm-pam lən b-knən
I IMPERF-feel want HAVE-child
I am feeling (like) (I) want to give birth.

5.4.3.2 *b(r)- + N* 'to be utilising N', 'USE'

This construction is semantically equivalent to coda copy, §5.4.1.1 above.

- | | | | | | |
|--------|-----------|---|----------|---------------|-----------------------|
| dəl | 'house' | → | br-dəl | (USE-house) | 'to house oneself' |
| baju? | 'clothes' | → | b-baju? | (USE-clothes) | 'to wear clothes' |
| kreta? | 'car' | → | b-kreta? | (USE-car) | 'to use a car' |
| kayuh | 'paddle' | → | b-kayuh | (USE-paddle) | 'to paddle' |
| layar | 'sail' | → | b-layar | (USE-sail) | 'to sail' |
| crmin | 'mirror' | → | b-crmin | (USE-mirror) | 'to look in a mirror' |
- (202) ?mot praho? b-kayuh tə?en jŋ̩ dak
get.onto boat USE-paddle TO:down foot water
(He) got into the boat (and) paddled down to the lower reaches.
- (203) b-layar=cə? kəh
USE-sail=EM 3S
He sailed.

The following example, which could be classed either under a) as 'HAVE' or here as b) 'USE', bridges the semantic link between the two meanings.

- (204) smə? knək kmpən br-dəl de=ba?
- person husband wife USE-house 3plA=be.alone
The married couple had a house by themselves.

Verbs derived with *b(r)-* can be used in a serial construction with the verb *swak* 'to go', *swak* + *b(r)-N* 'to go N-ing', deriving a process (§11.3). This parallels the construction involving *suwak+* imperfective verb, *swak* ?h-?oh (to go, IMPERF-shoot with a blowpipe) 'to go blowpiping' (§§5.3.1-2), and with verbs derived from nouns: *swak* + *mN-V* 'to go V-ing' *swak* + *m-rapah* (to go, PERFM-forest.litter) 'to forage, be nomadic' (§5.4.2).

- swak b-prajkpəp (go HAVE-trap) 'to go trapping'
swak b-klubi? (go HAVE-Salacca.conferta) 'to go seeking Salacca fruit'
- (205) ?le?le? ?are? muy ga=swak b-dre
then day one IMM=walk HAVE-rattan
Then one day (he) was going to go rattan collecting.

Comparative note *br-* 'HAVE/USE' is from Malay *ber-*. In Malay it functions to derive verbs from nouns 'to express possession or use of NP' (*basikal* 'bicycle' → *ber-basikal* 'to cycle'); 'to emit NP' (*telur* 'egg' → *ber-telur* 'to lay an egg'); 'to work in NP' (*kedai* 'shop' → *ber-kedai* 'to work in a shop') and 'to address as NP' (*adik* 'younger sibling' → *ber-adik* 'to call a younger sibling') (Liaw 1985: 118-20).

5.4.3.3 Reciprocal *br- + N*

The reciprocal is formed from verbal roots, by reduplicating the verb and prefixing *b-* 'MID' (§5.3.3.2). This process is also used with nominal forms to derive reciprocal verbs 'USE'. It has only been attested with three forms – the second person pronouns *kd* '2f', and *ji* '2' to derive reciprocal verbs 'to address each other with the second person form': *b-kd-kd* 'to address each other informally with *kd*' ← *kd* '2f' → and *b-ji-ji* 'to address each other formally with *ji*' ← *ji* '2'. With the noun *rodəŋ* 'companion' it derives an idiomatic form: *rodəŋ* 'companion' → *b-rodə-rodəŋ* 'to be engaged to marry each other'.¹⁶

5.4.4 Causative 'EQUIP'

Causative morphology, usually associated with verb to verb derivation, is also used to derive verbs from nouns. The range of causative affixes available to this process are: *par-*, *cər-* and *p-*, with selection based on the syllabic structure of the nominal. In a single example the manipulative causative prefix *tar-* derives: *tar-muy* 'to marry a couple' ← *muy* 'one'.

The derived verb is transitive and means 'to equip or provide with N', or 'to cause s.o. to have or use N', where N is the nominal which feeds the derivation. It is the causative equivalent of 'to HAVE/TO USE N'.

The derivation of causative verbs from nouns does not feed further derivations.

The correspondence between the form of the morpheme and the semantic groups in the following list is coincidental. The first three words are Semelai, the remainder are loanwords from Malay:

¹⁶ Cf. Pahang Malay *rodong* 'husband' (Wilkinson 1927).

rɔ?	'basket'	→ pan-rɔ?	'to provide with a basket'
suk	'hair'	→ par-suk	'to provide hair'
clon	'back'	→ ca<r>lon	'to turn one's back on (s.o.)'
muka?	'face'	→ mu<r>ka?	'to show one's face'
srwpl	'pants'	→ p-srwpl	'to put pants on (s.o.)'
ragi?	'pattern'	→ p-ragi?	'to decorate (s.th.) with a pattern'

There are two potential patterns for coding the causee, which are not interchangeable: a) as an object, or b) as a prepositional complement. Examples of the first pattern are provided in (206)–(207), where the causer is encoded as the subject and the causee as the object.

- (206) kadeh p-kasot knən?
who EQUIP-shoe offspring
Who put shoes on the child?

In (207) the subject has zero representation.

- (207) loc pa-dak perej
already EQUIP-water plate
(I've) already washed the plates.

The second pattern is to encode the causer as the subject, and the causee/beneficiary as a prepositional complement (208)–(209). Despite the prepositional phrase, the derived verb functions like a normal transitive verb, attracting subject proclitics:

- (208) de=pa-dəl tʃah ʔatə ha? kəhn
3A=EQUIP-house amidst earth AT 3S
They housed him down on the ground/out in the open. (Lit. They housed him amidst the ground.)
- (209) *She equipped them with a stick. The blind husband held onto one end, she took the other in order to lead him home.*
ki=p-dəŋ, mə=pŋəl tet knək, mə=pŋəl tet hn=kəhn
3A=EQUIP-stick one=side TO:spec husband one=side TO:spec O=3S
She provided (them) with a stick, one end to him, the other end to herself.

Causatives derived from nominals do not feed further verbal derivations, for example the prefix t(r)- 'HAPP': *t-pa-dak (HAPP-EQUIP-water) 'to happen to wet' is unacceptable. With verb to verb causatives, on the other hand, the t(r)- 'HAPP' prefix is available: cēc 'to wet' → t-pr-cēc 'to happen to splash' (§5.3.7).

5.5 Verbal clitics

There are three verbal clitics encoding categorial distinctions of subject, person and number (§5.5.1), mood (§5.5.2) and aspect (§5.5.3). The morphological specifics of clitics are discussed in §3.3.

The following sections provide a summary of each of the morphemes with respect to the verb. Although they form a phrasal constituent with the verb, they have wider scope, and therefore the main discussions of these categories are located elsewhere as indicated below.

5.5.1 The pronominal proclitic

The pronominal proclitic encodes person, number, and deference (§6.1.1). A table of the forms and the corresponding free forms can be found in Table 6.1. The domain of attachment of the proclitic is discussed in §3.3.1.

The pronominal proclitic forms an integral part of the transitive verb phrase where its primary function is to represent the agent NP, which may also be realised by an external post-verbal agent NP (§9.1.3.1):

$$\text{Pro}_1=V(A=\text{NP}_1)(O)$$

In (210)–(212) the proclitic cross-references the NP **k**<r>**dor** (be.female<NMZ>) 'woman', itself procliticised by the agent-marking proclitic **la**= 'A'. The cross-referenced constituents are in bold:

- (210) **ki=pŋke?** hn=dəŋ **la=k**<r>**dor**
3A=pick O=wood A=be.female<NMZ>
The woman picked the branch.
- (211) **sma?** **ki=gəp** **la=kubun**
person 3A=bite A=flying.lemur
The flying lemur bit the person.
- (212) **de=ca** **la=dəh**
3plA=eat A=3pl
They ate (it).

Although the proclitic is associated primarily with the transitive clause, it may also occur on a subset of lexically intransitive verbs, expressing motion, emotion or body actions (§5.5.1.2). This construction signals an event which is induced by an indirect external cause (§9.3.3).

- (213) **ki=cən**
3A=be.happy
She is happy.

Coding strategies are discussed in §9.1.

The presence of pronominal cross-referencing is not dependent on the external NP being referential, as shown in the following example:

- (214) **de=?ye-i?** **la=sma?**
3A=see-ITER A=person
People will keep on seeing (you).

However there is also a pronominal clitic **ko**= '3UA', which indicates an unidentified agent or agents, that may be either human or non-human, animate or even inanimate, although autonomous agency must prevail.¹⁷ There is no corresponding free pronominal form available for intransitive verbs. No external NP is possible in conjunction with this pronominal form.

¹⁷ This category is not to be confused with the indefinite pronouns (Schachter 1985: 30) which have 'unspecified' reference like German *man* or English 'one'. This is conveyed either by zero representation, or the use of the generic noun *sma?* 'person'.

- (215) *ko=tikam imbu?* hitam
3UA=stab bovine be.black
Someone stabbed the black bull.

- (216) *ko=glnq* ?luc
3UA=swallow pass
(Something) swallows (you) down.
[This is a phrasal lexeme from the avoidance style meaning 'to bathe'.]

In contrast to the free pronouns, which can only be used with animate reference, the referent of the third person bound form can be inanimate provided it is an entity which is potent, i.e. has the ability to impact upon something else, e.g. (216) above. As a result, other than humans (217), including their manifestations as ghosts, and animals (211) above, the proclitics are generally restricted to references to anthropomorphised inanimate entities (216) above, potent natural phenomena (218)–(219), bodily excreta like faeces and saliva (220), and blades and bullets (221).

- (217) ?le? da? hc=ca jsok
long.time NEG 1&2A=eat sambhur
We haven't eaten sambhur deer (in a) long time. (It's a long time (since) we ate sambhur.)
- (218) *ki=kantop* la=dloŋ pat^{bir} ke
3A=close,over A=tree tree sp. that
The pat^{bir} tree closed over (him).
- (219) *ki=ca* krntah
3A=strike lightning
The lightning (of the taboo) struck (it).
- (220) ?eç knon ye, da? ki=moh kph
faeces child 1 NEG 3A=want come.out
My child's faeces don't want (to) wash out (of the nappy).
- (221) *ki=ca* pruluh
3A=penetrate bullet
The bullets penetrate him.

Metaphorical agency is restricted to the third person singular form *ki*= '3A='. The metaphoric usage is associated with certain potent natural phenomena, including *ribut* 'wind', *krntah* 'lightning', *?us* 'fire', *mata?are?* 'sun' and *dak* '(flood)water', but excludes entities such as *2are2* 'rain', *2are2.b-ribut* 'storms', *guruh* 'thunder', *kabut* 'clouds', *?litol* 'mist' and other celestial bodies like *bulan* 'moon' and *bintang* 'star'. The verbs associated with natural phenomena are *tut* 'to blow', *p^{bil}* 'to rise, of wind', *ca* 'to consume (fire)', 'to strike (lightning)' and 'to penetrate (bullets and blades)', and *ryoh* 'to make a noise'.

- (222) *xi=tut* la=ribut
3A=blow A=wind
The wind blew (it).

The cross-referenced post-verbal NP, whether plural (223), or singular (224), is not always coded by *la*= 'A':

- (223) *de=ca* knte?
3plA=cat flea
The fleas had a meal (of me). (Lit. The fleas ate (me).)

- (224) *ki=ca* ?us
3A=consume fire
The fire consumed (it).

Occasionally the number of the agreement marker in the third person is not the same as that of the external NP which it cross-references; the agreement marker will be singular, the NP plural. This often occurs with inherently plural NPs:

- (225) *ki=beŋ* la=mntuhp?
3A=check.on A=parent.in.law
The parents-in-law checked on (her).

Compare (226) where the same NP is cross-referenced by a plural proclitic:

- (226) *de=lsh* ber la=mntuhp?
3plA=go:across check.on A=parent.in.law
The parents-in-law went across to check on them).

5.5.1.1 The proclitic and the transitive verb

Essentially, the presence of the proclitic on the transitive verb is indicative of an individuated event. Compare the clause in (227) where there is no pronominal, with that in (228), in which the verb hosts the proclitic. In the first example where the event is not individuated the clause has a universal or generic reading:

- (227) co jəl jkɔs
A V O
dog bark.at porcupine
Dogs bark at porcupines.

In the second example, an actual specific event is understood as taking or having taken place. The individuation of the event also implies an individuated object.

- (228) jkɔs ki=jəl la=co
porcupine 3A=bark.at A=dog
The dog barked at the porcupine.

5.5.1.2 The proclitic and the intransitive verb

Two classes of intransitive verbs may mark the subject with a pronominal proclitic. In the unmarked situation the single argument of an intransitive predicate is not cross-referenced on the verb (§9.3.2).

- (229) ?yam k<r>dor
cry be.woman<NMZ>
The woman is crying.

The presence of the proclitic on an intransitive verb marks the event as atypical. It appears to express events which are induced by an indirect external cause.

- (230) ?are? ke da? ye=dos, ?are? ke b-khon
day that NEG 1A=come day that HAVE-offspring
The day we didn't go, that day (she) gave birth.

Verbs which take the proclitic in this function are listed below. They are stative verbs expressing either motion or body actions and human emotions and conditions. The list is exhaustive in terms of data collected so far.

<i>Motion verbs</i>	<i>Emotion verbs</i>	<i>Body action verbs</i>
brhiuh 'to run'	c̪en 'to be content'	d̪m-d̪em 'to be lying down,
paloh 'to flee'	ŋren 'to be angry'	to lie down'
grek 'to fall'	risaw 'to be sad'	km-kʰom 'to be sitting, to
swak 'to walk'	kali 'to be brave'	sit down'
dos 'to arrive'	malu? 'to be shy'	?yam 'to cry'
?yot 'to return'	shjeh 'to be reluctant'	gl̪k 'to smile, laugh'
plec 'to emerge'		k2-kʰu?
		'to vomit'
		jr-jor 'to urinate'
		wh-wəh 'to rise, get up (from lying)'
		jtek 'to sleep'
		kʰbas 'to be dead'
		?jih 'to be ill/in pain'

This construction is used when the action is instigated by an external source. It is not itself a causative construction (§5.3.5), given that the verb lacks causative morphology and the valence of the clause is not increased. Examples (231)–(232) compare the 'intransitive imputed cause' construction (231), with the morphological causative in (232). In (231), the person's singing co-occurs with the child falling asleep. In (232), the agent is understood as having patted or rocked the child to sleep.

- (231) ye bpapi?, ?arən ki=jtek
I sing new 3A=sleep
I sang, (and) just then the child (went) to sleep.
- (232) ye=j<r>tek knon
1A=sleep<CAUS> child
I put the child to sleep.

The motivation cited by a speaker for the use of this construction is explained in the following context: this form would be used for instance, if one had invited someone to come to one's place. When they consequently arrived, one would express it in the construction in (233) or alternatively that in (234).

- (233) ki=plec (234) ki=dos
3A=emerge 3A=come
He emerged. He came.

Had the person just turned up out of the blue, at their own whim, the unmarked intransitive construction would have been employed:

- (235) kahn dos
3S come
He has come.

The cross-referencing post-verbal NP is present in (236). In this clause the purpose of the woman's actions are dictated by some requirement external to her.

- (236) A woman, who has just given birth, leaves the hut to get boiling water to bathe her belly.

ki=ainte la=krdor ke, dak jnraq ke
3A=exit.house A=woman that water boiling that
The woman went down out of the house (on account of) the boiling water there.

- (237) A man's employment has taken him to many locations throughout the country, as his niece explains:
s<n>broh ki=dos
all<NMZ> 3A=arrive
He has been everywhere.

5.5.2 Irrealis ma= 'IRR'

The irrealis is expressed by the proclitic ma= 'IRR' on the transitive verb. The morpheme occupies the position normally held by the pronominal proclitic, the two morphemes being mutually exclusive, as illustrated in §3.3.2, and discussed in §10.2.

The irrealis mood is not marked morphologically for intransitive verbs, with the exception of the class of intransitive verbs which may take transitive proclitic marking, and some stative intransitive verbs expressing inherent qualities of inanimate entities.

5.5.2.1 Irrealis with transitive verbs

Any semantically compatible transitive verb can host the irrealis proclitic ma= 'IRR', including a) causative, b) iterative and c) happenstance stems:

- a) ma=pan-ca (IRR=CAUS-eat) 'would want to eat'
- b) ma=tut-i? (IRR=blow-ITER) 'would keep on blowing'
- c) ma=t-genen (IRR=HAPP-disturb) 'would happen to disturb'

Unlike the pronominal proclitic, the irrealis morpheme does not co-reference an external NP, the presence of ma= 'IRR' being associated with the suppression of the agent in the clause: ma=V (O).

Clauses marked for the category of irrealis ma= 'IRR' may contain a single participant, the object (§10.2.1):

- (238) no2no2, ma=kʰom spoloh ryal
this IRR=get ten dollars
(For) this, (I) might get ten dollars.

Verbs marked for the irrealis are employed in the expression of deontic modality in transitive clauses. These are always embedded under a negative expression (§10.2.3):

- (239) da? n=jon ma=yɔk crsh
NEG IfA=permit IRR=fetch fish
I did not permit (you) (to) collect fish.

5.5.2.2 Irrealis with intransitive verbs

There are two sets of intransitive verbs which may be marked for the irrealis. The two groups exhibit slightly different consequences.

a) A subset of intransitive verbs which may take the pronominal proclitic without any overt morphological change to the verb, may also be marked for the irrealis. In this case, there are no arguments expressed (§10.2.1).

- (240) ɲren, da? dapt ma=ɲren.
be.angry NEG able IRR=be.angry
Angry, they weren't able (to) get angry.

Some of the verbs which are eligible in this construction are:

bt ^b ɔŋ	'to be afraid'	→ ma=bt ^b ɔŋ	'would be afraid'
risaw	'to be sad'	→ ma=risaw	'would be sad'
kali	'to be brave'	→ ma=kali	'would be brave'
ɲren	'to be angry'	→ ma=ɲren	'would be angry'
crdek	'to be clever'	→ ma=crdek	'would be clever'

b) The irrealis marker is also used with certain stative intransitive verbs to express an inherent potential quality of an inanimate object (§10.2.2.1). In this case, there is no reduction in the number of expressed arguments in the clause, giving the appearance of an attributive rather than a predicated phrase: N ma=V.

- (241) ppakit ma=k^bbəs
illness IRR=die
fatal illness

The quality must be inherent, for instance, tea is naturally bitter, and thus tea, even if it were sweetened with sugar, could not be described as ma=spek 'to be able to be sweet'. On the other hand, certain fruits are inherently sweet, (at least they have the potential once ripe), and therefore can be described as being ma=spek 'to be able to be sweet'.

Some of the verbs which are eligible in this construction are listed below. It is not always easy to provide an elegant gloss, but the sense here is of 'having the potential to be V', although the likelihood of the event taking place is more certain than the potential suggests.

The relations which hold between the predicate and the participant in the clause are complex – the NP may be an object (ca 'to eat'), a subject (susah 'to be difficult'), or a loosely involved entity (k^bbəs 'to die').

ca	'to eat'	→ ma=ca	'to be edible'
bul	'to be intoxicated'	→ ma=bul	'to be toxic'
k ^b bəs	'to be dead'	→ ma=k ^b bəs	'to be fatal'
susah	'to be difficult'	→ ma=susah	'to have the potential to be difficult'
guc	'to burn'	→ ma=guc	'to be flammable'
srut	'to smoulder'	→ ma=srut	'to have the potential to smoulder'

- layur 'to be withered' → ma=layur 'to have the potential to wither'
ʔpih 'to be ill, in pain' → ma=ʔpih 'to have the potential to produce illness'

Two transitive verbs of perception, ʔyəŋ 'to hear' and ʔye 'to see' also host ma=IRR with this function:

- ʔyəŋ 'to hear' → ma=ʔyəŋ 'to be audible'
ʔye 'to see' → ma=ʔye 'to be visible'

5.5.3 Imminent aspect ga= 'IMM'

The proclitic ga= 'IMM' expresses 'imminent' aspect, defined here following Frawley (1992: 322) as representing 'a point just prior to the beginning of an event, ... an event on the verge of obtaining'.

- (242) ʔare? ga=lsəm
rain IMM=fall
Rain is going to fall.

In clauses containing an agent, ga= 'IMM' takes on more of an intentive reading:

- (243) ga=kⁱ-jʔɔy sŋkalan
IMM=3A=make grinding.board
He was going to/intending to make a grinding board.

Given the intentive nature of the aspectual clitic, it cannot co-occur with verbs marked for happenstance (§5.3.7), or external imputed causative clauses (§5.5.1.2).

The aspectual clitic occurs in verbal clauses containing activity or process verbs. In the transitive clause ga= 'IMM' is proclitic to the verb, preceding the pronominal proclitic:

- (244) kəh, da? ga=ki=tɔŋoy wɔ? ha? del
3 NEG IMM=3A=wait.for longer AT house
Her, she wasn't going to wait for (him) at the house any longer.

In the intransitive clause, ga= 'IMM' is proclitic to the verb:

- (245) kəhn ga=hūm
3S IMM=bathe
He is going to bathe.

- (246) knən ga=grək
offspring IMM=fall
The child is going to fall.

- (247) y^te ga=kmɔŋ mn-(t)ulls
1 IMM=finish IMPERF-write
I'm going to finish writing.

The aspectual proclitic is also attested in non-verbal clauses where it is hosted either by a predicate nominal (248), or a predicative locative prepositional phrase (249). In these clauses the imminent aspect marker ga= 'IMM', functions to predict the identity, or location of the subject in a time frame subsequent to that of the current discourse (§10.1.1.1).

- (248) "jake" k^bəŋ ʔmaʔ, "[ga=knlək kp sarek]pred"
 then QUOTE mother [IMM=husband 2f future]pred
 "In that case," said (her) mother, "(he) will be your husband in time to come."
- (249) ga=tē hɔ̄n ji s<n>wak?
 IMM=TO:unspec where 2 walk<NMZ>
 Where are you going to (with) (that) walking?

In the verbal clause, the aspectual clitic marks an event on the verge of obtaining. Subtleties in meaning are achieved in combination with the aspectual properties of the host verb. Take for instance the example in (251), where the morpheme is proclitic to a transitive verb. The speaker contextualised the example thus: the protagonist has everything ready and is about to leave the house in order to go planting.

- (250) ga=k1=ptəm
 IMM=3A=plant
 He is about to plant.

The point which the speaker is making in relation to (250) is illustrated clearly in (251), where the woman has both the knife and the mango and is about to peel the fruit.

- (251) ki=roc, ki=gɔ̄ŋ masuk, ga=k1=gɔ̄s hmbacaj
 3A=pull.out 3A=take enter IMM=3A=peel *Mangga foetida*
 She pulled (the knife) out, she took (it) inside, (and) was going to peel the horse mango.

In conjunction with a verb in the imperfective form, ga= 'IMM' indicates an event which is merely contemplated. The event is yet to take place. The same speaker provided the following context for the next example. The protagonist is sitting at home intending to plant, but nothing has been prepared toward his execution of the activity:

- (252) kəhn ga=p<m>təm
 3S IMM=plani<IMPERF>
 He is going to be planting.

ga= 'IMM' can co-occur with the aspectual adverb ?areh 'just, recently'. When the adverbial ?areh 'just' modifies a predicate with the aspectual proclitic ga= 'IMM' it marks a point just prior to the beginning of the subsequent event (§10.3.1.1).

- (253) ?le?le?spn ?areh ga=tunes
 then=SC just IMM=shoot
 Then (they saw some plants) just about going to shoot.

5.6 The modal category

The predicate may be modified by a modal category, expressing either necessity or possibility. This is achieved either by an adverbial like msti? 'must', a verb in either a complement construction or a serialising construction, dependent on the particular

verb involved. In light of the diverse syntactic means employed to express this category, the modals are treated systematically here.

Modal auxiliaries are positioned in relation to the verbal word, generally directly preceding it as in the following examples:

- (254) kəhn lən [ca]
 3S want [eat]
 He wants to eat. (i.e. He is hungry.)
- (255) lən [ki=ca]
 want [3A=eat]
 He wanted (to) eat (it).

The modal must occur either in conjunction with a predicate, as in the first clause in (256), or where it has been previously implied, as in the second clause in (257).

- (256) beh, da? de=məh mnj-glor
 NO NEG 3plA-want IMPERF-bestow.name
 No, they weren't willing to bestow a name.

- (257) *The Sultan's grandchild has indicated that the Pauper is his father, and so the Sultan (A) declares that he has to marry his daughter to him. The Pauper replies in B.*

- A: "msti?", k^bəŋ, "n=n̩i<r>ka?" rəm hn=kp"
 must QUOTE 1fA=marry<CAUS> WITH O=2f
- B: "?ec beh" k^bəŋ, "da? məh"
 oh ND QUOTE NEG want
- A: "I must," (he) said, "marry (her) to you."
 B: "But no," (he) replied, "(I'm) not willing to."

5.6.1 Necessity

The category of necessity is expressed either by msti? 'have to' (§5.6.1.1), various uses of lən 'to desire', or məh 'to want, be willing' (§5.6.1.2).

5.6.1.1 msti? 'have to, must, be obliged'

msti? 'have to' (← Malay *mesti* 'must, needs must') expresses the expectation of an outcome given the set of prevailing circumstances.

- (258) ?mot ?pəs, msti? ki=gɔ̄ŋ grək
 mount be.high must 3A=bring fall
 (If) (he) climbs up high, he must make (himself) fall.

msti? 'have to' never occurs as a main verb, cannot be negated, nor stand alone without a predicate. As the example indicates, it does not immediately precede the verbal word. The presence of the quotative marker (see §13.1.1 for quotative placement) between msti? and the predicate in (257) above, suggests that it does not form a constituent with the predicate, but rather should be considered as a type of adverb.

Expressions of lack of obligation are discussed in §10.2.3.2:

5.6.1.2 Verbs of wanting as modal auxiliaries

lən 'to desire' and *məh* 'to want' both function as transitive main verbs, taking either an object NP (259) or a clausal O (260).

- (259) daʔ n=lən hn=kəh
NEG 1fA=want O=3
I don't desire him.

In (260), the main verb is *lən* 'to want', evident from the fact that it hosts the proclitic pronominal, and the second verb, *?yot* 'to return', does not determine the transitivity of the clause (see complement clauses §11.2).

- (260) sampay ki=məh [?yot leg bapa?]
so 3A=want [return TO:up father]
So he (was) willing (to) return up to his father.

Similar considerations apply to the following clauses.

- (261) daʔ kə=jən=cə? [ca savor knkən]?
NEG 2fA=want=EM [eat vegetable children]
Don't you want to eat the childrens' vegetable dish?

In their function as modal verbs, the verbs *lən* 'to desire' and *məh* 'to want' exhibit different syntactic behaviour, as demonstrated above in (260)–(261) where they precede the main verb. It is difficult to say how this should be interpreted, given that it is difficult to establish a sharp semantic contrast. The motivation behind the different patterns of distribution appears to relate to the realisation of the proposition.

a) In the first set of clauses, the proposition is either realised (262), or has been attempted, but not succeeded (263). The sense of the verb in this case appears to be related to an expression of intention, or lack of it in (262), or even 'necessity', as in (263).

- (262) daʔ məh [he=kna?]
NEG want [1&2A=catch]
We didn't want/intend (to) catch (it).
We didn't willingly catch (it).

- (263) A man wants to see his estranged wife, and needs to get up into the house to do so.
He tries, but to no avail.
lən ki=?endəl-i?, daʔ ki=kəm
want 3A=enter.house-APPL NEG 3A=succeed
He wanted to/needed to enter the house, (but) he didn't succeed.

In an intransitive clause, it is unclear whether the clause should be interpreted as the desire to drink, which may also be understood as the need to drink, if in fact the two are different:

- (264) A woman and tiger are travelling through the forest, without provisions. She expresses a desire to drink.
"beh, lən j<h>zəh," kʰləy "ʔəp"
NO want drink<IMPERF> QUOTE If
"No, I want to/need to drink," (she) said.

In (265) the modal is embedded in a causal clause, and expresses an unrealised desire.

- (265) When a child was born, he slipped off the verandah and disappeared. His parents have summoned seers to retrieve the child.
kira?, dom, lən tʰar-luc knkon kke,
see AFF want CAUS-pass child that
la=lən de=krwpi-i?
BCS=want 3plA=emerge-APPL
See, indeed (they) wanted to release the child there, because they wanted to bring (him) out.

In contrast to the preceding examples, in the following clauses which express probability, the narrator surmises that the event is likely to be realised:

- (266) ki=lagkah tupay rəm kʰoy, ma=swak jor-i? he?
3A=step.across squirrel WITH head IRR=move urinate-APPL AT
kʰoy, ttap lən ma=gren, pʰia? da? pasal!
head be.definite want IRR=be.angry animal NEG proper
The squirrel bounded across via her head, in order to go (by) to urinate on her head.
Definitely (she) must have wanted to get/got angry, (such) an improper animal!

b) The auxiliary *lən* 'to want' is used in the following construction to express purpose, where it expresses intention: 'if you want, need to do X, do Y':

- (267) The speaker is explaining to me that in order to speak Semelai, the word for a birth-vampire is matiʔana?, and not the Malay term puntianak.
lən cəl smlay, matiʔana?, tida?=hn pontianak?
want speak semelai matiʔana? not=CONN puntianak
(If) you want to speak Semelai, (it's) matiʔana?, not puntianak?

- (268) sggoŋge bmban wants to go up to the home of the sma? ʔatas ʔagen (person above wind) 'supernatural being'. The old man has given him a set of clothes which will allow his flight.
daʔ pakayan, lən ?yot leg ke
EXIST clothes want return TO:up that
He had the clothes (necessary) to return up there.

c) *lən* also functions with its modal sense to introduce the complement in ablitative and knowledge predicates (269), see §11.2.2, and in some information interrogatives (270), see §10.5.2:

- (269) ki=nam lən jtək
3A=feel want sleep
He felt (he) wanted to sleep. (He felt sleepy.)

- (270) "mandeh=pa hns lən b-c?ɔŋ?" kʰləy "knlək"
why=FACT THEN want MID-burn QUOTE husband
"What reason does your husband want it burnt?" (he) pondered.

5.6.2 Possibility

Possibility may be expressed in a number of ways, either by the auxiliary sot 'can, be able', by a complement-taking verb *dapat* 'to get, to manage to', or *k^bom* 'to get, to manage to'.

5.6.2.1 sot 'can be able'

The lexeme sot has three senses relating to the expression of possibility: permission, (or more usually the denial of it, i.e. prohibition), possibility and ability. It is used in contexts where an external force determines the situation, e.g. a taboo (271), or social obligation. Elsewhere, sot occurs in a minor clause type, with the meaning 'to metamorphose into' (§10.1.2.2).

Denial of permission is expressed in terms of what one must not do, and this is achieved by sot embedded under a negator. da? sot 'not permitted to' is the most frequent construction involving sot. It is discussed in §10.2.3.2 in relation to deontic modality. The non-negated use of sot is infrequent.

The construction is usually used to express general prohibitions, and so there is no subject expressed (271)–(272):

- (271) ma=tulah, da? sot cal glar gade?
IRR=taboo NEG permit utter name parent
It's the tulah taboo, (one's) parents' names shouldn't be uttered.

The following example, which expresses denial of permission, shows how the proclitics of the transitive verb do not percolate up to sot, indicating that this is an auxiliary verb:

- (272) da? sot ma=pcre?, ki=c^bɔŋ 2en t^bɔŋ
NEG permit IRR=disturb 3A=roast LOC ear
(One) is not permitted to/should not disturb (the crow), it (will) cause scabs in the ear. (Lit. It will roast the ear.)

sot is also used in specific contexts, the subject either preceding or following the clause:

- (273) 2əŋ da? sot mn-jamah
I NEG permit IMPERF-taste
I am not permitted to taste.

- (274) ye 2əŋ b-khon, da? sot t^bɔŋ-təŋ ye
1 recent HAVE-offspring NEG permit IMPERF-scrve.food 1
I had recently given birth, (and so) I was not able to serve food.

The discontinuity in the following example, where kira? 'figure' intervenes between the modal and the main verb, indicates that this, like msti? 'must', is not a serial verb. The non-negated use of sot expresses 'permission or ability':

- (275) A man is able to marry the woman now that he has gotten rid of her husband.
sot kira?, j^bɔy kmpan kəh
permit figure make wife 3
(He) was able, see, to make (her) his wife.

5.6.2.2 *dapat* 'to get to'

dapat 'to get to' (← Malay *dapat* 'to manage to, to get to). In Semelai it overlaps semantically with *k^bom* 'to get; succeed', a complement-taking achievement verb (§11.2.2). Unlike *k^bom*, *dapat* is not attested with proclitic marking. *dapat* 'to get to' is usually used in conjunction with a negator where it indicates a lack of possibility due to an external source. In (276) the reason behind the inability is a pantaj 'prohibition' and it is translated as possibility in accordance with this.

- (276) Following the birth of a child, there are prohibitions against certain activities. The woman is telling how she was confined as a result of this:

- da? dapat ye=gɔŋ masak
NEG get.to 1A=bring cook
(It) wasn't possible (that) I engage (myself) cooking. (i.e. I didn't get to cook.)

- (277) da? dapat ma=j^bɔk
NEG able IRR=observe
(She) would not have got to witness (that).

- (278) "da? dapat t-paloh," k^bləŋ
NEG succeed HAPP-flee QUOTE
"(They) weren't able to flee," (he) said.

6 Pronouns: personal, ignorative, and demonstrative

In §6 three types of pronouns are discussed. In §6.1 the discussion begins with personal pronouns. Ignorative pronouns are introduced in §6.2. The term ignorative is adopted following Wierzbicka (1980) for the multifunctional class of wh- words and indefinite, negative and free-choice pronouns. The chapter closes with a section on demonstratives in §6.3. Semelai has a basic two-term proximal/distal distinction, with finer distinctions conveyed by the adnominal use of deictically specified prepositions.

6.1 Personal pronouns

Personal pronouns are used for animate entities, human and supernatural beings, 'anthropomorphised' animals and personified inanimates. Noun phrases understood from context do not need to be expressed: there is a high occurrence of zero representation of noun phrases (§9.4). This is particularly true for inanimate entities which cannot be expressed pronominally. Alternative strategies for encoding inanimate entities are to use a demonstrative pronoun or a nominalised locative preposition (§6.3.2).

6.1.1 Semantic distinctions in the pronominal system

Semelai has both free pronouns, and bound pronominal proclitics, both of which are set out in Table 6.1 below. The bound pronouns cross-reference the subject in the transitive clause, and while free personal pronouns are necessarily referential, this is not the case with the third person bound forms; the unidentified agent clitic *ko*= '3UA' is evidence of this.¹ The pronouns are distinguished for person, number, deference and identifiability, although none of these categories applies across the full paradigm: for example, deference is only relevant to the first and second person minimal forms, both bound and free, and number is often ambiguous in first and second person bound forms.

¹ Although I follow tradition and refer to the bound forms as 'pronouns', they should in fact be labelled differently. (Lazard (1997: 247) suggests 'verbal actant markers' or 'actant indexes'. See Evans (2001) for similar observations.) Although the bound and free pronouns share common forms, and make similar semantic distinctions, the function of the two is very different. It is for this reason that the functions of the clitic forms are discussed in §5.5.1 in the chapter on verbal morphology and in §9.3.1.1.

6.1.1.1 Person

There is an inclusive/exclusive distinction in the first person: the exclusive *ye*=?en '1=AUG' is the augmented form of the minimal pronoun *ye* '1'; the inclusive forms are *he* '1&2' and *he*=?en '1&2=AUG'.

The third person clitic pronominal forms make a distinction between identifiable and unidentifiable agent reference, *ki*= '3A' and *de*= '3ptA' express the former, with *ko*= '3UA' used to express the latter.

The proclitics are discussed in §5.5.1.

TABLE 6.1 THE PRONOMINAL PARADIGM

		min familiar	min	aug	sg	pl	
PRONOUNS	1	?əŋ ₁	<i>ye</i>	<i>ye</i> =?en			
	2	<i>ko</i>	<i>ji</i>	<i>je</i> =?en			
	1&2			<i>he</i>	<i>he</i> =?en		
	3					<i>kəh</i>	<i>deh</i>
VERBAL 'A' PROCLITICS	3S					<i>kəhn</i>	<i>dehn</i>
	1	?əŋ ₁ = ²	<i>ye</i> =				
	2	<i>ko</i> =	<i>ji</i> =	<i>he</i> =			
	3					<i>ki</i> =	<i>de</i> =
	3UA						

6.1.1.2 Number

Number is clearly distinguished for free pronouns and the third person identified agent bound forms, but it is neutralised for first and second person bound forms. It is typologically unusual for number to be obligatory for third person but not for first or second (see Croft 1990: 111; Aikhenvald and Dixon 1998: 66). The plural forms of the free pronouns for first and second person are marked by the enclitic =?en 'AUG'; note that the first and second person familiar pronouns do not have a plural form. The inclusive first person form *he* '1&2' is inherently dual and speakers define it as *he b-dwa?* (1&2 BE-two) 'we are in a group of two'. The encliticised free form *he*=?en '1&2=AUG' indicates an augmentation of the basic form, which speakers explained as *he*=?en *ramay* (1&2=AUG be.many) 'we are many'. The third person has distinct singular and plural free pronominal forms: *kəh* '3' and *deh* '3pt'.

- (1) *ye*=?en *ye*=?en *ma=duga?* *req*
 1=AUG 1A=want IRR=try seek
 We want to try to look for (him).

The third person singular can be used for a group of people performing the same action, rather than the third person plural:

- (2) *sparoh* *paloh*. *ki=b?* *hn=kno?*
 some flee 3A=carry.on.back O=offspring
 Some fled. They carried (their) children on (their) backs.

6.1.1.3 Deference

Deference is encoded for first and second person minimal pronouns, but not for third person pronouns. It is based on the relative age and familiarity of the speaker and

² The first person singular familiar form ?əŋ₁ has the allomorphs *m*=, *n*=, *p*=, *g*=, a reduction of the final consonant of the pronoun to a nasal homorganic with the initial consonant of the verb (see §3.3.1).

addressee. The use of the deferential forms differs within the community. The standard usage is to employ the politer forms with anyone more than one generation older than ego, related or otherwise, and with unrelated individuals of casual acquaintance. The familiar is used within ego's age group, and with family, taking into consideration the stipulation regarding age. In some families, the polite form is only used with people with whom one is not on familiar terms, so that family elders are addressed with the familiar forms, the polite forms being specifically for outsiders and casual acquaintances. Some families use only the familiar form in all situations.

There are derived reciprocal verb forms, b-*kø*-*kø* (USE-2f-2f) 'to address each other in the familiar form', and b-*ji*-*ji* (USE-2-2) 'to address each other in the polite form' (§5.4.3.3).

The third person singular *køh* '3', rather than a second person pronoun, is employed to address one's parents-in-law in accordance with an avoidance taboo (§1.3.3.2).

Use of free personal pronouns is often avoided in favour of address terms based on kinship classification, tekonyms or nick-names.

6.1.2 Functions of the free personal pronouns

Pronouns are always referential. They are used in the following cases, their distribution parallel to that of nouns.

a) To draw attention to a person (3) or as a vocative to address someone:

- (3) *køh!*
 3
 (It's) him!

- (4) *?ɔy*, *ji!*
 VOC you
 You!

b) As an external NP_A co-referential with the pronominal proclitic in the transitive clause (§9.3.1). This is the least common environment, given that the person is already coded by the proclitic:

- (5) *da?* *ki=kø?* *la=køh*
 NEG 3A=know A=3
 Him, he doesn't know.

c) As a first or second person argument of an intransitive clause (§9.3.2), either pre- or post-verbally:

- (6) *?əpi* *swak*
 If go
 I went.

- (7) *p?ot* *kø* *ha?* *no?*
 stay 2f AT here
 You stay here!

In the third person the distinctive S-marked forms *køhn* '3S' and *dehn* '3plS' have a semi-transparent relationship to the unmarked forms. Note that in (8) the pronominal post-verbal NP_S is expressed by the unmarked form.

- (8) *dehn* *swak*, *pðor* *køh*
 3plS go follow 3
 They went (and) he followed.

d) The post-verbal O in the transitive clause optionally hosts proclitic *hn=* 'O' (§9.3.1.2):

- (9) *ki=gøŋ* *køh*
 3A=bring 3
 He brought him.

- (10) *ki=bø?* *hn=køh*
 3A=carry.in.sling O=3
 She carried him in a sling.

e) The complement in a PP (11) optionally in conjunction with proclitic *hn=* 'ABS' (12) (§8.2 and §9.3):

- (11) *gøŋ* *tet* *ye!*
 bring TO:spec 1
 Bring (it) to me!

- (12) *huc* *ki=pan-ca* *ha?* *hn=køh*
 rice 3A=CAUS-eat AT ABS=3
 She fed rice to her.

f) The subject in a non-verbal clause (§10.1.1):

- (13) *køhn* *sma?* *jakun*
 3S person Jakun
 She (is a) Jakun.

g) In an associative construction (§7.5) as the possessor:

- (14) *kke* *kñon* *he*
 that offspring 1&2
 That (is) our child.

When the third person possessor is a core NP, or co-referential with a core NP, the possessor is expressed by enclitic =*hn* '3POSS' on the possessee:

- (15) *købas* *?ma?=hn* *ke*
 be.dead mother=3POSS that
 His mother died.

Free pronouns can be modified directly by a demonstrative (§6.3): *kø no?* (2f this) 'this you, you here'; *køh no?* (3 this) 'this he', 'him here'. Speakers consider it coarse to use a second or third person form modified by the proximate term.

The distal demonstrative is used to track reference to a past time, generally one with which the hearer cannot identify: *ye ke* (1 that) 'I (back) then'.

The third person forms can also be modified by complex demonstrative expressions *nɔ?* *neŋ* (this before) 'aforementioned', e.g. *kəh nɔ? neŋ* (3 this before) 'he who was just mentioned'; or *ha?* *neŋ* (AT before) 'present just now', e.g. *kəh ha? neŋ* (3 AT before) 'he who was just here now' (§6.3.2).

6.1.3 Determiner ?a= 'DET'

The third person free pronouns can host the focus clitic ?a= 'DET' to give the forms ?a=kəh 'DET=3' and ?a=dəh 'DET=3pl'. This form is used for post-verbal arguments encoding absolute pronominalis. Speakers described the form as meaning 'actually' or 'in fact him/them', suggesting its function is to mark the pronominal as referential, specific, etc., hence the gloss determiner 'DET'.³

- (16) ga=jɔy praho?, ga=b-bhan, ?a=kəh
IMM=make boat IMM=MID-hew DET=3

In fact it was him, (he) was going to make a boat, was going to hew.

Like free pronouns this form can also be modified by *nɔ?* *neŋ* 'aforementioned'.

- (17) ki=gɔŋ lumbat, pałoh ?a=kəh nɔ? neŋ
3A=take jump flee DET=3 this before
He carried (it) jumping, (and) fled, in fact the aforementioned him.

This cliticised form of the pronoun is frequently employed in response to a question. In this context, the pronominal can be encliticised with the interjection *hɔ?* 'ATTN': ?a=kəh=hɔ? and ?a=dəh=hɔ?. *hɔ?* was characterised by speakers as being polite or pleasant. It expresses an appreciation on the part of the respondent towards the other parties' expression of interest. Elsewhere, *hɔ?* is uttered on presenting someone with something, or as an interjection 'look out!', e.g. if someone is about to inadvertently step on something (§13.2).

- (18) *The person in question has gone to stay downstream.*

- Q: te hɔn knon ji?
TO:unspec where offspring 2
Q: Where is your child?

- A: swak ?a=kəh=hɔ?
go DET=3=ATTN
A: He in fact has left.

- (19) lən ?yot ?a=dəh=hɔ?
want return DET=3pl=ATTN
They in fact want to come home.

6.1.4 Possessive də= ~ ddə= 'OF'

də= ~ ddə= 'OF' attaches to free pronominals to form possessive pronouns. The clitic is a reduced form of the possessive preposition *də* 'OF' (§8.9), which is used with nouns or demonstratives; də= ~ ddə= is used exclusively with pronouns (§3.3.6). Some examples are listed:

³ In Temiar there is a prefix ?a- described as a definite marker (Benjamin 1976b: 162). It attaches to demonstratives, kin terms and personal names. In Khmer there is a pronominalising affix ?a- (Huffman 1970). Closer to home, ?amay, a Semelai speaker who had at one time lived with the neighbouring Southern group the Temoq, reported that in Temoq the free pronominal forms were prefixed with ?a '-; ?a-ye '1', ?a-ji '2'.

de=kə	(OF=2f)	'yours'	ddə=ʃəp	(OF=1f)	'mine'
de=yɛ=ʔen	(OF=1=AUG)	'ours, excl.'	ddə=kəh	(OF=3)	'his, hers'

These forms express possession or ownership of an entity, which has zero representation in the clause.

- (20) *A man has been looking for lice in his wife's hair. He suggests changing places:*
ye=wər reŋ de=kə
1A=instead seek OF=2f
I'll look for yours in (your) stead.

This form is usually used in response to the question:

- (21) Q: də kadeh də ? A: də?, ddə=yɛ=ṣən
OF who OF AFF OF=1=SC
Q: Whose is (it)? A: Indeed, (it's) mine.

6.1.5 Possessor focus pə= 'POC'

The proclitic pə= 'POC' is hosted by second and third person pronouns. The attention drawing enclitic =hɔ? 'ATTN' is optional. The precise function of the proclitic is unclear, but it appears to convey a sense of 'pertaining to ...', as a possessor focus marker. Speakers suggested it is an alternative to the possessive proclitic ddə= 'OF'. pə= is only used in conversation, not traditional narrative.

- (22) r<1>məl nehneh p=kəh
be.male<NMZ> only POC=3
Only boys hers (=babies).

- (23) ye=jɔpək creh ji nintəm. guc pə=ji c<ŋŋ>ɔŋ
1A=observe fish 2 yesterday burn POC=2 roast<NMZ>
I saw your fish yesterday. Yours was burnt (from your) roasting (of it).

6.1.6 dri? 'self'

dri? from the Malay *diri* 'self' functions pronominally either as a collective pronoun 'Iself', or to a lesser extent as a reflexive pronoun 'self' (26). As a collective pronoun, it may be used with either inclusive (24), or exclusive (25) reference.

- (24) *The hearer in this exchange, like the speaker, is also a mother.*
susah ?en dri?
be.difficult LOC 1self
(Motherhood) is difficult for us (women).

- (25) *The hearer, the same person as the hearer in (24), is not a Semelai:*
sma? dri?
people 1self
my/our people (as opposed to yours)

dri? 'self' may also be a reflexive pronoun. It is not widely used in this function, and is only attested with third person reference:

- (26) wen dri⁷ tom ?en k⁸oy taja? den ke
throw self SRC LOC head stair like that
(She) threw herself from the top (of) the stairs like that.

It is more common to use *kba?* 'torso' (27), or *sac* 'flesh' (28) for a reflexive action:

- (27) ki=gantong kba? ?en l⁹e?
3A=hang body LOC neck
He hung himself by the neck.
- (28) ki=pjay sac
3A=neglect flesh
He neglected himself.

6.2 Ignoratives

The term ignorative is employed following Wierzbicka (1980).⁴ Ignoratives form a series representing six basic ontological categories: THING, PROPERTY, PERSON, PLACE, MANNER, QUANTITY and TIME. They cover a range of functions. They are used as interrogative pronouns in information interrogatives (§10.5.2), in non-interrogative contexts as indefinite pronouns (§6.2.1.1), and also to mark certain types of ontological complements.

Ignoratives are a functional class, and as such some members of the class are identifiable uniquely as ignoratives, while others are drawn from other word classes (§4.2.1.5).

6.2.1 Distinctions in the ignorative class

Table 6.2 sets out the basic ontological categories and the forms used across the range of contexts. As demonstrated, for some categories the basic interrogative-based lexeme is multifunctional,⁵ and the majority of the range of ignorative functions are expressed by a single lexical item per category, e.g. h⁵n WHERE, i.e. 'where; somewhere' (§6.2.5). For others, different lexical items and strategies are employed across the range, with the interrogative and generic noun sharing the functional load. For example, the category of PERSON is expressed in the following ways: by an interrogative *kadeh* 'who'; an indefinite pronoun *sma?* (person)/ko= (3UA) 'someone'; in negative contexts, by a negative pronoun *da?* *da?* *sma?* (NEG EXIST person) 'no-one', and by a free choice pronoun *b⁹r⁹n-kadeh* (any who) 'whoever', (§6.2.4).

The categories of PROPERTY and REASON shown in Table 6.2 share the same form, but are distinguished by differing syntactic contexts (§6.2.2). The ontological category of MANNER is expressed by a complex ignorative, h⁵n-mande 'in what manner, how', derived from a compound of the basic terms for PLACE and THING. The shaded areas in the table indicate the lack of a form for the particular category.

⁴ Other terms used for similar categories are epistemological classifiers (Durie 1985) and epistememes (Mushin 1995).

⁵ Here I follow Haspelmath (1997: 59) in considering the range of possible meanings as the result of different contexts, rather than necessarily a case of polysemy.

TABLE 6.2 IGNORATIVES: PRIMARY ONTOLOGICAL CATEGORIES

ONTOLOGICAL TYPE	FUNCTION			
	interrogative pronoun	indefinite	negative	free choice
THING	mandemoh			b ⁹ r ⁹ n-za=den
PROPERTY	mande	(mande-)mande	b-mande	
REASON				
MANNER	h ⁵ n-mande			
PERSON	kadeh	sma?	kadeh-kadeh b ⁹ r ⁹ n-kadeh	sma?
PLACE		h ⁵ n	b ⁹ r ⁹ n-h ⁵ n	tmat
TIME	mre?/bila?		bila-bila?	
QUANTITY	brapa?			

6.2.1.1 Ignoratives as interrogative pronouns

Interrogative pronouns are the prototypical function of ignoratives, and for each ontological category it is true that if a specific lexeme exists (i.e. which is not found in other word classes) it will include interrogative in its functional range. Further distinctions are made within the categories of THING, PERSON and PLACE. These forms are restricted to the interrogative context. The category 'THING' distinguishes a special form used to enquire about identity in terms of an entity's class membership: 'what class' (§6.2.2.1). PERSON distinguishes possession: 'whose' (§6.2.4.1). The category PLACE combines with directional notions to distinguish 'whither' and ' whence' (§6.2.5.1). The forms are set out in Table 6.3.

TABLE 6.3 IGNORATIVES: DERIVED CATEGORIES

ONTOLOGICAL TYPE	DERIVED CATEGORY	FORM	FUNCTION (INTERROGATIVE)
THING	class	nmoh	'what class'
	sound	j ¹⁰ nh-moh	'hear what'
PERSON	possession	d ¹¹ kadeh	'whose'
	direction	te h ⁵ n	'whither'
PLACE		t ¹² m ha? h ⁵ n	' whence'

The complex forms, those of Table 6.3, as well as the category MANNER in Table 6.2, are distinguished by their general lack of corresponding indefinite pronoun function, so that a correlation may be drawn between an increase in morphological complexity corresponding to a decrease in functional range.

6.2.1.2 Ignoratives as indefinite pronouns

As revealed in Table 6.2 above, there are no lexemes which function uniquely as indefinite pronouns in Semelai. These concepts are expressed either by: a) a basic/simple ignorative; b) a generic-ontological-category noun functioning as an ignorative, or c) a combination of the generic noun and a basic ignorative. The generic nouns employed most frequently in this manner are *sma?* 'person' and *b⁹r⁹n* 'thing' ← Malay *barang* 'things in general'.

It is not always easy to ascertain whether the noun is actually grammaticalised as an indefinite pronoun, although when used in the indefinite pronoun function, the generic noun never co-occurs with markers of referentiality, e.g. demonstratives, or classifiers. In contrast, the negative indefinite pronoun may have a classifier:

- (29) me=?ikur sma? da? da?
one=CLF person NEG EXIST
There wasn't a single person. (Or: There was no-one.)

A generic noun is not used to express unspecified place or time.

Cross-linguistically, indefinite pronouns express a range of complex semantic distinctions (Haspelmath 1997: 3) and this appears to be borne out in many, but not all, of the ontological categories in Semelai. Whilst the main context of indefinite pronoun usage is to express indefiniteness, they are also used here in contexts where they may express free choice, i.e. having a semantically non-specific referent, as well as contexts which are characterised in terms of existence, or more usually, non-existence, i.e. negative indefinite pronouns.

The lack of specificity associated with the free-choice and negative-context indefinites may be distinguished formally by the reduplication of the root, e.g. *mande-mande* 'anything'.

Free-choice pronouns actually represent free choice within a restricted domain, where the referent must be a token of the same type, for example in (30), the token must belong to the domain of culinary dishes.

- (30) ki=masak pay nina, buc, mande-mande
3A=cook for.now food rice RDP=what.sort
She cooked the food for now, rice (and) whatever.

Two free-choice pronouns take the form of a compound combining the lexeme *bərəŋ* 'thing' and PLACE *hɔn*, *bərəŋ-hɔn* 'wherever', and with 'person' *bərəŋ-kadeh* 'whoever'. *bərəŋ* is from the Malay *barang* 'things in general', as noted above. The compound appears to represent a calque of an archaic form of question words in Malay, e.g. *barang apa* 'whatever', *barang siapa* 'whoever', but which are no longer common, at least in the standard variety of Malay (Mintz 1994: 118).

There are no distinct negative pronoun forms, this function being achieved by placing an ignorable in a negated existential construction:

- (31) "zeh" k^haq, "da? da? mande me=kli=ca la=bsi?"
ab QUOTE NEG EXIST what.sort REL=3A=penetrate A=metal
"Ah," (he) thought, "there's nothing that the metal can penetrate."

It also suffices to use the negative existential without an ignorable as in the following example where it is not included in the response:

- (32) Q: "zən te hɔn?" k^haq A: "beh, da? da?"
oh TO:unspec where QUOTE NO NEG EXIST
Q: "Oh, where (are you going)??" (he) asked. A: "No, nowhere."
(Lit. "Oh, where to??" (he) asked.)

Finally, it should be noted that generic nouns do not function in the free-choice category, and they tend to be found in negative contexts, see also Haspelmath (1997: 182-3).

We will now turn to an examination of the individual types from Tables 6.2 and 6.3. The following description is organised around the basic ontological categories outlined in the introductory section above, taking the main interrogative function as

the starting point, and then examining the manner in which the range of indefinite functions is expressed.

Although the discussion of interrogative pronouns precedes the indefinite pronouns, their function in the interrogative clause is but one of these functions, and while it is the most basic and accessible function, it should not necessarily be considered primary. Note that cross-linguistically, indefinite pronouns tend to be the more marked category relative to the forms used in the interrogative function (Haspelmath 1997: 27), see §6.2.1.2 above.

6.2.2 THING

The range of choices available for expressing the ontological category THING are *mandehmoh*, and its variants *hmoh*, *dmoh* and *jmh*. These are used with non-human animate, inanimate and abstract referents in interrogative contexts to seek knowledge about identity. Only *mandehmoh* is used in indefinite contexts.

Although *mandehmoh* appears to be morphologically complex, a compound of *mande* 'what sort of' and *hmoh*, it is difficult to provide an insightful analysis, whilst being unable to draw a semantic distinction between *mandehmoh* and its variant forms. The general unavailability of the form to be used in indefinite contexts does suggest that it is complex, although this observation may reflect a gap in the corpus rather than being a valid claim about the functional possibilities of the term.

6.2.2.1 Interrogative 'what'

The most general term used for 'THING' is *mande(hmoh)* 'what'. *jmh* is a form only used by the elderly and, according to speakers, is the equivalent of *dmoh*. However, I was unable to establish any motivation in terms of the function of the various forms. A specific form is used to seek information regarding identity in terms of class membership: *nmoh* 'what class' (see below). The derived interrogative *j<nh>moh* (WHAT<NMZ>) 'what say' ← *jmh* 'what' is used to ask a speaker to repeat something that the hearer has not heard properly:

- (33) j<nh>moh neŋ?
what<NMZ> before
What was that (you said) before?

i. *mandehmoh* 'what' *mandehmoh* and its variants exhibit a tendency to be used when the referent is the subject of a non-verbal clause (34), or a subcategorised core argument of the verb (35), including clausal objects.

- (34) *The protagonists need boiled water and are forced to improvise:*
mandehmoh ga=jrajan dak?
what IMM=kettle water
What (is to be) the water kettle?

- (35) *dmoh me=kə=reŋ-i??*
what REL=2fA=seek-ITER
What do you keep on seeking?

In (34) and (35) the referent of *mandehmoh* and *dmoh* is an entity; in the following clauses, the ignoratives have an event as the referent:

- (36) ?ec, mandemoh mə=je=ʔen jʔoy ha? nəʔ?
 eh what REL=2=AUG do AT here
 "Eh, what are you doing here?"

- (37) "hməh mə=ma=ən" kʰləŋ "ha? hn=ʔəp?" kʰləŋ
 what REL=IRR=want QUOTE AT ABS=if QUOTE
 "What would (you) want," (she) asked, "with me?" (she) asked.

mandehməh ~ dməh are also used when the domain of an entity is identified, but not the individual entity in question; roughly as English 'which'. In (38), the superordinate class is identifiable by the speaker, but not the lower-order term. Note that the form of nominal used is that which indicates individuation or a lower-order member, affixed with the nominalising morpheme +n+ 'INDIV' (§7.6.1.1).

- (38) Q: mandemoh nk-dak?
 which INDIV-liquid A: dak ʔen
 water just
 Q: Which liquid (is it)? A: (It's) plain water.

- (39) Q: dməh p<n>ʔlaʔ na-ha?? A: dkan
 which animal<INDIV> DEM-AT Rhyzomis.sumatrensis
 Q: Which animal (is) this one here? A: (It's a) bamboo rat.

In (40) the ignorative is in second position, apparently motivated by the preference for ordering old information before new. ?aseŋ 'to be unusual' or 'outside the norm', is used here in a classificatory sense in conjunction with sma? 'person' comparable to the construction sma? putih (person be.white) 'European', and not as a property. This point can be clarified by referring to (47) in §6.2.3.1.

- (40) Some women protest when their husbands offer to fetch their disabled mother:
 A: "ʔma? sma? ?aseŋ"
 mother person be.unusual
 B: "sma? ?aseŋ mandehməh?"
 person be.unusual what
 A: "ʔma? ye=ʔen," kʰləŋ, "kʰoy nehneh, da? da? kbe?"
 mother 1=AUG QUOTE head only NEG EXIST head
 A: "(Our) mother is an unusual person."
 B: "What kind of 'unusual' person?"
 A: "Our mother," (they) replied, "(is) only a head, (she) doesn't (have) a body."

In (41) the speaker wants to know which race in particular the person belongs to.

- (41) bəŋsa? dməh?
 race which
 (To) which race (do you belong)?

ii. nməh 'what class' nməh is used when the individual type is identifiable, but not the superordinate, eliciting specific type in terms of class membership. This is the only construction in which nməh is used. The ignorative appears to be derived with the nominalising prefix n- 'INDIV': nməh (n-hməh INDIV-what). Note here that it is not the NP, but the ignorative which hosts the affix, compared to examples (38) and (39) in the previous section.

In (42) the question being asked is "To which domain does that animal belong?" The answer would be in terms of a superordinate classification (a), or a basic classification (b). Note that here in this construction the known information precedes that which is unknown, the ignorative occurring in second position.

- (42) Q: pʔla? nməh?
 animal INDIV:what A: a) cloy
 Q: Of which class (is that) animal? A: monkey.lesser.apes
 A: b) kane?
 rat.lesser.gymnure
 A: Rats and lesser gymnure.

6.2.2.2 Indefinite and negative pronoun functions

Apart from its function in the complement of verbs of perception (43), the ignorative mandehməh 'something' is not used in indefinite contexts.

- (43) *The protagonist has just become aware of something which has appeared at the far end of the clearing around the house:*
 ki=jpək mandehməh loc he? kʰoy,
 3A=observe what already AT:above head
 he? kʰoy halaman
 AT:above head clearing.around.house
 He observed something already up at the far end, up at the far end of the clearing around the house.

In negative contexts, either a nominal referring to the non-existing entity is used (44), or a headless relative clause (45):

- (44) da? da? ʔisi?
 NEG EXIST content
 There is nothing (inside). (Lit. There are no contents.)

- (45) da? da? mə=ən br-ca
 NEG EXIST REL=want MID-eat
 There was nothing to be eaten. (Lit. There was nothing that wanted to be eaten.)

Note the following idiomatic phrase which does not indicate non-existence, but rather, an abundance:

- (46) da? da? mə=ga=da? da?
 NEG EXIST REL=IMM=NEG EXIST
 There was everything in abundance. (Lit. There wasn't what there wasn't going to be.)

6.2.3 PROPERTY; REASON

mande is the only ignorative which encodes more than one ontological category. It expresses the category PROPERTY, 'what sort of' and also REASON 'because of what'. The choice between these interpretations is dependent on the context. The distinction is apparent in its syntactic distribution: in its PROPERTY function mande exhibits the syntactic distribution of the referent, e.g. as the adnominal constituent in

an associative phrase (47), whilst in the REASON function it is always clause-initial, an adjunct to the proposition (48).

- (47) *A woman is curious as to how her estranged husband has managed to reach her home which is located in the lug 'the area between the earth and the sky'.*

2endol, ki=tapa? rom mande troj.
enter.house 3A=ask WITH what.sort path
(He) came inside (and) she asked via which path (he had come).

- (48) mande da? b-pri?to? hayam na?ha??
why NEG MID-put.away chicken DEM-AT
Why haven't these chickens here been put away?

The interrogative functions are examined first in §§6.2.3.1–2 and the indefinite function in §6.2.3.3.

6.2.3.1 PROPERTY *mande* 'what sort of'

mande 'what sort of' is used to seek information about the general properties of animate (both human and non-human), inanimate and abstract referents: 'what are the attributes of X', as opposed to the identifying domain or class membership discussed in §6.2.2.1. As mentioned previously, the ignorative is normally placed following the head in the position occupied by the adnominal.

- (49) *The information sought is a description of the person in terms of distinguishing features.*

Q: "sma? mande?"
person what.sort

A: "sma? k^hoy me=bjs?, kb^he? da? da?, k^hoy nehneh"
person head one=CLF body NEG EXIST head only
Q: "What sort of person (is she)?"
A: "(She is) a head-person, (she) has no body, only a head."

Note the contextual ambiguity which arises between the translation 'of what' and 'what kind of' when inalienable possession is involved, as in the following two examples, which both happen to be riddles.

- (50) ?ec mande me=da? s?it?
excreta what.sort REL=NEG stink

What sort of faeces (is) the one that doesn't stink?

The answer to the riddle is ?ec dre (?ec 'excreta, bodily waste products', dre 'rattan') 'rattan shavings', the by-product of peeling rattan.

- (51) mot mande da? ?ye sma??
eye what.sort NEG see person

What sort (of) eyes don't see people?

The answer to the riddle is mot taga? (eye stair) 'step'. More riddles can be found in §14.1.

In (52) the ignorative is the complement of the modifying locative prepositional phrase. Interestingly, the ignorative associated with the attributive is used here rather than the locative PLACE form.

- (52) Q: kudes ?en mande?
skin.disease LOC what
Q: A skin disease on what (part of the body)? A: A skin disease of the foot.

(53) is an atypical example containing *mande* on its own. This may be due to the fact that the referent 'existence, life' is abstract. The question is rhetorical.

- (53)^o 'pdor-i7 hn=k^ho," k^hləŋ, "mande ma=gə=?əp=reŋ?"
follow-JTER O=2f QUOTE what.sort REL=IMM=1fA=find
"If I accompany you," (she) said, "what will become of me?" (Lit. "What sort of (existence) will I find?")

6.2.3.2 REASON *mande* 'because of what, why'

mande 'why' expresses reason, both purpose and cause, and does not distinguish between the two. It is only attested in interrogative contexts.

mande is always clause-initial when the intended ontological category is REASON, possibly due to the fact that it has scope over the whole proposition and not simply one constituent within it.

- (54) "mande" k^hləŋ "k^ho jtek ?en karom?" k^hləŋ kəh
why QUOTE 2f sleep LOC underneath QUOTE 3
"Why," (he) asked, "are you sleeping in the (space) underneath the house?" he asked.

- (55) "?oh, yo? mande da? ma=pdor-i?" k^hləŋ sma? gdo
oh but why NEG IRR=follow-JTER QUOTE person be.old
"Oh, but, why couldn't (you) accompany (them)?" the old man asked.

In circumstances where the use of *mande* REASON may be ambiguous with a PROPERTY reading, speakers claimed that the REASON reading could be distinguished by employing la= 'BCS': la=mande (BCS=why) 'because of what'. However this construction was only ever obtained when attempting to directly elicit ignorative forms.

6.2.3.3 Indefinite and negative pronoun functions

As we have already seen in relation to the interrogative function, *mande* 'what sort of' is used when the focus is not the identity of the entity, but rather the distinguishing features or properties of the entity. This is also apparent in the indefinite function. The indefinite function is not available to the REASON function of *mande* 'because of what'.

The ignorative is used in ontological complements of the verb of perception *jŋok* 'to look at', where the perceiver observes the properties of an entity:

- (56) *The perceiver's wife has died in childbirth and turned into a birth-vampire. He observes how red her eyes are:*
ki=jŋok mande mot mirah
3A=look.at what.sort eye be.red
He observed what red eyes she had.

The following example highlights the property aspect, not identity:

- (57) ki=coc kloc lpəc, jŋok mande kmpən, sot
3A=gut inside stomach look.at what.sort wife metamorphose.into

kmpen mə=ki=ca, mə=j?j? da?, mə=cin da?
 wife REL=3A=eat REL=be.raw EXIST REL=be.cooked EXIST
 He (=the husband) gutted the contents (of) the stomach (and) saw what sort of thing
 (his) wife now was, transformed into the wife that he (=the tiger) had eaten, the raw
 was there (and) the cooked was there.

- (58) An explosion has just taken place, and a woman is astounded by the accompanying transformation of her environs.
 mande ma=j?pk, brapa? ?i?ok!
 what.sort IRR=look.at how.much be.good
 Whatever (she) looked at, how fine it was!

As we saw above with mandehmoh, the ignorantive is rare other than in negative or free-choice contexts.

When the negative indefinite meaning 'nothing' is required, the ignorantive is used in a negated existential clause. Either mande ~ mande-mande '(RDP)-what' may be used.

- (59) An opponent's weapon is ineffectual:
 " eh" k?ləj, "da? da? mande mə=ki=ca la=bsi?"
 ah QUOTE NEG EXIST what.sort REL=3A=penetrate A=metal
 "Ah," (he) thought, "there's nothing that the metal can penetrate."
 (60) "bah" k?ləg, "da? da? mande-mande"
 NO QUOTE NEG EXIST RDP-what.sort
 "No!" he protested, "there's nothing/it's nothing."

The reduplication of the ignorantive mande 'what sort of', e.g. mande-mande emphasises the lack of specificity, as in (60) above. The reduplicated form may also be used in free-choice contexts, where it indicates an increased variety of alternatives, 'what-ever'.

- (61) Her husband has long gone, but each day the wife dutifully cooks for him, and sets the food aside:
 ki=masak pay ninca, huc, mande-mande
 3A=cook for.now food rice RDP-what.sort
 She cooked the food for now, rice (and) whatever.

The ignorantive mande 'what sort' may be prefixed with the middle marker b- 'HAVE' where the referent is a possessed property (§5.4.3). An alternative to the free-choice translation here would be the universal 'everything'.

- (62) na?ha? nɔ? sot kmpoj tijo,
 DEM-AT this metamorphose.into finish snake
 b-sisik, b-mande=hn
 HAVE-scale HAVE-what.sort=3POSS
 All (of) this one here had transformed into a snake, with scales (and) with whatever.
 (63) kampoj sma? ramay.
 settlement person be.many

kira?, b-raja?, b-mande, b-baten
 see HAVE-sultan HAVE-what.sort HAVE-headman
 (It was) a settlement of many people, sec, with a sultan and with whatever, with a headman.

Other forms used to express free choice are the compound b?r?ŋ 2a=deŋ (thing DET=manner) 'all manner of things' (64), and the verbal derivative b-məcɔm ~ b-məcɔb-məcɔm (HAVE-(RDP)-like) (§5.4.3). The difference between these forms and mande-mande 'whatever' is that the former emphasises the similarity or likeness of the intended referents.

- (64) sec h?gi? bkəl=hn,
 surreptitiously give provisions=3POSS
 ?anu?, baba t̪mpoh lutɔ, b?r?ŋ 2a=deŋ
 [HES] rice seven grain thing DET=manner
 (She) surreptitiously gave (him) his provisions, umm, rice, seven grains, all manner of things.

6.2.4 PERSON

The ignorantive kadeh 'who' is used for the category of PERSON in interrogative contexts. The possessive form də kadeh 'whose' is derived by addition of the preposition də 'OF' (§6.2.4.1). Indefinite contexts utilise either the ignorantive kadeh 'who', a reduplication of this, kade-kadeh 'whoever' or most frequently the generic noun sma? 'person'.

The ignorantive must have as its referent a human participant or any metaphysical extension thereof, or an 'anthropomorphised' animal character in a traditional narrative.

6.2.4.1 Interrogative 'who'

The ignorantive is used in the interrogative context to seek information about a person's identity:

- (65) kadeh nɔ?
 who this
 Who (is) this?
 (66) kadeh na?tə?n?
 who DEM-TO:down
 Who (is) the one (coming) down?
 (67) kadeh p-kasot ji?
 who CAUS-shoe 2
 Who put (your) shoes (on) you?

The ignorantive is also used to form interrogatives relating to personal names:

- (68) kadeh ga=glo? he?
 who IMM=be.named 1&2
 What are we going to be named?

kadeh 'who' can be modified by the quantifier ramay 'be.many (of people)' and function as an alternative to the ignorantive brapa? 'QUANTITY':

- (69) kadeh ramay p?ot ha? no?
who be.many stay AT here
How many (people) live here? (Lit. Many who live here?)

Possession 'whose' The simplest way to form a possessive interrogative where the possessor is human is to place the PERSON ignorative kadeh after the possessee utilising the associative construction (§7.5):

- (70) ?no?no? dol kadeh?
RDP-this house who
This (is) whose house?

Interrogatives pertaining to human possession are also expressed by an ignorative construction derived from the possessive preposition *dɔ* + the ignorative kadeh 'whose'. *dɔ* 'OF' is discussed in §8.9.

- (71) dɔ kadeh rokok no?
OF who cigarette this
Whose (are) these cigarettes?

- (72) dɔ kadeh ni-cɔ?
OF who INDIV-dog
Whose dog (is it)?

6.2.4.2 Indefinite and negative pronoun functions

Several different forms are used to express this category in different contexts, including kadeh, kade-kadeh, and the rare free-choice compound bɔrɔŋ-kadeh 'who-ever'.

kadeh 'who' is used in the complements of verbs of cognition to express lack of knowledge:

- (73) da? ye=kʰe? kadeh no?
NEG 1=know who this
I don't know who this is.

The indefinite pronoun, used when a specific referent 'someone' or a free choice like 'anyone' is intended, is an extension of the generic noun *sma?* 'person' without further modification. In (74), an ontological complement, it is used with the meaning 'someone'.

- (74) *The shaman has had someone come to him in a dream:*

- masuk po sma? ga=krwɔl
enter dream person IMM=emerge
(He) had a dream (that) someone was going to reveal (himself).

In (75) the speaker does not have a specific referent in mind.

- (75) cəl tet cɔn, mɔcɔm sma? ga=kʰbes
say TO:spec back like person IMM=die
Talk behind (their) back, as if someone were going to die.

(76) encompasses the free-choice sense.

- (76) *The speaker knows who dwells in the house, and hence knows the possible range of referents.*
da? sma?
EXIST person
Is someone/anyone there? (Anyone home?)

A further possibility exists in transitive clauses, where the third person unidentified pronominal clitic *ko=* '3UA' functions as an indefinite pronoun. The speaker does not know the identity of the agent (§6.1.1.1).

- (77) ko=sec tarek cɔ? ʔate
3UA=surreptitiously pull AT:below ground
Someone surreptitiously pulled (at it) (from) down on the ground.

The generic noun is also used in negative contexts to express non-existence:

- (78) da? da? sma?
NEG EXIST person
There is noone.

Alternatively, a headless relative clause may be employed in this context:

- (79) da? da? me=dɔs
NEG EXIST REL=arrive
Noone came. (Lit. There wasn't (one) who came.)

The interrogative kadeh 'who' forms a free-choice indefinite pronoun either as a reduplication kade-kadeh, or in a compound with bɔrɔŋ 'thing', bɔrɔŋ-kadeh 'whoever'. The latter is infrequent.

- (80) ga=de=yɔk kade-kadeh
IMM=3pl=take RDP-who
They'll take whoever.

- (81) "ʔɔh, bɔrɔŋ-kadeh," kʰlɛj
oh thing-who QUOTE
"Oh, whoever," (he) replied.

6.2.5 PLACE

PLACE is expressed by the ignorative hɔn in both interrogative and indefinite contexts. Synthetic compounds derived from hɔn and directional prepositions express 'whither' and 'whence' and in the indefinite pronoun function, a free-choice compound 'whoever'.

6.2.5.1 Interrogative hɔn 'where'

hɔn 'where' is used to seek information about PLACE. hɔn is inherently locative and therefore does not require a locative preposition (82). In contrast, a locative preposition must be used to express 'whence' and a directional preposition must be used to express 'whither'.

- (82) "hɔn ?ma? hne?" kʰlɛj
where mother THEN QUOTE
"Where (is) mother then?" (they) asked.

- (83) A tiger protests that he can't possibly help a woman take care of her new born baby. She retorts:

Q: "mande? hōn t̄i kb?"
why where hand 2f

A: "t̄i ʔap da?, yb?=hn t̄<wp>loj da? n=k̄e?"
hand If EXIST but=CONN hold<IMPERF> NEG IfA=know

Q: "Why? Where (are) your hands?"

A: "My hands I have, but (as for) holding, I don't know (how)."

A secondary use of hōn is to express 'how' in rhetorical-type questions. This appears to parallel the Malay usage of *mana* 'where'.

- (84) hōn ga=than?

where IMM=endure

How could it be possible? (Lit. How is (it) going to be endured?)

In (85) the first clause is a rhetorical question; in the second the speaker asks a direct question.

- (85) "hōn?" k̄loj "l̄en b-pri?to?" k̄loj. "ha? kadeh" k̄loj
where QUOTE want MID-put.away QUOTE AT who QUOTE
"kb?=ur pri?to? kmpen kb" k̄loj, "beh, da? da? wɔ?"
2fA=instruct put.away wife 2f QUOTE NO NEG EXIST more
"How," (she) asked "would (it) be put away?" (she) asked. "Whom," she asked,
"did you instruct to put (it) away? Your wife," (she said), "no, (she) isn't (here)
anymore."

'Whither/whence' hōn 'where', together with the relevant preposition, forms an ignorative construction used to seek information about the goal or source of movement.

To express 'whither' the neutral unspecified directional preposition te 'TO:unspec' (§8.6.1) is combined with hōn. The deictic directionals lej 'TO:up' and tə?en 'TO:down' are naturally never used in this function, presumably because they already presuppose some knowledge.

- (86) te hōn ?us?
TO:unspec where lighter
Where (has) the lighter (gone)?

- (87) This is the customary way of greeting someone encountered on a path:

Q: ga=te hōn? A: tet gombak
IMM=TO:unspec where TO:spec Gombak
Q: Where are (you) going? A: To Gombak.

With source of movement 'whence', the preposition is tom ha? (SRC AT) (§8.7); ha? 'AT' is the general deictic locative term (§8.5.1).

- (88) tom ha? hōn ns-dos ji nej?
SRC AT where NMZ-come 2 before

Where did you come from before? (Lit. From where (was) your arrival before?)

6.2.5.2 Indefinite and negative pronoun functions

Only hōn 'where' functions as an indefinite pronoun. The synthetic forms for 'whither/whence' are not available in this function.

hōn 'where' is used in the complements of verbs of cognition and perception to express lack of knowledge (89), or surprise (90).

- (89) da? ki=k̄e?, hōn ?arah ye swak
NEG 3A=know where direction 1 walk
He didn't know which direction I went.

- (90) "yah" k̄loj, "da? gama?", k̄e? hōn?"
ah QUOTE EXIST possible know where
"Ah," (he) said, "(he) is possibly (somewhere), (but) (who) knows where?"

- (91) sampay no? tr-ye=ce? hōn, 2anu?, blahan
then this HAPP-see=PM where [HES] blowpipe
Then now (he) happened to see where (there was), umm, a blow-pipe.

It is also used in a negated existential clause to express 'nowhere'. This would constitute a negative response to the question above in (87):

- (92) "da? da? hōn"
NEG EXIST where
"Nowhere."

A compound of the word bōrōj 'thing' and hōn 'where' is used to express an unspecified location or locations, corresponding to the English free-choice pronoun 'wherever'. Again, with the free-choice pronoun there is a degree of ambiguity, and a universal sense, 'everywhere', is also possible.

- (93) kəh no?, bōrōj-hōn jr-jor
3 this thing-where IMPERF-urinate
He pisses wherever/everywhere.

- (94) ʔəmu?, jampi?, b-məcb-məcm,
esoteric.knowledge incantation HAVE-RDP-like
ki=tontot bōrōj-hōn
3A=seek.esoteric.knowledge thing-where
Esoteric knowledge, incantations (and) whatever, he sought wherever/everywhere.

Finally, it is worth mentioning an idiomatic compound which is used to express realisation in response to visual stimuli 'to look like'. Interestingly, this is not embedded under a verb of perception or cognition, the presence of gayə 'appearance' from Malay *gaya* 'form, manner' rendering it implicit.

- (95) A woman climbs down from a tiger's back to take a drink from the stream. For the first time he realises that she is pregnant.

"hōn-gayə k̄at," k̄loj "kb no?!"
where-appearance be-pregnant QUOTE 2f this
"(It) looks like you here are pregnant!" (he) exclaimed.

6.2.6 MANNER

mande 'what sort of', in combination with hōn 'where', produces the complex ignorative hōn-mande 'do in what manner, how'.⁶ It is used to question the manner in which an action is performed, MANNER OF DOING. This ignorative is only available in the interrogative function.

- (96) A type of plant is being described. The listener enquires as to how the fruit is eaten, e.g. whether straight from the plant, or if it first requires preparation.

hōn.mande br-ca?
how MID-eat
In what manner/How is (it) eaten?

- (97) hōn.mande br-poy?
how MID-make
How is (it) made?

- (98) In this story the son-in-law, who has never seen his bodiless mother-in-law, is curious and asks:

Q: "ga=k'oy nehneh, hōn.mande lən ?ris?"
IMM=head only how want be.alive
A: "zə? ?ris" k'ləŋ=hn
oh be.alive QUOTE=3POSS
Q: "(If) (she) is just a head, how is (she) alive?"
A: "Oh (she) is alive," they exclaimed.

6.2.7 QUANTITY

The quantitative category is expressed by brapa? 'how much/many'. brapa? 'how much/many' is used for both animate and inanimate referents, the same form being used regardless of whether count or mass nouns are involved.

brapa? 'QUANTITY' is a borrowing from Malay *berapa* 'how much/many'. There is no indigenous equivalent for this term.

In both interrogative and indefinite functions, the ignorative is always followed by a noun or stative verb which defines the quantity in question.

- (99) brapa? tən ji=k<m>ləm?
how.many time 2A=carry.sth.long.on.shoulder<IMP>
How many times were you carrying (it)?

Some of these terms are listed below, including dimension and measure nouns, and classifiers (§§7.3.2-3):

k<1>jeh	'weight'(be.heavy<NMZ>)	brapa? k<1>jeh	'how heavy'
ny-t'ay	'size'(<NMZ>be.big)	brapa? ny-t'ay	'how big'
n2-le?	'time'(<NMZ>be.long)	brapa? n2-le?	'how much time'
?ikur	'animate being, CLF'	brapa? ?ikur	'how many people/animals'

⁶ This form is possibly influenced by the Malay terms for 'how': *macam mana* (like where) and *bagaimana* (how) which also contain the 'place' category *mana* 'where'.

?are?	'day'	brapa? ?are?	'how many days'
?rey	'be many'	brapa? ?rey	'how many things'

6.2.7.1 Interrogative brapa? 'how much/many'

The ignorative construction is in clause-initial position:

- (100) brapa? ?le? ləc ha? nɔ?
how.many be.long.time already AT here
How long have (you) already (been) here?

There are two alternatives in the quantification of humans. The ignorative can be modified by ramay 'to be many (humans)':

- (101) brapa? ramay ji dəs?
how.many be.many 2 come
How many (of) you came?

Alternatively, kadeh 'who' can replace the quantitative ignorative, the concept of quantity being expressed by the modifier ramay 'be many (of people)', see (69) above.

6.2.7.2 Indefinite pronoun functions

The only attestation of brapa? as an indefinite pronoun is in relation to expressions of unspecific duration of time.

- (102) kde? brapa? ?le?
dwell how.much be.long
(She) dwelt however long (there).

- (103) cəŋ tom ha? kke, tah brapa? bulan ki=goy
pass.through SRC AT there Q how.much month 3A=carry
From then on, (for) however many months he carried (hcr).

The reduplication of the ignorative brapa-brapa? is ambiguous between a free-choice or a quantifying reading, where brapa-brapa? means 'several':

- (104) de=papel nojom, brapa-brapa? nojom
3plA=summon seer RDP-how.many seer
They summonsed seers, several seers.

- (105) cəŋ=cə? ki=tənoy brapa-brapa? ?are?
pass.through=EM 3A=wait RDP-how.many day
From that point on, he waited for some/several days.

6.2.8 TIME

mre? is an indigenous term, bila? 'TIME' is a borrowing from Malay *bila* 'when'. Both mre? and bila? are fairly infrequent, the more usual means of enquiring about time being a combination of the ignorative expressing QUANTITY plus a temporal expression. bila? or the reduplicated bila-bila? have an indefinite pronoun function 'whenever, anytime' in (107) and (108).

- (106) mre? ga=?yot?
when IMM=return
When will (you) return?

- (107) sot bila? ʔəp b-croh
should when If MED-encounter
(I) should when I (am) met.

The free-choice form *bila-bila?* 'whenever' is from Malay *bila-bila* of the same meaning.

- (108) bila-bila? dos, jtek ha? nɔ?
RDP-when come sleep AT this
Whenever/anytime (you) come, sleep here.

6.3 Demonstratives

Demonstratives provide deictic information in relation to the establishment of identifiability, or reference of NPs.

There are two basic demonstratives in Semelai: the proximal *nɔ?* 'this' and the distal *ke* 'that', based on simple distance orientation from the speaker, or the speaker's adopted point of reference. Finer distinctions may be achieved by the reduplication of the terms, *nɔ-nɔ?* 'this here', *kke* 'that there', or modification with *neŋ* 'before' (§6.3.2.1). This type of information can also be conveyed by the adnominal use of locative prepositions, or pronominals derived from the set of locative prepositions (§6.3.1.1).

The demonstratives may function either adnominally or pronominally (Himmelmann 1996: 214), the same form being used for both functions (§6.3.2). They are used in three main contexts: spatial, discourse deictic, and recognitional (these categories are from Himmelmann 1996: 218-40). A summary of these usages and the forms correlated with them is provided in §6.3.3.

6.3.1 The demonstratives

As stated above, the demonstratives encode two degrees of distance: the proximal *nɔ? ~ nɔ?* 'this', and the distal *ke* 'that', based on simple distance orientation from the speaker, or the speaker's adopted point of reference:

- (109) knɔn nɔ?
offspring this
this child
- (110) wɔy ke
knife that
that knife

A third term, *ha?* 'AT', is a proximate deictic locative preposition which expresses a non-specific proximate location: 'in the immediate vicinity, but not at any particular point within it'. It functions adnominally only (see §8.3.2 for the attributive function of prepositions):

- (111) sma? ha?
people AT
people hereabouts

From the two basic terms are derived the forms *nɔ-nɔ?*, 'this, right here', and *kke* 'that, over there'. These forms are usually accompanied by a gesture of the hand or lips directed at the intended referent. Note that reduplication of the demonstrative serves to intensify the gesture, whereas reduplication in nominals usually implies a multiplicity of non-specific referents. Note that the proximal locative *ha?* 'AT' cannot be reduplicated. The reduplicated forms occur in both adnominal (112), and pronominal (113) functions.

- (112) da? k^bbes=cə? knkn ʔnɔ-nɔ? lagi?
NEG be.dead=EM child RDP-this still
This here child still isn't dead.

- (113) kira?, ma=gə=kł=turjɔ?, ʔah dom, kke ga=bapa?
figure REL=IMM=3A=give, ah AFF, that IMM=father
See, the one he gave it to, ah yes, that would be the father.

Deictic locative prepositions derive special locative demonstrative forms which function both adnominally and pronominally.

6.3.1.1 Locative demonstrative pronouns

Locative prepositions (both the neutral and deictically specified forms (§§8.4-5), and the deictically specified directionals (§8.6.1)) derive spatial demonstrative forms which function either adnominally or as the pronominal head of an NP.

The prefix used in this derivation is *na?* - ~ *na-* 'DEM' which is only found in this context. The forms are shown in Table 6.4.

TABLE 6.4 DERIVED DEMONSTRATIVES

PREPOSITION		DEMONSTRATIVE	
ʔen	'LOC'	na?-ʔen	(DEM-LOC)
ha?	'AT'	na?-ha?	(DEM-AT)
tɔ?	'AT:across'	na?-tɔ?	(DEM-AT:across)
he?	'AT:above'	na?-he?	(DEM-AT:above)
cɔ?	'AT:below'	na?-cɔ?	(DEM-AT:below)
leŋ	'TO:up'	na?-leŋ	(DEM-TO:up)
teʔen	'TO:down'	na?-teʔen	(DEM-TO:down)

This prefix is also attested on locative nominals: *na?-kloc* (DEM-inside) 'those inside', and *na?-cop* (DEM- yonder:down) 'those down yonder' ← *cop*. 'a locale down yonder' as the base.⁷

The locative demonstratives always have a deictic centre anchored in relation to the speaker, either in the actual utterance situation or in the narrated situation, and are always visible to both the speaker and hearer. In contrast, visibility is not a factor in their prepositional function.

- (114) hubi? na?-ha?
elephant.yam DEM-AT
this elephant Yam tree here

The derived pronouns function either adnominally following the head (115), or pronominally (116):

⁷ *cop* 'yonder:below', *heŋ* 'yonder:above, yore' deserve mention as they appear to be related to the prepositions *cɔ?* 'AT:below' and *he?* 'AT:above'.

- (115) do kadeh do, n2-go? na7-he??
OF who OF NMZ-fell.tree DEM-AT:above.
Whose are those traces of swidden-making up there?

In (116) the demonstrative pronoun is the complement of a PP. The preposition ?en 'LOC' (§8.4) is normally deictically neutral, but here it has a proximate interpretation. In all the examples in the corpus, forms based on the neutral locative are used specifically to refer to inalienable entities like the body parts in this example and in (117).

- (116) *The speaker is indicating the fingers of her right hand:*

sma? dr1? ca rom na?7-en
people 1self eat WITH DEM-LOC
My people eat with these.

- (117) *The narrator indicates the circumference of her thigh, na?en 'DEM-LOC', which she uses as the standard of comparison in describing the circumference of the mouth of a jar in the narrative.*

?ie?ie?, ?areh krwbl takar r?k?t,
eventually just.then emerge vessel COMP-be.small

[mə=mur na?en]
[REL=circumference DEM-LOC]

Eventually, just then out came a smaller water vessel, that was the circumference (of) this here.

6.3.2 Adnominal and pronominal functions

The demonstratives may function either adnominally or pronominally. The same form is used for both functions, although there is a tendency for the demonstrative to be reduplicated when used pronominally.

6.3.2.1 Demonstratives as adnominals

The adnominal use of demonstratives is attested more frequently than the pronominal use.

Demonstratives follow the head when used attributively:

- (118) "yo?", k^bləŋ, "nastip kp rom knkon ke" k^bləŋ
but QUOTE fate 2f WITH child that QUOTE
"But," he said "your destiny is with that child."

- (119) de=jon dom, kira?, la=kraba-krabat=hn ha?
3plA=give AFF figure A=RDP-official=3POSS AT
They gave (it), indeed, (I) reckon, his officials' thereabouts.

In the presence of an attributive, the demonstrative occurs in phrase final position:

- (120) tom.dəŋ t^bey ke
base.tree be.big that
that big hole

Although pronominals are implicitly referential, they may be modified either by the proximal (121) or distal (122) terms:

- (121) de=tapa? dom la=jpon ha?, "kp" k^bləŋ, "bansa? dməh?"
3plA=ask AFF A=Japanese AT 2f QUOTE race which
"ʔəŋ nɔ?, bansa? smilay" k^bləŋ
If this race Semelai QUOTE
The Japanese thereabouts indeed asked, "You," (they) asked, "(are of) which race?"
"Me here, (I) am Semelai," (he) replied.

- (122) dom, kira?, tukar.masak deh ke
AFF see cook 3pl that
Indeed, see, (he was) the cook (of) them there.

In imperative or directive constructions, the demonstrative may precede the head as in (123) and (124). In this construction, the demonstrative frequently precedes an overt mention of the noun as in the previous example above, and in the one which follows:

- (123) "nɔ?=spn," k^bləŋ, "nɔ? cəŋ masuk kloc rban!"
this=SC QUOTE this mousedeer enter inside pen
"This, see," (he) said, "this mousedeer, put (it) in the pen!"

- (124) "nɔ?antat kke ?ɔsok!" k^bləŋ
bring that placenta QUOTE
"Take that placenta!" (she) instructed.

Proximate demonstratives modified by neg 'before' neg is a temporal adverb expressing relative proximate time 'before in the immediate past, just before'. It functions here as a shifter, indicating something which has just previously occupied the deictic centre, either spatially, or within the discourse context. This interpretation is only possible with the proximate terms nɔ? 'this' and ha? 'hereabouts': nɔ? neg 'aforementioned' and ha? neg 'present just now'.⁸

This distinction is only available to the adnominal function of the demonstrative, and most uses are actually 'adpronominal' (Himmelmann 1996: 214), where the head is a third person pronoun, e.g. kəh nɔ? neg (3 this before) 'he who was just mentioned', kəh ha? neg (3 AT before) 'he who was just here'.

The following example demonstrates clearly that the speaker is certain that the referent, here an inanimate entity, is still present in the hearer's mind.

- (125) I know there is a lighter where I am sitting, and that noone has taken it, but I can't seem to find it either:
te hɔn ?us nɔ? neg?
TO:unspec where lighter this before
Where's the lighter here before (gone)?

nɔ? neg 'aforementioned', is used in discourse deictic contexts. The discourse deictic use of 'here before' demonstrates how the spatial location is contiguous.

⁸ Interestingly, neg has a variant form kneŋ, which is possibly ke + neŋ, although I was unable to find semantic evidence in support of this: kneŋ ca mi, iŋ ca lag! (before eat noodles, want eat again) "(He) ate noodles before, (and now he) wants to eat again."

- (126) *The shaman is returning to his body following an unsuccessful exorcism. The exhausted drummers stop drumming before he reaches his body, and he falls to the ground outside the house.*

cəg sot grəbəq, kəh nə? neŋ
immediately metamorphose.into cricket 3 this before
Immediately, he, the aforementioned, metamorphosed (into) a cricket.

This is not a case of ambiguity resolution – in the example above there has been no other protagonist in the story at this point. A textual example below illustrates this point:

- (127) pasuh ki=bək rəm pos=hn ki=gəj lumpət,
jar 3A=fasten WITH tail=3POSS 3A=bring leap

paloh 2a=kəh nə? neŋ
flee DET=3 this before
He fastened the jar with his tail. He took (himself) leaping and fled, him, the aforementioned.

The fundamental difference between the two expressions, is that in the second, ha? neŋ, the referent is recent, but time and space have intervened between a previous mention or the individual's actual physical presence, and so there is a lack of contiguity and the possibility that the hearer may be unsure of the referent's identity.

- (128) *We have been whiling away the afternoon on the verandah and in that time a number of passers-by have come and gone. The speaker now makes reference to one of them:*

kəh ha? neŋ, kmən ye
3 AT before sibling's.offspring 1
Him hereabouts before, (is) my sibling's offspring.

6.3.2.2 Demonstratives as pronominals

As pronominals, demonstratives are only employed in a spatial context and the referent must be located relative to the deictic centre. For similar observations see Himmelmann (1996: 214-15).

Demonstratives can be used pronominally for presupposed animate or inanimate entities. Pronominals are restricted in terms of the possible referent, which is usually animate. For animates, the demonstrative is restricted to occurring as the subject of an equative non-verbal clause.⁹

- (129) ?nə?-nə? knən kv
RDP-this offspring 2f
This here (is) your child.

- (130) *Two boys have been spotted by the human-eating ogre rəmgəst, who exclaims:*
na2-ha? lawok ?əp!
DEM-AT game If
These here (will be) my meal!

⁹ The ability of demonstrative pronouns to encode human referents as the topical part of an equative clause is attested across a range of languages, e.g. English, Chinese, Indonesian and Tagalog (Himmelmann 1996: 214).

Demonstratives can also be used to refer to inanimate entities, to make up for the fact that pronominal representation is generally not permitted for such entities. Typically, inanimate entities only occur in the role of a patient or as the subject of an equative or an existential clause, and in the corpus they are restricted to occurring as the O or subject NP. This is the case in (131) where the pre-verbal demonstrative is the O referring to the man's ears.

- (131) *This is the first mention of the ears. The narrator, shifting from the narrative situation to the utterance situation, establishes reference by tugging at her earlobe, ?nə?-nə? 'this here'. The same means is used to establish reference for the second pronominal na2-ha? 'these ones here'. The narrator is waving her fingers in the air as she speaks, drawing attention to her fingernails.*
?nə?-nə?, ki=kət kməpən rəm na2-ha?
RDP-this 3A=pinch.off.bits wife WITH DEM-AT
These, the wife pinched bits off (his earlobe) with these ones here.

Both the proximate and distal basic terms (?nə)-nə? 'this' and (k)e 'that', may function as pronominals. The entity is always either present in the spatial context, or prominent in the discourse context.

Some further examples, illustrating this, are presented below:

- (132) "da? ?nə?-nə?," kəleg, "tmpat jrajan dak"
EXIST this QUOTE place boiled water
"There is this," (he) said, "a thing for boiling water."

The distal term is used primarily in a contrastive role (133), or to single out an entity as in (134). Note that its function here is predictive.

- (133) *The speaker, a squirrel, has thrown some betel nuts down to the woman. As he threw them down, he indicated to her how to tell if they were sweet or intoxicating. He now draws her attention to the ones he believes should be sweet, as opposed to the others which are intoxicating, and gives her some directions:*
kke, ma=snek, p-su<>ku?
that IRR=be.sweet CAUS-section<CAUS>
Those, (that) should be sweet, cut (them) into sections!

- (134) kira?, mə=ga=kı=tur jo?, 2ah dom, kke ga=bapa?
figure REL=IMM=3A=give, ah AFF, that IMM=father
See, the one he gave it to, ah yes, that would be the father.

When used pronominally, the locative demonstrative may be modified by the proximate demonstrative nə? 'this':

- (135) na2-ha? nə? sot kməj tiŋ
DEM-AT this metamorphose.into completely snake
This here had completely metamorphosed into a snake.
All of him had metamorphosed into a snake.

- (136) *The husband has buried his wife alive, and stamps repeatedly on the earth piled up on top of the grave, before walking off.*
ki=kəm, ki=tt-t'ət na2-ha? nə?
3A=bury 3A=IMPERF-stamp.on DEM-AT:above this
He buried (her) (and) stamped (and) stamped on this up here.

The demonstrative may occur as the complement of either a locative or directional preposition. In this case 'here' or 'there' is an appropriate translation for the locative adverbial:

- (137) ?ah, reg tet ke, da? da?; reg tet nɔ?, da? da?
ah seek TO:spec that NEG EXIST seek TO:spec this NEG EXIST
Ah, (they) sought there, nothing; sought (here), nothing.

There is one other use which deserves mention. Demonstratives are also used in conjunction with *deŋ* 'manner' as manner adverbials, to express manner or similitude in the performance of an event (§10.3.2.2): *deŋ nɔ?* 'like this', *deŋ ke* 'like that'. The first example is of an actual situational use:

- (138) *The speaker is demonstrating how to do something.*
pɔy deŋ nɔ?!
- do like this
Do (it) like this!

In (139), it is used within the discourse context, and refers back to an immediately preceding portion of the text.

- (139) *The speaker has been telling me about the interaction between the European forces and the Semelai during the Communist Emergency.*
ʔleʔle? baysa? je=?en, dom deŋ ke=jia
so race 2=AUG AFF manner that=CL
So your people, yes, (they treated us) like that.

6.3.3 The major functions of demonstratives

The following is a summary of the major functions of the demonstratives identified so far: spatial, discourse and recognitional.

i. *Spatial Deixis* In the spatial function, demonstratives point out a referent, either a thing or a location. This may take place either in the actual utterance situation, or it may be in a narrated situation. The proximate term is used for entities located either with respect to the speaker, or alternatively, both the speaker and hearer may be placed at the deictic centre. The distal term is used for something which is located away from the deictic centre.

In spatial deixis situations, the demonstrative may be the complement of a locative prepositional phrase (§8.3.1), in which case the locative adverbial 'here' or 'there' is a more appropriate gloss:

- (140) wøy tɔ? ke
knife AT:across that
the knife across there

ii. *Discourse Deixis* In the discourse or linguistic context, the demonstrative is used to refer to an entity, either a thing, place or time within the text, as in (141), (142) and (144), or to an actual portion of the text itself, as in (143) and (145).

Both proximate and distal terms are used for tracking the reference of entities in discourse contexts, but in general the distal form *ke* 'that' is the unmarked term in

this function, as in (141), (142) and (144). The proximal term is also attested, as in (143).

- (141) *Mousedeer, caught and locked up in the pen, has convinced Macaque to swap places. The king returns and, infuriated, thinking that Macaque has eaten Mousedeer, stabs him to death.*

ki=tikam, ?ah, hn=kɔ? ke, kʰbəs
3A=stab ah O=macaque that be.dead
He stabbed, ah, that macaque (and it) died.

- (142) *An old woman is recalling her feelings over an incident when she was young. She indicates the remoteness from the present by modifying the pronoun with *ke* 'that, there'.*

yε ke, sampay da? moh ca, la=t-pekar hn=kəhn
1 there so.that NEG want eat BCS=HAPP-think O=3S
I (back) then, got so that I didn't want to eat, because of thinking of him.

The demonstrative may refer to the actual text. For instance, ?luc tom ke (pass SRC that) 'after that' indicating a single event or a larger scene, or kmɔŋ dɔ ke (finish OF that) 'that is the end', a typical way of concluding a traditional narrative. The distal and proximate terms are both used in this context, although the distal is the more common of the two.

- (143) *The narrator refers back to an event which has just taken place within the immediately preceding discourse. In an aside, she makes the following assertion:*

ns-gɔs dɔ gɔp nɔ?=pa
NM2-peel OF malay this=FACT
This, in fact, is the Malay way of peeling.

The distal form is also the unmarked choice for presenting participants within a narrative (144), as well as tracking them through the discourse (141). The first full nominal mention usually takes place only once the story is underway. It is difficult to tell if this is due to audience familiarity, which would place it under the recognitional use (see iii below), or some other reason. Only the distal term *ke* 'that' is available for this function.

- (144) kehn ke da? da? pn-reg ?en boh,
3S that NEG EXIST NMZ-seek LOC elsewhere
tpi ptom mn-jala?
daylight night PERFM-cast.net
ʔleʔle?=hn mn-jala? mi=tən ha?,
so=CONN PERFM-cast.net one=time AT
praw! te keh, praw! te nɔ?,
move.through.air! TO:unspec there move.through.air! TO:unspec here
ki=tibar jala?, beh da? ga=?ɔlɔ? ki=kʰɔm crɛh!
3A=draw.in casting.net NO NEG IMM=joke 3A=get fish
ʔleʔle?, beh da? ki=kʰe?, da? ki=?ye k<r>dor
so NO NEG 3A=know NEG 3A=see bc.female<NM2>

k<c>bac ki=jala?hn, kna? btol ?en k<r>dor,
fish<IMPERF> 3A=cast.net=O strike be.correct LOC be.female<NMZ>

2ah dom ph-poh k<r>dor ke, "ay" k'lej,
ah AFF IMPERF-struggle be.female<NMZ> that aye QUOTE

"k'om creh t'ay" k'lej sma? rabon

get fish be.big QUOTE person be.short-sighted

Him there, didn't have duties elsewhere, (so from) morning (until) night (he) was fishing. So, (he) was fishing one time hereabouts, (the net) flying though the air there, flying though the air here (as he cast it). He drew his casting net in. No, it wasn't a joke, he'd got a fish! So, no, he didn't realise, he hadn't seen the woman fishing. He had caught her (in the casting net), struck the woman with a direct hit. Ah indeed, the woman there was thrashing about. "Aye" (he) thought, "(I) got a big fish," thought the short-sighted man.

Another important function is to give prominence to an entity that has been presented in a series. This is particularly common in contrastive situations, and may be either pronominal or adnominal. In (145), the woman is summarising a contrasting series of entities that have formed the topic of the discussion. The distal term *ke* 'that' refers to the black millipedes which the woman has mentioned prior to talking about the red ones. Here it contrasts with the proximate term *ha?* 'AT', which is used in relation to the red ones she has only just mentioned:

- (145) me=mirah ha? ra?-kit, ke=spn ra?-t'ay
REL.=be.red AT COMP-be.very.small that=SC COMP-be.big
The aforementioned red ones are really small, these, see, are bigger.

For further examples of this function see (147) below.

The demonstratives can also be used to express relative time: sampay no? (until now) 'until now, at this point in time'; ye ke (1 that) 'me, back then'.

iii. Recognitional function In this function the demonstrative is used to point out an intended referent, something which is neither in the situational or discourse context, but with which the speaker judges the hearer to be familiar. In this case the entity is usually a person:

- (146) *The speaker knows that I am acquainted with the particular sister, however, she realises that I am going to require further information about the referent in order to be able to recognise her. Recall that the term kaka? 'EZ' includes not only one's elder female siblings, but also one's parent's siblings' offspring, as is the case here.*

dom kaka? ye, [kaka? mə=tuli? ke], dom gato?e?

AFF EZ 1 EZ REL=be.deaf that AFF [name]

Yes my elder sister, that elder sister who is deaf, yes gato?e?.

Forms used in this function are the distal *ke* 'that' (146), or the proximate locative *ha?* 'AT', in conjunction with *nej* 'before', see example (128) above.

There is a degree of ambiguity, and it is often unclear as to precisely which context holds. For instance in the following example, the distal demonstrative modifying *sma? rabon* 'short-sighted person', could be interpreted either as spatial or recognitional usage.

- (147) *A girl rushes home to tell her mother what has just happened to her. This is a continuation of the narrative in (144) above.*

?əŋ nej ki=jala? sma? rabon kke
If before 3A=cast.net person be.short-sighted that
Me, just now a blind man there caught (me) in his casting net.

- (148) *A wife is identifying a child in a cradle as hers and her husband's. It is the first time her husband has seen the child, because he abandoned her when she was pregnant.*

ki=tupuk, "kke knən he" k'lej
3A=point.out that offspring 1&2 QUOTE
She pointed, "That (is) our child," (she) said.

7 The noun phrase

This chapter examines the structure of the noun phrase (NP). The first three sections contain an overview of the phrase type. §7.1 describes the structure of the NP, a summary of the functions is set out in §7.2 and in §7.3 there is a description of potential heads. The discussion then turns to a description of the constituents within the NP in §§7.4–5. In §7.6 nominal derivation is examined: noun to noun in §7.6.1 and verb to noun in §7.6.2.

7.1 The internal structure of the noun phrase

The nominal forms the head of the noun phrase. Nominals may be modified within the NP by a quantifier which occurs pre-nominally, or by a demonstrative, an attributive or an associative which occurs post-nominally. The following diagram represents the unmarked or preferred order of constituents in the NP.

FIGURE 7.1 THE INTERNAL STRUCTURE OF THE NOUN PHRASE

Quantifier	Head	Attributive/Associative	Demonstrative
------------	------	-------------------------	---------------

Some examples illustrating the possibilities set out in Figure 7.1 are provided below.

The quantifier, often realised as a phrasal constituent, can be a numeral (1), or a numeral and either a classifier (2) or a measure noun (3). Quantifier phrases usually precede the head, but they can also follow it. The classifier phrase is frequently found post-nominally (3) and the ramifications of this are discussed in §7.3.2.

- (1) t̩mpoh t̩re?

seven day

seven days
- (2) p̩? ?ikur sma?

three CLF person

three people
- (3) g̩r̩m m̩=<pn>t̩m

salt one=wrap<NMZ>

one packet (of) salt

The attributive can be a locative prepositional phrase (4), verb phrase (5)–(6) or a relative clause (7). The occurrence of attributives is generally infrequent. The use of multiple attributives as in (6) is exceedingly rare.

- (4) sma? ha? tasik

person AT lake

the people at the lake

- (5) kn̩n̩n̩ k̩t̩

child be.small

a little child
- (6) sma? g̩d̩, b̩-j̩ypt̩, b̩-misay, b̩-huban

person be.old HAVE-beard HAVE-moustache HAVE-grey.hair

an old person, bearded, moustached (and) grey-haired
- (7) t̩jal̩=ce? ki=ba?i, tu?wan me=p̩?t̩ ha? dol

remain=CL 3A=be.alone grandmother REL=stay AT house

She remained (there) alone, the grandmother who stayed at home.

An associative always follows the head:

- (8) dol [sma? g̩d̩]NP

house person be.old

an old person's house

Demonstratives are always phrase final:

- (9) kn̩n̩ m̩=t̩ey ka

offspring REL=be.big that

that child who is big

7.2 Functions of the noun phrase

A noun phrase may function as:

a) the argument of a verb:

- (10) c̩ j̩l̩ jk̩s

dog bark.at porcupine

Dogs bark at porcupines.

b) the subject of a non-verbal clause:

- (11) [bapa?]Subj sma? cina?

[father]Subj person chinese

(His) father (was) Chinese.

c) the complement of a preposition:

- (12) da? ki=m̩h he? k̩mp̩n̩.

NEG 3A=want AT:above wife

He didn't want (to be) up at his wife('s).

d) an adverbial adjunct:

- (13) t̩ni pt̩m co? balay

day night AT:below verandah

Day (and) night (he was) down on the verandah.

Noun phrases also function as unattached units, e.g. vocatives, or afterthoughts as in (14). The NP kn̩n̩ 'offspring' is not predicated nor is it an adjunct. It is added as an afterthought to clarify the cause of the parents' crying. Normally this would be

expressed as an adjunct in the structure *la*= 'BCS', *la*=*knon* (BCS=offspring) 'on account of the children':

- (14) ?yam ?mabapa? ha? neŋ, knon
cry parent AT before offspring
The parents there cried, (on account of) their children.

Nominals can also occur as predicates (§10.1.1):

- (15) kəh ha? neŋ [kmn] ys_{Pred}
3 AT before [siblings'.offspring] 11_{Pred}
He, who was here just now, (is) my nephew.

7.3 The quantifier phrase

The quantifier phrase consists of a numeral (16), a numeral plus classifier (17), a numeral plus measure noun (18) or a non-numeral quantifier (19).

- (16) msɔŋ ten
five time
five times

- (17) tm̥poh bje? c̥ŋ
seven CLF hill
seven hills

- (18) dwet sams
money one. ringgit
money (to the value of) one ringgit

- (19) sbroh p?la?
all animals
all animals

Quantifier phrases usually precede the head as in the examples above. Numeral and classifier phrases can also follow the head like an attributive as in (20). Note in (21) that classifiers are not used for spears or blowpipes, only dogs.

- (20) ki=yok tyək, pe? bje? ?nom
3A=take banana three CLF be.ripe
He took bananas, three ripe (ones).

The ordering of the noun and numeral in the translation reflects the pattern in the original:

- (21) de=jon lmeŋ, muy, blahan, muy, co mə=2ikur
3A=give spear one, blow-pipe one dog one=CLF
They gave him a spear, one, a blow-pipe, one, (and) a dog, one.

Omission of the head is possible as in the following example where the phrase functions pronominally:

- (22) gi?=?ja 2əŋ mə=bje?!
- take=CL If one=CLF
Give me one!

The numeral can be used without either the classifier or the nominal being expressed:

- (23) yt=yɔk=?ja muy
1A=take=CL one
I'm taking one, okay.

7.3.1 Numerals

The first seven cardinal numerals, with the exception of the number two, are of Mon-Khmer origin.¹ They exist alongside the Malay set, shown in Table 7.1.

TABLE 7.1 THE NUMERAL SYSTEM

SEMELAI	MALAY	GLOSS
muy	satu?	'one'
dwa?	dua	'two'
hmpe?	tiga?	'three'
hmpon	empat	'four'
msɔŋ	lima?	'five'
pru?	enam	'six'
tm̥poh	tujuh	'seven'

Numerals higher than seven follow the Malay system: *lapan* ← *lapan* 'eight', *smilan* ← *sembilan* 'nine'. The system is extended by combining the Malay terms shown above, plus the following formatives: *puluh* ← *puluh* 'ten', *blas* ← *belas* 'for numerals eleven to nineteen', *ratus* ← *raios* 'hundred', *ribu?* ← *ribu* 'thousand', and the reduced form of *satu* 'one', *se-/sa-/* 'one': *dwablas* 'twelve', *tiga? puluh lima?* (three ten five) 'thirty-five', and so forth, following the Malay system exactly. Large quantities are not usually expressed precisely, the extended system being only marginal in a society that is not numerate.

Contracted forms of the numerals three and four occur before a noun: *pe?* ← *hmpe?* 'three', and *pon* ← *hmpon* 'four', *pon* *2are* (four day) 'four days', see (20).

The numeral *muy* 'one' has a reduced clitic form *mə=* (§3.3) which attaches to classifiers and measure nouns – *mə=2ikur* *co* (one=CLF dog) 'one dog', *mə=2<n>uta?* *mancis* (one=box<NMZ> match) 'one box (of) matches' – and also attaches to the numeral units *ratus* 'hundred' and *ribu?* 'thousand', see (24). *muy* 'one' also has the contracted form *mi=*, attested in *mi=tən* 'once' ← (*muy* 'one' + *tən* 'instance').

There is no intermixing of Semelai and Malay numerals, e.g. **tiga?* *puluh msɔŋ* 'thirty-five' is unacceptable and *tiga?* *puluh lima?* is used instead. An apparent exception is in the expression 'one hundred' where it does not enumerate, but instead quantifies an amount as in (24). The numeral 'one hundred' is *seratus* in Malay, but here the Semelai reduced form *mə=* 'one' is used rather than the Malay *se-*.

¹ The presence of Mon-Khmer numerals above three is a feature peculiar to the Southern Austronesian languages (Blagden 1906, Diffloth 1976d, Kruspe fieldnotes). Williams-Hunt (1952: 27) recorded Semelai numerals for eight, nine and ten: *kitwit* 'eight', *kantim* 'nine' and *kumai* 'ten'; these forms are no longer attested.

- (24) mə=ratus=jia de=jon
one=hundred=CL 3plA=give
They gave one hundred (ringgit).

There are constraints on the usage of the two systems. To name the months the Malay numerals are used: bulan tiga? (month three) 'March', bulan tujuh (month seven) 'July'. To enumerate months the indigenous system is used for the numbers 1-7, e.g. pru? bulan 'six months'. Likewise with time 'five o'clock' is Malay pukul lima? (hour five), but 'five hours' is msɔŋ jam?.

The use of ordinal numerals is exceedingly infrequent. The formation of ordinal numerals follows the Malay system. They are formed either by using the word namba? 'number' ← Malay *nombor*, followed by a Malay cardinal numeral, namba? tiga? (number three) 'third'; or the use of the Malay prefix ke-/kə-/ followed by a Malay cardinal numeral, k-tiga? 'third'.

Semelai numerals have a reduplicated form that is used to express 'a group of X number of people', e.g. sŋ-sŋɔŋ 'a group of five people' ← msɔŋ 'five' (§5.4.1.2). This form functions as an alternative to the classifier ʔikur for people (§7.3.2).

To express a sole participant the phrase consisting of a pronominal proclitic and baʔi? 'to be alone' is used, see (7) above.

Numerals may also be prefixed with kn- 'ENM' in order to enumerate: kn-dwa? ← dwa? 'two'. The numerals three and four prefix an allomorph kna-, e.g. kna-mpe? 'three'.

Some speakers only use these forms with inanimate entities; others use them for people, but only up to five, after which it is considered impolite.

- (25) ye=yɔk kna-mpon
IA=take ENM-four
I'll take the four (of them).

7.3.2 Classifiers

Classifiers are typically associated with the quantification of entities. Two systems of classification operate in Semelai. One is the system of numeral classifiers which are mostly terms borrowed from Malay;² the other is an indigenous system which can function either with a numeral as a numeral classifier, or independent of numeral classification as an 'individuator' (§7.6.1.1).

The set of classifiers in Semelai is small and rigid, and their use is not obligatory. The small inventory does not result in many entities employing the same classifier, but rather that many entities are simply not associated with a classifier. Although this topic requires further examination, it would appear that the optionality of the classifier in Semelai is based in its function. Hence, the assertion can be made that

² There are no restrictions on the combinatorial possibilities of borrowed and indigenous terms in Mah Meri where the two combine freely: muy balias (one ten) 'eleven' (Kruspe fieldnotes). Restrictions are attested in Jah Hut where the two indigenous numerals nl 'one' and nar 'two' can only be used in conjunction with indigenous terms. With numbers greater than two, the indigenous noun is replaced by a borrowed term: nar knɔ? 'two days', but tiga? harti 'three days' (Diffloth 1976c: 99-100).

³ The optionality of Mon-Khmer classifiers has led to a belief that their presence in Mon-Khmer is due to the influence of surrounding languages like Chinese and Thai. For classifiers in Mon-Khmer languages see (Adams 1986, 1989).

the function of classifiers in Semelai is not only to individuate entities in order to count them – this is in fact marginal given that not all entities are associated with classifiers, and of those that are, flexibility obtains – but individuation is a means of establishing referentiality, and providing anaphoric reference,⁴ see (3) and (22) above. The numeral mə= 'one' co-occurs most frequently with the classifiers.

We will examine the borrowed classifiers first, followed by the indigenous terms. Some examples of borrowed classifiers are listed below:

ʔikur	'animates: human, non-human'
bje?	'three dimensional objects: fruits, eggs, houses, hills, rivers'
bidaj	'things that are spread out: mats, thatch'
hlay	'length of fabric, strand of hair'
buku?	'lump: earth, congealed rubber'
rntaq	'oral narrative'

The two most striking differences from Malay usage of these classifiers are the use in Semelai of ʔikur to denote 'any animate entity, including people', see (26), whereas in Malay *ekor* 'tail' is used only for 'animals', e.g. dog, monkey, bird, fish, insect;⁵ secondly, the classifier bje? ← *biji* 'seed', the Malay classifier for small round objects, is collapsed with the Malay classifier *buah* 'fruit', used for objects of indefinite size or shape, resulting in one basic classifier bje? for three dimensional objects, see (17), (20) and (27).

In (26), the classifier follows the noun and the two numeral expressions function anaphorically.

- (26) kəh, ki=k^bəm pen ʔikur, dwa? kdor, dwa? rmɔl sdɔm
3 3A-get four CLF two be.female two be.male only
Him, he got four (of them), just two (who) were female, (and) two (who) were male.

The fact that the classifiers are based on the semantics of shape allows them to be used to express relative dimension as in the following example:

- (27) ny-t^bey mə=bje? dɔl
NMZ-be.big one=clf house
(It was) the width (of) a house.

Indigenous classifiers are derived from nouns expressing the typical container, using a process of affixation of the nominalising onset morpheme +.n+ 'UNIT' 'an N-FUL' (§7.6.1.3), or using the same derivational process, from the object which they enumerate. The terms which express contained measures are ambiguous between mensural classifiers and measure terms (§7.3.3), a distinction that is not always easily drawn (Craig 1992: 280). They are included here because their

⁴ The use of the classifier as a referent-tracking device is argued for in literary Malay in Hopper (1986). Hopper claims that the function of classifiers is to "allow the object to be deployed in the discourse" (1986: 316). The presence of classifiers marks the entity as integral to the discourse, rather than just being a 'prop' (1986: 313). This would make an interesting approach to the data in Semelai, and my impression is that Hopper's claims would be borne out. For further discussion on the function of classifiers, see Craig (1992).

⁵ Skeat (1896) records ʔikuch for humans in Mah Meri. It is no longer attested (Kruspe fieldnotes). The use of *eko* ← *ekor* for humans is also found in the Austronesian Iban communities on the island of Borneo (Asmah 1975: 243).

primary function is to express a subset of referents, rather than a precise quantity, although some may be used in this manner; secondly, because measure terms usually form a system of related quanta.

n ₂ -r?	'basketful'	← r?	'basket'
s<n>udu?	'spoonful'	← sudu?	'spoon'
k<n>opi?	'tinful'	← kopi?	'small tin can'
k<n>bə?	'bodyful'	← kba?	'body'

- (28) mə=k<n>amponj sma?
one=settlement<UNIT> people
one settlement! (of) people

The classifier b<nh>nih 'a measure of rice seed for planting', is derived not from the container, a basket, but the typical contained entity bnih 'seed'.

- (29) mə=b<nh>nih baba
one=UNIT-seed unhusked rice
one basketful (of) unhusked rice

Using the same morphological process, the classifier may be derived from the entity to which it refers: ns-das 'a panel' ← das 'a panel of pattern in mat-weaving', m<n>alaj 'a side' ← malaj 'side, behind'. The classifier for long thin objects – plants with woody stems, posts, cigarettes, teeth – is derived irregularly: d<ŋ>loŋ 'a felled, cleaned tree, a post; CLF' ← dloŋ 'tree'.

- (30) loc n=tap mə=ns-das
already 1fA=weave one=UNIT-panel
I've already woven one panel.

- (31) ki=kłm bəs mə=d<ŋ>loŋ
3A=carry.on.shoulder sugarcane one=<UNIT>tree
He carried a length of sugarcane on (his) shoulder.

The classifier for leaves of the betel-vine, nȳ-təŋ 'a leaf of the betel-vine, CLF' is derived from təŋ 'ear', the avoidance term for the leaf, see §1.1.

7.3.3 Measure nouns

Measure nouns name an intrinsically quantified amount. They do not co-occur with classifiers, given the general mensural nature of the classifier (§7.3.2). The measures dealt with in this section are: i) conventional measures, including local, imperial and metric, adopted from current Malay usage, and ii) the use of intrinsically quantified terms to express time. The occurrence of the conventional measures is infrequent. The more usual method for conveying information regarding dimensions or volume is to express it not in terms of a conventional measure, but in relation to a familiar object, placing the two nouns in apposition:

- (32) ... do mur no?, ppokel sksek, da? cukup
... OF circumference this pick plaque NEG enough
... a circumference of this size, a tooth-pick, isn't enough.

i. *Conventional Measures – Local, Imperial and Metric* Traditional local measures are based on the length of body parts to express length, or the capacity of a coconut shell to express volume. These terms are borrowed from Malay.

Some expressions of length are: jn>re? 'fingers width' ← jare? 'finger' (§7.6.1.1), jŋkol 'span from thumb to fore-finger' ← jengkal 'a span', ?sta? 'an ell, the span from middle finger-tip to elbow'. t<ŋ>kəl 'a cut length, rattan, wood' ← tkəl 'to cut' was used in external trade as a measure term for approximately nine feet of cut rattan.

Volume is expressed in terms of a local traditional standardised measure. The measures in Malay are equivalent to those given for Semelai: leng 'the capacity of a small half-coconut-shell' ← leng, cupuk 'the capacity of a half-coconut-shell, or two leng' ← cupuk, gantan 'the capacity of four cupuk' ← gantang.

The following terms are not regularly used: ?inci? 'inch' ← inci 'inch', batu? 'mile' ← batu 'mile', keti '600 grams, (approx.)' ← kati 'a catty'.

ii. *Monetary Terms* The term mos (emas ← Malay 'gold') is an archaic term for a unit of currency. ryal 'dollar' is an old Portuguese currency term, but it is also a current colloquial form in some regions of the Peninsula for the standard ringgit 'dollar'. Amounts above five ringgit use a Malay numeral and the term ryal. The etymology of the suppletive form sphah 'four ringgit' is uncertain, although it may be related to the Malay prefix se- 'one' and paha 'the ham of an animal, a quarter'.

sams	'one ringgit'	sphah	'four ringgit'
dwa? mos	'two ringgit'	lima? mos	'five ringgit'
tiga? mos	'three ringgit'	?ənam ryal	'six ringgit', etc.

There are similarities between the local system and the one used in Malacca during Portuguese rule (Gianno 1990: 37).

ii. *Time* Time is expressed in terms of the name of a period of the day, month or year. These are directly enumerated. Some examples are:

som	'morning, daybreak until approximately 10am'
day	'afternoon, approximately 2pm–6.30pm'
ptom	'night, dusk till dawn'
?are?	'day, period between dawn and dusk' ← Malay hari 'day'
tahon	'year' ← Malay tahun 'year'

Twenty-four hour days are counted in terms of ?are? 'day' and ptom 'night', the two being viewed as separate entities, not two parts of a whole:

- (33) pe? ?are?, pe? ptom
three day three night
three days (and) three nights

The lexemes ptom 'night' and som 'morning' derive mensural classifiers p<n>tom (night<UNIT>) 'night-long' and nən-som (UNIT-morning) 'morning long' expressing duration:

- (34) mə=p<n>tom da? jtek cəŋ
one=night<UNIT> NEG sleep at.all
The whole night long (I) didn't sleep a wink.

7.3.4 Non-numeral quantifiers

Quantifiers are used to express approximate quantity, and consequently never co-occur with numerals. They tend to precede the nominal. Quantifiers are drawn from the class of stative verbs, which are used predicatively (35), ignoratives (36), and a few terms which function solely as non-numeral quantifiers like sdom 'only' (37), and sbroh 'all' (38).

- (35) ki=jpk jrks ple klub?
3A=observe be.plentiful fruit *Salacca conferta*
She saw the Salacca fruits were plentiful.
- (36) de=cəŋ papel nojom, brapa? - brapa? nojom
3A=immediately call seer RDP-how.many seer
They immediately summoned seers, several seers.
- (37) sdom kəh, ki=k^be? simpy dwablas
only 3 3A=know braid twelve
Only him, he knows (how) to braid twelve-strands.
- (38) he b<paloh>an sbroh
1&2 TOG <flee> all
We were all fleeing together.

A quantifier can be used without the head:

- (39) sparoh swak
some walk
Some walked.

sbroh 'all' has a nominal form s<m>broh 'all' which functions pronominally:

- (40) s<m>broh ki=k^be?
all<NMZ> 3A=know
She knows (it) all.

The following are examples of terms that function as non-numeral quantifiers:

sdekēt	'a little'	← sedikit	'a little'
tyap	'each'	← tiyap-tyap	'each, every'
cukup	'be enough'	← cukup	'complete'
?ray	'to be many'		

Some terms are specific to certain entities; the following terms all mean 'to be many, plentiful', but are restricted to quantifying the following:

ramay	people	↔ ramai	'be lively, populous'
jrks	tubers, Salacca fruit		
ssy	rambutan, langsat and pulasan fruits; hair		
smol	segments of fruits such as durian, jackfruits		

7.4 The attributive phrase

The attributive phrase always follows the head. It can be a prepositional phrase (§7.4.1), a relative clause (§7.4.2) or a verb (§7.4.3).

7.4.1 The locative prepositional phrase

Prepositional phrases and prepositions can function attributively to locate a nominal. In (41) the attributive is a prepositional phrase, in (42) the preposition functions deictically (§8.3.3).

- (41) The speaker is making reference to the Bera River Malays, as opposed to others:
gɔp cɔ? bra
malay AT:down Bera.River
Malays down at Bera River
- (42) 2eh, papel sma? tɔ?
eh summon person AT:across
Ah, (he) summoned the people across (there).

7.4.2 The relative clause

Relative clauses function to modify an NP. Their structure is discussed in detail in §11.1. Some examples of relative clauses are given below, with the relative clause bracketed:

- (43) tu?wan [mə=p?ot ha? dɔl]
grandmother [REL=stay AT house]
The grandmother who stayed at home.
- (44) projkol! wah-wah tom ha? ?ate [ki=bɔy]
jump.up! IMPERF-rise SRC AT earth [3A=dig]
Jumping up suddenly, (he) rose from the earth (into which) he had dug (himself).

Adjectives preceded by the relativising/nominalising prefix mə= 'REL' can add further information about the head (45). This form is often used to name a category as in (46) and is commonly used to form lexical items in the avoidance vocabulary (47). In both cases the head is not represented.

- (45) knkon ke, mə=t^bey ke
offspring that REL=be.big that
that child, that (one) who is big
- (46) mə=gdo
REL=be.old
the elderly
- (47) The avoidance word for mouse deer is based on its physical attribute of a small hoof-size.
mə=kurus jɔŋ
REL=be.thin foot
(the one) who is thin (of) hoof

7.4.3 The verb

Verbs used attributively follow the head nominal. Intransitive verbs, unlike transitive ones, do not require modification, as in (48).

- (48) This is the Semelai name for illegal Indonesian immigrants. No sooner than they are repatriated, they return, as if they have the ability to fly.
 sma? per
 people fly
 people (who) fly/flying people

In (49) the verb is modifying the associative *rəh dlon* (branch tree) 'branch (of) a tree'.

- (49) resnes [rəh]N dlon k'bəs]
 twig [branch tree be.dead]
 Twigs (are) tree branches (which) are dead/dead tree branches.

Transitive verbs require the detransitivising middle voice prefix *br-* 'MID' (§5.3.3) in order to function attributively:

- (50) zate br-bəy
 earth MID-dig
 dug up earth

Adjectives The following points refer specifically to the adjective class of verbs. In the following example two adjectives describe the one entity:

- (51) co putih hitam
 dog be.white be.black
 a white dog (with) black (markings)

The adjective may be prefixed with the comparative morpheme *ra?* 'COMP' (§5.3.8):

- (52) cəŋ ra?-pəs
 hill COMP-be.high
 a higher hill

The head may be repeated with each adjective in a distributed reading (53), or the adjective may occur without the head (54).

- (53) ?le?je? de=kde? dol tʰay, dol kət
 eventually 3plA=stay.in house be.big, house be.small
 Eventually they stayed (some) in big houses (and) (some) in small houses.

In the following example there is the added dimension of contrast, expressed by the particle *ra?* 'COMP':

- (54) dol ra?-tʰay, ra?-kət, 2lok, buruk=spn ...
 house COMP-be.big COMP-be.small be.good be.worthless=SC
 (whether) a bigger (or) smaller house, (in) good (condition, or) bad, see ...

The adjective precedes the post-nominal classifier, as in (20) above.

If the head is elided, the classifier precedes the adjective, functioning pronominally:

- (55) mə=?ikur ra?-kət
 one=CLF COMP-be.small
 one smaller one

7.5 The associative phrase

In Semelai, the associative construction is Head-NP, the normal order of constituents of the Semelai NP. This construction serves to denote a specific relationship between two entities, such as the possessive in (56). An alternative construction to the associative phrase is the *do* 'OF' construction, where the morpheme *do* 'of' formally marks the relationship (57).

- | | |
|--|--|
| <p>(56) co 2i=təŋ
 dog NM=[name]</p> | <p>(57) co do 2i=təŋ
 dog OF NM=[name]
 the dog of Mr. ton</p> |
|--|--|

This construction is discussed in §8.9.

A second type of associative construction is the locative associative which is interpreted as Head-NP, or as NP-Head, and is parallel to the locative prepositional construction P-NP as discussed in §8.3.1. Two interpretations are possible, part-whole as in a), or locative as in b):

- (58) kloc rɔŋ
 inside basket
 a) the inside (of) the basket
 b) inside the basket

The difference between an associative construction and a compound noun is that in an associative construction each constituent in the construction retains its individual meaning. This is the case in (59) where the speaker describes a woman's situation. She has chosen to live amongst the tigers rather than with her own people:

- (59) k<r>dor pəpəŋ
 be.female<NMZ> tiger
 the tiger woman

In contrast, in a nominal compound the meaning of the compound is never simply the sum of the two parts. In (60), the word *djədəŋ* 'log' does not have any relationship to 'bone'.

- (60) djədəŋ muh
 log nose
 (a) nose bone
 (b) the bone of the nose (anatomical)

Secondly, in an associative construction, individual constituents of the phrase may be independently modified, evidence that it is not a cohesive syntactic unit:

- (61) co ra?-tʰay kəh
 dog COMP-be.big 3
 his bigger dog

In a compound, an attributive can only occur following the compound. If the attributive intervened the meaning of the compound would not be maintained as in (b):

(62a)	[tom dloŋ] ra2-t̥ay	SRC tree COMP-be.big the bigger bole
(62b)	*tom ra2-t̥ay dloŋ	SRC COMP-be.big tree *bigger source (of) the tree

Thirdly, compounds cannot represent the second N as enclitic =hn '3POSS' which is illustrated in §7.5.1.

7.5.1 Possessive enclitic =hn '3POSS'

An associative construction usually consists of two nominals or a nominal and free pronominal as in the examples above. In the possessive construction the third person human or anthropomorphised possessor, whose identity is contextually recoverable, can be represented by enclitic =hn '3POSS' attached to the possessed nominal.⁶ Only core arguments of the clause can be represented in this manner (§9.1.3).

(63)	cɔ=hn kl=g<n>lp̥-i7	2i=nac̥t
	dog=3POSS 3A=name<CAUS>-ITER	NM=[name]
	His dog, he named Mr nac̥t.	

(64)	kl=slet wɔy=hn 2en sɔŋ bñul	
	3A=insert knife=3POSS LOC end floor.beam	
	He slipped his knife into the end (of) the floor beam.	

7.5.2 Functions of the associative phrase

The following examples illustrate the range of semantic relationships expressed by the associative construction. The construction can be semantically vague, e.g. in kmbəc tu?ak1? (fishing.rod grandfather) 'grandfather's fishing rod', the grandfather may be either the owner or the user, or both.

- a) *relationship-person* knɔn kp̥ (offspring 2f) 'your offspring', 2adi? hc=2en (ys 1&2=pl) 'our younger sibling';
- b) *entity-owner* 2amɔj 2əŋ (backbasket 1f) 'my back-basket', wɔy=ha (knife=3POSS) 'his/their knife';
- c) *entity-user* kmbəc tu?ak1? (fishing rod grandfather) 'grandfather's fishing rod', wɔy kəh (knife 3) 'his knife';
- d) *product-creator* rɔ? kaka? (basket EZ) 'elder sister's basket';
- e) *activity-performer* s<n>wak kəh (walk<NMZ> 3) 'his walking', nji-taj kaka? (NMZ-weave EZ) 'elder sister's weaving';
- f) *part-whole* lher rɔ? (neck basket) 'neck of a basket', t̥bi sma? (hand person) 'person's hand';
- g) *object-time* bintaj kmraw (star dry.season) 'stars of the dry season', d̥ɔh padəy yuŋ (swidden past (unknown))⁷ 'five year old swidden';
- h) *object-purpose* dak j<nh>2ɔh (water drink<NMZ>) 'drinking water'; t̥laga? nm-hūm (well NMZ-bathe) 'bathing well';
- i) *entity-source* mbam t̥əŋ (blood ear) 'blood from the ear', 2əc t̥əŋ (excreta ear) 'ear wax', gtah dloŋ (exudate tree) 'exudate from a tree'.

⁶ =hn '3POSS' is probably related to Mah Meri han '3POSS'.

⁷ yuŋ has no known meaning other than here in this compound.

(62b)	*tom ra2-t̥ay dloŋ	SRC COMP-be.big tree *bigger source (of) the tree
-------	--------------------	--

The semantic distinctions between b)–e) above are best clarified using the following paraphrases: b) 'the X which Y owns'; c) 'the X which Y is using, but does not necessarily own'; d) 'the X that Y created', and e) 'the X that Y performed'. The language does not make a formal distinction.

- j) *entity-beneficiary* The associative construction is used to express the benefactive relationship.

(65)	ji rajin beh, tan rɔ? 2əŋ?	2 industrious NO weave basket 1f
		Would you please weave my basket/a basket (for) me. (Lit. Are you industrious or not, to weave a basket (for) me?)

There is no means of distinguishing whether this is a [N N]_{NP} construction, or [N]_{NP} [N]_{NP} where the second is a benefactive. The beneficiary NP cannot precede the entity. Some further examples follow:

- (66) *The speaker is going to the shop to buy a tin of milk for ʔamay.*

ye=reŋ susu?	ʔamay
1A=seek milk [name]	
I'm getting ʔamay's milk.	

(67)	dos ha? ke, j?oy ponoy krðor	reach AT there make hut woman
		(He) arrived there (and) made a hut (for) the woman.

7.5.3 Locative nominals in the associative phrase

Locative nouns in an associative-type construction provide further specification of location (68), or express a part–whole type relationship (69) as discussed in f) in the previous section.

(68)	t̥ah b<k>tək	middle be.overgrown<NMZ>
		amidst the undergrowth

(69)	joŋ tip̥?	the foot (of) the betel service

(70)	ki=pte? hn=bras kloc pasuh	3A=put O=rice inside jar
		He put the rice inside the jar.

The locative associative can function as the complement of a preposition to express differentiated coincidence (71)–(72). The preposition alone fails to make these finer distinctions (§8.1).

(71)	hc? dyal c̥ɔŋ	AT:above top hill
		on top (of)/above the hill

- (72) krwl̩ tom ha? karom prahō?
 come.out SRC AT underneath boat
 (He) came out from under the boat.

The following list contains the possibilities of the location nominal:

kloc	'inside'	sawcl	'left-hand'
karom	'underneath'	satom	'right-hand'
dyal	'above'	bjŋ	'foot, lower extremity'
lwar	'outside'	kʰoy	'head, upper extremity'
tjah	'middle'	clən	'back, behind'
malan	'side'		

7.5.4 The associative phrase and compounds

A particular lexicalised type of associative phrase is what I will label 'nominal compounds'. Though phonologically indistinguishable, I distinguish them from regular phrases because of the syntactic cohesiveness of the constituents as mentioned in the introduction to §7.5 above.

Nominal compounds are formed by the syntactic juxtaposition of two nouns, in a relationship Head-NP. The majority of nominal compounds in Semelai are syntactic compounds of the associative phrase type.

7.5.4.1 Criteria for distinguishing compounds

It is difficult to find criteria for distinguishing compounds from simple associative phrase constructions. In a sense, compounds are 'institutionalised' examples of the associative phrase which, by virtue of being institutionalised, become cohesive syntactic units.

As mentioned above, it is difficult to establish any firm phonological criteria for differentiating compounds from the associative phrase, although perceptually compounds are less likely to have an intervening pause.

In the text, compounds are indicated by a point between the lexemes, and are glossed lexeme by lexeme in the interlinear gloss, but translated as a compound.

The most important point is that compounds are inaccessible to the rules of syntax. In an associative construction, individual constituents of the phrase can be independently modified, see (61) above.

In the case of compounds which have the semantic relationship of possessee-possessor, relationship-person, or part-whole (73), the possessor cannot be marked by the preposition *do* 'OF'. In (74) the compound does not retain its meaning, in this case the meaning is nonsensical, as indicated by the question mark.

- (73) *djedəŋ.muh*
log.nose
nose bone

- (74) ? log of nose

In contrast, there is no semantic difference between the associative phrase in (75) and the alternative possessive construction in (76).

- (75) *A woman has chosen to live amongst the tigers rather than with her own people:*
 k<r> dor lpcq
 be.female<NMZ> tiger
 the tiger-woman

- (76) *k<r> dor* *dō* *pedjep*
 be.female-<NMZ> OF tiger
 the woman of the tigers

An attributive can only occur following the compound. If the attributive intervened the meaning of the compound would not be maintained, see the examples in (62) above.

Compounds cannot represent the second N as enclitic =hn '3POSS' (§7.5.1).

7.5.4.2 Compositional and non-compositional compounds

The semantic relationships between the head and the NP are set out in detail in §7.5.2. In general, the same relationships expressed in the associative phrase are also attested in compounds.

There are two types of compound, compositional and non-compositional, which are discussed below.

Endocentric compounds denote a sub-grouping of the class of entities named by the head. The second lexeme in the compound modifies the head by specifying its meaning. Usually the modifying lexeme is a noun in the associative construction, but sometimes another eligible nominal modifier may fill the slot, e.g. a verb as in (77) and (78).

- (77) *?ubat.cher*
antidote.have.hunger.pains
breakfast

Here the modifier is a derived verb *b-ribut* (HAVE-wind) 'be windy':

- (78) ?are? b-ribut
rain.HAVE-wind
storm

Occasionally a compound may consist of more than two lexemes, particularly terms expressing cognatic or affinal relationships: ?adi?bradi? sadara? ?anje? (have.sibling relation dog) 'half-sibling, different fathers',⁸ ?adi?bradi? bdiri? gad? (have.sibling standing parent) 'first cousins', sma? knlə? kmpən (person husband wife) 'married couple'.

The two types of compound, compositional and non-compositional are examined in turn below.

i. Compositional compounds The meaning of the compound is semantically compositional, i.e. it is either identical to the component parts, or directly related to them:

- | | | |
|-----------|---------------|-----------|
| gadə? jəŋ | (parent foot) | 'big toe' |
| tahŋ muh | (hole nose) | 'nostril' |

* *Tanjen* 'dog' is the Malay term *anjing*; it is not used in Semelai other than in this context. The compound is probably a direct loan from Malay (cf. Wilkinson 1927).

bri kmuc	(jungle ghost)	'underworld'
dak jor	(water urinate)	'urine'
huc kuniq	(rice yellow)	'yellow glutinous rice'
tpuj tawar	(flour antidote)	'flour antidote ceremony'

Certain lexemes occur frequently as formatives, for example *knøn* has a primary meaning 'offspring', e.g. *knøn eo* (offspring dog) 'puppy'; in compounds it may also mean 'dependent; smaller component' as in *knøn dak* (offspring water(way)) 'tributary'. Another lexeme used frequently is *?ec* 'faeces; exudate; waste-product': *?ec rokok* 'cigarette ash'; *?ec lo?* (exudate penis) 'semen'.

Some compounds are possibly the result of a direct calque on Malay compounds, e.g. Malay *mata* 'eye' is polysemous also meaning 'source' or 'orifice', although, like the pattern involving 'child' this pattern may be more widely attested (for instance in some Australian aboriginal languages 'eye' is used for 'spring'). Semelai *mot dak* (eye water) 'spring' = Malay *mata ayer* with the same meaning; *mot taja?* (eye ladder) 'space between rungs in a ladder, steps' = Malay *mata*.

In terms used in cosmology, Malay lexemes are frequently employed (reflecting a general trend in the use of Malay for specific domains, such as incantations), and institutionalised in these compounds: *lawot darah* (sea blood) 'name of a specific location in the underworld', *hantu?ayer* (ghost water) 'water spirit'.

ii. Non-compositional compounds In the second type of compound, which is less frequent than the first, there is no apparent or easily identifiable link between the meaning of the constituents and that of the compound. This may due to the fact that:

- a) the compound is metaphorical, requiring cultural knowledge in order to derive the meaning of the compound from its component parts:

tijo rboi	(snake aroid)	'rainbow'
tunku? cny	(hearth-post mousedeer)	'a termite mound'
bkaw muh	(flower nose)	'the lower flared portion of the nose'

- b) the meaning of either one or both of the component parts is no longer recognised by speakers: *skol k^hoy* ((unknown) head) 'headache'.

7.6 Nominal derivation

7.6.1 Noun to noun derivation

The function of noun to noun derivation is specifically one of individuation, either: a) a generic noun derives a specific 'individual type' noun 'INDIV', or b) a noun that names a container derives a mensural classifier 'UNIT'. For the morphological specifics, see §3.2.2.3.

7.6.1.1 Individuated nouns 'INDIV'

These nouns refer to 'an individual type or specimen' as opposed to the superordinate class of entities. Some examples are:

nm-cim	'kind of bird'	← cim	'bird'
nk-dak	'kind of watery liquid'	← dak	'watery liquid'

s<n>kø?	'kind of ghost'	← skø?	'ghost'
k<nān>yen	'kind of cloth'	← kayen	'cloth'

(79) dmoh d<ng>lo?

what tree<INDIV>

What kind of tree (is that)?

7.6.1.2 Dysphemistic nouns

The nominalised nouns can host a morpheme *?i=* 'NM', a clitic usually used to form male reference terms. When affixed to an individuated noun, it derives a dysphemistic noun (§3.3.7). Dysphemistic nouns⁹ are used to express dissatisfaction, frustration or contempt towards the referent, either an animal or an object.

?i=n̩-təŋ 'damned leaf of the betel-vine'

← n̩-təŋ (INDIV-betel-vine leaf) 'leaf (CLP)', 'type of betel-vine leaf'
← təŋ 'ear' 'betel-vine leaf (av.)'

?i=ni-co 'damned dog'

← ni-co (INDIV-dog) 'type of dog'
← eo 'dog'

(80) *This was uttered by an unusually disgruntled hunter after a fruitless expedition.*

da? da? ?i=j<n>alu

NEG EXIST NM=pig<INDIV>

There wasn't a damned pig (to be had).

(81) *A mother is losing patience with her child who for some reason will not come down from the veranda:*

sar! sar! da? da? ?i=p<n>2la?

descend descend NEG EXIST NM=animal<INDIV>

Come down! Come down! There aren't any damned animals!

7.6.1.3 Classifiers 'UNIT'

Nouns that name a container derive mensural classifiers expressing the following:

a) volume, 'an N-ful' n^h-ro? 'basketful' ← ro? 'basket', p<n>ereŋ 'plateful'
← pereŋ 'plate';

b) area m<n>alay 'a side' ← malay 'side', ny-k^hoy 'area of the head' ← k^hoy
'head';

c) duration p<n>tom 'duration of a night' ← ptom 'night-time';

d) a numeral classifier k<n>asaw 'rafter' ← kasaw 'rafter'. Numeral classifiers are often ambiguous between an individuating and a classifier sense. Classifiers are discussed in §7.3.2.

(82) me=p<n>ereŋ huc

one=plate<UNIT> rice

one plateful (of) rice

⁹ A dysphemism is defined by Allen and Burridge as "... an expression with connotations that are offensive either about the denotatum or to the audience, or both, and it is substituted for a neutral or euphemistic expression for just that reason" (Allen and Burridge 1991: 26).

7.6.1.4 Nominal reduplication

Nominals may be reduplicated utilising the process of light syllable reduplication (§3.2.2.4); glossed here as RDP 'reduplication', e.g. r̩im̩o-r̩im̩o 'various men' ← r̩im̩o 'male human'.

The derived form expresses a plurality of reference, either a range of possible referents from the domain of intended referents or it may signify a lack of specificity of the intended referents, a distinction which is contextually determined. Although the reduplication signifies plurality, it does not function as a means of expressing number alone.

- (83) jadi²=son, dom, kraba-krabat de=ca<r>ln
happen=SC AFF RDP-official 3plA=back<CAUS>
And so it happened see, indeed, the various officials turned their backs (on him).

The reduplication of ignoratives may derive indefinite pronouns: brapa² 'how many' → brapa-brapa² 'how ever many; several' (§6.2.1.2).

7.6.2 Deverbal nouns

It will become apparent that nominalisation is idiosyncratic in terms of applicability with respect to both form and function.

Lexical nominalisation in Semelai is a relatively complex and infrequent process, and as a result, the precise categorial status and functions are not fully understood. Just as verbs typically refer to events and nouns to entities, the expression of an event as an entity results in a deverbal noun which both exhibits features typical of the nominal class, and at the same time retains features of the class from which it originates (see Koptjevskaya-Tamm for Russian, 1993: 8-9).

The following discussion is divided into three parts. In §7.6.2.1 we begin by examining the various morphological processes of nominalisation and attempt to capture the semantic generalisations associated with each form. The structural aspects of the morphological processes are discussed in §3.2.2. The syntactic behaviour of nominalised forms is surveyed in §7.6.2.2, and in §7.6.2.3 the relations between action nominalisations and the arguments associated with the deverbal nouns are briefly described.

Two types of deverbal noun are identified: action nominals, forming the majority which denote an action or process, and less frequently, nominals which refer to a concrete entity, the result of the action or process.

An action nominal is a deverbal noun, denoting a general action or process, which may occur as the head of a nominal construction. Action nominals are used to express an act, a fact, an event, or the manner of performance of the act or event as in (85). It is only context which determines the different interpretations (for similar observations, see Koptjevskaya-Tamm 1993: 5, 18-21).

- (84) taŋkol, c<n̩ŋ>teŋ r̩om taŋkay
carry.by.stem carry.at.side<NMZ> WITH stem
taŋkol (is) the act of carrying (s.th.) by the stem at one's side.

- (85) nl-cəl kəh da? btəl
NMZ-pronounce 3 NEG be.correct
Her manner (of) pronunciation isn't correct.

Included in 'action' nominals are forms derived from stative verbs, which are used to express a fact, or the manner of being in a state. The derivation of stative verbs is less productive than with verbs denoting either actions or processes. For example, the eight adjectives expressing dimension only derive one of each of the four antonymic pairs (§5.3.8), only a subset of physical property adjectives feed nominalisation, and colour adjectives do not feed the process (Kruspe forthcoming). Stative intransitive verbs expressing human physical conditions as in (86) derive nominalised forms, while those expressing emotional states do not.

- (86) jd̩ə? pt-k̩ət n̩ə?
be.beautiful NMZ-be.pregnant this
(She) looks well (during) this pregnancy.

Action nominalisations are only used with reference to a specific instantiation of an event or a specific state, with a particular subject – see (85). They are never used as 'subjectless' abstract nouns, e.g. as in the English 'Felling is hard work.' (This is expressed by an unmarked verb: snaj ca pri (be.pleasant eat free) '(It was) nice to eat (for) free'.

Nominals which have concrete entities as their referents usually denote the object or result of the action or state, p<t>ret 'heat (of the day), sunshine' ← pret 'to be hot'. Interpretations of this sort are less productive than action nominalisations.

7.6.2.1 Nominalisation processes

The following morphemes derive nominalised forms from verbs. The same gloss is given to all forms because of the many to many relationships between morpheme choice and meaning category, the exception being the excessive agent nominalisation par- ~ pr- 'xs' (Table 7.2). The affix +n+ 'NMZ', the most productive and predictable morpheme, is used primarily to derive action nominals. The remainder exhibit differing degrees of productivity and predictability.

TABLE 7.2 LEXICAL NOMINALISATION PROCESSES

FORM			FUNCTION
prefix	infix	suffix	
+n+(<C>)			Action/Location/State
<C>			State
-	-(?)an		Action/State/Result
+pn+			
+pC+	<n>(<C>)		Process
<n>(<C>)			
par-			Agent
			Excessive Agent

The presence of coda copy +C+, in the derivational process, associated with the formation of the imperfective (§5.3.1), is not to mark an aspectual distinction, but rather to signal a reduction in valency.¹⁰ The deverbal noun always has a valency of

¹⁰ Benjamin notes a similar situation in Temiar. Although nominalisation can co-occur with the various morphologically expressed aspects, the aspectual meaning is not carried over into the nominalised form (1976b: 174). In Semelai, nouns which do not usually have coda-copied forms often have freely alternating nominalised derivatives, one derived from the root: sma? 'person' → s<n>ma? (person<INDIV>), and the other derived from a coda-copied base s<n>ma? (person<INDIV>) 'kind of person'.

one. Verbs which derive a coda-copy base, i.e. most transitive verbs, utilise this base to feed nominalisations as in a); intransitive verbs nominalise the bare root form as in b). The process is not glossed independently, but considered part of the general nominalisation process of reduced transitivity.

- | | | | | |
|----|----------------------|----------------------------|----------------------|----------------------|
| a) | nr-c ^b or | 'firing' | ← c ^b or | 'to treat with fire' |
| | s<np>maj | 'request for (s.th.)' | ← smaj | 'to request (s.th.)' |
| b) | p<n>tɔt | 'act of residence' | ← p ^b tɔt | 'to stay' |
| | k<m>req | 'alcoholic beverage (av.)' | ← kreq | 'to be dry' |

i. *Action nominalisation +n+(<C>) 'NMZ'* Action nominalisations are derived from both active and stative verbs. Active verbs derive nouns expressing 'the occurrence or mode of V-ing':

- | | | | |
|----------|------------|-------|--------------------|
| j<nh>zɔh | 'drinking' | ← jɔh | 'to drink' |
| nk-yɔk | 'taking' | ← yɔk | 'to take' |
| nj-tap | 'weaving' | ← tap | 'to weave' |
| ns-gos | 'peeling' | ← gos | 'to peel (rattan)' |
- (87) nk-yɔk kəh da? btol
NMZ-take 3 NEG be.correct
His manner (of) taking isn't correct.

Result nominalisation Some active transitive verbs have an interpretation which names 'the thing which is typically V-ed', or 'the evidence of the occurrence of V-ing'. There remains a strong link between the action and object reflected by the fact that in some instances an action nominalisation is also a possible interpretation. These derived forms do not denote the entity in isolation from the act of creation. To talk about 'peeled rattan' in general, one would use the middle voice construction drc br-gos (rattan MID-peel), not the nominalised form ns-gos (NMZ-peel). Some examples follow:

- | | | | |
|----------------------|--------------------------|---------------------|-------------------------|
| nk-c ^b ak | 'projectile'; 'throwing' | ← c ^b ak | 'to throw, with an aim' |
| t<nk>bok | 'a carved out space' | ← tbok | 'to carve out a space' |
- (88) c<nn>man sma? g^b-go?
retell.myth<NMZ> person IMPERF-fell.tree
The retelling of the myth of the person (who was) tree-felling.

Location nominalisation Some verbs derive a form which is ambiguous between an action nominalisation and a location nominalisation, naming 'the place where V happens'. Forms with this interpretation are derived from verbs naming postural positions and eliminatory functions performed by animates, and some process verbs such as fishing, all of which are intransitive verbs that typically take place in a static location. Again ambiguity between this and a mode nominalisation is attested.

- | | | | |
|----------|--------------------------|------------|-------------------------|
| nm-dəm | 'location of lying' | ← dm-dəm | 'to lie down, be lying' |
| k<nc>bəc | 'fishing spot' | ← k<nc>bəc | 'to fish with a pole' |
| nr-jɔr | 'place for/of urinating' | ← jr-jɔr | 'to urinate' |
| j<nr>tək | 'sleeping area' | ← jtək | 'to sleep' |
| p<n>aloḥ | 'escape route' | ← paloḥ | 'to flee' |

- (89) tɔm ha? tmpat nm-k^bom sunɔt kəh
SRC AT place NMZ-sit circumcise 3
from the sitting place (of) his circumcision

State nominalisation Monosyllabic adjectives derive nouns that express 'the name of the dimension/attribute associated with V':

- | | | | |
|----------------------|---------------|---------------------|--------------|
| ns-pes | 'height' | ← ?pes | 'to be tall' |
| ny-t ^b ay | 'size, width' | ← t ^b ay | 'to be big' |
- (90) brapa?
- how.much NMZ-be.many
How many (things)?

ii. *Action/state nominalisation -an 'NMZ'* This is basically a default morpheme, used when phonological rules would otherwise exclude nominalisation from taking place (§3.2.3.2). The derived nominal may be a state, e.g. j?ji?-an 'filth' ← j?jil? 'to be dirty', or an action or result, pan-cin-an 'cooking, cookery' ← pancin 'to cook'.

iii. *Nominalisation by coda copy <C>* Disyllabic adjectives derive a deverbal noun meaning either 'the dimension of V' as in a), or 'the one associated with being V' as in b). This form of nominalisation fed by disyllabic roots utilises coda copy alone (§3.2.2.2), without the infixation of a separate nominalising affix.

- | | | | | |
|----|---------|-----------------|--------|----------------|
| a) | j<q>ləŋ | 'length' | ← jləŋ | 'to be long' |
| | s<c>dəc | 'coolness' | ← sdəc | 'to be cool' |
| b) | r<l>mol | 'male, man' | ← rmol | 'to be male' |
| | k<r>dor | 'female, woman' | ← kdor | 'to be female' |

These nominalisations have either a concrete entity or an abstract state as the referent. The deverbal noun is only attested as a constituent in nominal constructions where it functions as the head. In (91) the deverbal noun is functioning as the head of an associative phrase, and in (92) as the complement in a prepositional phrase.

- (91) t<q>rag bulan
be.light<NMZ> moon
moonlight
- (92) boy ma=gɔt ʔen p<t>ret!
NEG:IMP IRR=bring LOC be.hot<NMZ>
Don't take (her) in the heat!

iv. *Object nominalisation +pn+ 'NMZ'* This form of nominalisation names the object of a transitive verb, deriving nouns which have concrete entities as the referent, the exception being the intransitive verb swak 'to go'.

- | | | | |
|--------|---------------------|-------|-------------------|
| pn-reŋ | 'thing sought' | ← reŋ | 'to seek' |
| pn-bər | 'lashing' | ← bər | 'to lash (s.th.)' |
| pn-ȝye | 'thing seen, sight' | ← ȝye | 'to see' |
| pn-pay | 'left-over food' | ← pay | 'to set aside' |
- (93) de=pðrɔr tntay s<pn>wak kəh
3plA=follow distance walk<NMZ> 3
They followed his route with their eyes.

Irregular forms are *pŋ-kɔl* 'head-rest' ← ***kɔl* (rest head) and *p-nam* 'feeling, opinion' ← *nam* 'to sense'.

- (94) *I stumbled, misjudging the drop in the level of the floor, and a bystander made the following observation:*

p-nam ʔma? ʔame, lantay

NMZ-sense mother Amelai floor

Amelia's mother thought (it was) the floor. (I.e., what Amelia's mother sensed was the floor.)

The following form is ambiguous between affixation of *pn-* 'NMZ', or the affixation of *p-* 'NMZ' in conjunction with coda copy: *pn-lən* 'desire' ← *lən* 'to desire'.

v. *Process nominalisation +pC+ 'NMZ'* The *+p+* 'NMZ' derived noun expresses the 'event of V-ing'. It is formed by the affixation of *+p+* in conjunction with coda copy (§3.2.2.2).

<i>p?</i> - <i>gɔ?</i>	'process of felling'	← <i>gɔ?</i>	'to fell (a tree)'
<i>pt-sot</i>	'origins'	← <i>sot</i>	'to metamorphose into'
<i>pl-bul</i>	'intoxicant'	← <i>bul</i>	'to be intoxicating'
<i>ph-jih</i>	'illness'	← <i>jih</i>	'to be ill, in pain'
<i>k^b<ps>bəs</i>	'death'	← <i>k^bəs</i>	'to be dead'
<i>r<p>mɔ?</i>	'harvest, (by reaping)'	← <i>rmo?</i>	'to reap'
<i>j<p>me?</i>	'drinking session'	← <i>jme?</i>	'to drink alcohol'

- (95) *siʔsa? k<p>ru? nintom*
remains fish.poison<NMZ> yesterday
the remains of yesterday's fishing

The derived action noun refers specifically to the process, and cannot be used to name an entity, in which case a non-process noun is used. For example, *ph-jih* 'illness' has the event as its referent, while *pjakit* (< Malay *penyakit*) has the entity as the referent.

vi. *Agent/attributive nominalisation with <m>(<C>) 'NMZ'* This is the least frequently occurring affix; its infrequency suggesting only marginal productivity.¹¹ (An alternative and productive means of deriving this type of nominalisation is with a headless relative clause (§11.1)). The structural process is discussed in §3.2.2.3.

Transitive roots derive an agentive noun with the meaning: 'the N who exhibits the trait of V-ing'; *s<m>rðər* 'a good rememberer' ← *sdər* 'to remember'; *k<m>t-k^bet* 'the members of the taxon of insects which sting' ← *kt-k^bet* 'to be stinging' ← *k^bet* 'to sting, of insects'.

This derivation also results in nouns which name 'the instrument used to perform V'. In the case of the instrumental interpretation the referent is inanimate: *k<m>bəc* 'fishing pole' ← *kbəc* 'to fish with a pole'; *p^b<m>ar-lət* 'apparatus used to extinguish a flame' ← *p^bar-lət* 'to extinguish a flame' ← *lət* 'to go out'.

¹¹ The same contrast is attested in Jah Hut (Diffloth 1976: 99), and in the Nicobar language Nancowry (Radhakrishnan 1970: 55). The morpheme <m>, according to Diffloth (1984: 264-6), is representative of a very early stage in Austroasiatic. In Proto-Monic it functioned to turn transitive verbs into agent nouns and stative verbs into attributive adjectives, e.g. the stative verb **tee?* 'short' → **j-m-tee?* 'short' [Attr.]. The morpheme *+p+* is found sporadically in other Mon-Khmer and Munda languages (Diffloth 1984: 264-6).

In the following the derived form fits semantically with 'the instrument used to perform V' as illustrated by the paraphrase in (96b):

- (96a) *ye=tap s<mai>wel* (96b) *ye=tap rom sawel*
1A=weave left.hand<NMZ> 1A=weave WITH left.hand
I weave (as) a user (of) my left hand. I weave with the left hand.

Locative nominals derive a locative nominal: *k<m>a<m>rom* 'low lying place' ← *karon* 'underneath, below'.

Property adjectives derive a noun denoting something which bears the attributes: 'the N which has V attribute'. These forms, which are infrequent, are used to derive nouns in the avoidance jargon (§1.1): *p<m>das* 'chilli (av.)' ← *pdas* 'to be hot, spicy'; *j<mr>hor* 'sambhur deer (av.)' ← *jhor* 'russet'.

vii. *Excessive par-/pr- 'xs'* This form is used to express the excessive actions of an agent of both transitive and intransitive events. These nominalisations are formed predominantly from volitional verbs and indicate 'one who does V to excess'. The morphological specifics are set out in §3.2.3.1.

<i>par-ca</i>	'glutton'	← <i>ca</i>	'to eat'
<i>pr-pre?</i>	'excessive teaser'	← <i>pre?</i>	'to tease'
<i>par-kən</i>	'one who has time for other's children'	← * <i>*kən</i>	(offspring)

In the following example, the form *par-t^bh* (xs-spit) 'excessive spitter' is formally indistinguishable from the causative form *par-t^bh* (CAUS-spit) 'to make spit':

- (97) *The speaker is chewing betel, and as a result of his salivating is spitting excessively.*
par-t^bh ʔəp!
xs-spit 1f
(What) an excessive spitter (am) I!

7.6.2.2 The syntactic behaviour of nominalised forms

Lexical nominalisations, like other nouns, form the heads of NPs, and occur in environments typical of this word class, e.g. as objects of prepositions (98), subjects of minor clauses (99), constituents in the associative phrase, either as the head (100) or the modifier (101), and in the formation of some interrogative clauses, e.g. 'whose' interrogatives (§10.5.2).

- (98) *dm-dəm kloc bs<t>prt*
IMPERF-lying.down inside be.dark<NMZ>
(She) was lying down in the darkness.

- (99) *par-kən, sma? sukbə? knkən bʔlu?*
XS-offspring person like child friend
(A) parkən is someone (who) likes other people's children. (I.e. not just children who are kin.)

- (100) *soloj ye cəb, sampay gdo, no?, knʔde? ye ha? tasek*
beginning I born until be.old this dwell<NMZ> I AT lake
All my life, this (has been) my place of residence, at the lake. (Lit. (Ever) since I was born, until now (when I) am old ...)

- (101) tlaga? dak j<nh>?oh
well water drink<NMZ>
well (for obtaining) drinking-water

Action nominals are infrequent as the arguments of verbs, although they do occur as the object of verbs of perception (96), or where there is the implication of perception:

- (102) da? ye=k^bm s<pn>wak keh
NEG 1A=succeed walk<NMZ> 3
I didn't succeed (in locating) her path.

There is no such restriction on derived nouns that have a concrete entity as their referent:

- (103) ki=yok k<m>boc
3A=take fish.with.rod<NMZ>
He took the fishing-rod.

7.6.2.3 Clauses containing action nominals

Essentially all the features of the verbal clause remain the same, only the form of the verb and the manner in which the syntactic relations between the verb and its argument are mapped is different. Compare the nominalised form in (104) with its underived form in (105):

- (104) loc pn-lən cu? ʔəŋ [rom hn=k^b]NVC
already NMZ-want grandchild If [WITH ABS=2f]NVC
Already my grandchild's desire (is to be) with you.

- (105) loc cu? ʔəŋ lən [rom hn=k^b]NVC
already grandchild If want [WITH ABS=2f]NVC
Already my grandchild wants (to be) with you.

This single argument of the deverbal noun is expressed as a free pronoun in an associative phrase. In the event that it is a third person pronominal it is in the unmarked form, never the S-form associated with intransitive clauses (§9.3.2). Clauses containing action nominalisations also maintain verbal features like aspect marking (104), and preserving the structure of serial constructions. Compare the two clauses in (106) and (107):

- (106) cəŋ p?st deh ha? ke
immediately stay 3pl AT there
Immediately they stayed there.

- (107) cəŋ p<n>?st deh ha? ke
immediately stay<NMZ> 3pl AT there
Immediately their place of staying (was) there.

The following clause contains two nominalisations, one expressing an action, the other, with a concrete entity as its referent:

- (108) la gi?2 ki=ca, sampay k<nj>?məŋ gərəm mə=c<pn>?ūm
still 3A=eat until finish<NMZ> salt one=wrap<NMZ>
She went on eating until finishing the packet of salt.

8 Prepositions and the prepositional phrase

Chapter 8 deals with prepositions and the prepositional phrase. In §8.1 the prepositions are introduced. §8.2 and §8.3 present an overview of the internal structure of the phrase and its functions. The discussion of the individual prepositions begins with the neutral locative in §8.4. Of particular interest is the set of deictic locative prepositions and the set of associated directionals (§§8.5–6). The remaining sections describe the morphosyntactic characteristics of the prepositions used to express source (§8.7), instrumental/comitative (§8.8) and the possessive (§8.9).

8.1 The prepositions

Semelai prepositions form a small closed class. The prepositions are monosyllabic, with the exception of tə?en 'TO:down' (§8.6.1). They have the phonotactics of a final syllable, but unlike final syllables they are not stressed.

The following prepositions are found in Semelai:

Locatives 'in, on, at'

Neutral locative (§8.4)

ʔən 'LOC' 'in, on, at'

Deictic locatives (§8.5)

ha?	'AT'	immediate location of Speaker
he?	'AT:above'	above Speaker
tɔ?	'AT:across'	same level, lateral to the Speaker
cɔ?	'AT:below'	below Speaker

Directionals TO 'to' (§8.6)

Neutral directionals

tet	'TO:spec'	'to', neutral, specific
te	'TO:unspec'	'to', neutral, non-specific

Deictic directionals

ləŋ ~ ləŋ	'TO:up'	upward, upstream, uphill
tə?en	'TO:down'	downward, downstream, downhill

Source tom 'SRC' (§8.7)

Instrumental/comitative rom 'WITH' (§8.8)

Possessive dɔ 'OF' (§8.9)

8.2 The structure of the prepositional phrase

The preposition typically forms a phrasal constituent with an NP. Some examples are provided in (1)–(8).

- (1) ʔen ʔate
LOC earth
on/in the ground
- (2) rom podɔŋ
with tiger
with a tiger

In (3) the PP hosts the relativiser mə= ‘REL’.

- (3) sampay=hn ki=ca mə=ha? joŋ tipɔ?
so=CONN 3A=eat REL=AT foot betel.service
So he ate the one at the foot (of) the betel service.

The preposition *tom* ‘SRC’ which expresses source (§8.7), takes a locative phrase as its complement when used spatially, as in (4) and (11).

- (4) de=ʔye tom ha? dʔoh masuk bri
3plA=see SRC AT swidden enter forest
They saw (him) enter the forest from the swidden.

The locative preposition *ha?* ‘AT’ (§8.5), *tom* ‘SRC’ (§8.7) and instrumental *rom* (§8.8), can take a verb as the complement:

- (5) ma=jŋok sma? mə=mirah ha? nəkol ʔen mham
IRR=look.at person REL=be.red AT play LOC blood
(She) observed (to her horror) people who were red at play in the blood.
- (6) məhn br-kis rom br-ʔɔh, da? mnjadi? knt̚ɔŋ
if=CONN MID-scrape WITH mid-blow NEG happen fall.off
If (it) is scraped and blown on, (it) won't come to fall off.

Personal pronouns occurring as prepositional complements host the proclitic hn= ‘ABS’:

- (7) da? dos ha? hn=hc
NEG come AT ABS=1&2
(The floodwater) didn't reach us.
- (8) ʔəŋ təʔen hn=ko
If TO:down ABS=2f
I (came) down to you.

With the neutral locative preposition ʔen ‘LOC’, only third person pronominal complements take the proclitic (§8.4.2).

Prepositions can occur in a ‘sandwich’ construction (P_i NP P_j) simultaneously preceding and following the NP.

- (9) ht? ʔmaʔbapa? he?
AT:above parents AT:above
up at (his) parents' (place) up (there)

This construction is also found with the deictic locatives (§8.5), the directional *tet* ‘TO:spec’ (§8.6), the comitative *rom* ‘WITH’ (§8.8) and the possessive *do* ‘OF’ (§8.9).

8.3 Functions of the prepositional phrase

The PP relates the noun to a larger syntactic unit, either as a peripheral expression, a subcategorised non-core argument, a non-subcategorised adjunct (§8.3.1) or as a subcategorised core argument of the verb, namely the indirect object (§8.3.4). Prepositions also function as predicatives (§8.3.2) and attributives (§8.3.3), with a purely locative interpretation.

8.3.1 Peripheral expressions

Peripheral noun phrases express adjuncts to the event: spatial location (10), source (11), goal (12), instrumental (13), comitative (14) and possessive (15).

- (10) ki=ʔye=cə? syok ha? mankok, ha? kwali?
3A=see=EM trace AT bowl AT wok
She saw traces of food in the bowl (and) in the wok.
- (11) cu?=hn grɔk tom kloc rɔ?
grandchild=3POSS fall SRC inside basket
Her grandchild fell (out) from inside the basket.
- (12) “lah tet nɔ?!” kʰləŋ, “he b-reŋ-reŋ ci” kʰləŋ
HORT, TO:spec here QUOTE 1&2 MID-INTNS-seek lice QUOTE
“(Come) here!” (she) said, “we'll look for each others' lice,” (she) said.
- (13) pasuh ki=bək rom pos=hn
jar 3A=tie WITH tail=3POSS
He tied the jar (on) with his tail.
- (14) ʔyot rom knon
return WITH offspring
(15) ʔma? do lagkaj
mother OF [name]
the mother of Langkang

8.3.2 Predicative function

Locative prepositions (16) (§§8.4–5), directionals (17) (§8.6) and comitatives (18) (§8.8), may function as the main predicator in non-verbal clauses (§10.1).

- (16) “co? ke paya?,” kʰləŋ=hn,
AT:down there stream QUOTE=3POSS
“Down there (is) a stream,” he said.
- (17) da? ki=ɛn wɔ?, wi leŋ nɔ?
NEG 3A=desire more [name] TO:up here
wi doesn't want to come up here any more.

The comitative is infrequent in non-verbal clauses:

- (18) “loc pn-lən cu? ʔəŋ” kʰləŋ, “rom hn=kb”
already NMZ-desire grandchild If QUOTE WITH ABS=2f
“Already grandchild's desire,” (he) said, “(is to be) with you.”

Directionals used as the main predicator can co-occur with the imminent aspect marker ga= ‘IMM’ as in (19), a feature typical of non-verbal clauses (§10.1).

- (19) ?yot dɔs ha? kke, k^məŋ bapa? "ga=tet kmpən kb?"
 return reach AT there QUOTE father IMM=TO:spec wife 2f
 He returned, (and on) arriving there the father enquired, "Are you going to your wife?"

The locative-based predicative is used as a lexicalisation strategy for avoidance lexemes (§1.1):

- (20) *The avoidance word for 'dog' is based on their location of residence.*
 mə=2en karom dɔl
 REL=LOC under house
 (the one) who (is) under the house

8.3.3 Attributive function

Locative prepositions (21) and prepositional phrases (22)–(23) (§§8.4–5) may function as attributives to noun phrases (§7.5.1). Some examples are given below, where it can be seen that the presence of the neutral locative results in non-referential or generic statements:

- (21) ki=ləwah, malaj tɔ? ki=jŋök b̥rəŋwɔŋ!
 3A=look.around side AT:across 3A=look.at tiger.stripes!
 She looked around (it), (and) across (on) the (other) side she saw the stripes (of a tiger's coat).
- (22) yok=cə wŋy tɔ? ke!
 fetch=EM knife AT:across there
 Fetch the knife over there!
- (23) sma? ha? tasik
 people AT lake
 the people at the lake, i.e. the Semelai
- (24) gpal, ple 2en dloŋ
 Canarium.littorale fruit LOC tree
 The canarium (is) a fruit (that grows) on trees.

Locatives and deictic prepositions derive demonstrative pronominals (§6.3.1.1).

8.3.4 Core function

Locative prepositions also signal the subcategorised core recipient argument 'indirect object'. The status of the PP as a core constituent is based on its syntactic behaviour (see §9.1.3). The subcategorised core argument encoded in this manner is limited to the recipient of verbs of oral communication (25) and verbs of giving, as in (26) and (27).

Although the preposition is marking the grammatical relation 'indirect object', the semantics of the preposition are retained, and it continues to make an independent contribution to the meaning. This is especially clear in the case of the deictic locatives (27).

Further examples of core recipients are given in the relevant sections below (§8.4.2 and §8.5.4).

- (25) de=kabar 2en gdo
 3plA=tell.news LOC be.old
 They told the news to the old (people).
- (26) h?gi? 2en hn=kəh!
 give LOC ABS=3
 Give (it) to him!
- (27) *The giver is on the verandah and the recipient is standing down on the ground:*
 rəkək ki=jon cɔ? bapa?
 cigarette 3A=give AT:below father
 She gave the cigarettes down to her father.

8.4 The neutral locative

2en 'LOC' is a deictically neutral locative preposition which functions to express the topological relation of 'coincidence', 'the near or total overlap of the located object and the reference object' (Frawley 1992: 255). It codes the general location of an event or participant, but does not provide specific information on deictic orientation. This can only be achieved with the deictic locatives (§8.5).

2en 'LOC' does not allow further spatial differentiation, hence it does not occur with the demonstratives (?nɔ? ~ nɔ?nɔ? 'this' or ke ~ kke 'that'). The lack of specificity means that 2en can function in the expression 2en boh (LOC elsewhere) 'elsewhere'; the deictic locatives cannot replace it in this function. It can however co-occur with a locative nominal (§7.5.3) as in the following example:

- (28) krwl tɔm 2en kloc lpeč
 come.out SRC LOC inside stomach
 came out from inside (her) stomach

The neutral nature of 2en 'LOC' is further reinforced by the fact that it cannot be used temporally, unlike the deictic locative ha? 'AT' (§8.5).

2en 'LOC' is used to code peripheral expressions (§8.4.1), predicates and attributives. These three uses are only available with a locative interpretation. It also signals the core recipient of verbs of giving and verbs of oral communication (§8.4.2).

8.4.1 Peripheral functions

A large number of examples are provided in this section to illustrate the broad range of specific sub-senses expressed by the locative preposition. 2en 'LOC' expresses simple spatial coincidence with the exact relationship determined by the pragmatic/semantic context of the situation. It is translated as 'at', 'on' or 'in'.

The reference object may either directly support the located entity (29), or it may be conceptualised as sharing the same location (33). As the examples show, the reference object may be inanimate (29)–(33), animate (34)–(35), or a body part (36)–(38).

- (29) kəhn pon b̥sanar 2en tɔm.dloŋ
 3S THEN lean LOC base.tree
 She thereupon leant against the bole (of) a tree.
- (30) km-k^bom 2en k^boy taja?
 IMPERF-sit LOC head steps
 (He) was sitting down at the top (of) the steps.

- (31) *ki=slet w^{ay}=hn ?en s^{ap} bnul*
3A=insert knife=3POSS LOC end floor.beam
He inserted his knife into the floor beam.
- (32) *kl=c^{um} ?en dawon kdah.*
3A=wrap LOC leaf palm
She wrapped (it) in a palm leaf.
- (33) *k^{ehn} plec ?en kampoj raja?*
3S emerge LOC village prince
He emerged at a prince's village.
- (34) *2le?le? masuk bri, dos ?en k<r>dor*
so enter forest arrive LOC bc.woman<NMZ>
So he entered the forest (and) came to a woman.
- (35) *pakay ?en knkn*
use LOC child
use (it) for children

?en 'LOC' can be used to indicate the body part serving as the locus of effect:

- (36) *luka? ?en tⁱ /blu /jŋj*
wound LOC hand /thigh /foot
wounded on the hand/thigh/foot
- (37) *k^{eh} ga=dm-d^{em} ?en blu knlek*
3 IMM=lie.down LOC thigh husband
She was going to lie down in (her) husband's lap.
- (38) *ki=tloj ?en suk*
3A=grab LOC hair
She grabbed (me) by the hair.

If the patient is expressed, it is coded as the core relation O as in (39), and not as a peripheral expression, as in (38).

- (39) *ki=tloj ?ma?*
3A=grab mother
(He) grabbed the mother.

?en 'LOC' is always used to signal the affected body part of verbs used to express the maleficent feeding on or taking of souls by supernatural beings, e.g. evil shamans feeding on the souls of people (40), or spectre huntsmen taking people's souls (41).

- (40) *kl=ca ?en bayan* (41) *ma=yok ?en bayan*
3A=eat LOC soul IRR=take LOC soul
They (evil shamans) feed on the souls (of people). would take (her) soul

8.4.2 Core function

The case-marking function of the neutral locative was introduced in §8.4.2 above. Further examples of the subcategorised core recipient argument of verbs of giving (42)–(43) and verbs of oral communication (44)–(46) are provided below. The status of these as core arguments is argued for in §9.1.3. With deictically neutral ?en 'LOC'

the location of the recipient in relation to that of the Speaker is not encoded. ?en 'LOC' is usually selected when the two parties are understood to be in the same location, or when the speaker simply chooses not to impart specific location. Compare this with the deictic locatives in §8.5.4.

- (42) *ga=y^e=jon ?en ?amay*
IMM=1A=give LOC [name]
I'll give (it) to ?amay.
- (43) *de=tanre? ?en hn=k^{eh}*
3plA=show LOC ABS=3
They showed (it) to her.
- (44) *ki=?ur ?antat ?en knon*
3A=instruct bring LOC offspring
She instructed (her) child (to) bring (it).
- (45) *ki=p-c^{ek}k^{ep} ?en kmpan*
3A=CAUS-converse LOC wife
He told his wife.
- (46) *h^{ir}-i? ?en dri?*
growl-APPL LOC 1self
(It) growled at us.

8.5 Deictic locatives

There are four deictic locative prepositions in Semelai. These are arranged in a 'person' or 'speaker' oriented system. The location of a referent is encoded deictically with respect to the location of the Speaker, or from the Speaker's adopted point of view. The system consists of one proximate and three distal locative prepositions:

ha?	'AT'	immediate location of Speaker
he?	'AT:above'	space above Speaker
to?	'AT:across'	space on the same level, lateral to the Speaker
co?	'AT:below'	space below Speaker

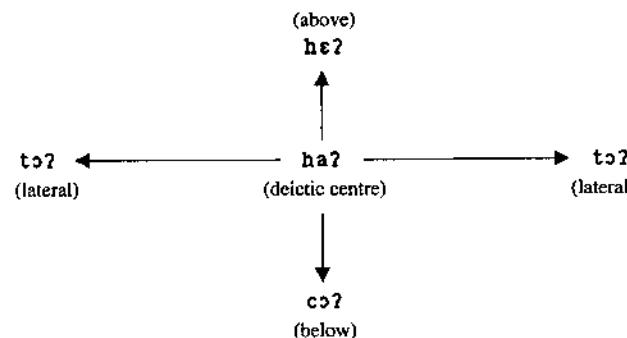
The proximate ha? 'AT' refers to the immediate location of the Speaker and constitutes the 'zero-point'.

The distal system encodes further specification with respect to topological features. It distinguishes three 'levels' of height in relation to the Speaker's actual or adopted location. The three distal terms are: a) he? 'AT:above', above the location of the Speaker; b) to? 'AT:across', lateral to, or across from the Speaker's location, and c) co? 'AT:below', below the Speaker's location.

The preposition to? 'SRC' (§8.7) which signals source is not distinguished for 'person' orientation. This information is encoded in the accompanying deictic locative expression:

- (47) *tom co? dol*
SRC AT:below house
from down at the house

FIGURE 8.1 DEICTIC LOCATIVES



The deictic locatives cover the same range of functions as the deictically neutral *zen* 'LOC' (§8.4). They are used with a locative interpretation as peripheral expressions (§8.3.1), predicatives (§8.3.2) and attributives (§8.3.3). The non-locative recipient reading is available to the core indirect object argument (§8.3.4).

The preposition can be repeated after the NP, as in example (48), giving rise to the 'sandwich construction' (P₁ NP P₂) (§8.2).

- (48) ki=go? mə=d<n>loj ha? chŋ̩ nɔ?; he? paya? he?, tuy
3A=fell one=tree<UNIT> AT hill this AT:above stream AT:above one
He felled one tree on this hill; up at the stream up there, (another) one.

Determined location is equated with referentiality of the NP, consequently the deictic locatives co-occur with both the proximate and distal demonstratives, and the zero-form preposition *ha?* 'AT' can also occur alone post-nominally as a deictic modifier (§6.3.1).

- (49) bnda? ha?, da? lag?i
thing AT EXIST still
That thing hereabouts, (it's) still (here).

The following sections describe the functions of the deictic locatives as peripheral expressions and to express the indirect object.

8.5.1 Location and orientation

The deictic system is not only relevant to one's immediate personal space, but is translated across to encompass all aspects of the spatial environment. In this section I briefly describe some extensions and elaborations of the deictic system.

The location of places and direction of motion (§8.5) is mapped out in accordance with the direction of the flow of water in the lake and river systems. Places and the people who inhabit them are located *he?* 'AT:upstream' (50), or *co?* 'AT:downstream' (51) from the speaker.

- (50) he? puyŋ̩ kāk
AT:above shaman [name]
up at shaman kāk's

- (51) ga=kī=tar-dos cɔ? paʔitam jahudi?
IMM=3A=CAUS-reach AT:below male.4th.born [name]
He will take (it) down(stream) to fourth-born uncle jahudi.

Traditional locations outside the immediate area of the lake are located as if the mode of travel was still by river, rather than road. Locations situated away from the lake and river system, or outside of the area traditionally traversed by the Semelai, are located with *to?* 'AT:across':

- (52) to? ʔayer.hitam
AT:across [top]
across at Ayer Hitam

Within the settlement, the lake and the wells which are situated on its banks are referred to as 'below', downhill:

- (53) co? dak /tlaŋa?
AT:below water /well
down at the lake/the well

The inland forest and the swiddens located within it are always 'above', uphill from the settlement:

- (54) he? bri /dʒoh
AT:above forest /swidden
up in the forest/the swidden

Homes in the village are likewise located with respect to the lie of the land: *he?* 'AT:above', *co?* 'AT:below' or *to?* 'AT:across'.

- (55) pen's shop is located across the field from where the conversation is taking place.
to? kday pen
AT:across shop [name]
across at pen's shop

In addition to topography, the system is also used to encode architectural features. The house, a raised structure, is always referred to as *he?* 'AT:above' if the speaker is on the ground and the ground is *co?* 'AT:below' if the speaker is in the house.

- (56) ʔendol he? dɔl, tar-koh baju?
enter.house AT:above house MCAUS-undo clothes
He went up into the house (and) removed his clothes.

Within the house, the kitchen and verandah are always on a lower level, *co?* 'AT:below', whilst the sleeping quarters are at a higher level, *he?* 'AT:above'. The differences in height within the house may only be the depth of a beam, or one step. These areas are often fixed in this manner regardless of the actual location or adopted viewpoint of the speaker:

- (57) ki=tabor he? dɔl, he? j<n>tck,
3A=sprinkle AT house AT:up sleep<NM2>

- co? habu?, co? ?ate?
AT:down kitchen AT:down ground
He sprinkled (it) in the house, up in the sleeping area, down in the kitchen (and) down on the ground.

he? 'AT:up' may be used for an adjacent, bounded non-visible area, e.g. in a house or outside:

- (58) ki=kunci? pintu? he? lwar
3A=lock door AT:above outside
She locked the door from outside.

Topographic features dictate the choice of preposition regardless of whether it is marking a peripheral locative or a subcategorised core argument. Few verbs or complements require a particular preposition, the choice being determined by the level of specification that the speaker wishes to impart.

Further specification of location is achieved by the use of a locative nominal in an associative phrase (§7.5.3):

- (59) iji-ziŋ ha? karom ponton
imperf-hide AT underneath firewood
(They) were hiding underneath the firewood.

The locative nominal can limit the choice of deictic locative, or vice versa. Those nominals which are distinguished for the dimension of height, *dyal* 'above', *kʰoy* 'upper end', *karom* 'below' and *jon* 'lower end', can only co-occur with the neutral *ha?* 'AT', the corresponding deictic locative *he?* 'AT:up' (60), or *co?* 'AT:down', respectively:

- (60) laiu?, ki=capay he? kʰoy tipo?, ki=ca
then 3A=reach AT:above head betel.service 3A=eat
Then, he took (one) (from) up at the head (of) the betel service (and) he ate (it).

The combination of the lateral preposition and the 'head' of the entity is unacceptable:

- (61) *lalu?, ki=capay to? kʰoy tipo?, ki=ca
then 3A=reach AT:lateral head betel.service 3A=eat
*Then, he took (one) (from) across at the head (of) the tray (and) he ate (it).

8.5.2 Distributed reading

The head may occur with a demonstrative complement indicating non-specific reference. This pattern is particularly common with the combination of prepositions *to?* 'AT:across' and *he?* 'AT:above', and *ha?* 'AT' and *he?* 'AT:above'. These combinations can be glossed as 'here and there'. This reading is available both to peripheral expressions (62)–(64) and indirect objects (65).

- (62) tugal ma=b<nh>nih ha? no?, ma=b<nh>nih he? ke
plant.rice one=seed<UNIT> AT here one=seed<UNIT> AT:above there
(She) planted, a seed here (and) a seed there.

With this particular construction the nominal may be omitted where it is understood from context. In (63) the two locative prepositions are in apposition.

- (63) knon ke b-kawan rom to?, he?
offspring that HAVE-friend WITH AT:across AT:above
The child made friends with (people) here (and) (people) there.

In (64) there is no overt reference of the located object 'tree'.

- (64) solog de=tbok to?, tbok he?
first 3plA=tap.resin AT:across tap.resin AT:above
At first they tapped (trees) here for resin, tapped (trees) there for resin.

This reading is also available to indirect objects. In (65) recipients are implied but not identified.

- (65) knon ki=kʰe? swak, kʰe? jon to?, jon he?
child 3A=know walk know give AT:across give AT:above
The child knew (how) to walk, (how) to give (to people) here (and) give (to people) there.

8.5.3 Peripheral functions

Here the spatial relationship is one of coincidence between the located object and the reference object, as defined in §8.4 for the neutral locative *ʔen* 'LOC'. The location can be inanimate (66)–(69) or animate (70)–(72):

- (66) ki=masuk dak ha? mpak.hitam
3A=enter water AT oil.be.black
He put water into the oil.

- (67) de=jŋɔk he? para? kbe? sima?
3plA=observe AT:above rack body person
They observed up on the rack a human body.

- (68) de=yɔk rmpʰəl co? ?ate
3plA=take throw.down AT:below ground
They took (it and) threw (it) down on the ground.

- (69) ?ie? dos to? ke
be.long reach AT:across there
(In) a while, he arrived over there.

- (70) ki=jŋɔk tabon ?ahom ha? hn=deh
3A=look.at vessel breath AT ABS=3pl
He looked at (his) vessel (of) breath in their (possession).

- (71) ha? kbe? sndir?
AT body self
to oneself

In (72) the back is expressed as an 'upper' part of the body, by selecting the preposition *he?* 'AT:above'.

- (72) ki=tampɔŋ he? elɔn
3A=hit AT:above back
He hit (him) up on the back.

8.5.4 Core function

As noted in §8.3.4, the interpretation of the locative in a core role function signals the indirect object of a verb of giving or oral communication. A deictic locative, used instead of neutral *?en* 'LOC', allows the speaker to encode deictic information about the location of the recipient in relation to the speaker.

The following examples with the verb *jon* 'to give' illustrate the contrasting usage. In (73)–(74) the recipient is in the immediate location of the donor; in (75) the donor is upstream.

- (73) ki=jon dom, ha? cu? kke
3A=give AFF AT grandchild that
He gave (it), yes, to the grandchild there.

- (74) cəl ha? hn=?əŋ!
say AT ABS=1
Tell (it) to me!

- (75) *The shaman in question, puyŋ kâk, lives upstream from where the speaker is located. The statement is understood as a reference to that particular shaman as opposed to any other of the shamans who reside in the area.*
ga=yε=jon he? puyŋ
IMM=1A=give AT:above shaman
I'll give (it) to the upstream shaman.

The same pattern is attested for verbs of communication:

- (76) *The speaker plans to ask someone who lives downstream:*
ye=tapa? cɔ? busu? wər
I=ask AT:below youngest.born instead
I'll ask youngest born instead.

8.6 Directionals

There are four terms which express the direction toward a location, two neutral terms distinguished for specificity of location and two deictically specified terms. The term used here to describe these is 'directional'.

tet	'TO:spec'	'to', neutral, specific
te	'TO:unspec'	'to', neutral, non-specific
leŋ-leŋ	'TO:up'	'upward, upstream, uphill'
taʔen	'TO:down'	'downward, downstream, downhill'

In contrast to the locative prepositions, the directionals only express orientation in relation to a goal and not actual 'coincidence' between the theme and the endpoint. Typically, the directionals co-occur with verbs of locomotion, either with verbs like *b-kayuh* 'to paddle' and *swak* 'to go', which do not encode the path of motion (77), or with verbs which do express the path, e.g. *yər* 'to ascend' (78), in which case path is expressed twice.

- (77) b-kayuh taʔen joŋ dak
USE-paddle TO:down foot water
(He) paddled down to the lower reaches of the water.

- (78) yər leŋ no?!
- ascend TO:up here
Come up here!

Directionals are also used to express the orientation of an action where the motion is projected from the actant in a static location. In the following example tet 'TO:spec' co-occurs with the utterance verb *cəl* 'to say' indicating manner of speech:

- (79) cəl tet cən, məcom sma? ga=k̥baš
say TO:spec back like person IMM=die
Talk behind (their) backs, as if someone were going to die.

In (80) the speaker was describing the orientation of one's face when swimming backstroke.

- (80) muka? leŋ maŋ
face TO:up sky
face toward the sky

The directional may occur as the main predicate of a clause where the verb is not expressed, as in (17) and (80) above and (81) below (§8.3.2). The directional tet 'TO:spec' can occur in the 'sandwich construction':

- (81) plto? tet tgah tet
explode TO:spec middle TO:spec
(It) exploded toward the middle.

Each of the directionals is now described in turn: tet 'TO:spec', te 'TO:unspec', leŋ ~ leŋ 'TO:up' and taʔen 'TO:below'.

i. tet 'TO:spec', te 'TO:unspec' The directionals tet 'TO:spec' and te 'TO:unspec' are the default directionals, used where the direction is neither up nor down. The two morphemes tet and te are not in free variation as indicated by the glosses. tet 'TO:spec' is used for a specific destination – it indicates the direction of motion where the goal is located on the same level as the Speaker, or it can function to encode direction where the topography of the trajectory or destination is unknown. te 'TO:unspec' is used where the destination is non-specific, either the speaker doesn't have a destination in mind, or doesn't want to disclose it.

- (82) goŋ tet no?!
- carry TO:spec here
Bring (it) here!

- (83) "dawn, ɿəŋ ɿyot tet dol," k̥leŋ
never.mind If return TO:spec house QUOTE
"Never mind, I am going back home," (he) said.

- (84) mə=hiay da? b-tipal, ki=goŋ tet k̥r̥dor ke
one=sheet NEG MID-remain 3A=take TO:spec be.female<NMZ> that
(There) wasn't a single sheet left, he took (them all) to the woman.

In order to disambiguate the lateral and default senses speakers often select the verb *dloh* ~ *lsh* 'to go across' for sites located laterally in the immediate vicinity:

- (85) ga=lsh tet k̥day
IMM=go.across TO:spec shop
(I) am going across to the shop.

te 'TO:unspec' is used for an unknown or a non-specific direction as in the interrogative in (86). In response, the speaker may use tet 'TO:spec' to encode a specific location outside of traditional Semelai territory.¹

- (86) *This is the customary way of greeting someone encountered on a path, or showing signs of going somewhere.*

ga=te h̥n?
IMM =TO:unspec where
Where are (you) going?

- (87) *The speaker indicates to the hearer to move out of the way:*

te ke, te ke, kp!
TO:unspec there TO:unspec there 2f
(Move) over there, over there, you!

The unspecified form is also used in the expression te keh, te no? (TO:unspec that, TO:unspec this) 'here and there', representing a distributed reading like that discussed for the deictic locatives in §8.5.2.

ii. *leg ~ ləŋ 'TO:up'* *leg ~ ləŋ 'TO:up'* signals the goal of upward motion. The two forms are in free variation. *leg ~ ləŋ 'TO:up'* is possibly linked etymologically to *ləŋ* 'the space above the horizon between the earth and the sky', or *malen* 'sky'.

- (88) ki=cək hn=hayam tom cō? ʔate leg dol
3A=throw O=chicken SRC AT:below ground TO:up house
He threw the chicken from down on the ground up to the house.

- (89) paloh leg knon dak tʰmajər
flee TO:up offspring water Tembangau
(They) fled up to the Tembangau tributary.

The directional is also used to indicate orientation where there is no change in location:

- (90) ki=hadap! leg ponoy ke
3A=face TO:up hut there
He faced up toward the hut.

iii. *təʔen 'TO:down'* *təʔen 'TO:down'* codes the goal of downward motion, see also (77).

- (91) "de? troj təʔen bri.kmuc," kʰləŋ
indeed path TO:down forest.ghost QUOTE
"Indeed, the path down to the underworld." (he) teased.

- (92) ?ie?le?, loh ?antat mə=rəmol təʔen jŋj dak
then go:across bring REL=be.male TO:down foot water
Then, he went and brought the male down to the lower reaches.

8.6.1 Directionals and the deictic locatives

The directionals are distinct from, yet related to the set of deictic locatives. Direction is calculated with respect to the location of the endpoint of the trajectory: *leg ~ ləŋ*

¹ The negative response to this question is "da? da? h̥n" (NEG EXIST where) "Nowhere."

'TO:up' is used for a goal which would be located using *he?* 'AT:above'; *teʔen* 'TO:down' corresponds with *cō?* 'TO:below'; *tet* 'TO:spec' and *te* 'TO:unspec' encode both the lateral plane *to?* 'AT:across', as well as being the default term. The following passage exemplifies the interaction of the two systems.

- (93) *A man's son has been down to the bri kmuc (forest ghost) 'underworld' visiting his dead sister. The bri kmuc is located beneath the world of the living.*

sampay ki=moh ʔyot leŋ bapa?
then 3A=want return TO:up father

dəs he? bapa?, "te h̥n kp swak?"
arrive AT:above father TO:unspec where 2f move

"ʔəŋ təʔen bri.kmuc," kʰləŋ
If TO:down forest.ghost QUOTE

So then he decided to return to his father.
(He) arrived at his father's ... "Where have you been?"
"I went down to the underworld," (he) replied.

The choice of directional is dependent on the verb type. A verb like *swak* 'to go', or *gɔŋ* 'to carry' is not specified for direction, so any of the directionals can co-occur with it. Verbs with inherent direction can also co-occur with the corresponding deictic locative. In this case, the choice between directional and deictic locative reflects a variation in the perspective on the event. In (94a) the use of *teʔen* 'TO:down' puts the focus on the motion of falling in terms of the trajectory of the event; in contrast in (94b) the focus is on the location of landing, or the endpoint.

- (94) (a) kəhn grək təʔen ʔate (94b) kəhn grək cō? ʔate
3S fall TO:down ground 3S fall AT:below ground
He fell down to the ground. He fell down onto the ground.

8.7 Source

The primary function of *tom* 'SRC' is to denote the source or origin of an entity. It is also used to mark the source of an opinion and the source of comparison. In temporal clauses *tom* 'SRC' is a subordinator where the complement is a verb or verb phrase.

Etymologically *tom* 'SRC' is derived from a noun. It occurs in compound nouns with the meaning of 'source, base': *tom dloj* (base tree) 'bole of a tree', *tom ʔare?* (source rain) 'rain-bearing clouds'.²

i. *Source* The source can be a location, animate (95)–(96) or inanimate (97), or a time (98). As the following examples illustrate, source is not restricted to motion

² There are also the words *pəm* 'night' and *aŋtəm* 'yesterday' which seem to be etymologically linked to *tom*, the former having the locative nominaliser *pə-*, the latter the nominaliser *a-*. The Semelai perceive night as preceding day, or the source from which day originates. Against this etymology, however, are the apparently unrelated forms found in Nyah Kur *tom* 'tree-base and trunk', and *pətam* 'night-time' (Diffloth 1984).

predicates, but also occurs with stative predicates (97) and in temporal expressions (98).

- | | |
|---|-----------------------------------|
| (95) timol tom ha? dak
emerge SRC AT water | (96) tom ha? kadeh?
SRC AT who |
| (He) emerged from the water.
From whom? | |
| (97) da? ʔlep tom ha? dol
NEG be.far SRC AT house | |
| not far from the house | |
| (98) knlek=hn tom som bsŋst mudik ɬej k ^b oy dak
husband=3POSS SRC morning be.dark go.upstream TO:up head water | |
| From early morning, her husband went upstream to the head waters. | |

When expressing origin from a location, *tom* 'SRC' has a locative prepositional phrase as the complement as in (95), (96) and (97). In (99) a locative preposition is not required because the nominal *hej* 'yonder:up'³ is inherently locative.

- (99) tom hej ji?
SRC yonder:above 2
You (come) from over there.

Note that temporal complements, e.g. (98) above, do not require a locative preposition either.

ii. Point of view *tom* 'SRC' is used to indicate speaker judgment based on sensory perception. Note that the locative preposition is not required in this construction and that the pronominal complement does not require the proclitic *hn*= 'ABS'.

- | | |
|---|-------------------------------------|
| (100) ʔiem tom ji?
be.pleasant SRC 2 | (101) kjeh tom ye
be.heavy SRC 1 |
| Is (it) tasty to you?
(It) is heavy for me. | |
| (102) solɔŋ tom ye=?en brapa? ʔlep pɔ?
beginning SRC 1=AUG how.much be.far very | |
| At first for us, (it was) so far. (Lit. At first for us, how much very far (it was).) | |

iii. Comparison *tom* 'SRC' also encodes the entity which is the basis of comparison:

- (103) iben tom pe? ʔare? pe? ptom, sma? pok rbana?,
more SRC three day three night person beat drum
(For) more than three days (and) three nights people beat the drums.

iv. Sequential clauses *tom* 'SRC' functions in a special construction type where the complement is a verb, rather than an NP. In these cases *tom* 'SRC' is preceded by the verb ?luc 'to pass' (104); ?yot 'to return' (105), or the aspect marker ?areh 'just, recently' (106). This construction is discussed in (§11.5.3.3).

³ *hej* 'yonder:above' and *ben* 'yonder:further:above' are possibly related to the locative preposition *ne?* 'AT:above'. There is also a form *cɔj* 'yonder:below'. This is the only instance of a suggested morphological process located in the final consonant.

- | | |
|---|--|
| (104) ?luc tom j<h>ʔɔh, ?yot ?mot cɔn
pass SRC drink<IMPERF> return get.on back | |
| After drinking, (she) got back up on (his) back. | |
| (105) ?yot tom b-layar sirah ha? dol ?ma?bapa?
return SRC USE-sail stop.off AT house parents | |
| Returning from sailing, (he) stopped off at (his) parents. | |
| (106) ?areh tom hɔm
be.new SRC bathe | |
| (I) just bathed. | |

8.8 The instrumental/comitative

rom 'WITH' expresses the instrumental 'with, by', the comitative 'with' and path 'via'.

rom 'WITH' can occur twice in a prepositional phrase, in the 'sandwich construction' as in (107) (§8.2).

- (107) Q: rom kadeh rəy te?en nɔ?
WITH who companion TO:below here
Q: "With whom did (you) come down here?"
- A: rom kaka? rom
WITH EZ WITH
A: "With elder sister."

i. Instrumental The instrumental codes the instrument (108)–(109), or body part used as an instrument in performing an action (110)–(111).

- | | |
|--|--|
| (108) ki=tikam pɔdɔŋ rom tomok
3A=stab tiger WITH dagger | (109) phəl rom kapl.trbəŋ
come.down WITH aeroplane
(They) came down in aeroplanes. |
| He stabbed the tiger with the dagger. | |
| (110) pasuh ki=bək rom pos=hn
jar 3A=tie WITH tail=3POSS | |
| He tied the jar with his tail. | |
| (111) ki=cəc rom wəy, tida? rom nɔ?nɔ?, ki=huris rom cros
3A=gut WITH knife NEG WITH this 3A=slit WITH nail | |
| She gutted (it) with the knife. Not with this, she didn't slit (it) with (her) nails. | |

rom 'WITH' is also used to express the material with which a process is performed (112), or an object is manufactured (113).

- | | |
|---|--|
| (112) de=pəl rom ?ate mur gadət.t'i
3plA=plug WITH earth circumference parent.hand | |
| They plugged (it) with (a piece of) earth the circumference (of) a thumb. | |
| (113) br-tajn rom sake?
MID-weave WITH pandanus | |
| (It is) woven from pandanus. | |

ii. Comitative The complement of the comitative phrase may be animate (114), or inanimate (115)–(116).

- (114) c<nn>man sma? ?adi?bradi? rom romgasi
reteli.myth<NMZ> person sibling WITH [name]
The retelling of the myth of the siblings and Romgasi (the ogre).
- (115) ki=cəj jon bəs rom wby
3A=immediately give sugarcane WITH knife
He immediately gave (him) the sugar-cane and the knife.
- (116) krwai rom luluk, rom rapah, rom pate
emerge WITH mud WITH leaf.litter WITH earth
(He) emerged, (covered) with mud, forest litter and soil.

Although the comitative primarily conjoins NPs, it may occasionally conjoin verbs, see (6) above.

Certain verbs subcategorise for an oblique comitative argument, e.g. nikah 'to marry', tukor 'to exchange', cray 'to separate from' (§9.1.2):

- (117) ki=b-kawan rom pəpəŋ
3A=HAVE-friend WITH tiger
He befriended the tigers.

Pronominal complements subcategorised for by the verb can be procliticised with hn= 'ABS' as in (118).

- (118) *This is the standard formula when meeting up with someone the speaker hasn't seen for a long time.*
?le? da? c<h>roh rom hn=jí
belong NEG meet.up<IMPERF> WITH ABS=2
I haven't met up with you (for) a long time.

In contrast, adjuncts do not host the proclitic (119):

- (119) *The speaker, who is dead and dwelling in the underworld, promises her mourning brother who is still living, that if he returns to the real world he will be like her in seven days time.*
dɔs kɔ hə? ke, tmphn zare? mɔdm̩ rom zap
reach 2f AT:above there seven day like WITH If
You go up there, (and in) seven days (time) (you will be) like with me.

iii. Path of motion With verbs of motion rom 'via' encodes the actual path of the event (120), or locations encountered on the path (121)–(122).

- (120) rs-rus rom pate (121) rps tupay rom khoy
IMPERF-drag WITH ground jump squirrel WITH head
dragging along the ground the squirrel jumped, via her head
- (122) sinah rom ladaj gades
stop.off WITH Ladang Geddes
(We) stopped off at Ladang Geddes.

8.9 The possessive

Possession can be expressed by do 'OF' in a construction: NP[possessee] do NP[possessor]. The construction is less frequent than the equivalent associative construction (§7.5).

- (123) dwet do knlek
money OF husband
the money of the husband

The construction is used when the possessor is a noun or demonstrative, with the alternant də= ~ dde= proclitic to pronouns (§6.1.4).

- (124) ?ma? do laŋkang, de=korog kloc rban kamej
mother OF Langkang 3plA=shut.up inside pen goat
The mother of Langkang, they shut up inside the goat pen.

The head can be a derived nominal expressing the manner of performing an activity:

- (125) ns-gos do gop nɔ?=pa
NMZ=scrape OF Malay this=FACT
This, you know, (is) the manner of (fruit) peeling of the Malays.
This, you know, (is) the Malay way of peeling fruit.

The possessed head is frequently elided:

- (126) *A boy and his friends have been collecting fruit. On arriving home, the boy discovers that the ripe fruit he has picked is now unripe, while the unripe fruit of his friends has ripened.*
ki=curah. dom, pie sot muda?. do b?lu?, ?nom
3A=tip.out AFF fruit metamorphose.into be.unripe OF friend be.ripe
She tipped (it) out. Indeed, (his) fruit had turned unripe. (That) of the friends (was) ripe.

When the complement of do 'OF' is the ignorative kadeh 'who', it forms the possessor interrogative 'whose' (§6.2.4.1). do 'OF' is often repeated following the question word (127) in a 'sandwich construction' (§8.2).

- (127) do kadeh do?
OF who OF
Whose (is it)?

The response to this question is to use do 'OF' and the possessor noun, or the contracted form prefixed to the pronoun.

The distal demonstrative ke ~ kke 'that' (128), and the deictic locative ha? 'AT' (129), may occur as complements of do 'OF'. The phrase functions as a discourse deictic indicating something previously mentioned (§6.3.3). The proximate demonstrative nɔ? 'this' is never used in this way.

- (128) yok sarek, dom, sbəŋ, dom d?oh t̩ay do kke
take tomorrow AFF be.full AFF swidden be.big OF that
The next day, yes, (it was) full, indeed the big swidden from before.

- (129) me=tom malen, huc kunerj dɔ ha?
 REL=SRC sky rice be.yellow OF AT
 'The one from the sky', that there (is the avoidance term) for yellow rice.

Example (130) is a standard way of ending a traditional narrative.

- (130) kmoŋ dɔ ke. da? ye=sdər wɔ?
 finish OF that NEG 1A=remember more
 (It) finishes there. I don't remember any more.

9 Grammatical relations, constituent order and coding strategies

In Chapter 9 the discussion centres around grammatical relations, constituent order and coding strategies. In addition to the three grammatical relations A, S, O, a fourth member, indirect object (IO), is identified. The IO exhibits dual membership – it is coded as an oblique in a prepositional phrase, but exhibits the syntactic behaviour associated with a core role (§9.1.4).

The most frequently employed constituent order is verb initial, with either A and O, or S placed after the verb. Various pragmatically determined permutations are possible, and these are outlined in §9.2.

The coding of grammatical relations is dependent on the transitivity of the clause, and it is in accordance with this distinction that they are discussed in §9.3.

The minimal clause in Semelai consists of just the predicate. The means of representation of NPs in the clause and zero anaphora are discussed in §9.4.

9.1 Grammatical relations

Grammatical relations can be divided into core and oblique (Andrews 1985). In Semelai, the distinction is validated by their differing morphological and syntactic behaviour. Core grammatical relations are central to syntactic processes such as causativisation and relativisation, and have distinctive patterns of NP-marking and verbal cross-referencing. A clear distinction between the core and non-core constituents can be established by examining relativisation, interrogative formation, fronting and the distribution of possessor =hn '3POSS' (§9.1.3).

9.1.1 Core relations

Core grammatical relations are necessarily subcategorised by the verb. These are characterised in terms of grammatical relations as A, O and S (Dixon 1994), as well as IO. A is the argument of a transitive clause which is prototypically an agent; O, the argument of a transitive clause which is prototypically a patient; and S, the single argument of an intransitive clause and the 'subject' of a non-verbal clause. An S takes A-like marking under conditions of external causation, to be discussed in §9.3.3. Semantically core relations tend to be unrestricted in that they may reflect a range of semantic roles. Thus, the grammatical relation A may be realised as an agent (1), or an experiencer of perceptual/aural stimuli, as in (2), according to the verb lexeme involved. A is the most restricted of the core relations in terms of permitted reference: the A must always be animate, the only exceptions being certain potent natural phenomena (§5.5.1).

- (1) m=bunuh hn=kb
IfA=kill O=2f
I'll kill you.
- (2) ki=jam k'lem deh
3A=sense odour 3pl
He sensed their odour.

Some verbs which have an agent as the A are:

bɔy	'to dig'	sunɔt	'to circumcise'	jɔy	'to make'
pajəl	'to summon'	cjew	'to look down'		

Some verbs which have an experiencer as the A are:

?ye	'to see'	jam	'to sense'	trsɔk	'to overhear'
?yan	'to hear'	lən	'to desire'		

O may be a patient as in (1) above, a location (3), a theme (4), a goal of motion or communication as in (5), or an instrument in (6).

- (3) ki=kde? dol
3A=live.in house
She lived in a house.
- (4) ki=cək dlon
3A=chuck stick
She chucked the stick.
- (5) kl=2ajak knlek
3A=invite husband
She invited (her) husband.
- (6) de=pakay jarla
3plA=use thorn
They use thorns.

The grammatical relation S is highly unrestricted, although not totally as the examples below illustrate. Usually, if the verb is active, the S is an agent (7)–(8), or an experiencer (9) or patient if the verb is stative.

- (7) kehn pk-pok
3S IMPERF-wash.clothes
She is clothes-washing.
- (8) dehn paloh
3plS flee
They fled.
- (9) bapa?=hn k'bəs
father=3POSS be.dead
His father is dead.

Some verbs which have an agent as the S are:

pləc	'to come out'	?w-?aw	'to stand up'	həm	'to bathe'
dos	'to arrive'	blajar	'to study'		

Some verbs which have an experiencer or patient as the S are:

jtek	'to sleep'	jde?	'to be beautiful'
trɔ?	'to be feverish'	ʃəŋ	'to be long'

The fourth grammatical relation, indirect object, is always a semantic recipient. It is coded as an oblique locative and is introduced in §9.1.2.1 below.

9.1.2 Obliques

Oblique relations are of two types, the complement and the adjunct. Andrews (1985) characterises complements as participatory, as opposed to adjuncts which are circumstantial. Participatory can be defined as 'actual participants in the situation implied by the verb', whilst circumstantial roles are those that provide the setting for the event (Andrews 1985: 69). Hence, a participatory locative provides information about the location of a participant, whilst the circumstantial locative provides information about the location of the event.

9.1.2.1 Complements

Complements are semantically restricted in that they are constrained to a specific semantic role determined by the predicate, e.g. motion verbs require a direction or goal complement (DIR) (10a), and cannot co-occur with a locative complement (LOC) (10b):

- (10a) kehn ?yot [ley] ke_{DIR}
3S go.back [TO:up there]
She went back up to there.
- (10b) kehn ?yot *[he?] ke_{LOC}
3S go.back *[AT:up there]
*She went back up at there.

Verbs of encounter, union or separation require an oblique comitative complement with rom 'WITH' (§8.8):

- (11) t-haron rom dol sma? gdo
HAPP-come.across WITH house person be.old
(They) happened to come across an old person's house.

While all complements, core and oblique, share the feature of being subcategorised by the verb, one type of complement, the recipient, shares features of the syntactic behaviour typical of core relations such as relativisation, fronting and optionality of the case marker in certain contexts. The recipient is coded by a locative preposition (§8.3.3), and is considered a fourth core grammatical relation indirect object. Verbs which subcategorise for a recipient are verbs of transfer, e.g. jon 'to give' and h?gi? 'to give'; and of verbs of oral communication, ?ur 2en 'to instruct (s.o.)', paryəŋ ha? 'to inform (s.o.)'.

Recipients are similar to complements, in being semantically restricted, but also share features typical of an oblique, most importantly the ability to make an independent semantic contribution by encoding the spatial distinction contained in the deictic locative series (§8.5):

- (12) ki=jon ha? /he? /lo? bapa?
 3A=give AT /AT:above /AT:below father
 She gave (it) to/up(wards) to/down(wards) to/her father.

Secondly, they share with complements a lack of obligatoriness: the verbs of giving do not require the expression of the core recipient (13a,b), just as verbs of transfer of location do not require the expression of the non-core locative complement (14a,b).

- (13a) de=jon kweh ?en sma? dr? (13b) de=jon kweh
 3A=give biscuits LOC people self 3A=give biscuits
 They gave biscuits to my people. They gave biscuits.

In contrast to English, the locative complement may be left unexpressed (14b):

- (14a) ki=pte? ?ate dyal no?no? (14b) ki=pte? ?ate
 3A=place earth above this 3A=place earth
 He placed the earth above this. He placed the earth.

It is possible to elide the O in these clauses, expressing only the IO, or locative complement as in (15).

- (15) ki=pte? dyal no?no?
 3A=place above this
 He placed (it) above this.

9.1.2.2 Adjuncts

The non-subcategorised oblique function or adjunct, is not determined by the predicate, but may occur freely provided it is semantically compatible, as the locative adjunct in (16) and the reason adjunct in (17), see §11.5.3.1. Unlike complements, adjuncts do not share any syntactic features with core roles.

- (16) de=jgok [kloc takar]Loc t'm sma?
 3A=observe [inside earthenware.vessel] hand person
 They observed inside the earthenware vessel, a human hand.
 (17) da? kl=jon la=?pat
 NEG 3A=give BCS=be.ugly
 He didn't give (it) because (it) was ugly.

Adjuncts which express circumstantial information in a PP are discussed in (§8.3.1). Temporal and manner adjuncts (§10.3.2) and the benefactive are expressed as bare NP adjuncts in (18) and (19). The associative phrase is also used for the benefactive (§7.5.2). The benefactor (BEN) precedes the O.

- (18) ye=bey [1D]BEN kmpn
 1A=check.on [2f]BEN wife
 I checked on (your) wife (for) you.
 (19) yok [knkn]BEN ji]BEN kweh!
 take [child 2]BEN biscuit
 Take a biscuit (for) your child!

9.1.3 Syntactic behaviour of core relations

Syntactically, the core relations S, A, O and IO behave in an identical manner with respect to the processes of fronting (§9.2.3), relative clause formation (§11.1) and interrogative formation (§10.5) in the transitive clause. Further, only core roles may host possessor marking (§7.5.1).

i. *Fronting* Core NPs, which are potentially case-marked in post-verbal position, do not retain this marking when fronted to pre-verbal position (20)–(21). Non-core obliques retain the preposition when fronted (22). (The process is the same for relative clause and interrogative formation; see below.) Fronting is discussed in detail in §9.2.3.

- (20) ki=goy hn=tilam he? rabo?
 3A=bring O=mattress AT:above roof.ridge
 He brought the mattress up onto the ridge of the roof.
 (21) tilam ki=goy he? rabo?
 mattress 3A=bring AT:above roof.ridge
 The mattress, he brought up onto the ridge of the roof.
 (22) he? rabo?, ki=goy hn=tilam
 AT:above roof.ridge 3A=bring O=mattress
 Up onto the ridge of the roof, he brought the mattress.

ii. *Relative clause formation* When the NP to be relativised is a subcategorised core argument S, A, O or IO, it is fronted, and its role in the lower clause is not coded (23). The restrictive clause is formally marked by the proclitic *me* = 'REL'.

- (23) jks [me=ki=jel] la=eo] paloh
 porcupine REL=3A=bark.at A=dog flee
 The porcupine that the dog barked at, fled.

In contrast, non-core constituents are relativised by placing the restrictive clause adjacent to the head, which is not fronted, and remains in situ. There is no relative marker indicating the function of the clause, as in (24) (see §11.1).

- (24) knkn pon ?yam he? dol [de=goy raja?]
 children THEN cry AT:above house [3p]A=bring prince]
 The children were crying up at the house where the princes had brought (them).

iii. *Interrogative formation* The formation of the interrogative requires the fronting of the ignorantive and the relativisation of the clause. Typically, the fronted constituent does not retain the NP-marker. In the pair of examples below, the indirect object of the declarative clause (25a) is not retained in the interrogative clause in (25b).

- (25a) ga=yey=jon [?en ?amay]io (25b) [kadeh]io me=ga=yey=jon?
 IMM=1A=give [LOC [name]]io [who]io REL=IMM=1A=give
 I will give (it) to ?amay. (To) whom will I give (it)?

By contrast, comparing the adjunct in the interrogative clause in (26) with (25b) above, the adjunct (ADJ) retains the comitative preposition *rɔm* 'WITH' when pre-verbal:

- (26) [rɔm kadeh ray]ADI dos ha? no??
 [WITH who friend]ADI come AT here
 With what friends did you come here?

iv. Possessor marking Possessor marking on NPs codes the third person possessive with the enclitic =hn '3POSS' (§7.5.1). Morphologically, only core roles or NPs possessed by a core NP host the enclitic. The enclitic functions as an alternative to the use of a free pronominal, and in this context occurs in free variation with it. Possessor marking of core roles S, A, O and IO is shown in examples (27)–(30).

- (27) [2ma?=hn]S k^bbas
 [mother=3POSS]_S die
 His mother died.
- (28) xi=yəŋ [la=2ma?=hn]A
 3A=hear [A=mother=3POSS]_A
 Her mother heard (her).
- (29) ki=slet [wɔy=hñ]O 2en sɔŋ bñul
 3A=insert [knife=3POSS]_O LOC end beam
 He inserted his knife in the end (of) the beam.
- (30) ki=7ur [2en knon=hñ]O
 3A=instruct [LOC child=3POSS]_O
 She instructed her child.

A complement or adjunct containing an NP possessed by a core argument may host enclitic =hn '3POSS'. In (31) the agent uses part of his body as an instrument. The object of the instrumental PP hosts the enclitic.

- (31) pasuh ki=bək rɔm pɔs=hñ
 jar 3A=tie WITH tail=3POSS
 The jar he tied with his tail.

In (32) the NP in the comitative expression is in a kin relationship with S.

- (32) mulɔ? kde? rɔm knon=hñ, soloŋ lƿəc
 begin dwell WITH child=3POSS first stomach
 (In the) beginning, (she) lived with her child, the first-born.

In the following example, the adjunct NP is co-referential with the elided O:

- (33) *The tiger is carrying the afterbirth by its tali? 'cord'.*
 ki=tanggoŋ 2en tali?=hn
 3A=carry.in.mouth LOC cord=3POSS
 He carried (it) in (his) mouth by its cord.

If the object in the complement (34), or adjunct (35), is not possessed by a core argument of the clause, the possessor is expressed by the free pronominal kəh '3'.

- (34) de=pɔt [ha? tɔmpɔt kəh]
 3plA=stay [AT place 3]
 They stayed at his place.

- (35) ki=jipɔk mandeh [ha? tɔm kponj kəh]
 3A=observe what [AT bole Kepong 3]
 She observed something at the bottom of his Kepong (tree).

Possessor marking, in other words, is clause bounded. In (36) the possessor is represented by a free pronominal and not the enclitic, despite the fact that the possessor is co-referential with the A in the preceding clause.

- (36) da? de=ca=hñ, [la=sac bapa? deh]
 NEG 3plA=eat=O [BCS=flesh father 3pl]
 They didn't eat it, because (it was) the flesh of their father.

- (37) *da? de=ca=hñ, [la=sac bapa?=hn]
 NEG 3plA=eat=O [BCS=flesh father=3POSS]
 *They didn't eat it, because (it was) the flesh of their father.

9.2 Constituent order

Constituent order in Semelai is neither fixed nor free, but fluid, allowing a range of variations which are outlined below. Hence, constituent order alone does not encode grammatical relations. In fact, a degree of ambiguity is frequently encountered and semantic/pragmatic plausibility must be relied upon in order to determine grammatical relations.

There is, then, no single basic order in Semelai. Constituent order is determined by the transitivity of the clause and the internal temporal structure of the event, in conjunction with factors motivated by discourse organisation, e.g. the tendency for new information to precede old. In the case of intransitive clauses obtained through direct elicitation, S precedes the verb, while in transitive clauses there is a tendency for verb initial structures. This is not always reflected in spontaneous discourse.

The function of the variations in constituent order is to focus or foreground certain constituents. An example is the role-marked pronominal S used post-verbally to mark a change in direction in the discourse, or the end of a scene in a narrative. This is illustrated in the following example where the woman has just given birth and instructs the tiger to take the placenta and bury it. In the second line where the woman has finished her statement, and in the final clause in the scene, the S follows the predicate, rather than preceding it.

- (38) "tantat kke 7ɔɔk" k^bəŋ,
 take that placenta QUOTE
 "yɔ?, boy ki=ca hñ," k^bəŋ, "k^bbas 2əŋ" k^bəŋ.
 but NEG:IMP 2FA=eat THEN QUOTE die If QUOTE
 ki=7antat, ki=tanggoŋ 2en tali?=hn,
 3A=take 3A=carry.in.mouth LOC cord=3POSS
 tanggoŋ 2en tali?, dos he? ke, ki=bɔy
 carry.in.mouth LOC cord arrive AT:above there 3A=bury

- 2en ʔate. ki=kəm ʔen ʔatt, ki=kəm. ʔyot kəhn.
 LOC earth 3A=bury LOC ground 3A=bury return 3S
 'Take this placental!' (she) instructed. 'But don't you eat (it) now (or) I will die!'
 (she) warned. He took (it) (and) he carried (it) in (his) mouth by its cord. Carrying
 (it) by the cord, he arrived up there, (and) he dug (a hole) in the ground. He buried
 (it) in the ground, he buried (it). He went back home.

The clausal constituents, the categories NP, PP and VP, have fixed internal order, and any change in that order reflects a change in the meaning. In (39) the PP must remain adjacent to the NP, otherwise it loses its modifying function, and is interpreted instead as a circumstantial locative (40).

- (39) dom, [ʔmaʔitam [he? ke]PP]NP ki=bt^by
 AFF [fourth.born [AT:above there]PP]NP 3A=fear
 Indeed, (it was) (his) fourth-born female cousin from up there, he feared.
 (40) dom, [ʔmaʔitam]NP ki=bt^by [he? ke]PP
 AFF [fourth.born]NP 3A=fear [AT:above there]PP
 Indeed, (it was) (his) fourth-born female cousin, he feared up there.

The frequent partial representation and zero-representation of arguments further complicates a characterisation of a basic constituent order. Many clauses consist only of a verb, or verb and pronominal proclitic.

In the following section the main generalisations about constituent order are addressed, beginning with core arguments. Constraints on the ordering of case-marked NPs are dealt with in §§9.2.1-3. The pre-verbal 'fronted' position is discussed in §9.2.3.

9.2.1 Core arguments

The position of the arguments depends on the clause type. There are three basic structures: (i) the universal clause, (ii) the perfective clause and (iii) intransitive and non-verbal clauses.

i. The Universal Clause If the transitive clause is a universal or generic statement, constituent order is A V O, neither A nor O is case-marked and the verb does not host a cross-referencing proclitic (§9.3.1.1).

- (41) pədɔŋ ca sma?
 tiger eat people
 Tigers eat people.
 (42) par-kən suka? knən
 XS-child like child.
 A 'parkən' (is someone who) likes (other people's) children.

In (43) and (44), the NPs do not have specific referents; the clauses describe the setting or background against which events unfold.

- (43) sma? pok rbana?
 person beat drum
 People were beating drums.

In (44) the speaker is not talking about a specific instantiation of an event, but rather a habitual activity.

- (44) sma? putih wen resen təm ha? kapl.trban
 person be.white throw ration SRC AT aeroplane
 The Europeans threw rations from (their) aeroplanes.

A minor function of this clause type is to provide clarification. In the following example, the narrator clarifies who is speaking to whom:

- (45) cəŋ cəl 2en ko?
 mousedeer say LOC pig-tailed.macaque
 It is Mousedeer saying (it) to Pig-tailed Macaque.

ii. The perfective clause The second type of clause is the perfective transitive clause expressing a specific instance of an event. There do not appear to be any constraints on the ordering of post-verbal A and O, either one may precede the other. A and O are marked by la= and hn= respectively (§9.3.1).

- (46) ki=bukə? la=knək hn=pintu?
 3A=open A=husband O=door
 The husband opened the door.

- (47) ki=bukə? hn=pintu? la=knək
 3A=open O=door A=husband
 The husband opened the door.

In the case of trivalent predicates containing an IO, the IO tends to precede the O, perhaps motivated by the fact that the recipient is a person and the O is usually inanimate, although see (47) above, where the inanimate O precedes the A:

- (48) de=jon 2en kəh mə=k<n>uta? mancis
 3A=give LOC 3 one=box<NMZ> match
 They gave him a box of matches.

Trivalent predicates containing all three arguments A, O and IO, in post-verbal position, are not attested in spontaneous discourse, but are obtained under elicitation.

iii. Intransitive and non-verbal clauses In these clauses the NP may either precede or follow the predicate. In (49) the clause is intransitive.

- (49) kəhn [dʌk'əs dəl]Pred
 3S [be.close house]Pred
 He was close (to) the house.

In (50), the S of each verbal clause, the former active, the latter stative, follows the predicate.

- (50) [ga=dmdəm]Pred {ʔəp}S, [bhɪh jɪ?̩]Pred [lɒəc]S
 [IMM=lie.down]Pred {lɪl}S [be.full very]Pred [stomach]S
 I'm going to lie down, (my) stomach is really full.

The constituents are ordered so that old information, represented by the NP, precedes new information which is contained in the predicate. The foregrounding of

new information places the predicate before the subject as in (50) above. For non-verbal clauses the situation is analogous to the intransitive verbal clause:

- (51) kəh ha? neg [kmon] ye]Pred
3 AT before [sibling's.offspring 1]Pred
He, here before (is) my nephew.

- (52) In the discourse leading up to this example the topic of conversation is the various kinds of lice. The speaker describes a particular type which she has just previously named:

- [p?la? r?̄-k̄t]Pred deh
[animal COMP-be.small!]Pred 3pl
Small animals, (are) they.

9.2.2 Obliques

Obliques tend to follow core arguments, both post-verbally (53)–(54) and pre-verbally.

- (53) ki=c̄ek hn=hayam tom co? 7at̄ leŋ doi
3A=throw O=chicken SRC AT:below ground TO:up house
He threw the chicken from down on the ground up to the house.

- (54) ?mot kəh 7en cl̄n prah?̄
mount 3 LOC back canoe
She climbed onto the back of the dug-out canoe.

The oblique precedes the S and the O in (55) and (56) respectively. There is a slight intonation break before the post-posed S in (55).

- (55) dom, klk̄l ha? bantal, kəhn.
AFF lie.head AT pillow 3S
Indeed, she lay (her) head on the pillow.

- (56) ki=pte? he? d?oh knl̄k=hn
3A=place AT swidden husband=3POSS
She positioned (her) husband up at the swidden.

Temporal and manner adjuncts tend to occur at the periphery of the clause, see (§10.3.2).

9.2.3 Fronting: the clause-initial position

The pre-verbal position in the transitive perfective clause functions as the focus position. The function of fronting is to establish a known entity as topical in the text. In subsequent clauses the NP has zero representation, other than A which is cross-referenced on the verb. The fronted NP may be any core or non-core NP; most frequently the position is occupied by the O. Note that the process of fronting is not accompanied by a change in verbal morphology, and as stated above, the A is still cross-referenced on the verb. Hence, fronting does not represent a process of syntactic promotion, but rather is a pragmatic strategy.

Core NPs are constrained by the following factors:

- a) The pre-verbal position can contain at most one core NP, (a constraint which does not hold for the post-verbal position). In (57) the A is fronted, see (21) for an example of O.

- (57) sma?, ki=tampɔŋ co
person 3A=hit dog
Person, he hit the dog.

Simultaneous fronting of both the A and the O results in an unacceptable clause:

- (58) *sma?, co ki=tampɔŋ (59) *co, sma? ki=tampɔŋ
person dog 3A=hit dog person 3A=hit
*Person, dog he hit *Dog, person he hit

A core NP can share the fronted position with an oblique:

- (60) tilam, ha? raboŋ ki=gɔŋ
mattress AT:above roof.ridge 3A=bring
The mattress up onto the ridge of the roof, he took (it).

Coordinated NPs sharing the same core grammatical relation may be fronted:

- (61) t̄b̄, jɔŋ, ki=bak
hand foot 3A=bind
(Her) hands (and) feet he bound.

- (62) snt̄p sereh, gubet̄, ki=gɔŋ yor he? raboŋ
pouch sirih betel.pounder 3A=bring ascend AT:above roof.ridge
The sirih pouch (and) the betel nut pounder, he carried up onto the roof.

The fronted core NP may be complex, e.g. modified by a PP, see (39) above.

- b) Fronted core NPs are not role-marked, as in (63a) and (64a). Clauses with a fronted A require stress on the NP and a slight pause following it (63b).

- (63a) *la=kəh, ki=7ye creh (63b) kəh, ki=7ye creh
A=3 3A=see fish 3 3A=see fish
*Him, he saw the fish. Him, he saw the fish.

Compare the pre-verbal recipient in (64b) with the pre-verbal locative in (65). In (64b) the IO is fronted and the NP-marker, locative 7en 'LOC', is not retained.

- (64a) ki=ur 7en km̄pan (64b) km̄pan, ki=ur
3A=instruct LOC wife wife 3A=instruct
He instructed (his) wife. He instructed (his) wife.

The absence of NP-marking on fronted core arguments contrasts with obliques which retain the preposition (65). See also (60) above.

- (65) he? jnt̄k, ki=blah hn=pinan
AT:top sleeping.area 3A=split.in.two O=betel.nut
Up in the sleeping area, she split the betel nuts in two.

Fronted NPs may be encliticised by the third person possessor clitic =hn '3POSS' as shown in the following example:

- (66) wəc=hn ki=wən
 gizzard=3POSS 3A=discard
 Its gizzards she discarded.

9.3 Coding strategies

In this section two basic strategies for coding syntactic function will be examined in relation to the core roles S, A, O and IO: a) cross-referencing, and b) NP-marking. Table 9.1 summarises the marking patterns discussed in the following sections.

A, the subject of the transitive clause, is coded by the proclitic *la*= 'A' on all pronouns and nouns, with the exception of inanimate nouns which are not eligible to be an A. (Note that A is simultaneously cross-referenced on the verb; neither S nor O are cross-referenced.) S, the subject of the intransitive clause, is unmarked except for the third person pronominal forms, which have distinctive absolute forms. O, the object of a transitive clause, optionally hosts the proclitic *hn*= 'O' across all NPs including third person pronominals.

TABLE 9.1 ROLE MARKING

	PRONOUNS		NOUNS	
	3	1, 2	+animate	-animate
A	<i>la</i> =			
S	<i>kəhn</i> '3S' <i>dehn</i> '3plS'			
O	<i>hn</i> = / \emptyset			

Coding strategies are examined in the context of transitive (§9.3.1) and intransitive (§9.3.2) clause types.

9.3.1 The transitive clause

The marking of A, the agent of the transitive clause, and O, the patient, is dependent on the clause type. As seen above in §9.2.1, grammatical relations are not marked in universal type clauses. The marking discussed here holds for all other transitive clauses, including all derived transitive clause types.

9.3.1.1 A-marking

A is the only consistently coded syntactic relation, coded by both cross-referencing on the verb and proclitic marking on the external NP. This pattern is attested in the most frequently occurring transitive clause type [PRO_i=V *la*=A_j], where the cross-reference marker PRO_i= is an unstressed pronominal proclitic on the verb, and *la*= 'A' is a proclitic on the external NP.

i. *Cross-referencing* The proclitic encodes the grammatical properties of an A in terms of person and deference in the first and second person, and number and identity in the third person. The first and second person proclitics are formally indistinct from the free-form pronominals, except for the lack of stress and position within the clause, i.e. proclitic to the left edge of the verbal word. The third person singular and plural forms are distinct from the free forms, and from the third person

S-function pronouns. The pronominal proclitics are set out in §6.1.1, and discussed in relation to the verbal word in §5.5.1.

The proclitics do not function just as agreement markers, but also have an anaphoric function given the high frequency of zero representation of NPs in the clause.

The cross-referencing proclitics occur in the transitive/perfective declarative clause in both main and dependent clauses, proclitic to the main verb, or the initial verb of a serialised verb construction (69). Pronominal proclitics do not occur in the transitive imperative (§10.6.1) or irrealis constructions (§10.2). They are used in the external cause construction where the S of the intransitive verb is coded as an A (§5.5.1.2). Features of the proclitic are examined below.

The agreement marker cross-references the post-verbal A:

- (67) ki=gəŋ *la*=lslɒs
 3A=bite A=flying.ant
 The flying ant bit (him).

- (68) dom, de=ca *la*=dəh
 APP 3plA=eat A=3pl
 Indeed, they ate (it).

The third person unidentified agent proclitic *ko*= '3UA' (69), never cross-references an external NP_A (70).

- (69) *ko*=sec tarek cɔ? ?atɛ?
 3UA=surreptitiously pull AT:below ground
 Someone surreptitiously pulled (it) (from) down on the ground.

- (70) **ko*=sec tarek *la*=sma? cɔ? ?atɛ?
 3UA=surreptitiously pull A=person AT:below ground
 *Someone surreptitiously pulled (it) (from) down on the ground.

The agreement marker cliticises to the initial verb in serialised constructions (§11.3.1). In (71), the main verb *halaw* 'to chase' is the second in the series. See also (69) above.

- (71) de=cəŋ *halaw* *la*=cɔ
 3plA=immediately chase A=dog
 The dogs immediately chased (him).

ii. *la*= 'A' In the transitive clause the A normally occurs post-verbally and is coded by the proclitic *la*= 'A'. Not all post-verbal NPs marked with *la*= are coding the A; *la*= also expresses causation or reason, see (§11.5.3.1).

- (72) ki=gəŋ *la*=knɪək=hn bantəl
 3A=bring A=husband=3POSS pillow
 Her husband brought the pillow.

- (73) ki=ca *la*=kəh
 3A=eat A=3
 He ate (it).

There are no constraints on the ordering of post-verbal core NPs; see (46) and (47) in §9.2.1.

If the post-verbal NP_A is elided, the proclitic la= 'A' is likewise elided, as in (75). la= can never stand alone, unlike hn= 'O', (§9.3.1.3).

- (74) de=tampəŋ̩ co la=rāja? 3plA=beat dog A=prince
The princes beat the dogs.

- (75) de=tampəŋ̩ co 3plA=beat dog
They beat the dogs.

The NP_A does not host the proclitic in pre-verbal position:

- (76) rāja?, rāja?.muda?, rāja?.naga?, de=goy ʔyot
prince crown.prince prince.underling 3plA=bring return
The ruling prince, the crown prince and the deputy crown prince, they took (the mother) back.

Where NP_O has zero representation, it is only through context that the initial NP is interpreted as the A. If taken out of context, the fronted NPs in example (76) could also be interpreted as O, in which case the clause would have the following reading: 'The ruling prince, the crown prince and the deputy crown prince, they (=the captors) took (them (=princes)) back.'

In clauses obtained through direct elicitation the A-marker is always used; in spontaneous text it is frequently absent in the following environments.

a) Where it is contextually recoverable:

- (77) *It is general world knowledge that spectre huntsmen have the ability to abduct people, and not vice versa.*
sa?lɒn, ki=goy sk?z?
[name] 3A=take spectre.huntsman
(As for) sa?lɒn, a spectre huntsman took (him).

Alternatively, it is recoverable because it represents old information within the discourse. In (78) the husband's fear is mentioned for the first time and the NP_A is coded.

- (78) kbran de=?ye la=r<1>mɔl kmpən=hn
be.fearful.lest 3plA=see A=be.man<NMZ> woman=3POSS
(He) was fearful lest the men see his wife.

In a subsequent recapitulation the NP is not marked:

- (79) "dawen de=?ye r<1>mɔl" kʰləŋ=hn
never.mind 3plA=see be.man<NMZ> QUOTE=3POSS
"Never mind (if) men see (you)," he said.

b) If there is no overt post-verbal O, the post-verbal A can occur without la= 'A', (80) and (81). (80) is ambiguous if taken out of context; the NP could be construed as O.

- (80) *In this story there are three princes, see (76) above.*
de=goy rāja?
3plA=bring prince
The princes brought (them).

- (81) ki=tapa? lsips 3A=ask flying.ant
The flying ants asked (him).
[In the context of the story, there is only one flying ant present.]

In (82) the NP_A could not be confused as the recipient 'REC' because the preposition used for this argument is not ommissible post-verbally:

- (82) ki=tapa? [ha? lsips]_{REC} /lsips
3A=ask [AT flying.ant]_{REC} /flying.ant
He asked the flying ants.

c) If the number or person of the NP_A and NP_O are understood to be different the post-verbal A may be a bare NP. The O is pre-verbal in (83) and post-verbal in (84).

- (83) ʔəŋ̩ neg ki=jala? sma? rabon kke
1f before 3A=cast.net person be.blind that
Me before, a blind man there caught (me) (in his) casting net.

- (84) de=ca kntε?
3plA=eat flea
The fleas fed on (me).

d) Marking is also optional when the NP is non-individuated, as in (84) above, and in the following example:

- (85) "ha? nɔ?", kʰləŋ "de=?ye-l2 sma?"
AT this QUOTE 3plA=see-ITER person
"Here," he said, "people keep on seeing you."

9.3.1.2 O-marking

The O of the transitive clause may be marked by hn= 'O', proclitic to the NP:

- (86) ki=tikam hn=kɔ? ke
3A=stab O=pig.tailed macaque that
He stabbed that pig-tailed macaque.

O-marking is also used in two 'agent suppressed' single-argument constructions: middle voice (§5.3.3) and the unrealis (§10.2). The change in the number of syntactic arguments in the construction is not accompanied by a change in NP-marking, and consequently the same grammatical treatment is maintained for the O. The following examples are of the middle voice in (87) and the unrealis in (88).

- (87) b-ptε? hn=dloŋ
MID-place O=wood
Wood (was) placed (on it).

- (88) ma=kanɔŋ hn=pinaj dwa? bje?
IRR=conceal O=betel.nut two CLF
(She) would conceal the two betel nuts.

In the positive imperative construction the A is not coded, but the O is coded as in a prototypical transitive construction (§10.6):

- (89) pdər hn=?əŋ!
follow O=if
Follow me!

O is not systematically marked. Typical of core arguments, the fronted pre-verbal O is never marked; post-verbal O may host proclitic hn= 'O' on the NP. In the event where an O has zero representation in the clause, the proclitic may appear as an enclitic on the verb. Again as with A-marking, it is not possible to predict the presence or absence of marking on the O in post-verbal NPs. Direct elicitation never yielded examples with the proclitic, contrasting with direct elicitation of other clitics, however isolated clauses containing the marker, when rechecked, were always deemed acceptable.

The O-marking proclitic is used with a range of NPs and does not reflect properties of the NP such as animacy or affectedness of the O. In (86) above the O is animate, in (90) inanimate, and in (91) and (92) it is unaffected by the event.

- (90) ki=tarek hn=jala?
3A=pull O=casting.net
He pulled the casting-net (in).

- (91) kɔŋ ttawa?, papel hn=yε=?en!
beat gong call O=I=AUG
Beat the gong (to) call us!

- (92) ki=tr-ye hn=b2iŋ?
3A=HAPP-see O=friend
He happened to catch sight of (his) friends.

In pre-verbal position O is never case-marked. Compare examples (93a) and (93b).

- (93a) ki=bntaq hn=tłam
3A=unroll O=mattress
He unrolled the mattress.

- (93b) tłam ki=bntaq
mattress 3A=unroll
He unrolled the mattress.

A post-verbal NP_O cannot simultaneously host object-marking hn= 'O' and third person possessor =hn '3POSS':

- (94) ki=slet *[hn=wŋy=hn] ʔen sŋŋ bṇul
3A=insert [O=knife=3POSS] LOC end beam
*He inserted his knife in the end of the beam.

It is only possible to have O-marking as in (95) below, or possessor marking as in (29) above.

- (95) ki=slet hn=wŋy ʔen sŋŋ bṇul
3A=insert O=knight LOC end beam
He inserted his knife in the end (of) the beam.

Elided O may be represented by =hn 'O' enclitic to the verb. This is applicable to both animate and inanimate NPs. In (96)–(98), the O represents human participants, although not necessarily animate like the corpse in (98).

- (96) ki=prat(i?hn) ła=bapa?
3A=watch=O A=father
The father watched her.
- (97) da? tntu?=ca?, da? ki=bunuŋ=hn
NEG certain=EM NEG 3A=kill=O
(He) wasn't certain, (that) he hadn't killed her.
- (98) ki=bantay=hn
3A=butcher=O
He butchered her.

In the following examples the inanimate O – fruit in (99), and a grave in (100) – is represented by enclitic =hn. This is highly unusual. There is no pronominal anaphora available for inanimate NPs which must have either a full NP or demonstrative mention, or receive zero representation.

- (99) "ʔes beh," ki=təŋ pədəŋ, "da? n=jon=hn"
oh NO QUOTE tiger NEG IfA=give=O
"Oh no," said the tiger "I'm not giving (them)."

- (100) ki=bɔŋ=hn
3A=dig=O
He dug it.

In the case of serial verbs the enclitic representing O cliticises to the final verb. In example (101) =hn 'O' is enclitic to the final verb even though it is actually the object of the first verb.

- (101) ki=goŋ paloh=hn
3A=bring flee=O
He took (him) fleeing.

9.3.1.3 IO-marking

The NP-marker for the IO is a locative preposition, either the neutral locative ʔen 'LOC' (§8.4), or one of the set of deictic locatives (§8.5). Verbs of giving tend to select from either, whilst verbs of communication or address tend to use the neutral locative ʔen 'LOC', as in (48) and (103). In their function as markers of the IO, their semantic range is limited to that of recipient or addressee. This is discussed in (§8.3.4).

- (102) ye=tapa? cɔ? busu? wər
1A=ask AT:below youngest instead
I'll ask down at the youngest instead.

These locatives are also used to code the causee, the usurped agent, in causative constructions derived from transitive roots (see Figure 5.5, §5.3.5.1).

- (103) ʔare?·b-ribut p-bt'bŋ-i? ʔen sma?
rain.HAVE-wind CAUS-be.afraid-APPL LOC person
Storms frighten people.

The NP-marker of the IO is distinct from the other core role markers in that it is never omitted in post-verbal position. In this respect, it patterns like an oblique.

- (104) *de=jon kəh mə=k<N>utb? mancis
3A=give 3 one=box<NMZ> match
*They gave him a box of matches.

In (105), the pronominal IO has additional NP-marking in the form of the O-marking proclitic hn= 'ABS', discussed in the previous section.

- (105) ki=tanre? ʔen hn=kəh.
3A=show LOC ABS=3
She showed (it) to him.

Proclitic hn= is also attested in non-core PPs where, like here, it is optionally proclitic to a pronominal, but never a full NP. It is likewise glossed 'ABS(olutive)' to distinguish it from the O-marking function above. The oblique is subcategorised by the verb:

- (106) msti? n=nɪ<r>ka? rɔm hn=ko
must 1fA=marry<CAUS> WITH ABS=2f
I must marry (her) to you.

9.3.2 The intransitive and non-verbal clauses

S, the single argument of the intransitive clause, is not cross-referenced on the verb and lacks distinctive morphological marking, except for the third person pronominal forms kəhn '3S' ← kəh '3' and dehn '3plS' ← deh '3pl' (§6.1).

S can either precede the verb, or follow it (§9.2.1). Examples (108) and (109) show the unmarked NP in both pre- and post-verbal position.

- (107) kɔ? ke k'bəs
pig-tailed.macaque that die
The pig-tailed macaque died.

- (108) k'bəs pɔdɔŋ ke
die tiger that
The tiger died.

=hn is never enclitic to nouns to mark the function S. In (109) the enclitic in this position marks the third person possessor =hn '3POSS' (§9.1.3), translated: 'His tiger is dead.'

- (109) *pɔdɔŋ=hn k'bəs
tiger=S be.dead
*The tiger is dead.

S-function pronouns in pre-verbal position take the form kəhn '3S' and dehn '3plS' (110a) and (111a) respectively. Etymologically these pronouns are probably derived from the free pronominal form fused with enclitic =hn 'ABS'. When the third person pronominal follows the verb it is usually in the unmarked free form, kəh '3' or deh '3pl' (110b); infrequently it may be in the third person absolute form (111b). The motivation for this remains unclear.

- (110a) kəhn wh-woh
3S IMPERF-get.up
She got up.

- (110b) wh-woh kəh
IMPERF-get.up 3S
She got up.

- (111a) dehn swak
3plS go
They went.

- (111b) swak dehn
go 3plS
They went.

The subject, the single argument of the NVC, is coded in the same manner as the S of an intransitive clause, providing evidence that this argument is also an S. Nouns (112a), and first and second person pronouns are unmarked; the third person pronominal forms are the S-form pronominals kəhn '3S' and dehn '3plS' (112b):

- (112a) r<l>mol kke sma? kramat
be.male<NMZ> that person ascetic
That man (is) an ascetic.

- (112b) kəhn sma? kramat
3S person ascetic
He (is) an ascetic.

Post-verbal S-form pronouns only occur in connected discourse, see (38) §9.2, never in elicited data.

9.3.3 A-marking and the intransitive clause

The previous section focussed on the prototypical coding of core NPs. A subset of intransitive verbs may employ A-marking both on the NP and on the verb, cross-referencing the single argument with the pronominal proclitic and marking the post-verbal NP with la= 'A' – both features usually associated with transitive predicates.

The verbs which may mark S as A are human activity verbs such as jtek 'to sleep', swak 'to go', paloh 'to flee', dɔs 'to arrive', wh-woh 'to awake', ʔyam 'to cry', and stative verbs expressing human emotion: c2en 'to be happy', risaw 'to be sad', ʔnih 'to be ill', btɔŋ 'to be afraid', sʔrɔ? 'to miss (s.o.)', nyen 'to be angry'.

- (113) ki=c2en la=bapa? ʔyot
3A=be.content BCS=father return
She is happy because her father has returned.

This coding strategy is used when the action or state is caused by an external source, but it is not itself a causative construction (§5.5.1.2) – the verb lacks causative morphology, the clause is not transitive, and the NP_{causer} is not represented as an argument of the verb, but rather as a causal adjunct (114).

- (114) ki=ʔyam la=yε
3A=cry BCS=1
He cried because of me.

Example (115), like (114) above, illustrates clearly that the post-verbal NP is not marking A, given that it is not co-referential with the proclitic on the verb. The post-verbal NP is first person inclusive he '1&2', the proclitic is third person singular ki= '3A':

- (115) ki=malu? la=hc
3A=be.shy BCS=1&2
He is shy because of us.

The example in (116) contrasts this type with a causative clause in (117).

- (116) ye bnapit², ?arch ki=jtek
1 sing only.then 3A=sleep
I sang, (and) only then the child slept.

- (117) ye=j<r>t²ek knon
1A=sleep<CAUS> child
I put the child to sleep.

As with prototypical transitive constructions, the post-verbal A-marker la= 'A' is not obligatory. Given that this is a single argument verb, there is no potential for confusion with another argument:

- (118) ki=palo^h kni^h
3A=flee husband
The husband fled.

Examples with both an external agent NP and a reason NP only occur under elicitation:

- (119) ki=?yam la=kəh, la=kwch da? ye=jon
3A=cry A=3 BCS=biscuit NEG 1A=give
He cried because I didn't give him a biscuit.

9.4 Zero anaphora

Zero representation of nominal expressions is characteristic of Semelai. Any NP with an established identity may be anaphorically elided. Prior to discussing this, it is necessary to examine the means of representation of NPs in the clause. This is summarised in Table 9.2. In general, animate NPs have three means of lexical expression, NP, pronoun or clitic, although admittedly enclitic O is of very low frequency. Inanimate NPs lack a pronominal form, and their representation by clitic O is marginal. An alternative to using a lexical expression, or clitic pronouns which are unavailable with intransitive verbs, is the availability of zero representation of an argument.

Animate NPs which normally have a free or clitic pronominal form can be considered elided when they have zero representation in the clause. Animate NPs in the function of A have partial representation or ellipsis by virtue of the agreement marker, when the external NP has zero representation. Agreement marking is compulsory, except where the verb is repeated, so that in general NP_A is always represented in the clause.

These characteristics are driven by a general cross-linguistic preference that only the minimum number of arguments in a clause should have lexical representation (DuBois 1987). Hence, once the identity of participants is established, role and quantity constraints motivate the patterning of argument representation in the clause. This preference takes a number of forms. In (120) S is given zero representation.

- (120) ptom jtek
night sleep
(At) night (he) slept.

TABLE 9.2 AVAILABLE EXPRESSIONS OF CORE ARGUMENTS

NP	FULL NP	DEM	PRO	CLITIC	ZERO
A	+animate	+		+	+
	-animate	+			+
S	+animate	+	+	+	
	-animate	+	+		+
O	+animate	+	+	+	+
	-animate	+	+	?	+

In (121) one argument, O, has lexical expression in the first clause. In the following intransitive clause, where it is the single given argument, it receives zero representation.

- (121) ki=tikam, ah, hn=kō? ke, k'bəs
3A=stab ah O=pig-tailed macaque that die
He stabbed, ah, that macaque, (and it) died.

In (122) and (123) the A is expressed and the O, animate in (122) and inanimate in (123), is not expressed.

- (122) ki=tikam la=kəh (123) ki=pan-cin la=kmpən
3A=stab A=3 3A=CAUS-be.cooked A=wife
He stabbed (him). The wife cooked (it).

In (124) no external NP is represented, but the A is partially coded by the pronominal proclitic, in this case ki= '3A'.

- (124) ki=pan-cin
3A=CAUS-be.cooked
She cooked (it).

In (125), A and O are represented in the first clause; in the second clause where the verb is repeated, the A is recoverable and thus receives zero representation; in the final clause the A is represented where the action is new, but the O, which is given information, receives zero representation.

- (125) [de=yōk jatul], [yōk crēh], [de=ca]
[3plA=take pig] [take fish] [3plA=eat]
They took pork (and) took fish, (and) they ate (it).

In a clause with an understood O and IO usually only one of the two arguments is expressed, either the O as in (126), the IO in (127), or both the A and the O are expressed post-verbally in (13). Clauses where both the O and IO arguments are expressed are uncommon; an elicited example is given in (129).

- (126) ki=jon kwch
3A=give biscuit
She gave a biscuit.

- (127) de=h²gi? ʔen keh
3ptA=give LOC 3
They gave (it) to him.

- (128) ki=jon ia=koh hn=dak
3A=give A=3 O=water
He gave the water.

- (129) ki=jon kweh ha? knon
3A=give biscuit AT offspring
She gave a biscuit to the child.

(130) is a typical passage of text. There is an initial mention of the object, but it then receives zero representation.

- (130) keh hne swak brintay=ca?. ki=k³m [ma=?ikur jalul]_o
3 THEN go hunting=EM 3A=get one=CLF pig
ki=t³q, ki=gog ?yot
3A=grab.hold 3A=bring return
ki=7anjot la=krdor, ki=kr ja?-i? la=k<r>dor ke
3A=bring.in A=woman 3A=work-APPL A=bc.female<NMZ> that
ki=ca, ki=pan-cin, ki=ca
3A=eat 3A=CAUS-be.cooked 3A=eat
He soon went hunting. He got a pig, (and) he grabbed hold (of it) (and) he took (it) home. The wife took (it) in (and) she worked (=butchered and dressed) on (it). She ate (it). She cooked (it) (and then) she ate (it).

10 Basic clauses

This chapter explores basic clauses. In the first section, §10.1, non-verbal clauses are described. In §10.2 the unrealis constructions are discussed. The unrealis is used to express epistemic and deontic modalities, however the category is only expressed morphologically for transitive clauses. Before turning to the various non-declarative clause types, the modifiers of the simple clause are described in §10.3. Sections 10.4 to 10.6 discuss negation, the interrogative clause and the imperative clause, respectively. The markedness of relations holding between certain categories in the grammar is evident in the imperative construction (§10.6): transitivity is distinguished in the negative imperative, but not in the positive imperative. The morpheme is the unrealis proclitic ma= 'IRR(ealis)' introduced in §10.2, a category which is only expressed morphologically for transitive clauses.

10.1 Identificational, ascriptive and existential clauses

The clauses discussed in this section consist of a single argument, the subject and a predicate. The subject is defined as that which identifies the referent. The predicate is that which says something about the referent.

The clauses are referred to generally as non-verbal clauses (NVC). There are two non-verbal clause types in Semelai. The first type consists of the juxtaposition of two NPs, or an NP and a PP without an overt predicator. It is divided on semantic and to a lesser degree syntactic grounds into three sub-types: equative, ascriptive and locative. The second type of NVC, the existential and the metamorphic, have copula-like predators da? 'EXIST' and sot 'to metamorphose into'. The clause types are set out in Table 10.1.

TABLE 10.1 NON-VERBAL CLAUSE TYPES

PREDICATOR	PROPERTY	LOCATION
	Complement is NP	Complement is PP
NONE		
Equative/ascriptive	NP NP	
Locative		NP PP
OVERT		
Existential		
Metamorphic	NP sot NP	NP da? (PP)

Verbless clauses are discussed in §10.1.1; for the overt predicator clause see §10.1.2.

10.1.1 Clauses without overt predators

The clauses are defined and then a discussion reflecting the general homogeneity of the type follows in §10.1.1.1.

There are three clause types: equative, ascriptive and locative.

i. *Equative* This clause type identifies the referent of the subject as the referent of the predicate, where NP = NP.

The subject NP and predicate of the equative clause are freely permutable, distinguishing the equative from other clause types:

- (1) puyor kampung bapa? [waʔ.emboŋ] ʔap]Pred
shaman village father [EB 1f]Pred
The shaman (from) Kampong bapa? (is) my eldest brother.

- (2) waʔ.emboŋ ʔap [puyor kampung bapa?]Pred
EB If [shaman village father]Pred
My eldest brother (is) the shaman (from) Kampong bapa?.

ii. *Ascriptive* A particular property is ascribed to the referent. The complement cannot comprise proper nouns or pronouns other than as possessors. This distinguishes it from the equative and locative clause types.

- (3) bapa? saʔlan [pləsit]Pred
father [name] [malefic.shaman]Pred
saʔlan's father (was) a malefic shaman.

iii. *Locative* Locational clauses function to locate an entity. The locative clause resembles the equative in that it contains two referring expressions. In the locative clause the subject codes the locatee, and the predicate the location, expressed as a PP.

- (4) dom. kbəʔ=hn [he?] ke]Pred
AFF body=3POSS [AT:above there]Pred
Indeed, his body (was) up there.

Location is also expressed by the existential clause (§10.1.2.1).

10.1.1.1 General features of the clause type

Issues relating to constituent order, the subject, predicate, aspect and negation are examined in the following sections.

i. *Constituent order* In equative and ascriptive clauses constituent order is [Subject]-[Predicate] or [Predicate]-[Subject], determined by the constraints of the representation of old versus new information, see §9.2. In the locative type the ordering of constituents is fixed as location: locatee without regard to the new/old distinction.

- (5) "[cɔ? ke]Loc paya?" k'leŋ
[AT:below that]Loc river QUOTE
"Down there (is a) river," (he) said.

Interrogative ascriptive clauses display a tendency to place the old information after the ignorative in accordance with the usual ordering of interrogative clauses (see §10.5):

- (6) "zɔ? mandemoh nene?" k'meŋ
oh what noise QUOTE
"Oh, what (is) (that) noise?" he asked.

Interrogatives based on the ignorative nmoh 'what class' (§6.2.2.1) exhibit the reverse order [Predicate]-[Subject]:

- (7) p?la? nmoh?
animal INDIV:what
Of which class (is that) animal?

ii. *The subject* The subject can be any nominal expression (§7.1), e.g. a noun as in (8), or a deictic as in (9). The subject of the NVC is coded in the same manner as the S of an intransitive clause: nouns and first and second person pronouns are unmarked (8); third person pronouns are in the S-forms kəhn '3S' and dehn '3plS' (§9.3.2).

- (8) The identity of an individual is being discussed in terms of her ethnic group.
kmpən [sma?] jakun]Pred
wife [person Jakun]Pred
(His) wife (is) a Jakun.

- (9) A man's wife is identifying a child in a cradle as theirs. It is the first time her husband has seen his child because he abandoned his wife when she was pregnant.
ki=tupuk, "[kke [knən he]Pred]NVC" k'leŋ
3A=point.out [that [offspring 1&2]Pred]NVC QUOTE
She pointed, "That (is) our child," (she) said.

In (10) the subject is a complex nominal expression, an associative phrase modified by a temporal adjunct.

- (10) ien ca, [siʔsa? knən he ne]Subj [cɔ? habu?]Pred
want eat, [leftovers offspring 1&2 before] Subj [AT:below kitchen]Pred
(If you) want to eat, our children's leftovers (from) before (are) down in the kitchen.

iii. *The predicate* The predicate of equative and ascriptive clauses is realised as an NP: a noun, pronoun, deictic or a deverbal noun (§7.1).

As stated above, the predicate of an ascriptive clause cannot comprise proper nouns or pronouns, other than as a possessor as in (3) above, distinguishing it from the equative and locative clauses.

The predicate NP can display all the features of a complex nominal expression. In (11) the predicate is a recursive associative phrase.

- (11) koŋkoloh [cɔ [skɔ? [dloŋ],]lNP]Pred
nightjar [dog [spectre.huntsman [tree],]lNP]Pred
The nightjar (bird) (is) the dog of the tree(-dwelling) spectre huntmen.

The complement in (12) is modified by an attributive locative phrase.

- (12) sron [[tyək]_N [2en bri]_{PP}]Pred
 sron [[banana]_N [LOC forest]_{PP}]Pred
 The sron (is) a forest banana.

The ascriptive complement in (13) is modified by a relative clause.

- (13) dom bren {mesen gan me=ra?+t'ay}Pred
 AFF Bren {machine gun REL=COMP-be.big}Pred
 Indeed, the Bren (is) the machine gun (which) is the biggest.

The complement of a locative clause is realised as a locative (14) or directional PP (15), or if the location is a person, as a comitative PP (16) (§8.3.2).

- (14) dom 2ma? [he?] sbj padaj he?]Pred
 AFF mother [AT:above end field AT:above]Pred
 Indeed (his) mother (was) up at the end (of) the field up there.

- (15) [te h̥n]Pred 2us?
 [TO where]Pred lighter
 Where (has) the lighter (gone)?

- (16) nasip ka [rom knkon kkəl]Pred
 fate 2f [WITH child that]Pred
 Your fate (is) with that child.

When the object of the comitative complement is a pronominal it may host proclitic hn= 'ABS' (§9.3.1).

There is a potential problem with the analysis of locative and directional clauses.¹ In some instances it could be argued that the non-verbal predicate is just a locative complement of a verbal clause where the verb has been elided. The following clause provides an example of this: (17a) is non-verbal, (17b) is verbal.

- (17a) 2əŋ [ga=[te2en pŋkalən]_{PP}]Pred
 If [IMM=[TO:down landing.place]_{PP}]Pred
 I am going down to the landing place.

- (17b) 2əŋ [ga=swak [te2en pŋkalən]_{PP}]Pred
 If [IMM=go [TO:down landing.place]_{PP}]Pred
 I am going to go down to the landing place.

iv. Aspect Clauses lacking an overt predicator have limited possibilities with respect to aspect. In the equative and ascriptive clauses the imminent aspect marker ga= 'IMM' (§5.5.3), functions to predict the identity or a quality of the subject, in a time frame subsequent to that of the current discourse. This may be a long period, a matter of months as in (18), or virtually in the ensuing instant as in (19). The subject can be animate, as in (18) and (19), or inanimate, as in (20).

¹ This is a general problem with this construction which arises in many languages, e.g. Shakespearian English, 'I will to Dover'; and in the Australian language Kayardild, yada ñarn-kiring-ku (I beach-ALLATIVE-MODAL POTENTIAL) 'I will to the beach' (N. Evans p.c.).

- (18) "jake" kʰləŋ 2ma?, "[ga=knək kə sarek]Pred"
 then QUOTE mother [IMM=husband 2f future]Pred
 "In that case," said (her) mother, "(he) will be your husband in time to come."

Example (19) previews the metamorphosis of the 'person' back into the tiger at the moment when he is about to be slain by the father.

- (19) *The tiger has been masquerading as the children's mother.*
 yo? [sma?, songoh [ga=podəŋ]Pred]NVC
 but [person true [IMM=tiger]Pred]NVC
 But the person, as truth have it, would be the tiger (again).

- (20) t̥mp̥t [ga=ni-pʰui kəh]Pred kʰləŋ
 place [IMM=NMZ-come.down 3]Pred QUOTE
 "(This) place is going to (be) (the one of) his descent," (they) said.

v. Negation Non-verbal clauses employ b2en 'NOT' as the metalinguistic negator (§10.4.2.2). The negator is in initial position:

- (21) *An attempt to identify an animal in the trees is discredited by the responding speaker.*
 b2en, [he?] ke]Loc [musəŋ]
 NOT [AT:above there] [civet]
 On the contrary, (that) up there (is) a civet.

vi. Complex clauses The NVC can enter into complex constructions such as complement clauses (22), and concatenated constructions (23)–(24) (§11.2.1 and §11.4).

- (22) ki=kʰe? [ga=bapa?]NVC
 3A=intuit [IMM=father]NVC
 He intuited (he) was going to (be) (his) father.

In (23), the first NVC is locational, the second equative with the subject elided.

- (23) *A tiger has a human 'wife', as he refers to her. However, they do not live as man and wife, he lives on the ground; she lives up in the elevated hut.*
 [kəhn ke cɔ? 2ate]NVC, yo? "kʰməŋ]"NVC kʰləŋ=hn
 [3S that AT:below ground]NVC but, [wife]Pred]NVC QUOTE=3POSS
 He (was) down on the ground, but "wife" (he) called (her).

- (24) *A man responds in a jocular manner to an inquiry regarding the gender of his most recent born child – other than what he already has, what possibility is there?*
 la gi?=hn k<|> dor, la gi?=hn r<|> mol
 again=CONN be.female<NMZ>, again=CONN be.male<NMZ>
 Again (it's only) a girl, again (it's only) a boy.

10.1.2 Clauses with overt predicators

There are two clause types discussed in this section: the existential clause which predicates location (§10.1.2.1), and the metamorphic clause which predicates property (§10.1.2.2). They are characterised respectively by the presence of an overt predicator, da? 'EXIST' and sot 'to metamorphose into' (Table 10.1).

The predicates are unique in that generally they do not share the set of distributional properties found with other grammatical categories (§4.3.4). For example, da? 'EXIST' is not considered a verbal predicate on the grounds that it cannot be elided like a normal verb, does not co-occur with aspectual markers, and does not host any verbal morphology. Similar conditions apply to sot 'to metamorphose into', which cannot be elided, or negated either descriptively or metalinguistically. Hence, they are not considered verbs as such, although the possibility that they could be degenerate verbs must be acknowledged.

10.1.2.1 The existential clause

The existential clause states the existence or non-existence of an entity: 'There is/are (not) ...':

- (25) da? kbə?
EXIST body
There was a body.

The one feature the predictor does share with verbs and auxiliary verbs is that it is negated by the predicate negator da? 'NEG'. The predictor occurs to the right of the negator:

- (26) da? da? kbə?
NEG EXIST body
There wasn't a body.

The existential predictor da? 'EXIST' appears to be a loan from the Malay existential particle *ada* '(there) is/are', where it has parallel functions.

i. *General features of the existential clause* The existential clause consists of the predictor da?, the subject NP and an optional locative PP. The NP may precede or follow the existential marker; the locative phrase, which expresses specific location, usually occurs clause finally:

- (27) da? ninca ha? ponog ke
EXIST food AT hut that
There was food in the hut.

It is possible for the predictor to stand alone, with neither subject nor location expressed, as in response to the following question. Note that the question in (28) is also an NVC.

- (28) Q: te hən ?us ? A: da?
TO:unspec where lighter EXIST
Q: Where (has) the lighter (gone)? A: It's (around).

In (29), once established, the subject is elided, and the negated predicate stands alone.

- (29) bantəl da? da?.
pillow NEG EXIST

- req tet ke, da? da?. req tet no?, da? da?
seek TO:spec there NEG EXIST seek TO:spec here NEG EXIST
(There) was no pillow. Seeking there, (there) wasn't (a pillow), seeking here, (there) wasn't (a pillow).

Existential clauses with da? 'EXIST' can also be interpreted as expressing possession, i.e. 'have NP'. The ordering of constituents parallels that of the existential clause, where the NP preceding the predictor is specific/given. There is no formal distinction between the existential and the 'having' construction, although in the possessive the clause usually lacks a locative NP. Frequently, a clause can be interpreted in either way. As a result, da? is glossed as 'EXIST' in all instances and the choice is left to context. The possible alternate readings are exemplified in (30).

- (30) "da? no?no?", kʰəj, "tmpat jrajan dak"
EXIST this QUOTE place boiling water
"There is this," (she) said, "a thing to boil the water (in)." "We have this," (she) said, "a thing to boil the water (in)."

An alternative 'having' construction formed with the middle voice prefix, (§5.4.3), is exemplified in the second clause in the following example:

- (31) ?us ke da?, br-wṇy=hṇ, beh
lighter that EXIST HAVE-knife=3 NO
A lighter (he) had, a knife (he) didn't.

ii. *Constituent order* The two possible orderings of constituents in the existential clause are motivated by the representation of old/new information. If the existent NP represents new information, it follows the predictor da? NP (LOC); if the NP represents old information, e.g. in response to a question, it precedes the predictor: NP da? (LOC). In both instances, the LOC is in clause-final position.

NP final order is used for non-specific NPs to introduce a new NP (32) or, with an established participant, to introduce a new event or state (33).

- (32) "2ma? ye=?en" kʰəj, "kʰoy nehneh, da? da? kbə?"
mother 1=AUG QUOTE head only NEG EXIST body
"Our mother," (they) said, "(is) only a head. (She) doesn't have a body."

- (33) siraj jtək, ki=?yot kantup la=patʰir ke
while sleep 3A=return close.over A=tree.sp. that
jadi?, da? da? keh.
happen NEG EXIST 3
While (he was) sleeping, the pathir tree closed back over, (and so it) happened
(that) he wasn't there (anymore).

NP initial ordering is used when the NP represents old information.

- (34) A member of the family has returned home hungry, and someone enquires about food on his behalf. They get the following response:
huc=hṇ da? ha? kdej
cooked.rice=3POSS EXIST AT cooking.pot
His rice is in the pot.

- (35) *The protagonists, who live alone in the forest, return home to find that someone has mysteriously prepared a meal for them. This had happened once before.*
 loc ʔyot, dom, pn-ca da? ha? dɔl
 already return AFF NMZ-eat EXIST AT house
 (When) they returned, indeed (again) there was foodstuffs in the house.

- (36) *The narrator, who has been describing centipedes, summarises her account:*
 hitam da?, mə=mirah pon da?
 be.black EXIST REL=be.red too EXIST
 There are black (ones), (and) there are ones which are red too.

iii. Negation The internal negator da? 'NEG' negates the existential predicate (29). The existential clause, unlike the other NVC type, has the potential to co-occur with aspectual negators (§10.4.1). In (37) the aspectual negator is da? ... wo? 'NEG ... longer'.

- (37) pdaq da? da? wo? ha? təga?
 sword NEG EXIST more AT stair
 The sword was no longer on the stairs.

10.1.2.2 *The metamorphic clause*

The predicator sot 'to metamorphose into' represents a non-volitional change in the manifestation of an entity. Etymologically sot 'to metamorphose into' is related to Mah Meri sot a general verb meaning 'to become' (Kruspe in prep. a). In Semelai, however, 'sot' is restricted to the following: certain changes-of-state which form part of a natural process of development, e.g., the formation of a foetus in the womb as in (38), or the growth stages of certain plants as in (39); the consequence resulting from the transgression of a taboo as in (40), and the metamorphosis of a person at death as in (44).

- (38) *The Semelai believe that the foetus develops from menstrual blood.*
 mhām sot knən
 blood metamorphose.into offspring
 The blood transforms into the child.
- (39) solon jmɔ?, ʔleʔle? sot dloŋ
 first vine eventually metamorphose.into tree
 First (it is) a vine, eventually (it) metamorphoses into a tree.
- (40) *If one cuts one's nails at dusk, one's soul will turn into a tortoise and be in danger of being consumed by a malevolent shaman.*
 menka tkəl cros ʔare2.doy, sot banəŋ
 if cut nail day.afternoon metamorphose.into tortoise
 If (one) cuts (one's) (finger or toe) nails (at) dusk, (one) (will) metamorphose into a tortoise.

i. The subject In addition to the predicator sot 'to metamorphose into', the clause contains two NPs, which are understood to have the same referent; at least one NP refers to the new identity or 'state' of the metamorphee. The pre-predicate NP is termed the subject in line with the definition provided in §10.1.1.

The subject is expounded by an NP, realised either as a noun (41), or a third person pronominal (42).

- (41) puyɔŋ deh sot groŋj
 shaman 3pl metamorphose.into cricket
 Their shaman metamorphosed into a cricket.

As (42) illustrates, the third person pronoun is in the S-form, typical of the intransitive clause, and consistent with the other NVC types discussed above in §10.1.1 and §10.1.2.1.

- (42) kəhn sot tijɔ
 3S metamorphose.into snake
 He metamorphosed into a snake.

The subject can host the third person possessor enclitic =hn '3POSS' (43). The predicate can also host the enclitic, but not simultaneously with the subject as in (48).

- (43) ləc joŋ=hn sot pbs tijɔ ləc
 already foot=3POSS metamorphose.into tail snake already
 Already his feet (had) metamorphosed into a snake's tail already.

The subject may either precede the predicate as in (44), where it represents new information, or follow it as in (45) where it is old information.

- (44) A matiʔana? 'birth vampire' is a vampire which develops from a woman who dies in childbirth.
 "dom=cə?" kʰləŋ, "kmpən sot matiʔana?" kʰləŋ
 APP=EM QUOTE wife metamorphose.into birth.vampire QUOTE
 "Indeed," (she) said, "(your) wife metamorphosed into a birth vampire," (she) said.

- (45) kʰbəs b-knən, [sot matiʔana?]Pred kəhn
 die HAVE.offspring [metamorphose.into birth.vampire]Pred 3S
 (She) died giving birth (and) she metamorphosed into a birth vampire.

ii. The predicate The predicate consists of the predicator and an NP. The NP which follows the predicator, the complement NP, forms a constituent with the predicator; the complement NP cannot be fronted, suggesting that it is incorporated with the predicator:

- (46) *matiʔana?, sot [kəhn]Subj
 birth.vampire metamorphose.into [3S]Subj
 *A birth vampire, she metamorphosed into (one).

Further, the complement NP can never be elided:

- (47) *puyɔŋ deh sot
 shaman 3pl metamorphose.into
 *Their shaman metamorphosed into (one).

The NP is a simple noun (48), which may be encliticised with the third person possessive =hn '3POSS'.

- (48) *cəŋ [sot grələŋ=hn]Pred. kəh nɔ? neg*
immediately metamorphose.into cricket=3POSS]Pred 3 this before
Immediately he, the aforementioned, metamorphosed into a cricket-of-himself.

The presence of the enclitic has only been noted in clauses containing the resultative *cəŋ* 'immediately'. Its function is not fully understood, although it may formally mark the continuity of the identity of the referent of the two NPs; it is only the outward appearance that is altered, hence the slightly odd translation in (48) above.

The metamorphic clause cannot be negated, either internally or externally, consistent with the resultative nature of the construction. It also exhibits limited possibilities with respect to aspectual modification as described in the following section.

iii. Aspect The metamorphic clause is only used to express a resultative state, and cannot be used to express a change of state, except by presenting the transformation as a series of resultative stages. This signifies that the complete transformation is not instantaneous, as in (49) where the metamorphosis is described in stages.

- (49) *loc jɔŋ=hn sot pps tijə loc,*
already foot=3POSS metamorphose.into tail snake already
... siŋəŋ piŋəŋ loc, sot tijə
... as.far.as waist already metamorphose.into snake
Already his feet (had) metamorphosed into a snake's tail already ... Already up to his waist, (he had) metamorphosed into a snake.

Given that the predicate is resultative, there are limited possibilities for aspectual marking. Only resultative aspectual markers like *loc* 'already' (49), *cəŋ* 'immediately' (48) and *kmoŋ* 'completely', which function to emphasise the completedness of the event or sub-event, can occur in this clause type. *kmoŋ* 'completely, all' is the only constituent which occurs between the predicator and the complement NP:

- (50) *The metamorphosis complete, every part of the transgressor has become snake-like:*
naʔ-hə? nɔ? sot kmoŋ tijə
DEM-AT this metamorphose.into completely snake
This one here had completely metamorphosed into a snake.

10.2 Irrealis constructions

The irrealis mood is expressed by the clitic *ma= 'IRR'* on transitive verbs; it is not marked morphologically for intransitive verbs, with the exception of those stative intransitive verbs expressing human emotion or experience which can host pronominal proclitics (§10.2.2.1 and §10.2.2.3), or intransitive stative verbs expressing qualities of inanimate entities (§10.2.2.2). The morphological possibilities for *ma=* are discussed in (§3.3.2) and its functions are introduced in (§5.5.2). The formal features of irrealis clauses will be discussed in §10.2.1.

ma= 'IRR' functions to express both epistemic and deontic meanings. Epistemic modality² (§10.2.2) encompasses the following:

- a) Inferential
- b) Experiential
- c) Mirative
- d) Hypothetical

Deontic modality (§10.2.3) is restricted to the expression of:

- a) Negative transitive imperatives
- b) Denial of permission
- c) Prohibitions
- d) Lack of obligation

The expression of epistemic modality by the irrealis is optional, whereas with deontic modality it is obligatory.

Cross-linguistically, the irrealis mood is often associated with the expression of a range of unrealised or non-actual events, although it must be stressed that the difference between the two categories is not in terms of the logic-bound interpretation of modality, i.e. a contrast between real and unreal events (Givón 1994: 268-9). The essence of the morphosyntactic expression of modality in Semelai is as a means of providing information on the status of the transfer of information between the speaker and hearer, i.e. a communicative pragmatic device.

10.2.1 Formal features of an irrealis clause

This section focuses on the irrealis in epistemic clauses. Specifics relating to deontic clauses are discussed in §10.6.1.

The presence of the proclitic *ma= 'IRR'*, which occupies the position normally filled by the verbal pronominal proclitic, signifies the expression of a lower degree of certainty with respect to the expressed event. There is considerable neutralisation in the irrealis, so that categories such as person/number marking are not distinguished, and there are reduced possibilities for aspectual marking (cf. Palmer 1986: 82 on cross-linguistic parallels for this).

The argument frame for transitive verbs prefixed with *ma= 'IRR'* is as follows: *ma= 'IRR'* and the pronominal proclitics are mutually exclusive, so the agent is unexpressed, but the status of the remaining core arguments, O and IO, remains unchanged.

The NP aligned with A is not explicitly present either as a verbal proclitic or a core NP in the irrealis clause. It can however be represented externally as an unattached NP. In (51), it is a clause-initial NP *twanprti?* 'the princess'. Proof that the NP is not a core argument is attested by the fact that the NP aligned with the subject of an irrealis clause cannot be relativised. The same restriction does not apply to objects or indirect objects which may be the head in such clauses.

² Mood is not to be equated with modality. Following Palmer (1986), I draw a distinction between the grammatical category of 'mood', and the semantic category of 'modality'. See also, Frawley (1992: 386-7), and Bybee and Fleischmann (1995: 2).

- (51) dom twanpri?=?son, ma=gas dom ple hmbacaj ke
 AFF princess=SC IRR=peel AFF fruit *Mangifera foetida* that
 Indeed, (it was) the princess, see, indeed (to their) puzzlement (she) peeled that
 horse mango.

An O is marked the same as for a realis clause; it can host the clitic hn= 'O' as in (52).

- (52) ma=tɔrsh hn=gtah
 IRR=tap O=rubber
 As I recall, (I) would tap rubber. (Lit. Rubber would be tapped.)

The O can be fronted to pre-verbal position as in realis clause types. (53) is a rare example of animals given pronominal representation in a clause.

- (53) *The speaker is talking about a certain type of 'centipede', klay.*
 ?dspl deh ma=t-genen, wc-wc deh
 if 3pl IRR=HAPP-disturb IMPERF-curl.up 3pl
 If they would happen to be disturbed, they curl up.

The O can be relativised, as in the bracketed constituent in the next example:

- (54) "... kira? da? bnda? [mə=mə=lən]" k"ləŋ
 figure EXIST thing [REL=IRR=want] QUOTE
 "... you see there are things that (you) would desire," (he) said.

As stated earlier, there are reduced possibilities of aspectual marking in the irrealis. The only frequently attested marker of aspect in the irrealis is the imminent aspect marker ga= 'IMM'. See examples (55) and (63).

- (55) ga=ma=jon ?en ?amay
 IMM=IRR=give LOC [name]
 (It) would be given to ?amay.

The following example represents the only instance in my data of the aspectual loc 'already' in an irrealis clause:

- (56) ləc ma=?ye, ləc dos ptom, deh p?st ha? tmpat kəh
 already IRR=see already arrive night 3pl stay AT place 3
 Already, it seems (when) it would have been possible to see (it), already night had
 fallen, (and) they stayed at his place (not realising they were already there).

The discourse clitic =son (§13.3.2), expressing speaker conclusion ('SC'), is the only discourse clitic attested in clauses in the epistemic mood:

- (57) *The speaker remembers the arrival of the first white people, British forces making food drops from aeroplanes. The elders wouldn't let the young people approach the soldiers.*
 beh=hn da? ma=kali=?son, da? de=jon
 NO=CONN NEG IRR=be;brave=SC NEG 3plA=permit
 On the contrary, as I recall, it wasn't (that) (we) weren't brave, they wouldn't let
 (us).

10.2.2 Epistemic modality

The epistemic interpretation of the irrealis expresses the status of knowledge. This may reflect either the speaker's knowledge or experience, or it may reflect the speaker's assumptions of the hearer's knowledge or experience. Thus the epistemic mode can be characterised as a subjective view of events or states.

The exchanged information is filtered according to the subjective view of the speaker, and in a somewhat iconic manner, is expressed morphologically if it represents a marked situation, i.e. the status of the proposition is judged as to whether or not it is deemed concordant with expected conditions of normality, either speaker- or hearer-based. It is important to point out that much of what is conveyed in the epistemic mode could be equally expressed in the realis. A similar observation is made by Lazard in relation to the function of this category in Balkan and Middle Eastern languages: "The important point is that, at the morphosyntactic level, neutral forms implying nothing about the origin of the information conveyed are opposed to marked forms implying reference to the origin of the information without specifying it" (Lazard 1999: 95).

The epistemic function of the irrealis is used to express:

- The inferential: a conclusion drawn from evidence rather than direct perception.
- The experiential: a retrospective view of one's actual experience of previous states or events.
- The mirative: unanticipated observation, based on the direct perception of current events or states.
- The hypothetical: the speaker suspends their assertion, given that the truth-value of the protasis is unknown.

An example of each of these is given below.

Inferential:

- (58) *A man reassures his wife that she has the opportunity to sleep. In his opinion there is nothing that will disturb her.*
 "zəh jtek!" k"ləŋ knək, "da? da? ma=kacaw," k"ləŋ
 oh sleep QUOTE husband NEG EXIST IRR=disturb QUOTE
 "Oh sleep!" said the husband, "there's (nothing) (that) would disturb (you)," he said.

Mirative:

- (59) tlaŋa? ma=yɔk hn=dak j<n>h>zəh
 well IRR=take O=water drink<NP>
 (It is) the well (from which) the drinking water would be fetched.

Mirative:

- (60) *A woman, strolling in the garden, felt something leap momentarily onto her head. She looked up into the trees to see what it was.*
 ki=blo-bloj, ma=jɔk tupay
 3A=INTNS-look.upwards IRR=look.at squirrel
 She looked and looked upwards (and) to her surprise saw a squirrel.

Hypothetical:

- (61) A p?re? 'ghost of one who has died a violent death' tells his mates that they must not feed from the soul of his human wife, as their ilk are wont to do:
 da? ga=deŋ nɔ?, ma=yɔk ʔen bayaj
 NEG IMM=manner this IRR=take LOC soul
 (It) isn't going to be like this, (that you can) take (her) soul.

The following discussion will examine the various manifestations, according to predicate type, that occur within each function listed above. The semantic groupings of predicate types are: a) cognitive disposition (hear, see, observe, understand); b) affective states and bodily sensations (desire, be angry, be difficult), and c) activity predicates (transitive verbs expressing activities or processes). It is only through a combination of predicate type, person and context that the particular nuance is derived, so that within this category there are sub-types, some of which have more specific meanings when combined with certain person categories.

Take, for instance, predicates which express the inner feelings of the individual, for which the first person is the only truly reliable authority. In the first person this is experiential (§10.2.2.2) or mirative (§10.2.2.3). When used in the second and third persons, the function is skewed, and instead of the authority of first-hand knowledge the status of the information is of a lower ranking, as inferred or deduced knowledge (§10.2.2.1).

The essential difference between the two moods (realis and irrealis) is captured in the following pair of examples provided by a speaker to illustrate the difference between clauses containing the pronominal proclitic in (62) and the irrealis *ma*= in (63). In (62), the clause describes an actual event: the protagonist has some food and states that she is shortly going to feed it to the child, demonstrated by the use of the aspect marker *ga*= 'IMM', indicating that the event is on the verge of happening. In the second clause (63), the protagonist states what she would do in an expressed world: that if she had some food, the consequence would be that she would feed it to the child. (62) is a statement of fact, whereas (63) is a statement of intention in a hypothetical world.

- (62) ga=yɛ=pan-ca ʔen knɔn
 IMM=1A=CAUS-eat LOC child
 I am going to feed (it) to the child.
 (63) ga=ma=pan-ca ʔen knɔn
 IMM=IRR=CAUS-eat LOC child
 (If I had food), (it) would be fed to the child.

10.2.2.1 *The inferential*

Inferred information is that which is derived not from direct perception, but indirect evidence. The presence of the irrealis overtly marks a conclusion underpinned by lack of direct evidence. In (64), the irrealis is used to state a general statement about approaching rain. It is something that could be heard, if the right circumstances prevailed, but given that they do not, the hearer cannot verify the speaker's statement.

- (64) gagoh, swara? ʔare? ma=?yəŋ
 gagoh sound rain IRR=hear
gagoh is the sound of rain that would be heard/be audible (as it approaches).

The inferential reading is primarily used with the third person, where it provides a subjective evaluation of the narrative event, indicating that the speaker qualifies the information as inference. The inferential is glossed as 'it appears ..., it seems ...'.

In (65) the narrator makes an assertion using the irrealis, conveying that the statement is based on inference or indirect knowledge acquired, for example, through oral tradition.

- (65) gadə-gadə? yɛ=ʔen, ma=kʰe?
 RDP-parent 1=AUG IRR=know
 It seems (our) forefathers would have known (this).

The narrator frames a story:

- (66) ma=?ye=sən, dom kaka? ye dapat ma=jŋɔk
 IRR=see=SC AFF EZ 1 able IRR=observe
 I think (someone who) would have seen (it), indeed I think my elder female sibling would have got to witness (that).

In a similar vein in (67) the speaker states an opinion about what she feels would be the general attitude toward giving away something treasured.

- (67) da? da? [ma=ma=jon]
 NEG EXIST [REL=IRR=give]
 I don't think it is something that (one) would want to give (away).

- (68) A woman advises her son not to keep on washing his jeans, as they will certainly be slow to dry on account of the incessant monsoonal rains.
 boy ma=pɔk-i! susah ma=krey la=?are?
 NEG:IMP IRR=wash-ITER be.difficult IRR=be.dry BCS=rain
 Don't keep washing (them)! (It) will be difficult (for them) (to) dry because of the rain.

i. *The inferential and affective state predicates* When affective predicates occur in conjunction with the categories of second or third person, the irrealis expresses a lack of authority on behalf of the speaker to be certain of a person's inner state. The speaker can only make an inference based on their perception of the person's demeanour. Example (69) is in the second person.

- (69) A dog is whimpering repeatedly at the food that he and his master have discovered in their house. Based on the dog's behaviour, his master is of the opinion that the dog wants to eat.
 ki=spɔt-i! la=cɔ ke,
 3A=whimper-ITER A=dog that
 la=pn-ca da? dom ha? ponɔŋ dehn ke.
 BCS=NMZ-eat EXIST AFF AT hut 3plS that

"da? da? ma=lan mə=sŋat" k^bəŋ tuhən.
NEG EXIST IRR=want REL=whimper QUOTE master

"ma=lan ca, pn-ca he=cə?" k^bəŋ
IRR=want eat NMZ-eat 1&2=EM QUOTE

The dog kept on whimpering, because the food was indeed there in that hut (of) theirs. "There's nothing you could want (to) whimper (about)," said the master. "(If you) want (to) eat, indeed it's our food!" (he) said.

It is also possible to express the affected state of a third person (70)–(71), the narrator inferring how the protagonist must feel:

- (70) *A man's father is blind and stays in the house, unable to assist in the daily chores.*

ma=susah bapa?hn p?ot ha? dol
IRR=be.difficult father=3POSS stay AT house
(It would have been) difficult for his father staying in the house.

- (71) ki=lanakah tupay rōm k^boy.

3A=step.across squirrel WITH head

ttap lən ma=gren, p?la? da? pasal!
be.certain want IRR=be.angry animal NEG proper

The squirrel bounded across via her head. Certainly, I'd say (she) would have wanted to get angry, such an improper animal!

ii. The inferential and intransitive stative predicates An interesting extension of the irrealis is found in relation to stative predicates used to classify non-referential inanimate entities in terms of an inherent quality (§5.5.2.2).

- (72) ple ma=bul
fruit IRR=be.toxic

The fruit (is) poisonous. (poisonous fruit)

- (73) ple da? br-ca, ma=bul
fruit NEG MID-eat IRR=be.toxic
The fruit isn't eaten, (it's) poisonous.

The choice of the irrealis here appears to be conditioned by the fact that presumably one has not eaten the fruit and discovered that it is poisonous, but rather the knowledge is based on ethno-botanical knowledge which has been acquired from another source. Consider another case involving the verb bul 'to be toxic, of inanimates', which when applied to persons means 'to be intoxicated'. When used to express drunkenness, the realis mood is used. Obviously one has outward evidence if a person is drunk, their speech or movement is impaired, or perhaps one has observed them drinking and the amount consumed.

- (74) sma? bul
person be.intoxicated
(a) drunken person

Some further examples follow. The speaker categorises a set of taboos in terms of consequence:

- (75) banja? ma=?pih
group IRR=be.ill
the group (associated with) potential illness

This construction is also applied to the naming of transgressions, which have the potential to cause illness, or even death. In a sense one is uncertain of the outcome, but there is the possibility of such a consequence based on cultural knowledge and witnessed events. Note that in (76) it is not the NP which is affected by the verb, unlike the example of the drunken person in (74).

- (76) ma=cika?
IRR=diarrhoea
(food-mixing taboos resulting in) potential diarrhoea

Although this usage could initially appear to be somewhat lexicalised, the following examples suggest that it is indeed marginally productive with certain intransitive stative verbs, in particular those relating to cognitive disposition as in (77), and affective states as in (78).

- (77) *The speaker ponders the abilities of his forebears:*

da? məcəm tulis, de=k^be?=ja, ma=crdek
EXIST like write 3A=know=CL IRR=be.clever
(There) was like writing, they knew (it), (those) (who) were clever.

- (78) *The narrator describes how the protagonist's body must have felt as he transformed into a snake.*

beh da? ga=?olo?, məcəm ma=layur, g<p>hop kbe?
NO NEG IMM=joke like IRR=wither be.hot<NMZ> body
No, it was no joke, (it felt) like (he wanted) to wither (from) the heat (of his own) body.

10.2.2.2 The retrospective experiential

The co-occurrence of the irrealis with the category of first person expresses the retrospective experiential. This is the recollection of the speaker's direct experience of past states and events. The former state of affairs, which no longer exists, is represented in the irrealis. The retrospective account is distinct from a resultative meaning, which is what would be conveyed if the realis was employed in its stead. The remembering of previous states or events is based on the speaker's direct personal experience, and it is translated as '(as) I recall ...'. The experiential occurs with activity (79)–(81), perception and affective state predicates (82)–(83).

- (79) he=k^bəm dwet=sən, dom nɔ?, ma=tərə hñ=gtah
1&2A=get money=SC AFF this IRR=tap O=rubber
We'd get money, see, indeed (like) this, as I recall, (I) would tap rubber.

- (80) nɔ?nɔ?, ma=k^bəm spuluh ryal
this IRR=get ten dollars
(For) this, as I recall, (I'd) get ten dollars.

In the following examples the irrealis event is embedded under the negator da? 'NEG'. In (81) the speaker reminisces about the period of internment and the curfews imposed during the Communist Emergency.

- (81) da? ma=k^bom 2p_{es} ?are? deŋ nɔ?
NEG IRR=get be.high day like this
As I recall, (I) didn't get to (wander at large) in the middle of the day like this.

In (82) the clause is modified by da? prnah 'never'. The speaker aligns herself with the forebears, expressed as the collective first person. She was in fact present in the events which she recounts.

- (82) *In order to attract the Semelai to the internment camps, the British gave every family one hundred ringgit.*
me=c^bo?, da? prnah ma=?ye, drl?, dwet ma=ratus
REL=before NEG ever IRR=see 1coll money one-hundred
(Our) forebears, never had we (Semelai) seen (it), money (like) one hundred (dollars).

In (83) the two uses of the irrealis, with first and third person are illustrated. In the first instance, the narrator recalls that she did not witness the events herself. In the second instance of the irrealis she infers that those who probably would have seen the events were the people who lived in the area at the time. She herself was living downstream at the time and only heard the story, which she states in the final clause, using the realis.

- (83) da? ma=?ye crit_e, dom me=ha? nɔ?.
NEG IRR=see story AFF REL=AT this
ma=?ye, sma? m_e=kde? ha? nɔ?.
IRR=see person REL=stay AT this
ye=sən cɔ? ke, cɔ? krweŋ.
i=SC AT:below that AT:below keruing
ha? ke crit_e ye=?yen
AT that story 1A=hear
As I recall, (I) didn't witness this story here. (Those who) would have witnessed (it) were the people who lived here. I (was) down there, down at Keruing. That story there, I (only) heard (it).

10.2.2.3 The mirative

The mirative is employed in contexts resulting from direct perception, marking it as an unanticipated observation counter to expectation or normality.

One of the options for irrealis, cited by Chafe for the Amerindian language Caddo, is to mark a situation as contrary to the speaker's expectation, "a negation of normality" (Chafe 1995: 357) associated with an actual realised event. This is also attested in the Albanian 'admirative' (Friedman 1986: 180). For a discussion of the validity of mirativity as a cross-linguistically relevant category, see DeLancey (1997); and for further evidence of its recurrence across languages, see Lazard (1999).

In Semelai, the mirative occurs in first and third person in conjunction with predicates of visual perception and cognition and exclusively in the first person with affective state predicates.

The use of the mirative sense with the third person 'perceiver' contradicts the 'assertability constraint' applying elsewhere. The ability of the speaker to know if another is surprised by what they perceive suggests that the narrator is accorded special 'narrative authority'. All the examples of third person with perception verbs are drawn from traditional narratives. Although this requires further investigation, it is supported by the general use of the mirative with the third person in narrative discourse.

i. *The mirative and perception and cognition predicates* This construction, irrealis marking on a verb of perception or cognition, has the resultant meaning of observing something counter to expectation or normality, 'to X's surprise' indicating surprise, or even disbelief as in (84). Speaker A informs Speaker B of a death. Speaker B provides the standard response to this type of information, indicating a lack of prior knowledge.

- (84) A: kahn k^bbəs
3S die
B: da? ma=?yeŋ
NEG IRR=hear
A: He died.
B: I hadn't heard.

A speaker paraphrased the above response by using the restrictive negative modifier da?...cəŋ 'not at all', indicating that the speaker had no prior knowledge of the information just offered; it is new to her:

- (85) ye, da? ye=?yeŋ cəŋ
I NEG 1A=hear at.all
Me, I hadn't heard at all.

The mirative occurs with the cognitive predicate k^be? 'to know, understand', expressing an inability to comprehend something:

- (86) *The speaker doesn't understand the motive behind the instructions she has received from her husband:*
beh, da? ma=k^be? pn-?ur kəh
NO NEG IRR=know NMZ-direct 3
No, to my surprise (I) don't understand his instructions.

The perception predicates which occur with this meaning are ?yeŋ 'to hear' (87), jyɔk 'to look at, to observe' (88), and prati? 'to watch'. The verb ?ye 'to see', which has a more general sense of 'to perceive visually', is not used in this construction.

- (87) *The child, nicknamed bɔn tr̩k (bear mischievous) 'Mischievious Bear' is, under normal circumstances, a chatterbox. Today, as his grandmother comments, he is unusually quiet. She teases him gently, knowing that his silence is due to my presence.*

da? ma=ʔyəŋ cəŋ cokpɔ bɔn.tr̩k
NEG IRR=hear at.all speak bear:mischievious
(To my surprise), I haven't heard Mischievious Bear speak at all.

- (88) *A woman ventures down to the dwelling place of the ghosts of those who have died a violent death. She sees people who are red, amusing themselves in the blood.*

ma=jpɔk sma? ma=mirah ha? nkoi ʔen mham
IRR=look.at person REL=be.red AT play LOC blood
(She) observed (to her horror) people who were red at play in the blood.

The juxtaposition of the two predicates of perception emphasises the complete unexpectedness of the observation:

- (89) *A man looks up from wiping his eyes only to see someone long absent approaching him:*

ki=sapu? mot, ma=prati?, ma=jpɔk,
3A=wipe eye IRR=watch IRR=look.at

dom ma=sunot hn=kəhn ke
AFF REL=circumcise O=3S that

He wiped his eyes (and) to his complete surprise looked (and) observed, indeed, (it was) the one who had circumcised him.

The perception verb *jɔk* 'to look at' often co-occurs with the ignorative *mande* 'what sort'. The pair of examples in (90)–(91) represent what can be considered a near minimal pair contrasting the use of the irrealis.

- (90) *A group of siblings are searching for their youngest sibling, who they ostracised from the family home as a child. They are surprised when they discover his good fortune in life.*

"tah na? he? ʔadi? hn=ʔen?" sampay deh, dom,
Q NMZ-AT:above YS 1&2=AUG until 3pl AFF
dɔs deh ma=jpɔk mande dol, gdadɔ dol
arrive 3pl IRR=look.at what.sort house huge.house! house
"Maybe the one up there (is) our younger brother's?" So they, yes, they arrived (and) to their surprise saw a house, a huge, magnificent house.

In comparison in (91), the speaker is instructing the woman to open her eyes revealing to her that she will see his house if she should do so. Hence, what she sees is not counter to expectation. This is reflected in the fact that the verb of perception is not marked for the irrealis, compared to previous examples.

- (91) "kɔ=łən ma=ʔye dol ʔəŋ, sapu?=cəŋ mot!
2fA=want IRR=see house If wipe=EM eye

kɔ=ʔye," kʰlaŋ, "dol ye," kʰlaŋ.
2fA=see QUOTE house 1 QUOTE
ki=sapu?, ʔarch ki=jpɔk mande
3A=wipe just.then 3A=look.at what.sort
ha? tɔm kpon kəh,
AT base tree.sp 3
dol tharibən, gdadɔ dol
house grand huge.house! house
"(If) you want to see my house, wipe your eyes! You'll see," (he) said, "my house," (he) said. She wiped her eyes (and) just then she saw something at the base of his Kepong tree, a grand house, a huge (house).

ii. The mirative and affective state predicates The use of the mirative in the first person in relation to affective states may seem surprising given that one's inner feelings should perhaps not be unexpected to oneself. However, in Semelai culture inner feelings are considered autonomous, and their manifestation is not necessarily anticipated. The use of the mirative in predicates of this type occurs in Albanian where it expresses intensification of the feeling, and the expression of the state in an unmarked form would be considered cold (Lazard 1999: 100). Examples are provided below.

- (92) *A tiger is about to take a woman on his back across the water. She expresses her fear:*

"beh," kʰlaŋ k<r>dor "da? ma=kali"
NO QUOTE be.female<NMZ> NEG IRR=be.brave
"No," said the woman, "I'm not brave (enough)."

- (93) *A woman relates her discomfort whenever she looks at centipedes:*

ma=gren ja<k>jpɔk³
IRR=be.angry look.at<IMPERF>
(I) get uncomfortable (whenever) (I) look at (centipedes).

iii. The mirative in interrogative clauses The irrealis can occur in interrogative clauses of the information type. This is in contrast to examples of the interrogative examined in the literature which all involve polar interrogatives (Mithun 1995, Givón 1994), and not information interrogatives as is the case here. Givón attributes the presence of the irrealis in polar questions to the epistemic sub-mode of low certainty (1994: 273). Cross-linguistically, there are two options behind the motivation for the irrealis in interrogatives: a) the interrogative can show the same sort of sensitivity to mood as the declarative, real or unreal; or b) it can consistently mark all interrogatives as irrealis on the basis that they have not been realised. The latter is not the case in Semelai as all interrogatives with the exception of those discussed here are generally in the realis.

In Semelai the presence of the irrealis in an interrogative clause could express one of two things: a) mirative, a speaker-based expression of an event counter to normality, or b) inferred, where the speaker cannot infer the person's motive. Either

³ The motivation for the presence of /a/ in the imperfective form of *jɔk* 'to see' is unclear.

interpretation is plausible. The questions are based on the speaker's direct perception of an unanticipated situation, as illustrated below. When *ma=* is procliticised to the verb it forms an interrogative which speakers consider to be impolite, expressing vexation toward the referent. These are direct questions where the referent is the second person, with the exception of the rhetorical question in (94), my sole example of first person usage.

- (94) *The speaker, who is repeatedly scrubbing at a garment, is vexed by her effort, and asks herself out loud:*

brapa? tən lən ma=brus?
how many instance want IRR=scrub.cloth
How many times do (I) need to scrub (it)?

In the following example, someone has turned up at the speaker's house unexpectedly, and they enquire as to the visitor's purpose:

- (95) mandehməh mə=ma=rej dəs ha? nə??
what REL=IRR=seek reach AT here
I can't infer what (it is) that (you) would be seeking, turning (up) here?

This construction is also used in the relativisation of affective states:

- (96) *The speaker cannot imagine what possible difficulty could have prevented the addressee from achieving his goal.*
dməh mə=ma=susah?
what REL=IRR=be.difficult
What (was it) that could have been difficult (for you)?

In (97) a man encounters a group of monkeys who are fearful of him. This vexes him, for he can see no reason for their fear. In fact, he wants to invite them to join him in his travels. In the following clause, containing an affective state verb, he asks them why they fear him:

- (97) mənde mə=bt^bɔŋ-i? ye?
why IRR=be.afraid-APPL I
da? da? hal mə=bt^bɔŋ-i? ye
NEG EXIST reason IRR=be.afraid-TR I
Why are you afraid of me? There's no reason (to) be afraid of me.

iv. Discourse functions of the mirative in traditional narratives In traditional narrative discourse, the irrealis is used by the narrator as a tool to mark unexpected events which are deemed, for whatever reason, as being counter to expectation or normality. This is tied in with the fact that these events are also central to the unfolding story. The failure of the participants in the narrative to understand the importance of the events may provide the most salient reason for selecting the irrealis. Note that here the irrealis is always employed in narratives for events which are realised.

- (98) *The arrival of some travellers occurs out of the blue, counter to expectation.*

?əh=sən, kira?, mə=dəs deŋ ke.
EXCL=SC figure IRR=arrive manner that

beh, sɪput skadar, ?əh ki=pan-ca, ki=j<r>tek
NO snail [name] EXCL 3A=CAUS-eat 3A=sleep<CAUS>
Ah, you see, (they) came unexpectedly like that. On the contrary, sɪput sekadar, ah, he fed (them) (and) put (them) to bed.

In the following example the princess peels the skin from a mango. None of the other participants is aware of the chain of events. The posture maintained while the activity is carried out is counter to normality. This is also a pivotal scene in the story, for later we discover that she has impregnated herself with the handle of the knife. At the end of the scene, the realis is again adopted as the mode of narration.

- (99) twanptri?=sən, mə=gəs dom ple hmbacəŋ ke.
princess=SC IRR=peel AFF fruit *Mangifera foetida* that
ma=gəs=sən nm-k^bom deŋ ke,
IRR=peel=SC NMZ-be.sitting manner that
kira?, mə=pte? hn=dol wəy ke ʔən pṛijkaʔan,
figure IRR=put O=handle knife that LOC vagina
ma=gəs. ns-gəs də gəp nə?=pa.
IRR=peel NMZ-peel OF malay this=FACT

kna? syap, ki=ca
when be.ready 3A=eat.

The princess, indeed, peeled that horse mango. Surprisingly, she peeled it sitting like that, see, putting the handle of the knife into her vagina (to) peel. That's the Malay way of peeling you know. When (she) was done, she ate (it).

A further example of actual events expressed in the irrealis is given in the following example from another text. A father watches his daughter, but cannot understand her actions:

- (100) "ga=j^boy mənde daya?"
IMM=do what darling
"da? da?=ca?" k^bəŋ.
NEG EXIST=EM QUOTE
ma=yək hn=dloŋ, mə=cəŋket rəm dloŋ dom baju? kke,
IRR=take O=stick IRR=carry WITH stick AFF clothes that
ga=mə=cam^bɔŋ he? ?us
IMM=IRR=add AT:above fire
"What are you doing darling?" "Nothing," (she) replied. And to his puzzlement, (she) fetched a stick (in order to) carry that shirt with a stick, in fact (she) was going to put (it) on the fire.

10.2.2.4 The hypothetical

The presence of irrealis marking in conditional constructions is well attested cross-linguistically (cf. Chung and Timberlake 1985: 250-5).

The irrealis in the hypothetical clause is a direct reflection of the lack of evidence available to the speaker at the moment of utterance. The truth-value of the protasis is unknown, and given that the proposition is dependent on this, the result is a suspended assertion. This is in accordance with the assertability constraint.

In Semelai, the protasis is marked for the irrealis, and the consequent is expressed in the realis.

- (101) *The speaker is a p're? (a ghost of one who has died a violent death) who has a human wife. These ghosts prey on the souls of humans. Knowing what his ilk are like, the p're? warns his mates to stay away from her. Should they not abide by his wishes, things won't be the same between them.*

"ten," k'laq "beh, da? da? pasal, ma=yok ten bayan keh"
EXCL QUOTE NO NEG EXIST proper IRR=take LOC soul 3
"Eh," (he) warned, "no, it wouldn't be proper, (if you) were to take her soul."

In (102), the consequent precedes the condition:

- (102) *On a still day the speaker describes how the noise produced by a specific bbaleg 'windmill', located a couple of kilometres away on a hill, is faint if heard from their location. A bbaleg is a large windmill, constructed specifically for the enjoyment of the noise it produces.*

da? bhbeh, swara? bbaleg, ma=?yəŋ təm ha? nɔ?
NEG be.clear sound windmill IRR=hear SRC AT this
It isn't clear, the sound of the windmill, (if/when) it is heard from here.

In (103) the clause is negated by the external negator beh 'NO'.

- (103) *Each evening a woman puts food aside for her husband. In the morning she finds the plate empty, but she never actually witnesses him eat.*

beh, ma=?ye ki=ca, beh
NO IRR=see 3A=eat NO

No, (it was) not as if (she) witnessed (that) he ate (it), on the contrary, no.

The irrealis may be used with ?psal 'if, when(ever)' (§11.5.3.2). This is the only instance where a lexical item expressing modality is used in conjunction with irrealis marking; it is not used with other conditional connectives, or when ?psal 'if, when(ever)', hosts the enclitic =hn 'CONN'. Example (104) comes from a text where a woman is discussing centipedes. She teases the hearer, stating the possibility that should she be unable to light the cigarette and fulfil her desire, then the chances are that she will be chased by a centipede as a consequence of having unfulfilled desire.

- (104) ?oy, ?psal da? ma=k'om, ki=halaw klay
VOC if NEG IRR=get 3A=chase centipede
You, if (you) don't succeed, a centipede will chase (you).

- (105) *The speaker states how she must carry out the instructions which she has been given.*

"?ah beh" k'laq "da? ma=c?oŋ
EXCL NO QUOTE NEG IRR=roast

kəb, da? ga=kł=toroy wɔ? ha? dol
3 NEG IMM=3A=wait longer AT house
bt'ɔŋ kuma-kuman, dom yup"
be.afraid.of RDP-germ AFF blanket
"Ah, no," she said, "(if I) don't burn (it), he isn't going to wait any longer at the house, (for) he is afraid of the germs, indeed, (of) the blanket (itself)."

In some clauses there is evidence of 'double' components, where both the conditional and the inferential (see §10.2.2.1) are present.

- (106) *In former times the sultan's men would travel upstream to collect taxes from the Semelai. They understood that if they hadn't paid the tax, they would certainly have been killed. If they did pay, then perhaps they would not be killed.*
- məhn da? de=jon barang, ga=de=bunuh. ma=jon, beh
if NEG 3plA=give thing IMM=3plA=kill IRR=give NO
If they didn't give anything, they were going to kill (them). (If they) gave, on the contrary (they wouldn't be killed).

Another example of this is attested in (107). Here the irrealis is an inferential affective state predicate used with second person reference in a conditional construction (see §10.2.2.1).

- (107) *A man cannot be bothered taking his wife into the forest to search for Salacca fruit. He instructs her on her options:*
- "ma=kali, swak, da? ma=kali, p?st" k'laq
IRR=be.brave go NEG IRR=be.brave stay QUOTE
"(If you) are brave, go, (if you) aren't brave, stay," (he) said.

Clauses containing the irrealis are frequently employed in the vicinity of negative imperatives and denials of permission.

- (108) *A child is playing with a bunch of keys on the verandah. His mother warns him against dropping them as she does not want to have to bother with retrieving them from below.*
- bɔy ma=g<n>rek! susah ma=yk-yɔk
NEG:IMP IRR=fall<CAUS> be.difficult IRR=IMPERF-take
Don't drop (them)! (It) is difficult (for me) if (I) have to keep on retrieving (them).

10.2.3 Deontic modality

One of the features of the irrealis is its occurrence in clauses associated with statements expressing deontic modality. The irrealis is used in the following four instances:

- to convey negative transitive imperatives;
- to deny permission;
- to state prohibitions;
- to express lack of obligation.

All cases involve a transitive verb embedded under a negator, either the imperative negator *bɔy* 'NEG:IMP' (§10.2.3.1), or the declarative negator *da?* 'NEG' and the modal or matrix verb (§§10.2.3.2–3).

Negative imperatives, denials of permission and statements of obligation should be regarded as representing advisable or precautionary actions, which if followed will prevent an adverse consequence affecting the addressee. This act has to be balanced against the belief that fulfilment of desire is tantamount to maintaining good health; wanting puts a person at risk of falling ill and illness is associated with soul loss. As a result, there is an apprehension on the speaker's part toward preventing someone from what they intend to do, or are in fact doing, for the speaker cannot judge the inner state of the addressee. This argument leads back to the epistemic function of *ma=* 'IRR' in the earlier discussion in §10.2.2. In all four types, the speaker attempts to move an addressee to action by imposing a condition of obligation. It must be understood, as expressed in §10.6, that in terms of the organisation of Sembalai society the authority to command is relatively weak. A brief introduction of each type is provided below, but the main discussion is found in §10.6.

10.2.3.1 The negative transitive imperative

The clitic *ma=* 'IRR' is used to form the familiar form of the negative imperative of the active transitive verb. The irrealis is not used when the directive is aimed at someone with whom the speaker is not on familiar terms. A full discussion of the imperative can be found in §10.6.3.2.

- (109) *A pregnant woman, trying to satisfy her cravings, approaches a fruit-laden Salacca tree in the jungle. Salacca fruits are a favourite of tigers, and unbeknown to the woman the owner of the tree is nearby. She hears a voice:*
- "*bɔy ma=pŋke? klibi? ke!"*
 NEG:IMP IRR=pick Salacca.conferta that
 "Don't pick those Salacca (fruits)!"

It is important to note that once deontic force has been established, the irrealis marker can be replaced by the pronominal proclitic (§10.6.3.3).

10.2.3.2 Prohibition/lack of obligation

The internal negator *da?* 'NEG' is used in the following contexts where a denial of permission or prohibition is stated. In cases expressing denial of permission, the clause containing the irrealis is embedded under a negated matrix verb. The remaining cases involve a modal preverb and negation of the matrix clause. These constructions are discussed under Complementation in §11.2.2.3.

- i. *Denial of permission* Denial of permission is expressed by the negation of the matrix verb *jon* 'to allow, permit':

- (110) *da? n=jon ma=yɔk creh*
 NEG IfA=permit IRR=fetch fish
 I did not permit that you would collect fish.

Note that in giving a positive directive, e.g. using the verb *?ur* 'to instruct', the complement is in the realis and not the irrealis;

- (111) *The narrator relates how in earlier times, on account of the ghosts' bad behaviour, the elders disallowed the tradition of leaving the corpse in the abandoned house to be smoked. They directed that the dead be buried instead.*

da? de=jon wɔ? ma=jɔy deg ke, de=?ur kəm
 NEG 3plA=permit longer IRR=do manner that 3plA=direct bury
 (So the elders) they wouldn't permit that (the people) do it like that anymore, (and) they directed (them) (to) bury (the dead).

- ii. *Prohibitive* The modal auxiliary *sot* 'permit, can, able' (§5.6.2) expresses ability or permission in terms of the possibility of carrying out an action. Negated predicates containing *sot* are used to state general prohibitions, particularly in relation to general precautionary statements regarding the consequence of the transgression of taboos. The transgression of a taboo generally results in adversity for the transgressor. There are few situations where it has implications for the transgressor's family or wider community, other than in the case of first degree incest which results in flooding. Thus the obligation to abide by the taboo is weak, but again it is up to the individual. *sot* is glossed as 'permit', but translated in this context as 'should' to express weak obligation. General prohibitions are always presented as impersonal expressions as in the following examples. The consequence is always stated in the realis.

- (112) *The retribution for wilfully disturbing or drawing attention to the formation of scabs behind the ear.*
- da? sot ma=pɔr?, ki=cɔŋ ʔen təŋ*
 NEG permit IRR=draw.attention.to 3A=burn LOC ear
 (One) should not draw attention to (the crow), it (will) burn one's ear.

The irrealis is also used in the formation of the classificatory names of the act of the various taboos as seen in the second instance of *ma=* 'IRR' in (113) (§10.2.2.1). This particular transgression pertains to the avoidance speech style.

- (113) *The speaker states the taboo:*
- da? sot ma=cəl. ma=sener*
 NEG permit IRR=utter IRR=forest.jargon.tease.by.allusion
 (One) should not utter (it). (It would be) the avoidance speech style (taboo).

- iii. *Lack of obligation* The following constructions expressing a lack of obligation use the irrealis in conjunction with idiomatic expressions borrowed from Malay, which in Sembalai function like auxiliary verbs: *da? ?usah* 'never mind; don't', *da? payah* 'don't feel obliged to'.

- (114) *ki=psupjt ʔen bapa?*
 3A=avoid LOC father
 "bah," kʰlɛŋ "da? ?usah ma=psupjt"
 NO QUOTE NEG be.necessary IRR=avoid
 She avoided (her) father. "No," (she) thought, "it's not necessary that (he) be avoided."

- (115) *ga=ki=roc, "tabe? hne" kʰlɛŋ*
 IMM=3A=pull.out excuse THEN QUOTE

"bah" k^hləŋ=ha "da? payah ma=p-tabe?"
 NO QUOTE=3POSS NEG feel.obliged IRR-CAUS-excuse
 He was going to pull (it) out. "Excuse (me)" (he) said.
 "No!" he replied, "don't feel obliged to excuse yourself."

10.3 Adverbial modifiers of the basic clause

Adverbs are defined as "modifiers of constituents other than nouns" (Schachter 1985: 20), although as will be seen in this section, some of these forms in Semelai are also capable of modifying nouns with similar semantic consequences (§10.3.1.3). They are discussed here rather than separately in §7.

The class of adverbs is highly heterogeneous, both semantically and grammatically, and has perhaps the highest frequency of Malay loans. There is a small closed class of lexical adverbs, otherwise adverbial meanings are expressed by nominal adjuncts (§10.2), verbal adjuncts in the form of stative verbs lacking any distinguishing morphological marking (§10.3.3), and modifying serial verb constructions (§11.3).

The semantic categories involved can be categorised by the grammatical means of expressing the adverbial function: lexical adverbs express aspect and quantification; nominal adjuncts express time and similitude; verbal adjuncts and serial verb constructions predominantly express manner.

Peripheral members of the adverb class tend to have more restricted distribution, and variation in distribution is not attested as resulting in a change in scope or meaning. The true lexical adverbs, on the other hand, often have greater possibilities of distribution which is reflected in a change in scope or meaning.

10.3.1 Lexical adverbs

Lexical adverbs express aspect and quantification in terms of replay and extent.

10.3.1.1 Aspectual adverbs

These are loc 'already', ɻareh 'just, recently' and la giʔ 'still'.

i. **loc 'already'** The aspectual adverb loc 'already' defines an event as taking place prior to the reference time. The essential point about loc 'already' is that situations which occur prior to the reference time do not automatically require its use.

The correlation between position of the constituent and its scope in the clause is unclear. Unlike other aspectuals, loc 'already' may precede or follow the predicate apparently without any accompanying semantic change.

One factor present in the use of loc 'already' is that the event that it marks is still relevant at the moment of discourse. It does not merely mark the sequential nature of events, nor does it mark consequence. This is seen in (116) where the people have already left the house. The thief then comes and picks the lock while they are out. The aspectual only has scope over the first clause.

- (116) loc sma? swak, ki=goy bsi? cokel pintu?
 already people go 3A=bring metal pick.at door
 The people already gone, he took a piece of metal (and) picked the door.

- (117) A man abandoned his wife several years back to go cock-fighting. He returns home
 only to discover that she has remarried.
 kna? dɔs he? ke, loc b-knɛk kmpən
 when arrive AT:above there already HAVE-husband wife
 Alas, when (he) arrived up there, his wife had already (re)married.

The interpretation of loc 'already' differs with active and stative verbs. In a clause with an active verb in the unmarked form, the presence of loc indicates that the task has been completed in its entirety, and this took place prior to reference time. In (118)–(119) the speaker uses the scenario of planting out a swidden.

- (118) loc ptəm
 already plant
 (They have) already planted.

In a clause with an imperfective form of an active verb, a portion of the task has been finished for the time being, prior to the reference time (119). The remainder of the task will be completed at a future point in time.

- (119) loc p<m>təm
 already plant<IMPERF>
 (They have) already been planting.

In a clause with a stative verb, loc 'already' describes a change of state that came into existence prior to the point of reference, but is simultaneous with it. In (120), a pre-existing state which was arrived at in the past is still current at the moment of discourse. Note that loc 'already' follows the verb in this example.

- (120) sma? pok rbana?,
 people beat drum
 sampay b-?b-?bs loc, ja<>re? t^hi
 until MID-INTNS-be.swollen already digit<HAVE> hand
 People beat the drums until their fingers were really swollen.

- (121) The protagonist refuses to eat the fish prepared for him, explaining:
 ɻəp loc jnuh ca crəh
 If already be.surfeited eat fish
 I'm already sick and tired (of) eating fish.

In the non-verbal clause loc 'already' indicates that a current state came into existence prior to the reference time, but without the change of state implication as for stative predicates.

- (122) dak cin da? da? wɔ? loc
 water be.cooked NEG EXIST more already
 There's already no more boiled water.

loc 'already' can simultaneously precede and follow the predicate, in order to emphasise the completion of a state. This distribution is not attested for the other aspectuals. In (123), the change in the state of the husband's feet was complete prior to the wife entering the room.

- (123) *The wife is taken aback:*

loc	jɔj=hn	sot	pəs	tijɔ	loc!
already	foot=3POSS	transform,into	tail	snake	already
Already his feet had transformed into a snake's tail!					

loc 'already' may also stand on its own:

- (124) *Dinner is cooked and ready to be served:*

"loc,"	kMəj	kəhn,	"ca=cə?!"
already	QUOTE	3S	eat=IMP,
"Done," she said. "eat!"			

An enumerated temporal NP modified by loc 'already' gives the meaning 'ago, past' as in the following example:

- (125) hmpe? ʔare? loc
three day already
three days ago

The aspectual loc 'already' can host the relative marker mə= 'REL', to give mə=loc (REL=already) 'the prior/previous one':

- (126) ʔəb, mə=loc deg ke
oh REL=already like that
Oh, the previous ones (were) like that.

it. ʔareh 'just, recently' ʔareh is an adjective 'to be new':

- (127) ga=poy hatap ʔareh
IMM=make thatch be.new
(We're) going to make new thatch.

ʔareh has three functions: a) as an aspectual adverb meaning 'just, recently'; b) in conjunction with a predicate containing the imminent aspectual clitic ga= 'IMM' a prospective meaning 'just going to'; c) as a clausal/discourse connective profiling a new state of affairs and glossed as 'at last, only then', 'from that point on', see §11.5.3.3.

Most of the examples in this section are given in context, for without the context it is difficult to ascertain the true senses of ʔareh. Neither a) nor b) occur in negated clauses.

a) Retrospective ʔareh 'just, recently' As a predicate internal modifier with the retrospective meaning 'just, recently', ʔareh always precedes the predicate.

- (128) *The narrator has recently given birth and is in confinement. The placing of a spoon into a cooking pot by a new mother is forbidden at this time.*

ye	ʔareh	b-knən,	da?	sot	tj-təŋ	ye
1	just	HAVE-give.birth	NEG	permit	IMPERF-dish.out	1
I had recently given birth, (and so) I was not allowed to dish out (food).						

- (129) *The speaker declines the invitation to eat, stating that she has recently eaten:*

ʔareh	ca
just eat (I've) just eaten.	

b) Prospective ʔareh ga= 'just going to ...' When the adverbial ʔareh 'just, recently' modifies a predicate with the aspectual proclitic ga= 'IMM' (§5.5.3), it marks a point just prior to the beginning of the event. Unlike other uses of ʔareh 'just' where the event has taken place, the presence of ga= 'IMM' indicates that the event is 'arrested' and does not proceed at that precise moment. In (173), which was elicited, the speaker describes how his hand had been raised about to perform the action of opening the tin when I turned up.

- (130) ʔareh ga=tbok saden ji dos
just IMM=open.tin sardine 2 come
(I) was just about to open the (tin of) sardines (and) you came.

- (131) co=hn ʔareh ga=lintaq jalan
dog=3POSS just IMM=cross track
His dog was just about to cross the track.

iii. *lagi?* 'still', 'to V on'. The adverb lagi? in pre-verbal position means 'still', 'to keep on' with active verbs, and with stative verbs means 'more' in the sense of 'increasingly'. In post-verbal position it expresses exact replay 'again' (§10.3.1.2). lagi? also functions as a degree adverbial 'more' (§10.3.1.5), and modifies negated predicates, internal da? 'NEG' and external beh 'NO' (§10.4). In this section the pre-predicate use of lagi? 'still', 'to keep on' is discussed. When lagi? precedes a clause negated with da? 'NEG', the result is a temporal clause connective 'before ...', where it defines a prior existing state (§11.5.3.3). lagi? is a borrowing from the Malay *lagi* 'still' (Mintz 1994: 413).

As a pre-predicate modifier, lagi? functions to denote the continuation of an action or state which remains unchanged, 'to keep on doing/being something'. Both the participants and the verbal action are the same and the action proceeds without interruption:

- (132) *A woman is eating salacca fruit, which are eaten together with salt.*
lagi? ki=ca, sampay k<np>məŋ gərəm mə=c<ph>?um
still 3A=eat until finish<NMZ> salt one=wrap<NMZ>
She kept on eating, until finishing one package of salt.

- (133) lagi? msəŋ ʔare?
still five day
(There's) still five days left.

In a negated clause, the anticipated event has still not taken place (134), or an anticipated state has not been reached as in (135).

- (134) lagi? da? dos
still NEG come
(He) still has not come.

- (135) sma? k^bet, lagi? da? dos hmpon bulan.
person be.pregnant still NEG reach four month
da? sot cuk simply
NEG permit weave band
A pregnant person who hasn't yet reached four months, should not weave bands.

iv. Aspectual modifiers of negated predicates The internal negator *da?* 'NEG' and the external negator *beh* 'NO', form complex negators by placing an adverbial modifier in post-verbal position to mark extent. These are discussed in full in negation (§10.4.1).

The adverbials are *lagi?* 'still, yet', the otherwise unattested *wɔ?* 'longer' and the modifying serial verb *cəŋ* 'completely'.

beh ... lagi? 'not ... yet' and *beh ... cəŋ* 'not at all' form responses to questions (§10.4.1).

- | | | | |
|---------------|--------------------|---------------|------------------|
| da? ... lagi? | 'not ... yet' | beh ... lagi? | 'not ... yet' |
| da? ... cəŋ | 'not ... at all' | beh...cəŋ | 'not ... at all' |
| da? ... wɔ? | 'not ... any-more' | | |

Post-predicate *lagi?* has the sense 'again' (§10.3.1.2), however in this negated clause type it has the sense usually associated with its premodifying function, namely 'still' or 'yet'. The activity has not yet commenced:

- (136) da? p<m>təm lagi?
NEG plant<IMPERF> yet
He isn't planting yet.

wɔ? 'any-more' or 'longer' is only attested in this construction with the internal negator *da?* 'NEG', and is not attributed with any independent meaning.⁴

- (137) da? zyot wɔ? ha? kampog deh
NEG return longer AT village 3pl
(They) didn't return to their village anymore.

- (138) da? ziem wɔ? ca
NEG be.pleasing longer eat
(It) isn't tasty to eat anymore.

In non-negated clauses, *cəŋ* functions as a supplemental V₂ verb in a serial construction with the sense of 'completely' or 'all at once' (§11.3.2.1); in conjunction with a negator it means 'not ... at all':

- (139) da? ki=k^bom cəŋ
NEG 3A=obtain at.all
She didn't get (any) at all.

10.3.1.2 Adverbs of replay

Three adverbs, *lagi?* 'again', *pacə?* 'likewise' and *wər* 'too' can be characterised as expressing the repetition, or replay of an event. Three types of replay are identified, which correlate with the three Semelai terms above:

- i. *Exact replay* The same verbal action is performed by the same participants. This is expressed by post-verbal *lagi?*, glossed here as 'again'.

Types ii) and iii), known as 'token replacement' (Evans 1995: 239), represent only 'partial' replay, where the same verbal action is performed by different participants. The form depends on which tokens are replaced – A, S, O, or something else.

- ii. *Absolutive replacement* *pacə?* 'likewise' indicates that either a different S is carrying out the same or similar verbal action, or a different O is affected, or the state is applicable to a different subject in a non-verbal clause. In a second, less frequently attested usage, *pacə?* expresses 'event' replacement, where the participant is the same but the verbal action is different. The verbal actions however are linked, given that they constitute subparts of a complex event.

- iii. *Accusative replacement* The adverb *wər* 'too' signals the replacement of A, S, or any non-core constituent in the clause. It may also function to indicate event replacement.

Before proceeding, it is necessary to point out what is considered to constitute the 'same' event. The use of a term 'again' may only express a subset of similarities. As has been shown above, only part of the event needs to be repeated, the same participants, different participants, or a different or differing verbal action performed by the same participants. These points are expanded in the treatment of each form below.

- i. *Replay lagi? 'again'* *lagi?* 'again' occurs in post-verbal position. It represents a complete replay: the same verbal action executed by the same participants. However, like other types of replay, it represents a new stage of the verbal action as a new discrete, non-continuative event. This is in contrast to polysemous pre-verbal *lagi?* which expresses continuity or a durative action (§10.3.1.1). This is shown in the following examples.

In intransitive (140), transitive (141) and serial verb clauses (142), the adverb occurs immediately following the verb or verb complex.

⁴ In the Southern Austronesian language Muh Meri, there is an aspectual adverb *wɔ?* 'not yet' (Kruspe in prep. a.).

- (140) A woman, fleeing through the forest, meets up with different animals. She begs for assistance, but when the animal complies, it promptly dies. Each meeting is punctuated by the woman running on again, or alternatively, the woman's running is punctuated by these meetings.

brahluh lag? kəh
run again 3
She ran (on) again.

In (141), taken from a procedural text, the action must be performed again. The rice is pounded once and then winnowed. After winnowing, it is returned to the mortar to be pounded again.

- (141) ?luc.tom br-gum, buh lag? ha? lson
after MID-winnow put again AT mortar
After being winnowed, (it's) put in the mortar again.

- (142) Someone comments on the current phase of the moon:

?yot t'ey lag?
revert be.big again
(The moon) is waxing.

ii. *pacə?* 'likewise' This adverb occurs in clause-final position. Speakers translate *pacə?* with *samə?* from the Malay *sama* 'same', so it is glossed here as 'likewise' to reflect this. Etymologically, *pacə?* appears to be a combination of the two discourse enclitics =pa 'CL' and =ə? 'EM'. It is sometimes abbreviated to =pa 'CL'.

There are two meanings which were described in the introduction: absolute replay and event replacement.

Absolute replay This indicates the replay of an action, where the 'same' action is repeated on different objects in O-replacement (143), different agents replay the same action in S-replacement (144), or the same attribute pertains to replacement persons or entities (145).

- (143) ki=wes pinaj, ple hmbacan ki=wes pacə?
3A=peel betel.nut fruit *Mangifera foetida* 3A=peel likewise
She peeled the betel nut, (and) the horse mango, she peeled (it) likewise.

Example (144) illustrates S-replacement.

- (144) he p?st pacə?, p?st btol
1&2 stay likewise stay be.true
We'll stay too, stay for real.

In (145), a discussion is taking place about various ethnic groups at the time of the Communist Emergency. Someone responds that all the people had the same place of origin, the Chinese as well. This is an example of subject-replacement in a non-verbal clause.

- (145) Q: tom ha? hōn cina?
SRC AT where chinese.person

- A: cina? ha? malesye pacə?
chinese AT Malaysia too
Q: Chinese from where?
A: The Chinese (were) likewise from Malaysia.

Event replacement with *pacə?* 'too', is less frequently attested. It expresses a number of related but different actions which are performed as subparts of one event. This is illustrated in (146) where the event is 'someone sets up a camp', and the sub-events imply some sort of repetition of required activity, e.g. clear the area, make a shelter, fetch water, light a fire.

- (146) ki=pas?us pacə?
3A=light.fire too
He lit a fire too.

Referring back to example (143) above, this reading is understandable if the verbal action is characterised as 'remove the inedible outer layer of the fruit in order to consume it'. The woman peels a betel nut and a mango, both requiring different methods of peeling and different implements, clippers for the former and a knife for the latter.

iii. *wər* 'instead, rather' The simplest translation of *wər* is 'instead' or 'rather'. It is used for the repetition of a verbal action where an element is replaced – either an argument, or a verbal component. The replayed action is not simultaneous with the original action of the other person, but usually follows it. Often the desire of the second person to replay the event is based on a desire to emulate the other person, thus it is translated as 'X wants to V in Y's stead'. *wər* may be pre- or post-predicate. The correlation between position of the constituent and its scope in the clause is unclear, although the lexeme generally appears to directly follow the constituent it modifies.

In (147) *wər* carries the meaning mentioned above, 'X wants to V in Y's stead'. This is an example of S-replacement.

- (147) An elephant fights a ghost with the intention of killing him. Instead of killing him, the elephant dies.
pmalej wər k'bəs
elephant instead die
The elephant died instead.

In (148) the A is replaced.

- (148) A blind man adeptly takes over the task of cutting the cassava from his sighted wife.
ki=t'kəl la=knlək btol wər
3A=cut A=husband be.correct instead
The husband cut (it) correctly in her stead.

The indirect object can also be the target of replacement:

- (149) jon ha? hn=yə wər!
give AT ABS=1 instead
Give (it) to me instead!

In (150)–(151) a change in the course of action is expressed and it is the verbal argument which is replaced.

- (150) "beh, da? b-tojoy," k^bləŋ, b^bilak wər
NO NEG MID-wait.for QUOTE get.out.of.way instead
"No, (it) won't be waited (for)," (he) said, (and he) got out of the way instead.
- (151) A recently exorcised person replaces his old previously afflicted self:
kəh pade.hal, ki=b^bɔŋ wər p^bre?
3 in fact 3A=fear instead ghost
p^bluc.tom ke, ki=?yot b^bɔŋ p^bre?
after that 3A=revert fear ghost
He in fact, he feared spectre huntsmen instead. From then on he reverted to fearing (them).
- (152) A change in the site of the illness:
təŋ da? 2ŋih wə?, g<ŋŋ>log wər sat
ear NEG be.sore longer swallow<NMZ> instead have.laryngitis
(My) ear isn't sore anymore, (my) throat instead (has) laryngitis.

10.3.1.3 Restrictive adverbs

The restrictive adverbs *?en* 'just, nothing but' and *nehneh* 'just, only' modify both VP and NP constituents with meanings which are clearly linked. Hence both functions will be discussed in this section. A third restrictive adverb, *sdom* 'only', which is used specifically in relation to number, is also discussed below.

i. *?en* 'just, nothing but' *?en* functions as a post-verbal modifier that restricts the event to the single action described by the verb. It describes an action which is self-contained and performed specifically for that purpose, and is defined as 'just, nothing but'. As a nominal modifier it has the meaning 'merely, nothing more than'.

a) VP modifier The act is a discrete event; no other activity is engaged in. This is in contrast to *nehneh* below which means 'nothing but (V)'. The modifier always occurs in final position in the VP. A selection of contextualised examples clarifies the above definition.

- (153) The speaker explains how he only knows how to make a bamboo flute, not play it.
j<y>?y ?en mə=yɛ=k^be?
make<IMPERF> just REL=1A=know
Just (how) to make (it), is what I know.
- (154) The children just cook the food but, given the circumstances, not for the purpose of eating.
de=sayor ?en. da? de=ca=hn, la=sac bapa? deh
3plA=stew just NEG 3plA=eat=O BCS=flesh father 3pl
They just stewed (it). They didn't eat it because (it) was their father's flesh.
- (155) A man and his dog leave the house, but instead of going off, they do the following:
deh p^bot p^bgl? ?en
3pl stay lie.in.wait just
They just stayed lying in wait.

In (156) *?en* is further modified by the borrowed Malay equivalent *saja?* 'just, only' to add emphasis.

- (156) g<r>hop ?en saja?
hot<CAUS> just only
(She) just heated (it) up that's all.
- b) NP modifier As a post-nominal modifier *?en* means 'merely' or 'just' in the sense of 'nothing more than', signifying an object which has no intended use or function. This is in contrast to *nehneh* below which means 'nothing but'.

In (157), the water collected is not destined for a specific use, e.g. water stored for the household which may be used for washing, cooking or drinking. Alternatively, it may be used to refer to plain water as opposed to boiled drinking water or tea.

- (157) dak ?en
water just
just water

In (158) the speaker describes a piece of incense which has not had a specific healing incantation spoken over it.

- (158) kmyan ?en, da? b-jampi?
Styrax.benzoin just NEG HAVE-spell
(It's) just (ordinary) benzoin, without a spell.
- (159) The speaker is explaining that unlike other words, there are no dysphemisms (*kasar*) or euphemisms (*?alus*) for the word 'ghost'.
kmuc ?en, kasar ?alus da? da?
ghost just be.coarse be.refined NEG HAVE
(The word is) just 'ghost', there aren't dysphemistic or euphemistic (terms).

ii. *nehneh* 'just, only' As a modifier of the VP, *nehneh* means 'just, only', to engage in one activity to the exclusion of all others, 'to do nothing but V'. *nehneh* also functions as a nominal modifier with a very similar meaning.

a) VP modifier *nehneh* has an exclusive reading. It directly follows the verb phrase (160)–(161) or non-verbal predicator (162).

- (160) "g^b2-go?" nehneh, k^bləŋ
IMPERF-fell.tree only QUOTE
"(I'm) doing nothing but felling (trees)," (he) said.
- (161) The pregnant woman expresses her craving for the Salacca fruit.
ki=lən nehneh la=iyor
3A=want only A=saliva
(My) saliva wants nothing but (Salacca fruits).
- (162) barəŋ ?a=deŋ da? nehneh
thing DET=like EXIST only
There was just everything.

b) NP modifier As a post-nominal modifier *nehneh* has the meaning 'just, only', in the sense of 'nothing but ...'.

- (163) luka? nehneh, la=kli=ckit kmpen
wound just BCS=3A=pinch.off.bits wife
(There was) nothing but wounds because (his) wife had pinched bits off (him).

- (164) kira?, b-knən pə? bradi?, ptre? nehneh...
figure HAVE-offspring three sibling prince just
ki=p-bwai sma? "beh, ?pat manda=sbn, lasa?,
3A=CAUS-converse person NO be.ugly what.sort=SC boil
kudes nehneh, kuman," yŋ?, lən b-kmpen
skin.disease just germs but want HAVE-wife
You see, three siblings were born, all of them princes ... People said behind their backs, "No, how ugly, nothing but boils, skin diseases (and) germs." But they wanted to marry.

The following clause is ambiguous between a nominal modifier, or a verbal modifier, as illustrated in the translations:

- (165) nɔʔ=hn beh, pakay dɔl papan nehneh
this=CONN NO use house board only
Now no, they use houses (made of machined) boards only.
Now no, they only use houses (made of machined) boards.

iii. *sdom* 'only' *sdom* is a restrictive modifier which functions to modify VPs and NPs (220) expressing number. It means 'only' as in 'no more than the number stated'. It has varying distribution: as a verbal modifier it post-modifies (166)–(167); as an NP modifier it pre-modifies (168)–(169).

- (166) p?·pe? sdom
BE-three only
They were only a group of three.

- (167) sta? dos pon ?ikur, dwa? kdor, dwa? rmol sdom,
after reach four CLF two be.female two be.male only
sinap kke. da? b-knən wɔ?
stop there NEG HAVE-offspring more
After four (of) them arrived, only two female (and) two male, (she) stopped there.
(She) didn't give birth anymore.

- (168) sdom gorəm mə=b<2>l?
- only salt buy<IMPREF>
(It was) only salt that was bought.

- (169) knən deh, sdom mə=?ikur, tet ke
child 3pl only one=CLF TO:spec there
Their child the only one, (went) there.

10.3.1.4 Temporal adverbs

This group of adverbs refers to progressions and sequences rather than a period or duration of time (§10.3.2.1). Members of this group tend to occur in initial position, either clause or predicate initial. Variation in position may result in variation in scope, but not meaning. Some of these are lexical adverbs; others have functions as other parts of speech, e.g. verbs.

soloj	'initially, first'	← Malay	sulong 'first'
?le?le?	'eventually'	← ?le?	'to be long (of duration of time)'
c?co?	'first (in a series)'	← c?co?	'to be in front'
crner	'after, later'	← crner	'to be behind'
nəŋ ~ knəŋ	'before, earlier'		
hne	'soon; then'		
pon	'thereupon, then'	← Malay	pon 'emphatic particle'

The first two temporal adverbs always precede the predicate:

- (170) This describes the stages of growth of a plant:

soloj jmə?, ?le?le? sot dloj
first vine eventually metamorphose.into tree
Initially (it) (is a) vine, eventually (it) turns into a tree.

nəŋ 'before' is a retrospective adverb, expressing recency of a point in time prior to the speech moment. It can occur in either pre- or post-predicate position. It cannot co-occur with the retrospective aspectual adverb ?areh 'just' (§10.3.1.2).

- (171) A: bila?
when
A: When?
B: knəŋ, som nəŋ
before morning before
B: Earlier, this morning (just past)

- (172) nəŋ ga=kı=ptəm
before IMM=3A=plant
(She) was going to plant (them) just before.

In (173), the adverb is in post-predicate position. The second clause is marked with nəŋ 'before', marking it closer in relation to the moment of speech. The event in the first clause was prior to that of the second.

- (173) mə=j?l? ki=ca he? bri,
REL=be.raw 3A=eat AT:above forest
mə=cin ki=ca sayor nəŋ
REL=be.cooked 3A=eat stew.with.vegetable before
The raw he (had) eaten up in the forest, the cooked he ate stewed just before.

The adverbial use of c?co? 'to be in front' and crner 'to be behind' is always associated with a series of identical events involving a series of participants. The adverb follows the constituent under its scope; in (174) the adverb is in pre-verbal position and in (175) it is post-verbal.

- (174) kadeh c?co? 2ye, kmpən kah, kaka? deh kah?
who first see wife Q EZ 3pl Q
Who (was) first to see, the wife, or the elder sister?
- (175) A shaman is attempting to retrieve two people, a man and a woman. He tackles the female first.
ki=?antat c?co? me=k dor
3A=take first REL=be.female
He took the female first.

In imperative clauses the adverb is clause final, following the vocative, if present:

- (176) The speaker urges the addressee to eat now, for he, the speaker, will eat later:
sar ca je=2en c?co?
descend eat 2=AUG first
You go (down) (and) eat first!

The final two items to be discussed here are hne 'SOON, THEN' and pon 'THEREUPON, THEN', which are not so much temporal adverbs, as indicators of temporal sequence. hne occurs either in conjunction with the aspectual clitic ga='IMM', or without it, in which case its function appears to merge with that of pon.

hne 'SOON', indicates that the event will happen shortly, in the very near future. The adverb is glossed in small caps because the speakers were not able to provide a definition. Speakers generally agreed on a time of around two hours as the maximum time limit, which in most cases held true. It never co-occurs with ga='IMM' in a retrospective context.

- (177) bɔy swak! he hne ga=ca
NEG:IMP walk 1&2 SOON IMM=eat
Don't go (anywhere)! We're going to eat soon.

The position of hne is fixed, occurring immediately preceding the verbal word (§5.1). This is demonstrated in (178) where hne precedes ga=bpngi? (IMM=walk), and is itself preceded by pon 'THEREUPON'.

- (178) ?əŋi pon hne ga=bpngi?
If THEN SOON IMM=walk
I am going soon then.

hne, like ga='IMM', cannot be used alone, for instance in response to a question, or to indicate the status of an event.

In the following examples the function of hne seems more like a marker of consequence or assertion like pon 'THEREUPON' which is described below. It either precedes (178) or follows the predicate (179)–(180).

- (179) "bila? wen, col 'tik tak!'" kʰləŋ, "snek hne" kʰləŋ.
when throw utter bounce! QUOTE be.sweet THEN QUOTE
"When (I) throw (it and it) makes a noise 'tik tak!'" (he) said, "then (it) is sweet."
(he) said.

- (180) "ma=tlon hne" kʰləŋ, "kmɔŋ" kʰləŋ "tɔŋ"
IRR=help THEN QUOTE finish QUOTE hand
"If (you) were to hold (it) then," (he) said, "(your) hands would be finished."

pon 'thereupon, then; too' is a borrowing from Malay, *pun* 'also, too'. In Semelai it has two functions: a) a predicate modifier; b) an NP modifier. Essentially, it functions to indicate a successive focus of attention to a sequence of either events or entities. The two functions are discussed separately below.

a) pon functions as a predicate modifier with fixed pre-verbal location. It marks an expected sequential relationship amongst verbal actions, 'then, thereupon'. In the first example, it has been raining continuously for seven days and seven nights resulting in floods. The rain finally stops and the consequence, the subsiding of the floodwaters, is marked by pon:

- (181) ?are? prhnti?, dak pon ?yot sar
rain stop water THEN return descend
The rain stopped (and) thereupon the (flood)waters receded.

In (182) the husband's mission to collect the incantations has been completed and consequently he returns home as planned. The presence of pon emphasises or draws attention to the consequence, which is the result of a previous event.

- (182) sta? ki=kʰmampi?, kəhn pon ?yot
when 3A=obtain incantation 3S THEN return
After he obtained the incantations, he then returned (home).

b) pon is also used to modify an entity, with the meaning 'as well', 'too'. It follows the NP and suggests, as above, a successive focus of attention.

- (183) A woman who has given birth in the forest has only one piece of cloth, and calculates what she can make with it.
kayen me=hlay, baju?, kayen me=hlay, sy-suy pon
cloth one=CLF shirt cloth one=CLF USE-loincloth too
One piece of cloth (for) clothes. (and) one piece of cloth to wear a loincloth too.

- (184) A man has slain and butchered a tiger. Inside he finds the traces of his wife, both cooked and raw, which the tiger had eaten earlier.
ma=cin da?, me=jɔ? pon da?
REL=be.cooked EXIST REL=be.raw too EXIST
There was the cooked, (and) there was the raw, too.

10.3.1.5 Quantification

Under the general title of quantification adverbs degree and frequency are examined.

i. *Degree* Adverbs of degree modify verbs and verb phrases, other adverbs and NPs. They are generally of fixed distribution:

jɔ?	'very'
cukup	'enough' ← Malay
pala?	'enough' ← Malay pala 'enough'

These express the comparative degree:

- cmo? 'more, (increasingly)'
lagl? 'more, (comparative)'

jø? 'very' modifies verbs, both stative and active, with slightly different but related senses. It immediately follows the constituent over which it has scope:

- (185) "klan jø? ɻap" kɻəŋ
be.frightened very If QUOTE
"I (was) very frightened," (he) said.

In (186), the modifier is fronted to clause-initial position to provide emphasis.

- (186) "jø?" kɻəŋ "ji=sayaŋ cɔ"
very QUOTE 2A=feel.affection.for dog
"Very (much)," he said "you felt affection for the dogs."

With stative verbs the presence of jø? 'very' expresses the degree of intensity of a state:

- (187) cros ɻap jø?
claw If be.long very
My claws are very long.

- (188) kmpən pam jreh jø?
wife feel be.exhausted very
(His) wife felt very exhausted.

With active verbs jø? expresses degree in terms of frequency, and is translated as 'always'. The Semelai translation of this is selalu? ← Malay *selalu* 'always'.

- (189) gŋ-gŋ jø?
IMPERF-bite always
(It) always bites.

Example (189) may be paraphrased by the nominal par-gŋ (xs-bite) 'one which bites excessively' (§7.6.2.1).

- (190) boy ma=jon bɻiu? jø?
NEG:IMP IRR=give friend always
Don't always give (them) to friends!

cmo? means 'more' in the sense of increasingly. The standard of comparison is with a previous state of the S. This modifier only occurs with adjectives, which it always precedes:

- (191) cmo? tʰey
more be.big
increasingly big

- (192) cmo? lyar knkən kke
more be.wild child that
That child is increasingly wild.

An alternative to cmo? is the premodifying function of lagl? 'still, more':

- (193) cmo? mahal
more be.expensive
increasingly expensive
- (194) lagl? mahal
more be.expensive
increasingly expensive

The concept of 'enough' is expressed either by pala? or cukup, the latter being more frequently attested. Both are Malay borrowings (*pala* and *cukup*) and modify stative VPs (195)–(196) and NPs (197).

- (195) pala? bhth
sufficient be.full
sufficiently full
- (196) bhth jø? ɻap, ca sayor knkən, cukup ɻiem
be.full very If eat stew child enough be.pleasant
I'm very full (from) eating the children's stew. (It's) so tasty.
- (197) "nincə pala? ha? ke=cə?", kɻəŋ "na-ha?"
food sufficient AT there=EM QUOTE DEM-AT
"(There's) sufficient food there," (he) said "(with) these ones."

btol 'be true' ← Malay *betul* 'correct, true' expresses a value judgement, implying a sense of doing or experiencing something in the prototypical manner. In (198), there is a shortage of betel nuts (pinar). An onlooker remarks on the intensity of the addressee's search. He uses the verb mos 'to sniff out, of dogs', likening the person to a dog working frantically on the scent of an animal:

- (198) ji=mos btol ɻen pinar
2A=sniff.out be.true seek betel.nut
You are really sniffing about seeking betel nuts.

- (199) *The gossip is that someone recently widowed has remarried. Someone wryly comments:*
cpat btol=jə!
be.quick be.true=CL
(That was) really quick!

ii. Frequency Adverbs of frequency are few, with most borrowed from Malay. They occur in clause-initial position, with the exception of slalu? 'always' which can also follow the predicate. Some of these are listed here:

ten	'instance'
prnah	'ever'
da? prnah	(NEG ever) 'never yet'
kadaj, kadaj-kadaj	'sometimes'
cərə?	'always', 'at times, occasionally'

← Malay *pernah*
← Malay *tak pernah* (NEG ever)
← Malay *kadang/kadang kadang*

crom	'always'	
slalu?	'always'	← Malay selalu

Some examples follow:

- (200) mi=tən lagi?, btulok swak
one-time again set.off walk
Once again, (they) set off walking.

kadəŋ can be used disjunctively as in (201):

- (201) kadəŋ ?yot, kadəŋ beh
sometimes return sometimes NO
Sometimes (he) returned, sometimes not.
- (202) cərə? ki=gəp-i7
always 3A=bite-TR
It always bites.

Additional means of expressing frequency are: the iterative suffix -iʔ 'ITER' (§5.3.7); the nominalising prefix par- 'xs' expressing excessive agent (§7.6.2.1), and by modifying a nominal adjunct as in (205) below.

10.3.2 Nominal adjuncts

10.3.2.1 Time

Adverbs of time have scope over the whole clause, providing information about the point or period of time at which an event takes place. They occur in either clause-initial or clause-final position. Time adverbials range from the general, naʔdəy 'long ago', to the specific, ?are? nɔ? 'today'. An extensive but incomplete list of time adverbials follows:

soloŋ	'(in the) beginning'	← Malay sulong
naʔdəy	'(in the) distant past, long ago'	
dəzən	'(in the) recent past; within one's life-time'; 'the day before yesterday, or up to three previous days'	
sarek	'(in the) future'; 'tomorrow'	
bulan	'month, calendar or lunar'	← Malay bulan
tahun	'year'	← Malay tahun
?are? muy	'one day, (past or future)'	
nintom	'yesterday'	
?are? nɔ?	'today'	
ptəm nɔ?	'tonight'	
tahoy	'the day after tomorrow'	
tulst	'the day after the day-after-tomorrow'	← Malay tulat
bantin	'in four days time'	

* The range of meanings given here is defined by the context.

tpi	'day time'
ptəm	'night time'
som	'morning, from around 5 am to 11 am'
tjah ?are?	'midday, from 11 am to 2 pm'
ronoŋ	'around 2 pm'
doy	'afternoon, 2-5.30 pm'

'Time of day' adverbials have scope over the whole clause and can be either clause initial (203), or clause final (204) with no difference in meaning.

- (203) sarek da? ki=kʰe? wɔ? cakup smlay
future NEG 3A=know longer speak Semelai
(In the) future she won't know (how) to speak Semelai anymore.
- (204) ?yot ronoŋ
return mid.afternoon
(He) returned mid-afternoon.

Frequency is expressed by modification with a measure NP (§7.3.3):

- (205) tyap-tyap ?are? b-beŋ, yɔ?=hn buntar
every-every day MID-inspect but=CONN be.rotten
Daily (they) were inspected, but (they were always) rotten.

In order to express 'last' or 'next' week, month, or year, mə=?luc (REL=pass) 'the one passed', or mə=ga=dos (REL=IMM=come) 'the one coming' are used.

10.3.2.2 Similitude

There are two types of similitude, one expresses similarity to an NP, macəm NP 'like (NP)', the second expresses similarity of verbal action, V deŋ nɔ?/ke 'to V in this way, that way'. In (206), similitude of the NP 'hair', is expressed with 'macəm', and similitude of the verbal action is expressed with deŋ ke.

- (206) ki=?ulor hn=suk macəm dde=jəŋ deŋ ke
3A=let.down O=hair like OF=if way that
She let (her) hair down like mine (in) that manner.

i. *macəm* 'like, similar (to)' *macəm* expresses similitude, 'like, similar (to)'. It is a Malay borrowing, *macam* 'sort, type', but in Malay this function would be expressed with *seperti*, *sebagai*, or *serupa* 'like'. *macəm* is a nominal, it can be reduplicated as in *macəm-macəm* 'all kinds of things', and this can be prefixed with b- 'HAVE-' as in *b-macəm-macəm* 'to have/be all kinds of things'. The similitude adjunct is shown in brackets:

- (207) ki=jɔy b-ksum [macəm ksum klpkap]
3A=make HAVE-cocoon [like cocoon butterfly]
He made (himself) cocooned, like a butterfly's cocoon.

In (208), the predicate is nominalised.

- (208) "rumram" [macəm swara? bdli p<n>ito?]
bang.bang [like sound gun explode<NMZ>]
"Bang, bang," like the sound (of a) gun firing.

In (209) the complement is a PP *rom ʔəp* 'with me'.

- (209) *The dead sister down in the underworld urges her young brother, who has followed her there, to return home up to earth. She promises him that after seven days he will die and return to the underworld like her.*
 dəs ko he? ke, t̪impoh ʔare? [mbcom rom ʔəp]
 arrive 2f AT:above there seven day [like WITH 1f]
 (When) you arrive up there, (after) seven days (you will be) like with me.

ii) *dej no?* 'in this way, like this; *dej ke* 'in that way, like that', 'thus' These express similitude of verbal action. They occur in post-verbal position.

- (210) lən cokp bkrasch, poy dej no?
 want converse whisper do like this
 (If) (one) wants to talk (in a) whisper, (one) does like this.

- (211) *In the past, the ghosts' bad behaviour led the elders to forbid the tradition of leaving corpses in the mat racks of houses to be smoked. They directed that the dead be buried instead.*
 da? de=jon w? ma=j?oy dej ke. de=ʔur kəm
 NEG 3plA=allow longer IRR=do like that 3A=direct bury
 They wouldn't allow (the people) any longer to do it like that. They directed (them) to bury (the dead).

In (212) the verb is in the nominalised form.

- (212) ma=gps=sən, nm-k?om dej ke
 IRR=peel=SC NMZ-sit like that
 (It) was peeled, you know, sitting like that.

The adverbial is often used in conjunction with the modifier *pa-cə?* 'likewise':

- (213) ki=gə2, dom dej ke pacə?
 3A=fell.tree AFF like that likewise
 He felled the tree, indeed likewise like that.

10.3.2.3 Numerals and numeral expressions

Numerals and derived numeral expressions may function adverbially: *muy* 'one' functions adverbially with the meaning 'to do V together, in the manner of/as husband and wife' (214), or in relation to incest as in (215), where the actants are siblings.

- (214) de=ʔajak ca, ca muy rom kmpən
 3plA=invite eat eat one WITH wife
 They invited (him) to eat, eat together with his wife.

- (215) deh jtek muy, suwak muy
 3pl sleep one walk one
 They slept (together as) one, (and) went (about) (as) one.

Imperfective derivations of the numerals three to seven are used adverbially with the meaning 'to be in a group of X persons' (§7.6.2.1):

- (216) dəs t?ru?
 come BE-six
 They came (in) a group of six (persons).

10.3.2.4 Instrumental nominal adjuncts of orientation

To do something left- or right-handed, or to move in a direction projected from the body, a certain morphological form with the 'instrumental nominalising' infix <m> is required (§7.6.2):

- (217) ca s<mam>təm
 eat right<NMZ>
 eat right-handed

- (218) swak c<mas>res
 go rib<NMZ>
 (He's) moving sideways.

10.3.3 Verbal adjuncts

The most common means of expressing manner is for intransitive verbs, either stative or active, to function as an adverb of manner without any morphological change. This is a sequence of two verbs, where the second in the sequence is not simply predicated, but functions to modify the initial verb in terms of manner. Both verbs have the same S, but the S is never expressed for the second verb.

- (219) *The speaker tells how the cobra was moving alongside him, its head upright, as he rode on his motor-bike:*
 swak wh-woh
 move IMPERF-get.up
 (The cobra) was moving upright.

In the following example the subordinate verb is modified:

- (220) ye=k?e? reŋ mudəh
 1A=know seek be.easy
 I knew (how) to find (it) easily.

Adjectives can be reduplicated when used adverbially (§5.3.9). The reduplication signifies an intensive reading, indicating thoroughness or a high degree, 'really':

- (221) ki=brkas ʔiɔ-2ilək
 3A=tie.together INTNS-be.good
 She tied (them) together really well.

- (222) reŋ gli-glis!
 seek INTNS-be.thorough
 Search really thoroughly!

The verb *pəh* 'to become' co-occurs with the following reduplicated adjectives in an idiomatic construction expressing 'to become V': *panay* 'to be clever', *t̪ey* 'to be big', *jleŋ* 'to be long, tall', *ʔjəs* 'to be high', *dpeš* 'to be short', *gmuk* 'to be fat'.

- (223) pəh t^be-t^bay ?a=kəh=h^b?
become INTNS-be.big DET=3=ATTN
As for him, (he) grew bigger.

10.4 Negation

There are four basic negative operators in Semelai: the propositional negator *da?* 'NEG' (§10.4.1), the two metalinguistic negators *beh* 'NO' (10.4.2.1), *b?en* 'NOT' (10.4.2.2)⁶ and the imperative negator *b?oy* 'NEG:IMP'. Although the negators discussed here function in different domains, it is on account of their functional commonality that the first three are discussed together in this section. The imperative negator is discussed in (§10.6.3). Expressions of negative attitude are described in §13.2.

10.4.1 Internal negation *da?* 'NEG'

- *da?* 'NEG' is a descriptive or semantic propositional (truth-functional) negator.⁷ When used to negate active verbal predicates, it functions to negate an event; used with stative and non-verbal existential predicates, it functions to deny the existence of a state. *da?* 'NEG' is also used with certain adjectives to form antonyms negating constituents: *da?* ?ilok (NEG be.good) 'to be bad'.

da? 'NEG' is the only negator that must be used in conjunction with the predicate; it cannot stand alone as a minimal clause. Etymologically, *da?* 'NEG' appears to be a contraction of the Malay negator *tidak* (*tida?*), which in colloquial Malay is contracted to *dak* (Wilkinson 1927) or *tak* (*ta?*).

The following examples illustrate the use of *da?* 'NEG' in verbal clauses, and in non-verbal clauses containing either a nominal or an existential predicate. Examples (224)–(225) are intransitive and (226)–(227) are transitive.

- (224) da? ?yot kahn ke
NEG return 3S that
He didn't return.
- (225) da? kē-kēt lubuk
NEG INTNS-be.small open.water
(It was) not *small* (by any measure), the (stretch of) open water.
- (226) da? ki=j?oy sŋkalan
NEG 3A=make grinding.board
He didn't make a grinding board.

⁶ Propositional negation is also referred to as internal or descriptive negation; metalinguistic negation is referred to as pragmatic or external negation. The distinction between the two forms can be seen in the following examples from Frawley (1992: 431). The first is propositional and denies the truth of the proposition, the second is metalinguistic and rejects the assertability of the statement:

- a) Bill didn't paint the house.
- b) Bill didn't *paint* the house, he slapped it all over with cheap whitewash.

⁷ It is also referred to as 'sentential' or 'standard negation' (Payne 1985).

- (227) da? ji=sdər?
NEG 2A=remember
Don't you remember (it)?

The following examples are of negated non-verbal clauses. In (228) the clause has a nominal predicate and in (229) it is an existential clause:

- (228) b?en, da? musn̩j
NOT EXIST civet
No, (it's) not a civet.
- (229) da? da? mande-mandeh
NEG EXIST RDP-what
There wasn't anything.

As the examples above show, *da?* 'NEG' occupies predicate-initial position, preceding the verbal word, including modal auxiliaries (§5.6). The scope of the negation is over the whole predicate. The following examples illustrate the negator occurring before different predicates, containing the unrealis proclitic *ma=* 'IRR' in (230), an aspectual proclitic in an intonation interrogative in (231), a modal auxiliary in (232) and a serial verb in (233).

- (230) "da? ma=k^bc?", k^bləŋ
NEG IRR=know QUOTE
"I wouldn't know," (he) replied.
- (231) "ʔəŋ nɔ?", k^bləŋ, "da? ga=k^bp^br-luc?" k^bləŋ=hn
If this QUOTE NEG IMM=2fA=CAUS-pass QUOTE=3POSS
"Me here," (she) said, "Aren't you going to release (me)?" she asked.
- (232) da? məh de=reŋ
NEG want 3plA=seek
They didn't want to seek (it).
- (233) kmpən, da? ki=gəŋ swak
wife NEG 3A=bring walk
(His) wife, he didn't take walking.

The S of the intransitive clause in pre-verbal position always precedes the negator:

- (234) kahn da? ga=dɔs
3S NEG IMM=come
He won't be coming.

The negator can be preceded by any constituent which normally occurs in clause-initial position, e.g. a fronted argument as in (233) above, a time adverbial (235), the external negators *beh* 'NO' (236) and *b?en* 'NOT' (228), or a clausal or discourse connective.

- (235) sarek, da? ki=k^bc? wo? cbkəp smlay
future NEG 3A=know longer talk Semelai
In the future, she won't know the Semelai language anymore.

- (236) ?ah beh, da? ki=pakay la=nel
ah NO NEG 3A=accept A=[name]
But no, Nel didn't accept (that).

Internal negation is restricted by the aspect of the predicate. It may co-occur with the imperfective form (237), the perfective form (226) and (233), and with the imminent aspect clitic *ga*= 'IMM' (231) and (234).

- (237) ye da? j<y>?ɔy
I NEG make<IMPERF>
I haven't been making (it yet, but I intend to).

- (238) "beh," k^bəŋ, "da? n=ca"
NO QUOTE NEG IfA=eat
"On the contrary." (he) said, "I didn't eat (it)."

da? 'NEG' cannot coincide with the prospective expression *?arch ga*= 'just going to', or the prospective interpretation of *hne* 'soon'; nor does it co-occur with the various markers of past events: *?arch* 'recently', *i^bc* 'already' (239)–(240). This is consistent with the fact that *da?* 'NEG' expresses negation, and an event which is yet to happen, as in the case of clauses with prospective markers, cannot be negated. Likewise, an event that has already happened cannot be negated retrospectively.

- (239) *da? ?arch ga=swak (240) *i^bc da? ?yot
NEG just IMM=walk already NEG return
*(I) wasn't just about to go. *Already he hadn't returned.

There are three aspectual adverbs which are negated with *da?* 'NEG', giving rise to the following complex negators: *da? ... lag?* 'not ... yet', *da? ... wɔ?* 'not ... anymore/longer' and *da? ... cəŋ* 'not ... at all'.

lag? and *cəŋ* also co-occur with the metalinguistic negator *beh* 'NO' (§10.4.2.1), as *beh ... lag?* 'not ... yet' and *beh ... cəŋ* 'not ... at all'. The adverbial negators are discussed in terms of the aspectual component in §10.3.1.1.

The negator *da?* 'NEG' always immediately precedes the verbal word; the post-verbal aspectual modifier generally directly follows the verbal word.

- (241) da? b-ʔajkot wɔ? dak ...
NEG MID-carry longer water
(when) the water was no longer carried ...

(242) is the only example in my corpus where the O, knkon nɔ?nɔ? 'this child' precedes the post-verbal modifer, in this case *lag?* 'yet'.

- (242) "ʔah!" k^bəŋ deh, "da? k^bbəs=cə? knkon ?nɔ?nɔ? lag?"
ah QUOTE 3pl NEG die=EM child this yet
"Ah!" they said, "this child hasn't died yet."

The post-verbal modifier precedes the subordinate verb as illustrated in the denial of permission construction:

- (243) da? de=jon wɔ? swak
NEG 3A=allow longer walk
They didn't allow (us) to go any more.

The aspectual negators are discussed individually below.

i. *da? ... lag?* 'not ... yet' The pre-predicate aspectual modifier *lag?* means 'still', or 'yet'; post-predicate it means 'again' (§10.3.1):

- (244) lag? ki=jɔh
still 3A=drink
Still she drank.

In the negated clause type it takes the sense usually associated with premodifying *lag?*, hence 'not ... yet'.

- (245) da? swak lag?
NEG go yet
(He) hasn't gone yet.

ii. *da? ... wɔ?* 'no longer', or 'anymore' The combination of the negator *da?* 'NEG' and *wɔ?* 'longer' is translated as 'no longer' or 'anymore'.

- (246) "da? ʔəpi=tɔŋoy wɔ?," k^bəŋ ha? dɔl
NEG IfA=wait longer QUOTE AT house
"I'm not waiting at home any longer," (she) said.

- (247) da? ki=k^be? wɔ? jɔy prahə?
NEG 3A=know longer make canoe
He no longer knew (how) to make a dugout canoe.

iii. *da? ... cəŋ* 'not ... at all' *cəŋ* is an intransitive verb meaning 'to fit, pass through'. *cəŋ* as a modifying serial verb means 'completely' (§11.3.2.1):

- (248) ki=kamek cəŋ la=miskin
3A=carry.on.hip completely A=pauper
The pauper carried (it) all on his hip.

When accompanied by the negator *da?* 'NEG', it means 'not ... at all':

- (249) da? ma=ʔyəŋ cəŋ cakbp bɔn trək
NEG IRR=hear at.all speak bear [name]
To my surprise, (I) haven't heard Mischievous Bear speak at all.

10.4.2 Metalinguistic negation

Metalinguistic negation belongs to the pragmatic domain. Horn (1985: 121) defines it as:

... a device for objecting to a previous utterance on any grounds whatever – including its conventional or conversational implicata, its morphology, its style or register, or its phonetic realisation.

In the following discussion, the negation centres around implicata.

Sentence-external negation may be paraphrased as, 'The sentence S is not true' (Horn 1985: 143). It denies the truth of the statement S that has been either uttered or implied previously. It is used to REJECT S or its implicata (Horn 1985: 153).

In Semelai, the distinction between internal and external negation is formally marked by different morphemes, unlike English where 'not' serves both functions. A further distinction is made on the basis of the clause type: the negator beh 'NO' is associated primarily with verbal clauses (§10.4.2.1), and b?en 'NOT' with non-verbal clauses (§10.4.2.2).

An alternative analysis here is that beh 'NO' is simply an interjection 'no', and that interjections normally express external negation. For instance the English 'no' and 'not' exhibit comparable patterning. However, the following examples, where beh 'NO' functions alone without the internal negator da?, clearly show the distinction between the two forms.

- (250) Q: hūm knon? A: beh hūm, sar pk-pok
bathe offspring NO bathe, descend IMPERF-wash.clothes
Q: (Is) (your) child bathing? A: (He's) not bathing, he went down to wash (his) clothes.
- (251) Q: ji ga=tet bahaw? A: beh, tet triyan
2 IMM=TO:spec Bahau NO TO:spec Triang
Q: Are you going to Bahau? A: Not (Bahau), (I'm going) to Triang.

10.4.2.1 beh 'NO'

In its simplest use the metalinguistic negator beh 'NO' expresses denial of something stated or implied in a previous utterance, providing pragmatic contrast.

beh 'NO' co-occurs with verbal, nominal, prepositional and existential predicates, with the meaning 'on the contrary; but no'. It is not used in negated nominal, prepositional or possession clauses, in which case b?en (§10.4.2.2) is used. beh 'NO' occurs primarily in four environments:

- a) In verbal clauses whether affirmative (252), or internally negated with da? 'NEG' (253).

- (252) nɔ?hn beh, pakay dol papan nehneh
this=CONN NO use house board only
Now(adays) no, they only use houses (made of machined) boards.
- (253) Everyone is expecting the child to hand the sugar cane over, but he keeps on walking past the crowd:
beh, da? ki=jon la=knon kke
NO NEG 3A=give A=child that
On the contrary, that child didn't give (it to anyone).

- b) In non-verbal clauses where the predicate is either a non-negated nominal (254) or a prepositional phrase; or a negated (255) or non-negated existential predicate.

- (254) A p?re? 'ghost' is warning his mates not to eat his human wife. (This type of ghost usually feeds on the souls of humans.)
"kira?" k'lej, "kmpən ?əp, beh kmpən sma?"
figure QUOTE wife If NO wife person
"You see," (he) said, "(she) is my wife, not some human's wife."

- (255) A man wants a child, but his wife is yet to fall pregnant.
beh, da? da? knon
NO NEG EXIST offspring
On the contrary, there was no child.

- c) As a post-predicate tag in Alternative interrogatives (§10.5.1.2):

- (256) br-lən beh, ca hubi? na?ha??
MID-desire NO eat cassava DEM-AT
(Do you) want to eat this cassava here, or not?

- d) When encliticised by =hn 'CONN', beh functions as a clause connective 'not that ...' (§11.5.4):

- (257) beh=hn da? ma=kali=sən, da? de=jon
NO=CONN NEG IRR=be.brave=SC NEG 3plA=allow
Not that we weren't brave, see, they wouldn't allow (it).

Preceded by the internal negator da? 'NEG', beh functions as the clause connective da? beh 'it isn't because ...' (§11.5.4).

beh 'NO' can also stand alone as a minimal clause where the predicate is elided. In this function, it may be post-modified by the aspectual adverbs lagi? 'yet' and cəŋ 'at all'. This provides further criteria against the counter-analysis that beh is simply an interjection 'no', rather than a metalinguistic negator.

- (258) Q: ʔyot kəhn? A: beh lagi?
return 3S NO still
Q: (Has she) returned? A: Not yet.
- (259) Q: rəkok, da? A: beh cəŋ
cigarette EXIST NO at.all
Q: (Got) (any) cigarettes? A: No, (none) at all.

The responses in (258) and (259) could be paraphrased as 'It's not yet true that (she has returned)' and 'It's not the case (that there are any at all)'. Generally, beh 'NO' occurs in clause-initial position preceding the descriptive negator, if present. beh 'NO' may be preceded by the free pronominal as in (260), or by a temporal expression as in (261).

- (260) ye, beh, da? ga=dəs
1 NO NEG IMM=arrive
Me, no, (I'm) not going to come.

- (261) nɔ?hn beh, de=kəm
this=CONN NO 3plA=bury
Now(adays), in contrast, they bury (them).

In the following example beh is used in the response to the implication of the previous utterance, as illustrated in (262), where clause B negates clause A.

- (262) A: "instdi," kʰlaŋ, "n=nɪ<r>ka? rɔm hn=kw"
- must QUOTE IfA=marry<CAUS> WITH ABS=2f
- B: "ʔec beh" kʰlaŋ, "da? moh"
- oh NO QUOTE NEG want
- A: "I must," (he) said, "marry (her) to you."
- B: "But no," (he) replied, "(I) don't want to."

Some further contextual examples of the use of beh 'NO' with negated verbal clauses are provided below.

- (263) ki=paʔay swak ʔen boh. beh, da? ki=moh la=cɔ ke
- 3A=call go LOC elsewhere NO NEG 3A=want A=dog that
- He called (the dog) to go elsewhere. But no, the dog didn't want to.

- (264) Some hungry brothers are in search of food. Entering the house of an ogre they find a human body up on the rack in the kitchen.
 de=jŋɔk he? para?, kbə? sma?, beh, da? de=ɛn de=cā
 3ApI=observe AT:above rack body person NO NEG 3plA=want 3plA=eat
 They observed up on the rack the body (of) a person. No, they didn't want to eat (it).

beh 'NO' may be used with a clause negated by da? 'NEG', as we have seen in the examples above, or it can occur in relation to a clause containing a positive assertion. The presence of the external negator does not affect the truth value of the accompanying proposition. The fact that beh 'NO' does not negate the clause is evidence of its external nature, and clearly illustrates that it does not overlap with the internal negator da? 'NEG'. In the following examples beh occurs in predicates which do not contain the negator da?, but precedes an affirmative clause, verbal (265)–(266), or non-verbal (267), to negate or contradict that which has been stated or implied in the previous discourse. Here, beh 'NO' has the same function as with the negated clauses above.

- (265) The speaker describes the methods used to finish off graves, comparing the present with the past.
 no? pakay b-simin, cɔcɔ? beh, b-pṭe? hn=dloŋ tros
 this use USE-cement before NO MID-put O=wood inner.hardwood
 Now cementing is used, on the contrary before, inner hardwood was placed (on it).

In (266) the implicatum is a more remote assertion.

- (266) A man has returned to reclaim his estranged wife. He is not, however, out to seek conflict. Against his wishes, fate has decreed that he encounter her fiancé.
 ?are? muy ke, beh, ki=trmpa? la=tunɔjan k<r>dor
 day one that NO 3A=meet.up.with A=fiancé be.female<NMZ>
 That day, no (to his chagrin), the woman's fiancé bumped into him.

beh is used to negate an assertion in a contrastive sense: 'but it is not true for X that ...'. These clauses always follow the proposition with which it contrasts. In

(267) below, kanaw beh (flying,lemur NO) rejects the assertion that *kandau* lemurs have tails. beh occupies the position following the NP. The discussion, which centres on this assertion, concerns the various gliding mammals which do and do not have tails. The contrast is not in relation to an assertion of the identity of the species, in which case bʔen 'NOT' would be used (§10.4.2.2). beh is intonationally integrated with the proposition, indicated by the absence of a comma:

- (267) kubug ps-pbs, kanaw beh
 flying.squirrel HAVE-tail flying.lemur NO
 Flying squirrels have tails, flying lemurs don't.⁸
- (268) ?us ke da?, br-wøy=hn beh
 fire that EXIST HAVE-knife=3POSS NO
 (He) had a light (=matches), (but) was without a knife. (Lit. A light, he had, had a knife, no.)

10.4.2.2 bʔen 'NOT'

bʔen 'NOT' is a metalinguistic, or external negator, used specifically to negate or contradict propositions containing non-verbal nominal predicates (§10.1.1). In effect it functions as the non-verbal equivalent of beh 'NO'. It generally occurs in contradiction to a previous assertion regarding the identity or description of an entity. bʔen 'NOT' may occur in conjunction with da? 'NEG'.

The non-verbal nature of bʔen 'NOT' is reflected by its inability to be modified by the aspectual adverbs cɔŋ 'at all' and lagi? 'yet', or to function as a tag in an alternative interrogative. Some examples are given below:

- (269) Attempts at identifying an animal in the treetops are discredited by the responding speaker:
 bʔen, hc? ke musaq
 NOT AT:above there civet
 No, up there (is a) civet.
- (270) The wife has called the entity hanging from the beam a blanket. The husband corrects her that it is not a blanket. It is in fact the skin of his former identity.
 "ʔec bʔen" kʰlaŋ, "yup" kʰlaŋ
 eh NOT QUOTE blanket QUOTE
 "No (it's) not," (he) said, "a blanket," (he) said.

The difference between bʔen and beh can be seen in the following two examples. In (271) the nominal predicate which describes the locality dictates the choice of bʔen 'NOT', whereas in the second example (272), the preceding verbal predicate dictates the choice of beh 'NO'.

- (271) muy dɔh muy dol, dloŋ ddɔs; bʔen mɔɔcm no?no?
 one swidden one house tree be.close NOT like this
 (Back then it was) one house per swidden, timber (was) close; not like this (now).

⁸ kubug is the 'Red Giant Flying Squirrel *Petaurus petaurista*', known in Malay as *Tupai terbang merah*, Kubin or Kandau. In Semelai, kanaw is the 'Malayan Flying Lemur *Cynocephalus variegatus*', which in Malay is Kubung.

- (272) dol b-ra?wen, b-tukər dol ʔen d?oh;
 house MID-leave MID-change house LOC swidden
 beh mbe?om no?no?
 NO like this
 (Back then) the house was abandoned, the house in the swidden was exchanged (for another); not like this (now).
- Further examples illustrating the intrinsic differences of the external negators are given below. In (273) bʔen is used in relation to the nominal predicate which follows.
- (273) *The recipient opens a parcel which the giver has told him contains meat from the hunt. On opening the parcel he discovers otherwise.*
 ki=gog təʔen habu?, ki=buko?. ki=jgok,
 3A=take TO:down kitchen 3A=open 3A=look.at
 bʔen, kʰoy bapaʔ=hn, blu, bləŋ bapaʔ=hn
 NO head father=3POSS thigh arm father=3POSS
 He took (it) down into the kitchen and he opened (it). He looked at (it), no (it wasn't meat), (it) was father's head, father's thighs and arms.

bʔen co-occurs with da? 'NEG' in the following riddle (274) because it is again the identity of the possessor of the attribute, as in (273), rather than the existence that is being negated. Rather than contradict an actual assertion, here it contradicts a potential assertion on behalf of the listener.

- (274) b-sisi?, bʔen da? naga?, b-jambol, bʔen da? syol
 HAVE-scale NOT NEG snake HAVE-tuft NOT NEG bird sp.
 (It) has scales, but no (it's) not a spitting cobra, (it) has a tuft, but no (it's) not a syol
 (a bird sp. that has a crest).
- The answer to the riddle is 'a pineapple'.*

Comparative note The two metalinguistic negators in Semelai present an interesting contrast to the Malay negators *tidak* and *bukan*, which apply to descriptive rather than metalinguistic negation.

There are similarities between the Semelai bʔen and the Malay negator *bukan*. *Bukan* is used to negate nouns and noun phrases. *Bukan* can be used to negate verbs or adjectives where it relates not only to the verb or negative, but the implications associated with the statement as well. *Bukan* can be used to negate prepositional phrases where the noun is being negated rather than the verb associated with the prepositional phrase. Examples of a comparative situation are found in (268)–(269). *Bukan* is not used in Semelai, and *tidak* is only used as an external negator tidaʔ=hn 'it is not that ...'.

10.5 Interrogatives

There are three types of interrogative clause in Semelai: the intonation interrogative, the alternative interrogative (§10.5.1), and the information interrogative (§10.5.2).

The first two are grouped together as 'polar' interrogatives based on the required response. Responses are discussed in §10.5.1.3.

10.5.1 Polar interrogatives

Polar interrogatives are those that seek an expression of affirmation or negation of the truth of the questioned proposition. The response sought to a polar interrogative is either an affirmation or negation of the proposition. There are no words for 'yes' and 'no' in the interrogative context. This is achieved by means of an echo system, i.e. the partial repetition of the predicate. A sub-type is the alternative interrogative (Sadock and Zwicky 1985), which provides a list of constituents, or a proposition and its negation, to which the respondent must make a choice. There are three interrogative morphemes associated with polar interrogatives: tah 'Q', expressing doubt, 'who knows', 'maybe' (§10.5.1.1); kah 'ALT' expressing possible alternatives and beh 'NO' the external negator meaning 'or not?' (§10.5.1.2).

10.5.1.1 Intonation interrogatives

Any declarative clause may become a question by using rising intonation over the relevant constituent. In the first example it is over the verb ?ris 'to live':

- (275) ?ma?bapa? ?ris?
 parents be.alive
 Are your parents alive?

In (276) the rising intonation coincides with the verb br-kʰom (MID-get), and then falls away over k<c>bəc (fish.with.rod<IMPERF>).

- (276) br-kʰom k<c>bəc?
 MID-get fish.with.rod<IMPERF>
 Was (anything) got fishing?

In the non-verbal existential clause in (277), the rising intonation contour is over the existential marker da? 'EXIST'.

- (277) This is the greeting called out as one approaches a house:
 da? sma??
 EXIST person
 Is anyone there?

In the following clause the rising intonation is over the NP kmpən 'wife'.

- (278) The father cannot believe his son is going to attempt to visit his wife who has turned into a birth-vampire.
 ga=tet kmpən, kb?
 IMM=TO:spec wife 2f
 Are you going to (your) wife?

tah, from the Malay *tah* ~ *entah*, is a particle expressing doubt which can be added to an intonation interrogative. The doubt is that of the person asking the question, in relation to the assertion being made.

- (279) Brothers searching for their younger sibling catch sight of a distant swidden.
 tah na2-he? ?adi? he?en?
 Q DEM-AT:up VS 1&2=AUG
 Maybe that one up there (is) our younger sibling's?

tah can have scope over the whole clause, in which case it occurs in clause-initial position:

- (280) The speaker is enquiring whether the basket being woven is the one she has requested:
 tah ji=tap rɔ? ?əpi?
 Q 2A=weave basket If
 (Would) you perhaps be weaving the basket (for) me?

- (281) tah nɔ? dde=jl?
 Q this OF=2
 Would this perhaps (be) yours?

Alternatively, the morpheme may directly follow the questioned constituent, limiting scope to that constituent. In (282) tah follows the verb:

- (282) "sak tah ?" kʰləŋ "bɔs 7nɔ?nɔ?" kʰləŋ
 peel Q QUOTE sugar.cane this QUOTE
 "Aren't you going to perhaps peel this sugar cane?" (he) asked.

The morpheme tah 'Q' is also used with information interrogatives (§10.5.2).

10.5.1.2 Alternative interrogatives

There are two alternative interrogative particles: kah 'ALT' is used for NPs and beh 'NO' for predicates. The former requires the expression of each constituent, whereas the latter is elliptical, containing only the affirmative proposition and a tag.

i. kah 'ALT' The scope of kah 'ALT' is limited to the constituent it follows. kah expresses a choice of two or more possible alternatives, and is often used where two NPs are juxtaposed in an alternative relationship, 'is it A, or is it B?' It is repeated after each constituent. kah is a borrowing from Malay *kah*, a general interrogative post-position or particle (Mintz 1994: 238).

- (283) "kulat.kibebər kah," kʰləŋ "kulat sisir kah, nɔ?nɔ?"
 fungus.[name] ALT QUOTE fungus comb ALT this
 "Is this a *kibebər* fungus (or) a comb fungus?" (she) asked.

kah may be used in conjunction with the ignorative mande 'what.sort' (284)–(285) (§6.2.3). The construction parallels the Malay *apakah* (*apa* 'what' + *kah* 'interrogative particle') 'what'.

- (284) sma? kʰbas kah, mande kah, kloc lyang nɔ?
 person be.dead ALT what.sort ALT inside crevice this
 Is it a corpse (or) what in this crevice?

In (285), the NP does not stand in juxtaposition with another NP, but indicates that there could be a number of things which could be in contention for burning.

- (285) "ɔh yɔ?," kʰləŋ, "mande kah ga=b-cɔŋ?"
 oh but QUOTE what.sort ALT IMM=MID-burn
 "Oh but," (he) said, "what is going to be burnt?"

ii. beh 'NO' beh (§10.4.2.1) can be used as a clause-final tag forming an alternative question with the meaning 'or not?' Only positive assertions can contain the negative tag. The tag always follows the predicate expressing the activity in question:

- (286) ma=lən, beh?
 IRR=want NO
 Would (you) want to, or not?

The response is based on the clause nucleus rather than the tag (§10.5.1.3). Although the question presents two propositions to the respondent, there is an element of bias in interrogatives of this type, in that the interrogator appears to expect a negative response.

- (287) Q: "than beh, ko=kde? ?" kʰləŋ puyəŋ ke
 endure NO 2fA=dwell QUOTE shaman that
 A: "beh," kʰləŋ=hn
 NO QUOTE=3POSS
 Q: "(Can you) endure (it) or not, (if) you stay?" asked the shaman.
 A: "No (I can't)," he replied.

10.5.1.3 Responses to polar interrogatives

The response sought to a polar interrogative is achieved through an 'echo' system, the partial repetition of the predicate. This may be either the verb (288)–(289), aspect (290) or whatever constituent is relevant to the question (291).

- (288) Q: "ji=kʰɔm?" A: "yɛ=kʰɔm" kʰləŋ=hn
 2A=get 1A=get QUOTE=3POSS
 "Did you get (anything)?" "I got (something)," he replied.

- (289) Q: "tah ko ?pih?" A: "?pih" kʰləŋ k<r>dor
 Q 2f be.in.pain QUOTE be.in.pain QUOTE be.female<NMZ>
 "Are you in pain perhaps?" (he) asked. "I'm in pain," replied the woman.

- (290) Q: "ləc ca?" A: "ləc"
 already eat already
 Q: "(Have you) already eaten?" A: "(I have) already."

In (290) the aspectual ləc 'already' conveyed an affirmative response. The negative response is achieved by beh lagī? ('NO' and aspectual lagī? 'still') 'not yet'. In an existential predicate like (277), the affirmative response requires the existential marker da? 'EXIST'.

- (291) A: da?
 EXIST
 A: There is.

The affirmative marker *dom* 'AFF' is not used to mark a positive response to a question, generally being used to reinforce an affirmative statement. Example (292) contains the strong affirmative *toko?* 'of course, don't doubt (it)', which occurs in addition to repetition of the main clause predicate.

- (292) Q: "ji=k^he? py-poy?" A: "toko? ye=k^he?"
 2A=know IMPERF-do of.course 1A=know
 "Do you know (how) to do it?" "Of course I know."

The negative marker, *beh* 'NO' can be included in a response. In agreeing with the negative question in (293), and the negative portion in (287) above, the external negator *beh* 'NO' is used along with an internally negated clause.

- (293) Q: da? ga=k^ho=jon? A: beh, da? n=jon
 NEG IMM=2fA=give NO NEG IfA=give
 "Aren't you going to give (them)?" "No, I'm not giving (them)."

Intonation interrogatives may also have clausal answers providing additional information:

- (294) *The wife has been abandoned by her husband. The response is in the affirmative:*

- Q: kmpen ki=ba?i?
 wife 3A=be.alone
 A: 2ah, kmpen ki=ba?i tinal rom knon
 ah wife 3A=be.alone stay WITH offspring
 Q: The wife (stayed) alone?
 A: The wife alone stayed with (her) child.

Alternative NP interrogatives formed with *kah* require the repetition of the relevant option as a response.

10.5.2 Information interrogatives

Information interrogatives (Sadock and Zwicky 1985: 179) utilise an interrogative pronoun to specify that part of the proposition for which specific knowledge is sought.

The ignoratives that function as interrogative pronouns were introduced in §6.2, and the reader is referred there for a full discussion of the ontological categories. Table 10.2 presents a summary of the ignoratives attested in the interrogative function. The derived categories provide further distinctions within the basic ontological categories.

TABLE 10.2 INTERROGATIVE PRONOUNS

BASIC ONTOLOGICAL CATEGORY	GLOSS	FORM	DERIVED CATEGORY GLOSS	FORM
THING	'what'	mandehmoh	'what class of'	nmoh
PROPERTY	'what sort of'	mande		
REASON	'because of what'	mande		
PERSON	'who'	kadeh	'whose'	do kadeh
PLACE	'where'	hōn	'whither'	te hōn
MANNER	'how'	hōn-mande	'whence'	torn ha? hōn
QUANTITY	'how many'	brapa?		
TIME	'when'	mre?, bila?		

The word order of the interrogative clause depends on the syntactic relations between the constituent and the predicate. It is common cross-linguistically for the ignorative to be in initial position, and this is generally the case in Semelai, in particular amongst the basic ontological categories of THING, PERSON, PLACE (295), REASON, MANNER, QUANTITY (296) and TIME. For further examples see §6.2.

- (295) "hōn 7ma? hne?" k^həŋ
 where mother THEN QUOTE
 "Whre (is) mother then?" (they) asked.

- (296) brapa? n^h-le? ha? nō?
 how.many NMZ-be.long AT here
 How long (are you) here?

If the referent of the ignorative is a core argument, the following patterns are noted:

i. *Subject* In both verbal and non-verbal clauses the ignorative pronoun used in relation to the subject is in the 'fronted' clause-initial position.

The form used to seek information about the identity of inanimate entities is *mandehmoh* 'what' (297), and for animate entities it is *kadeh* 'who' (298).

- (297) "2oh, mandehmoh nene?" k^həŋ
 oh what noise QUOTE
 "Oh, what (is) the noise?" (he) asked.

Note that there is no cross-referencing of the agent on the transitive verb:

- (298) kadeh p-kasot ji?
 who EQUIP-shoe 2
 Who put (your) shoes (on) you?

ii. *O or IO* When the ignorative represents a core argument O or IO, the format of the interrogative clause is the same as that of the relative clause construction: the ignorative is fronted, and the verb hosts the relativising proclitic *mə=REL* (§11.1). Consistent with the pattern attested for fronting, core arguments do not retain NP-marking (§9.1.3). The result is an equative non-verbal clause. This strategy is only employed with the ontological categories of THING (299) and PERSON (300).

- (299) "Tec" k^bləj, k^bləj jsok "hmoh [mə=jɪ=ca] cəŋ?"
 eh QUOTE QUOTE sambhur,deer what [REL=2A=eat mousedeer
 "Eh," said, said Sambhur deer, "what (is it) that you are eating. Mousedeer?"

- (300) *The speaker asks for advice as to whom she should give the offering to.*
 kadeh mə=gə=yə=jən?
 who REL=IMM=1A=give
 (To) whom (is it) that I am (going to) give (it)?

The relativised format is also used where the O represents a clausal object:

- (301) "hmoh mə=mə=lən" k^bləj "ha? hn=2əŋ?" k^bləj
 what REL=IRR=want QUOTE AT ABS=1f QUOTE
 "What (is it) that (you) would want," (she) asked, "with me?" (she) asked.

iii. Adjuncts Ignoratives that substitute for the categories expressed as syntactic adjuncts within an NP or as the complement of a PP (303) display the distribution of the referent which they substitute. Adjuncts are PROPERTY, CLASS, POSSESSION (302):

- (302) qo kadeh, n^b-go? naʔ-he??
 OF who NMZ-fell.tree DEM-AT:above
 Whose (is that) tree-felling up there?

The ignorative *mande* expresses the ontological categories of PROPERTY and REASON (§6.2.3). As discussed in (§6.2.3.2), *mande* is always clause initial when the intended ontological category is REASON, i.e. it is an adjunct to the proposition. When it expresses PROPERTY it is embedded within the relative syntactic constituent, for example in (303) as the complement of the preposition.

- (303) Q: kudes ʔen mānde?
 skin.disease LOC what.sort
 Q: A skin disease on what (part of the body)?
 A: kudes ʔen joŋ
 skin.disease LOC foot
 A: A skin disease of the foot.

An unusual feature of the interrogative clause is the propensity for the nominalisation of verbal predicates, resulting in a nominalised clause. This construction is attested with the derived ignorative constructions expressing possession (302) above, and direction (304).

- (304) təm haʔ hōn, ns-dəs ji neŋ?
 SRC AT where NMZ-arrive 2 before
 From where (was) your arrival before?

In rhetorical manner questions formed with hōnmande 'how', where the ability to perform an activity is questioned, the predicate contains a complex verbal construction with the verb lən 'to want' (§11.2.2.2).

- (305) ga=k^boy nehneh, hōnmande lən ʔris?
 IMM=head just how want be.alive
 (If she) was just a head, how was she able to be alive?

- (306) *The woman in question is under a taboo, preventing her from putting her hand in a pot. The question is, how will she eat in such circumstances?*
 lən ca, hōnmande lən cluk kdeŋ?
 want eat bow want put.in pot
 (If) she wanted to eat, how was she able to put (her hand) in the pot?

10.6 Imperatives

The imperative refers to a sentence used to express a command. The imperative in Semelai is used to offer suggestions, invitations or requests, and in the negative, warnings. It would be misleading to characterise the imperative as either an order or command as this would require the compliance of the addressee. The speaker presents a proposition for the addressee to act upon. In an egalitarian-type society like that of the Semelai the type of authority required to achieve this does not exist between individuals at any level of social interaction, even between an adult and a child.

10.6.1 Formal features of an imperative clause

The form of the imperative depends on the transitivity of the clause: the positive imperative has one form (§10.6.2); whilst the negative imperative has two forms, one for intransitive clauses (§10.6.3.1) and one for transitive clauses (§10.6.3.2).

In the imperative construction, pronominals are in the free form, indicating in the transitive clause that they are not attached to the verb, but function as vocatives. They occur either pre- or post-verbally (307)–(308). The pronominal representing the addressee is the second person singular or plural, k^bo '2f(amiliar)', ji '2 non-familiar' or je=2en '2=AUG'.

- (307) boy swak, k^bo!
 NEG:IMP walk 2f
 Don't go, you!

Other vocatives are names, nick-names and address terms (396). All vocatives, regardless of their form, are optional.

- (308) busu?, ca=cə?!
 youngest.born eat=IMP,
 Youngest born, eat!

In the hortative construction (§10.6.2.3), used where the addressee is invited to perform an action which the speaker is also performing, the vocative is the first person inclusive he '1&2', or he=ʔen '1&2=AUG':

- (309) tə2, he dmpok 2adi?
 HORT₂ 1&2 follow YS
 Let's go follow younger sibling!

The imperative is also distinguished from the declarative by the optional presence of enclitic illocutionary particles – the clitics =cə? 'IMP₁', =ja 'IMP₂' – and the phrase na? yandeh 'IMP₃', or the negative imperative marker boy 'NEG:IMP' (312),

in combination with heavier final stress of the clause and a rising clause-final intonation pattern. The final point is important where there is no overt marking of the imperative.

The two enclitics occur frequently and appear to be freely interchangeable (310)–(311). These are not solely imperative clitics; they also function as discourse clitics (§13.3.1 and §13.3.3). In the positive imperative the enclitic attaches to the verb; in the negative to the negator:

- (310) *jtsk=jə!*
sleep=IMP₂
(Go to) sleep!

- (311) *pa-hūm=cə?* *?adid?*
CAUS-bathe=EM YS
Bathe (your) younger sibling!

The negator directly precedes the verb:

- (312) *bəy* *dish* *tet* *ke!*
NEG:IMP move:lateral TO:spec there
Don't go across there!

Any other information that the speaker wishes to include, like the O in (311), or the PP in (312), can usually only follow the verb.

The imperative construction is largely confined to active verbs, however stative verbs which are controlled by the subject, like adjectives of speed, do occur in the negative imperative:

- (313) *bəy* *?oyon!*
NEG:IMP be.slow
Don't be slow!

Usually stative verbs only appear as the complement of an active verb:

- (314) *bəy* *ləh* *dk²es* *kd!*
NEG:IMP move:lateral be.close 2F
Don't come close, you! (Lit. Don't move laterally (so as) to be close, you!)

10.6.2 The positive imperative

There are two types of positive imperative clause, a specific imperative construction (§10.6.2.1) and a normal declarative clause used with imperative intonation (§10.6.2.2). The hortative construction, which is essentially an inclusive imperative, is a sub-type of the imperative (§10.6.2.3).

10.6.2.1 The basic type

In addition to the standard features of the imperative clause, the positive imperative includes one of the illocutionary markers: the clitics =cə? 'IMP₁', =jə 'IMP₂', and the phrase *na?* *yandeh* 'IMP₃'. All three markers are optional.

- (315) *dlh=cə?*
go.across=IMP₁
Go across!

- (316) *yɔk=jə!*
take=IMP₂
Take (it)!

The phrase, which directly follows the verb, is infrequent:

- (317) *"bək* *na?**yandeh!"* *kʰləŋ*, *"gindəŋ* *kayen!"*
fasten IMP₃ QUOTE piece cloth
"Tie the cloth!" (he) said.

The form of the positive imperative is the same regardless of the transitivity of the clause, unlike the negative imperative where a distinction is made (§10.2.2). This is exemplified with an intransitive clause in (318) and a transitive clause with an O in (319).

- (318) *"?yot=cə?"* *bapa?* *kʰləŋ*
return=IMP₂ father QUOTE
"Come (home)!" the father said.

- (319) *"?ah,"* *kʰləŋ*, *"j?oy=cə?* *ponoŋ!"*
ah QUOTE make=IMP₁ hut
"Ah," (she) said, "make a hut!"

In the transitive imperative clause the case of the core grammatical relations O and IO (§§9.3.1.2-3) is not overtly coded, and it occurs, like the addressee, as a bare NP (319). In (320) the IO is *?əŋ* '1' and the O is *mə=bje?* 'one piece'. Usually the IO of a verb of transfer is coded by a locative preposition. Here, it resembles the benefactive which is not overtly marked (§9.1.2.2).

- (320) *gi?**=ja* *?əŋ* *mə=bje?*
give=IMP₂ if one=CLF
Give me one!

The word order of the imperative clause is usually verb initial, followed by core arguments, the addressee, and finally by any complements. The post-verbal NP in the following example is ambiguous between an O and an addressee:

- (321) *dom* *jgɔk=cə?* *??*
AFF look.at=IMP₁ EB
Yes indeed, look at (it) elder brother!

In (322) the intransitive imperative has an addressee followed by a locative complement.

- (322) *p?**öt=cə?* *kd* *ha?* *no?*
stay=IMP₁ 2F AT this
You stay here!

In (323) the O is *?əŋ* '1f', and the dependent clause is *reŋ klub?* 'to seek Salacca fruit'.

- (323) "tantat=jal!" k^blaŋ, ?əŋ, req klub?!"
 take=IMP₂ QUOTE If seek *Salacca conferta*
 "Take me!" (she) pleaded, "to look for salacca fruit."

10.6.2.2 Imperatives with declarative-like forms

It is possible to find examples of clauses that function as positive imperatives, but are not marked as such, given that they do not contain an imperative marker. However, neither do they look like a declarative because the addressee/agent is not coded.

In the following clauses a directive is given, but neither clause contains an imperative marker (324)–(325). Only the absence of grammatical marking suggests it is not a declarative. However (325), also a directive, has object marking, not usually present in the imperative. Note that in examples (324)–(326) the speaker, ?əŋ 'If', is the beneficiary of the directive. Recall that the benefactive is not overtly marked in Semelai. It may be significant that in all the examples given in this section the interaction is between a husband and wife or parent and child.

- (324) "yok ?əŋ bantal!" k^blaŋ
 fetch If pillow QUOTE
 "Fetch me a pillow!" (she) said.

- (325) "tbas ?əŋ hn=troŋ!" k^blaŋ
 slash If O=path QUOTE
 "Slash me a path!" (he) said.

- (326) *The speaker, up on a sleeping platform, wants something passed up from a lower level in the house.*
 yok he? ?əŋ!
 fetch AT:above If
 Pass (it) up (here for) me!

In (327) the intonation pattern is unusual, falling on 'dog' and rising sharply on 'hit'.

- (327) *A parent directs a child to hit a dog which is annoying her:*
 co, kon!
 dog hit
 The dog, hit (it)!

10.6.2.3 The hortative

The hortative resembles the imperative, but includes the speaker as an accomplice in the event, though not necessarily as subject (328). If a pronominal is used, it is the free pronominal he '1&2' or he=?en '1&2=AUG'. There are three hortative markers: lah 'HORT₁', tə? 'HORT₂' and hate? 'HORT₃' which are discussed below. The hortative markers lah 'HORT₁' and tə? 'HORT₂' may be used without a verb, each with a different meaning (see below). The pronominal is optional in all forms. Unlike the imperatives in (§10.6.2.1), the O may be marked, as in (328) where hn='O' is proclitic to the O; however this may simply be so because the O is an accomplice in the act.

- (328) "lah pdor hn=?əŋ! dom," k^blaŋ lsloŋ, tə?en bri.kmuc!
 HORT₁ follow O=1f AFF QUOTE flying.ant TO:down jungle.ghost
 "Come, accompany me! yes," said the flying ant, "down to the underworld!"

lah 'HORT₁' is used to make an invitation: 'Come, I am making an invitation, let's ...!' lah is possibly the Malay emphatic suffix -lah which has many functions, one of which is to take the edge off commands in order to render them more polite. As (329) illustrates, lah can stand alone, i.e. without a main verb, and in this context it simply means 'come!' In Malay, *lah* is never used in this manner.

- (329) lah!
 HORT₁
 Come!

- (330) ?ie?le2, ki=?ajak knɛk req ci,
 so 3A=invite husband seek lice
 "lah tet no?" k^blaŋ, "he b-reŋ-reŋ ci" k^blaŋ
 HORT₁ TO:spec this QUOTE 1&2 MID-INTNS-seek lice QUOTE
 So she invited (her) husband to look for lice. "Come here!" (she) said, "let's look for lice (for) each other."

tə? 'HORT₂'⁹ is used to summon accomplices to perform an action: 'let's ...' tə? precedes the pronoun, which is optional. tə? may stand alone.

- (331) ?ie?le2 k^blaŋ=hn, "dom tə? hedmpok ?adi?"
 so QUOTE=3POSS AFF HORT₂ 1&2follow YS
 So, (he) said, "Indeed, let's go follow younger sibling!"

- (332) tə?, he b<tbody>an hubi?
 HORT₂ 1&2 TOG<dig> cassava
 Let's go dig cassava together!

- (333) *A ghost realises that his wife, who is human, cannot endure life in the underworld.*
 "da? kp=ien kp=ca? tə?, jake ?yot tet dol kp!" k^blaŋ
 NEG 2fA=want 2fA=eat HORT₂, then return TO:spec house 2f QUOTE
 "Don't you want to eat (this)? In that case, let's go back to your place!" (he) said.

tə? 'HORT₂' and lah 'HORT₁' may be used together, in which case tə? precedes lah, and the latter functions to make a milder directive:

- (334) "tə?, lah!" k^blaŋ
 HORT₂ HORT₁ QUOTE
 "Come, let's go," (he) said.

A third marker hate? 'HORT₃' is brusque and expresses annoyance. Typical usage would be of a parent to a disobedient child.

⁹ In the Central Asian language Jah Hut, there is an exclamatory marker te? glossed as 'enough said, let's go' (Diffloth 1976c: 84).

- (335) hate? swak!
HORT, go
Let's move (it)!

10.6.3 The negative imperative

The imperative negator is *bɔy* 'NEG:IMP' 'refrain from V' (§10.4.3). The form of the negative imperative clause is dependent on whether the verb is transitive or intransitive. The intransitive form consists of *bɔy* 'NEG:IMP' and the verb (§10.6.3.1). The transitive imperative is formed by the negative imperative and a verb bearing the irrealis proclitic *ma=* 'IRR' (§10.6.3.2), or by a declarative clause (§10.6.3.3). People on familiar terms use the first form; the second is used where the addressee is non-familiar (§10.6.3.3).

The negator is pre-verbal, but may stand alone if the verb is elided (337), but it usually hosts either *=e?* 'IMP₁' or *=ja* 'IMP₂', or is followed by *na?* *yandeh* 'IMP₃':

- (336) bɔy na?*yandeh*!
NEG:IMP IMP₃
Don't!

Urgency can be communicated by lengthening the vowel of the negative imperative marker as in *bɔy* (337), or lengthening the vowel and inserting a glottal stop, *b?ɔy* [bɔy] (338). The choice is open to the speaker.

- (337) "bɔy!" kʰləŋ bapa?, "kɔ kʰbəs" kʰləŋ
NEG:IMP QUOTE father 2f die QUOTE
"Don't!" pleaded the father, "you will die."

- (338) "ɔh yɔ?=hn," kʰləŋ, "b?ɔy!" kʰləŋ,
oh but=CONN QUOTE NEG:IMP QUOTE
"kbaran kna? kuman"
be.apprehensive.lest strike germ
"Oh, but!" (he) warned, "don't! (I am) fearful lest (you) get (affected by) the germs."

10.6.3.1 The negative intransitive imperative

The negator *bɔy* 'NEG:IMP' followed by the verb, forms the negative imperative of intransitive verbs:

- (339) bɔy ?yot!
NEG:IMP return
Don't go back!

Inclusion of the addressee is optional; the addressee is clause final in (340), and in (341) it follows the negator.

- (340) "bɔy ryoh je=?en!, ?ap hne ga=?tck" kʰləŋ
NEG:IMP make.noise 2=AUG If THEN IMM=sleep QUOTE
"Don't make a noise you! I am going to sleep," she said.

- (341) "bɔy kɔ dɔs!" kʰləŋ
NEG:IMP 2f arrive QUOTE
"Don't you go (there)!" (he) warned.

Other constituents within the clause, for instance locative complements (342), are placed after the verb in the manner of the positive imperative construction.

- (342) "bɔy dm-dəm ʔen blu ?əŋ!" kʰləŋ
NEG:IMP IMPERF-lie.down LOC thigh If QUOTE
"Don't lie down in my lap!" (he) said.

10.6.3.2 The negative transitive imperative

The imperative negator *bɔy* 'NEG:IMP' is followed by the verb which hosts the irrealis proclitic *ma=* 'IRR' (§5.5.2).

The verb and addressee may be elided, in which case the imperative enclitics or the particle may follow the imperative negator. The enclitics and the particle are never used in the formation of the negative imperative when the verb is present. The absence of the particles is not ungrammatical, but speakers judge their inclusion to be better in a pragmatic sense.

In (343) the verb is elided and the imperative negator is encliticised by *=ja* 'IMP₂' in the first instance; in the second where the verb is present, the enclitic is not used.

- (343) bɔy=ja! bɔy ma=yɔk!
NEG:IMP=IMP₂ NEG:IMP IRR=take
Don't! Don't take (it)!

In (344) the O follows the verb.

- (344) bɔy ma=kɔj ?əŋ!
NEG:IMP IRR=hit 1f
Don't hit me!

As with other forms of the imperative non-core participants in the clause follow the verb. The following exemplify various clauses with a locative in (345) and an instrumental complement in (346).

- (345) "bɔy ma=kɔjʃt ha? g<ng>ləŋ!" kʰləŋ
NEG:IMP IRR=choke AT swallow<NM2> QUOTE
"Don't choke (my) throat!" (he) said.

- (346) bɔy ma=tɔŋ rɔm tʰi!
NEG:IMP IRR=hold WITH hand
Don't hold (it) by the hand!

The transitive negative imperative form is also used with intransitive verbs which may take pronominal proclitics (§9.3.3.1). Consider *malu?* 'to be shy', whose derived transitive means 'be shy (of s.th.)'. The use of this verb in a negative imperative is illustrated by (347), in which directions are given by an uncle to encourage a child to walk past someone he fears. The first verb, which is intransitive, is in the positive imperative, the second is a negative transitive imperative, the third is unmarked but expresses a positive transitive imperative.

- (347) di**lh=cə?** boy ma=malu?!
- move.lateral=IMP, NEG:IMP IRR=be.shy take thing AT:across
- Go over there! don't be shy (of them)! (and) fetch the thing over there!

10.6.3.3 boy 'NEG:IMP' and declarative-like imperatives

Frequently, in negative transitive imperative clauses *ma*= 'IRR' is not used and the clause has the form of a declarative, the addressee being coded by a verbal proclitic, either *kr*= '2f' or *ji*= '2'. Note that a free vocative pronoun would not normally occupy this position and, in fact, vocatives are rarely attested in this construction. The construction is used when the speaker is not on familiar terms with the addressee, but also in other instances where the motivating factor is not so clear. It is possible to account for two instances where this construction arises.

(a) The imperative mode is already established:

- (348) A tiger warns a woman not to touch his fruit in the first clause. She is insistent, and consequently he is losing his patience by the time he uses the proclitic construction in the second clause.

"boy ma=pjke? klub? xe!" ...
NEG:IMP IRR=pick salacca.conferta that
"beh," k^bləg, "boy kr=pjke?! ple ʔəp"
NO QUOTE NEG:IMP 2fA=pick fruit If
"Don't pick those salacca fruits!" ... "No," (he) said, "don't you pick (them)! (They are) my fruit."

With the next negative imperative in the text, he reverts to using the unrealis:

- (349) boy ma=p-kr ja?!
- NEG:IMP IRR=CAUS-work
Don't get (me) worked up! (Lit. Don't make (me) work!)

(b) To emphasise the importance or urgency of the addressee's compliance with the wishes of the speaker. Examples (350)–(351) both represent life and death situations.

- (350) A woman's life depends on a tiger's actions:
"zanta kke zosok!" k^bləg,
t
take there placenta QUOTE
"yb2, boy kr=ca hne," k^bləg, "k^bbəs ʔəp"
but NEG:IMP 2fA=eat THEN QUOTE die If
"Take that placenta!" (she) said, "but don't you eat (it) now!" (she) said, "(or) I will die."

- (351) The speaker is the shaman, warning that if the people should cease beating the drums during the exorcism his soul will be unable to find its way back to his body and he will perish.

"tmpoh zare?, tmpoh ptəm, boy ji=prhənti? rbana?"
seven day seven night NEG:IMP 2A=stop drum
"(For) seven days (and) seven nights, don't you cease (beating) the drums?"

11 Complex clauses

In this chapter, syntax at the level above the simple clause is examined, in other words, sentences comprising combinations of simple clauses. Relative clauses, where a clause modifies a noun, are discussed in §11.1. This is followed by a discussion of complementation, where a complement clause fills a core argument position, A, S or O, in a higher clause (§11.2). The discussion then turns to the three constructions where a clause functions to modify a VP or a proposition: serial verb constructions (§11.3), the concatenation of independent clauses (§11.4), and connective clauses (§11.5).

An areal feature of South-east Asian languages is the general absence of formal signalling to code the conjoining or embedding of clauses in complex constructions, resulting in a low level of structural difference. For the most part, in Semelai the various complex sentences simply consist of the unmarked paratactic and, to a lesser extent, the hypotactic concatenation of two or more clauses. Note that although the distinction can be difficult to establish in certain cases, there is in principle a distinction between the simple concatenation of independent clauses (§11.4) and serialised verb constructions (§11.3).

A note is in order on the analytic means of distinguishing the covertly combined clauses. Decisions were made in accordance with the intuitions of the Semelai speakers, phrase and clause level pauses, and intonation contours.

11.1 Relative clauses

The relative clause (hereafter RC) is a post-nominal restrictive clause, which may or may not have a head. There are two relativisation strategies in Semelai. The first overtly marks the RC by procliticising the invariant relative marker *ma*= 'REL' (§3.3, §3.3.5) to the initial constituent of the embedded clause. The role of the head in the RC is not encoded, other than by gapping. The RC may be headed (1), or headless (2).

- (1) sma? [ma=dəm-dəm]
person [REL=IMPERF-[ie.down]]
the person who (is) lying down

- (2) ma=sjək
REL=be.sweet
(the ones) that are sweet

The positions which can be relativised in the headed RC are the core arguments, S, A, O and IO (accessibility to relativisation is one of the tests used to determine core

roles (§9.1)).¹ In headless clauses, subjects of non-verbal clauses may also be relativised, in addition to the core arguments.

In the second relativisation strategy, used where the relativised NP is not a core argument, the two clauses are simply juxtaposed without the overt marker *mə=REL*:

- (3) knɔn pon 2yam ha? dol [de=goy raja?] children THEN cry AT:above house [3plA=bring prince]
The children were crying up at the house to which the princes had brought (them).

The role of the head in the main clause affects the position of the RC: core heads in the main clause are relativised in clause-initial position (4), RCs which have oblique heads always occur post-verbally at the periphery (3).

- (4) sma? [mə=k1=goy] la=kubur neŋ] moneŋ person [REL=3A=bite A=flying.squirrel before] [name]
The person who the flying lemur bit before (was) 'Morning'.

11.1.1 Relative clauses formed with *mə=REL*

Only the core arguments, A, S, O and IO are readily accessible to relative clause formation (with the subject of non-verbal clauses occasionally attested).

The RC has the same structure as an independent clause, except for the omission of the NP co-referential with the head, and the cliticisation of *mə=REL* to the initial constituent of the verbal word, irrespective of its role in the RC.

Examples (5)–(6) illustrate the formation of the relative clause with a head that is an S in the main clause. This set of examples is from directly elicited data. (5) shows the independent clauses corresponding to the RC given in (6).

- (5) jkɔs paloh
porcupine flee
The porcupine fled.

jkɔs ki=jəl la=cɔ
porcupine 3A=bark.at A=dog
The dog barked at the porcupine.

The head must be in clause-initial position if it is a core role in the main clause:

- (6) jkɔs [mə=k1=jə] la=cɔ paloh
porcupine [REL=3A=bark.at A=dog] flee
The porcupine that the dog barked at fled.

The similarity between the form of the relative clause in (6), and the sequential clause is illustrated in (7). The sequential reading is offset by a slight pause not present in the RC.

- (7) [jkɔs, ki=jə la=cɔ]s1, [paloh]s2
[porcupine 3A=bark.at A=dog]s1 [flee]s2
The dog barked at the porcupine (and) (it) fled.

¹ This distribution forms one of the defining features of core roles, as does the formation of interrogatives involving core roles O and IO, which incidentally are also relativised (§10.5.2).

The relativisation of various core arguments are illustrated in the following examples, beginning with S:

- (8) sdara? [mə=br-dɔl] he? kke?
relation [REL=HAVE-house AT:above there]
(my) relations who were housed up there

- (9) minggu? [mə=ga=dɔs] week REL=IMM=come
the week that is going to come (i.e. 'next week')

When the relativised position is A, both the NP and the pronominal proclitic are absent from the relative clause. Relativised As are infrequent.

- (10) kde? 'to live, dwell (s.wh.)' is a transitive verb in Semelai.
ma?ye=sən, sma? [mə=kde? ha? nɔ?] IRR=see=SC person [REL=live AT here]
(They) would have seen (it), the people who lived here.

The following examples illustrate relativisation on O:

- (11) gorɔm [mə=de=cəl] ki=lək ha? para?
salt [REL=3plA=pronounce] 3A=lick AT rack
Salt, that they called (it), he licked on the rack.

In negated clauses, the relative marker is proclitic to the negator *da?* 'NEG':

- (12) "pinan" k"ləŋ "[mə=da? ma=ən] bubuh jŋ tipo?
betel.nut QUOTE [REL=NEG IRR=want] place foot betel.service
"The betel nuts", (he) said, "that aren't wanted, place (at) the end of the betel service."

It is not possible to relativise on the coalesced object of an imperfective clause (§9.1.), as demonstrated in the unacceptable form below:

- (13a) kəhn g?go? dloŋ (13b) *dloŋ [mə=kəhn g?go?]
3S IMPERF-fell tree tree [REL=3S IMPERF-fell]
He (is) a-tree-felling. *The tree that he (is) felling.

This does not reflect a general condition disallowing RCs in the imperfective:

- (14) sma? [mə=g?go?] dloŋ ?adi? ye
person [REL=IMPERF-fell tree] VS 1
The person who is tree-felling (is) my younger sibling.

In the following example, the relativised NP is an IO:

- (15) kadeh [mə=ga=ye=jə]?
who [REL=IMM=1A=give]
(It is to) whom that I will give (it)?

The verb ?ur 'to instruct' encodes the recipient as an IO:

- (16) *The king commands his child to bathe the stranger. The child is the elder sibling of the stranger's future spouse.*
 knøn [mø=kø=?ur] pa-hūm] kaka?
 offspring [REL=3A=instruct CAUS-bathe] EZ
 (His) child who he instructed to bathe (the stranger) (was) the elder sister.

Headless RCs can relativise on the subject of a non-verbal predicate:

- (17) *The betel service has two ends.*
 sampay=hñ, ki=cø [mø=ha? jø] tipø?]
 so.that=CONN 3A=eat [REL=AT foot betel.service]
 So he ate the one at the foot of the betel service.

11.1.1.1 The headed relative clause

Any NP can function as the head of a relative clause. The head immediately precedes the RC:

- (18) *A woman has married a 'ghost' and descended into his homeland, the underworld where the ghosts of people who have died in accidents dwell. The lawot darah (sea blood) 'sea of blood'² is located there.*
 mæ=jøk sma? [mø=mirah] nøkø? ʔen mham
 IRR=observe people [REL=be.red] amuse LOC blood
 (She) observed (to her horror) people who were red, amusing (themselves) in the blood.

The RC can also be used to modify a nominal constituent of an NP as in (18). In this case, the head is not fronted. In (19) the head is the possessor NP of an associative construction.

- (19) ?adi? knøk mæ=?areh
 ys husband REL=be.new
 (It was) the younger sibling (of) the husband who was new.

The ignoratives mandehmøh ~ hmøh 'what', and kadeh 'who' may function as the head of a RC in an interrogative clause (§10.5.2) as in (20), and as the complement in a negated existential predicate (24).

- (20) "mandehmøh" k^høj, "[mø=kø=trmpa? ??e?]?"
 what QUOTE [REL=3A=encounter EB]
 "What (was it)," (she) wondered, "that elder brother had encountered?"

The head of a relative clause can itself be a headless RC as in (21). Headless RCs are discussed in §11.1.2.

- (21) *The narrator is talking about the tradition of placing carved wooden planks on graves, which has gradually been replaced by concreting the grave.*
 kira?, [mø=røj[n [mø=jøy]], pakay døj
 figure [REL=be.industrious [REL=dø]] use wood

² This compound is based on Malay lexemes. Note mham the Semelai term for 'blood' in the example.

møcdm papan kke
 like plank that

See, (the ones) who were industrious who did (it), used wood like those planks.

The presence of the RC tends to exclude any other modification of the head, although demonstratives are attested, like the recognitional use of ke 'that' (§6.3.3) in (22):

- (22) *The speaker realises that I am going to require further information about the referent in order to be able to recognise her, so she elaborates:*
 dom kaka? yc, kaka? [mø=tuli?], ke, dom gato?e?
 AFF EZ 1 EZ [REL=be.deaf] that AFF [name]
 Yes my elder sister, that elder sister who is deaf, yes gato?e?.

The head may host an enclitic like =pa 'FACT':

- (23) *The narrator recounts how her grandmother fooled the sultan's tax-collectors by making up the required weight in lead.*
 prom=pa [mø=kø=gø] la=kmpøn ke] junjun ha? cru?
 lead=FACT [REL=3A=bring A=wife that] heaped AT winnowing.tray
 (It was) lead, you know, which the wife brought heaped (up) on the winnowing tray.

The only constituent which can intervene between the head and the relative clause is the quotative marker k^høj 'QUOTE': as in (20) above.

The head and its RC are typically clause initial, except in the following two instances:

a) In negated existential predicates the RC follows the predicate:

- (24) da? da? mandeh [mø=kø=cø la=bsi?]
 NEG EXIST what [REL=3A=penetrate A=metal]
 (There) was nothing that the metal could penetrate.

b) When the head is an oblique in the main clause, it is relativised in clause-final position as in the following examples involving a locative in (25), and a causal phrase in (26).

- (25) knø? dom, hc? sma? [mø=mñ-jaga? dol]
 strike AFF AT:up person [REL=IMPERF-watch.over house]
 (It) struck indeed, the man who was up watching over the house.

- (26) *A woman is confronted with the fact that she has inadvertently committed bigamy when her first husband, who had abandoned her, suddenly reappears. She nervously confronts him in a clandestine meeting:*

kbran ki=bunuh la=knøk [mø=?areh].
 be.apprehensive.lest 3A=kill BCS=husband [REL=be.new]
 (She) was apprehensive lest he kill (her) because of the new husband. (Lit. because of the husband who was new.)

11.1.1.2 The headless relative clause

Headless relative clauses function as nominals rather than as nominal modifiers. In the absence of a head, the domain of relativisation is simply the class of objects for which the proposition holds true.

The formation of headless RCs constitutes a significant lexicalisation process in the formation of lexical items in the avoidance language (§1.1). Some examples of the resulting phrasal lexemes are set out below:

- (27)
- pədəŋ*
- 'tiger':

mə=kam ?ate
REL=hold.in.fist earth
(the one) who holds the earth in (his) fist

- (28)
- bageh*
- the 'Dusky Leaf Monkey' (
- Presbytis obscura*
-) has white markings around the eyes.

mə=putih mət
REL=be.white eye
(the one) who is white of the eye

Headless relative clauses can relativise on any core argument, (including the subjects of non-verbal clauses); note that the latter possibility is not attested for headed RCs.

- (29) [mə=ha? dol baten]
-
- [REL=AT house headman]

(It is) (the one) who (is) at the headman's house.

Unlike headed clauses, there is no tendency for the headless RCs to be in initial position; they simply fill the position of the constituent for which they are substituting. A clear example is (30) where the RC provides further delineation of the subject. There is a distinct pause between the initial NP and the headless relative clause, indicating knkon kke 'that child' is not functioning as the head.

- (30)
- The speaker is indicating a specific child amongst a group of children.*

knkon kke, [mə=t³ay], kke knon baten
child.of.other that [REL=be.big] that offspring headman
That child, (the one) who is big, that (is) the headman's child.

Headless RCs may have a specific referent, as in (30), or the referent may be non-specific:

- (31) da? da? [mə=cə] hn=deh]. [mə=cə hn=kəh]

NEG EXIST [REL=penetrate O=3pl] [REL=penetrate O=3]
There was nothing which could penetrate them, (nothing) which could penetrate him.

This use of a headless relative clause is often associated with a contrastive function as in (30) above, the example in (32) and in many of those that follow.

- (32)
- The speaker has mentioned that years ago cloth could be bought for twenty-five cents a piece. She now identifies the type of cloth:*

ye=sdər yε, kayen cəp=sən glər,
1A=remember I cloth chop=SC be.named

dom [mə=dwəpuluhlima? sen].
AFF [REL=twenty-five cent]

məhn kayen batek=sən, litapuluh sen

if cloth batik=SC fifty cent

Me, I remember (it) was called branded cloth, yes (the one) that (was) twenty-five cents. If (it was) batik cloth, (it was) fifty cents.

The headless RC functions as an NP in any one of its functions, e.g. an argument in a verbal clause, a nominal subject in equational and locative NVC, negated existential clauses and so forth. Examples of these are given below.

- a) In verbal clauses the relative clause simply functions as an NP, e.g. the subject in the first example where the RC is the possessee NP:

- (33)
- A boy has collected ripe fruits in the forest. On returning home he discovers that they are now unripe, while the unripe ones his friends collected are now ripe.*

[mə=muda?] də bʔiwa? nom
REL=be.unripe OF friend be.ripe
(The) ones that were unripe of (his) friends (were) ripe.

In (34) it functions as the O.

- (34)
- The shaman is performing an exorcism on two siblings, one female and one male, who have turned into snakes. He chooses the male:*

ʔeʔile?, ləh ?tantat [mə=rəməl] tə?cn jəŋ dak
then go:across take [REL=be.male] TO:down foot water
Then, (he) went across to take (the one) who was male down to the lower reaches of the water.

The headless RC can itself be the head of a relative clause, which is embedded underneath it as in (35). The headless RC is co-referential with the pronominal proclitic de= '3plA'.

- (35)
- The narrator is talking about her forebears.*

[mə=gdo [mə=c³cɔ?]], de=pakay kələh
[REL=be.old [REL=be.infront]] 3plA=use cup
(The ones) who were old, who were before, they used (the word) kələh.

- b) In equational and locative clauses the RC functions as the subject NP, where it is always clause initial:

- (36) [mə=c
- ³
- cɔ?] ʔəsnən, [mə=c
- ³
- rənər] zaynən

[REL=be.infront] Esnon [REL=be.behind] Zainon
(The one) who is in front (is) Esnon, (the one) who is behind (is) Zainon.

In the locative NVC, it is the PP which bears the relativising clitic as in (37). The RC is used pronominally as a discourse deictic, 'the one I was talking about just now', representing another alternative for pronominal reference for inanimate entities (§6.3.2.2).

- (37) duwapuluhlima? sen, beh da? mahal jə? , [mə=ha? nɔ?]
-
- twenty-five cent NO NEG be.expensive very [REL=AT here]
-
- Twenty-five cents, no (it) was not very expensive, that (one) there.

- c) In negated existential clauses the RC always follows the existential expression, a pattern attested above for the headed RC in §11.1.1.1:

- (38) da? da? [mə=n=croh]
 NEG EXIST [REL=lfA=meet]
 There wasn't (anyone/anything) that I met.
 I didn't encounter anyone/anything.

11.1.2 The juxtaposed relative clause

Juxtaposed relative clauses are formed by the simple juxtaposition of two clauses. There is no overt marking on the relative clause, other than the gap, and the clause is not fronted. The role of the head in the main clause is usually an inanimate oblique. (Recall that when the head is an animate oblique NP the relativiser mə= 'REL=' is used (§11.1.1).) The head is usually an O in the RC (39), but this pattern is also used in a few marginal examples where the head is not a core argument of the RC, see (40) below where the head is an oblique locative.

- (39) dos ha? kajaj [ki=rɔŋkop ha? pasir]
 arrive AT pandanus.thatch [3A=invert AT sand]
 She came to the palm-thatch (which) he had inverted on the sand.
- (40) knkon pon ?yam he? dol [de=goŋ raja?]
 children THEN cry AT:above house [3plA=bring prince]
 The children were crying up at the house which the princes had brought (them) to.

In (41) the head is an instrumental adjunct.

- (41) b-tray war rom jagon [de=təbur sma?]
 MID-try instead WITH corn [3plA=cast person]
 (It is) tried instead with corn (that) someone casts (at them).

In (42) neither the head nor its co-referential position in the RC are core arguments.

- (42) kəhn jagrom rom tmpat [deh mn-(t)ugpil]
 3S encounter WITH place [3pl IMPERF-sow.rice]
 He came across the place (where) they were sowing rice.

An alternative means is to employ a nominalised phrase:

- (43) tmpat nm-k'om sunot kəh ke
 place NMZ-sit circumcise 3 that
 the seating place for his circumcision

11.2 Complement clauses

Complement clauses are clauses which take the place of a core NP in a higher clause, either as the single argument S of an intransitive predicate, or the O of a transitive predicate. The clause which functions as an argument of another clause will be referred to as a complement clause (hereafter COMP), and the clause of which the COMP is an argument will be referred to as the main clause, regardless of the type of syntactic bond. Overt complementisers are not normally used to signal the complement's role, although the presence of the verb lən 'to desire' does appear to behave as a complementiser in certain COMP clauses (§11.2.2.4).

Complement clauses are classified according to the syntactic characteristics of the COMP. There are basically three complement types: a) those which maintain the integrity of two independent clauses and are simply juxtaposed in a paratactic manner (§11.2.1); b) those which display a hypotactic relationship so that the COMP clause has reduced features associated with clausehood and the two clauses are syntactically integrated (§11.2.2); c) those which are paratactic in structure, but have features associated with the hypotactic type. These are limited to desiderative and achievement predicates (§11.2.3).

The degree of syntactic fusion between the clauses is reflected by the semantic groupings of the predicates (see Croft 1990: 180–3 on the iconicity of this relationship). Paratactic complements express realised states or events, usually ones over which the agent has no control, whereas hypotactic COMPs are more likely to express unrealised states or events on which the agent may exert control. Further, the syntactic structure of the constructions is reflected in the fact that the former expresses two propositions, one per clause, whilst the latter expresses only one. In this respect, the hypotactic constructions display a semantic similarity with serial verb constructions (§11.3). The major difference between the hypotactic and serial clauses is the absence of syntactic and phonological integration in the hypotactic structure.

A further distinction in the syntactic and semantic relationship of complement-taking verbs with their COMP is evident in the degree of bondedness of the clauses in terms of the impact of the main agent over the outcome of the COMP. The stronger the influence of the agent, the less likely the COMP will be coded as an independent clause. Hence, immediate perception and cognitive action predicates like jŋk 'to observe' and da? sdər (NEG remember) 'to forget' are coded as independent clauses, while non-implicative verbs of manipulation like da? jon (NEG allow) 'to forbid' or ?ur 'to instruct' are hypotactically joined. Some verbs participate in more than one construction type, e.g. knowledge predicates.

In general, the number of predicates which take COMP clauses is small, and their usage infrequent. The following list provides an overview of the complement structures in terms of syntactic and semantic relationships.

Paratactic clause complements (§11.2.1):

- Assertion of fact
- Immediate perception/cognitive action
- Factive

Reduced clause complements (hypotactic) (§11.2.2):

- Utterance
- Permissive/denial of permission
- Cognitive

Desiderative predicates (§11.2.3):

- Desiderative
- Achievement

11.2.1 Paratactic clause complements

This is the simple paratactic juncture of two independent clauses where each clause displays the features of an independent main clause (cf. concatenated clauses §11.4.1). The discussion will begin with a general overview of the COMP clause.

11.2.1.1 General features of paratactic complements

Both stative intransitive (44) and transitive (46) main verbs take COMP clauses which maintain the integrity of an independent clause. The COMP functions in the S position in intransitive clauses, and in the O position in transitive clauses. When the main clause is intransitive, the construction is V [COMP]_S:

- (44) "ni^k *ye=ʔen br-ʔyot*" k^l*ŋay*
be.good [1=AUG MID-return] QUOTE
"(It's good we have returned (to each other)," (he) said.

When the main clause is transitive the construction is: A V [COMP]_O:

- (45) ki=j^pok=san {sma? masuk tom co? balay}
3A=watch=SC [person enter SRC AT:below verandah]
He watched someone enter from down on the verandah.

In (46) the A of the transitive COMP is represented by the pronominal proclitic on the verb.

- (46) ji=ʔyəŋ [de=g<n>ipr-i?] ʔi=nunjk̩t
2A=hear [3plA=call.by.name<CAUS>-ITER NM=[name]]
You've heard they call (him) by the name ʔinujk̩t.

The optional elision of a core role from the COMP is motivated by anaphoric discourse elision (§9.4). Discourse motivated elision can take place in the COMP, regardless of whether the subjects of the main and COMP clauses are co-referential or not, provided the elided argument is recoverable from context. In (47) co-referentiality does not hold, but the elided argument in the COMP is recoverable from context. A co-referential reading is not plausible as shown in the nonsensical starred translations below:

- (47) ronon de=j^pok ʔyot
late.noon 3plA=observe return
Late noon, they observed (them) return.
*Late noon, they observed (themselves) return.

Generic statements do not require the encoding of any participants:

- (48) *A woman is reminiscing about the rations the Semelai received during their internment at the height of the Communist Emergency.*
snaj ca p̩r̩
be.nice eat free
(It was) nice (to) eat (for) free.

In (49)–(50) the S in the COMP is post-verbal, reflecting the permissible main clause word order V S.

- (49) ki=j^pok [dɔs=ce? sma?]
3A=observe [arrive=CL person]
He observed someone arrive.

- (50) da? ki=than [masuk g̩r̩v̩m lada? zen rwaj]
NEG 3A=endure [enter salt chilli LOC stomach]
He couldn't endure salt (or) chilli entering his stomach.

i. *The scope of negation, aspect and mood* Each clause can be negated or aspectually modified independently of the other. Note the similarities between this construction, and concatenated clauses (§11.4.2). In (51) the scope of the negator is restricted to the main clause. Note that there are a number of embedded clauses contained within the bracketed COMP.

- (51) da? ji=k^he? [ʔɒsɒl ye=ʔyəŋ gr̩gr̩ bom ha? ke,
NEG 2A=know [whenever 1A=hear rumble bomb AT there
ye ke, sampay da? moh ca, la=t-p̩eker hn=kəhn]
I that so.that NEG want eat BCS=HAPP-think O=3S]
You don't know (that) whenever I heard the rumble of the bombs there, I there, (got) so that I didn't want to eat because (I would) happen to think of him.

In (52) the COMP is negated.

- (52) ye=k^he? [ji da? b-lawok]
1A=know [2 NEG HAVE-food]
I know you don't have (any) food.

The COMP can also be marked independently for aspect. In (53) it is marked by the aspectual clitic ga= 'IMM', and in (54) by the adverbial modifier ʔareh 'recently'.

- (53) susah [ga=m̩-jala?]
be.difficult [IMM=PERFM-casting.net]
(It) will be difficult (for) (me) (to) go fishing with the casting net.

- (54) ye=pek^her [ye ʔareh b-knɔn]
1A=think [1 recent HAVE-offspring]
I think I had recently given birth.

ii. *Nominal equative complements* This construction is used to express the recognition of the identity of an entity, usually a person. The COMP is a nominal equative predicate (§10.1), consisting of the imminent aspect marker ga= 'IMM' and an NP. The structure of the COMP is always: COMP → [ga=NP].

The referent, whose identity is stated in the COMP, is not represented in the main clause. Only two predicates are used in this construction: k^hnal 'to know, to recognise' and k^he? 'to know, to intuit'. (55) illustrates the use of k^hnal 'to recognise' in a simple clause with no COMP.

- (55) da? b-k^hnal kbə?
NEG MID-recognise body
It wasn't recognisable (that it was a) body.

In (56) the COMP is a nominal predicate. As one would expect from the semantics, it is only the main clause which is negated, that is, his inability to recognise the thing in front of him, and not the fact that it is a body.

- (56) *A man peers down into a grave, but it is impossible to distinguish what is inside it.*

da?	b-k ⁿ al	cəŋ [ga=sma?]
NEG	MID-recognise	at.all [IMM=person]
	(It)	wasn't recognisable at all (that it was) a person.

Although in (57) the aspectual *ga*= 'IMM' does not add 'immediate' semantics, it can in other instances of the construction, as in (56).

- (57) *The shaman has seen a person in his dream:*

ki=k ⁿ e?	[ga=sma?]
3A=intuit	[IMM=person]
	He intuited (that there) was going to be someone.

iii. *Predicates taking paratactic independent clause COMPs* Main clause predicates which take paratactic complements are listed below according to semantic categories. They are: a) assertion of fact including immediate perception and cognitive action predicates; b) assertion of recognition of identity, and c) factive, including fear predicates. Some further examples are included under each type.

a) Assertion of fact

sder	'to remember'	?ye	'to see'
da? ... sder	'to forget'	ma=2ye (IRR=sec)	'to be visible'
pek ⁿ er	'to think'	tr=ye (HAPP=sec)	'to catch sight of'
z ⁱ ŋat	'to think, hold an opinion'	prati?	'to look at'
z ⁱ ŋ ⁿ j	'to hear'	jŋok	'to watch, observe'
ma=?yan	'to be audible'	k ⁿ e?	'to know (a fact)'
	(IRR=hear)		

- (58) da? ki=sder [tu?aki?] kehn kⁿbəs]
NEG 3A=remember [grandfather] 3POSS be.dead]
He forgot his grandfather was dead.

b) Knowledge of identity

k ⁿ nal	'to recognise' (see (56))	k ⁿ e?	'to recognise, intuit' (see (57))
--------------------	---------------------------	-------------------	-----------------------------------

c) Factive

than	'to endure [COMP]' susah	snag	'to be nice, pleasant (to) [COMP]' 'to be difficult (to) [COMP]'
		kali	'to be daring, be brave (to) [COMP]'
zilok	'to be good (that) [COMP]'	bt ⁿ ŋj	'to be afraid of [COMP]'
payah	'to be troublesome (to) [COMP]'	kbran	'to be apprehensive of [COMP]'

- (59) *A father grumbles about his childrens' tardiness in returning from bathing:*

z ⁱ ŋ ⁿ sl	te?en	dak,	payah	ləŋ	?yot
when	go:down	water	troublesome	want	return
Whenever (they) go down to the water, (it is) troublesome (for them) to come home.					

- (60) ma=bⁿŋj 2mot kappl.trbaj
IRR=be.afraid get.in aeroplane
(I) would be afraid to go in an aeroplane.

See also examples (50) and (53) above.

11.2.2 Reduced complement clauses

We now turn to a discussion of the reduced or hypotactic clause complement type, first examining the general features of the COMP, and then the various idiosyncratic features of the predicate types.

11.2.2.1 General features of reduced complement clauses

In reduced complement clauses the COMP exhibits a loss of syntactic structure. The two clauses are bonded into a structure resembling a single-clause structure, which functions semantically to convey a single proposition.

The distinctive features of this clause type are:

- the obligatory deletion of the argument in the COMP which is co-referential with an argument in the main clause;
- the infinitival nature of the verb in the COMP. This is evident from the inability of the COMP to be marked independently for negation, aspect and mood.

Each of these points is taken up below.

i. *Co-referential deletion* The deletion of co-referential arguments is a consistent feature of clause combining processes. In complement clauses the deletion of an NP under co-reference with one in a higher clause is always obligatory, and is a distinct process from normal discourse motivated elision, though both deletion and elision may occur in the same construction.

The deletion of the pivot is distinguished from elision by two factors. Firstly, only arguments in A/S roles in the COMP are available for deletion. This NP can be co-referential with any core argument of the main clause. Secondly, co-referential deletion in complements is obligatory, whereas discourse-based elision is optional. The result of these two factors is a reduced-clause type as illustrated in the examples below.

In (61) the A of the COMP is co-referential with the A of the main clause.

- (61) ki=kⁿe? [ren ninca ptɔ:m]
3A=know [seek food night]
He knew (how to) seek food (at) night.

In (62) the A of the COMP is co-referential with the O of the main clause:

- (62) ki=?aŋak knlək [ren cl]
3A=invite husband [seek lice]
She invited (her) husband to look for lice.

In (63) the A of the COMP is co-referential with the IO of the main clause:

- (63) ki=?ur=cə? [jaŋa?] ten ?ma?
3A=instruct=EM [guard] LOC mother
He instructed his mother to guard (it).

The usual rules of discourse-motivated elision in conjunction with obligatory deletion often ensure an absence of overt arguments, or at most, the main clause subject has proclitic representation in this construction:

- (64) ki=?ur [bək k<r>mɔŋ]
3A=try [tether finish<CAUS>]
He instructed her (to) tether (them) all.

ii. The infinitival nature of the complement The verb of the COMP is in the neutral or infinitive form. This is easily ascertained in the case of transitive verbs, which fail to host the pronominal proclitic or allow the representation of the cross-referenced external NP in the COMP (65). The inclusion of the agent in the COMP would indicate that the second predicate is realised and the result would be a sequential reading (66).

- (65) ki=duga? [roc] (66) ki=duga?, ki=roc
3A=try [pull.out] 3A=try 3A=pull.out
He tried (to) pull (it) out. He tried (and) he pulled (it) out.

In the case of intransitive verbs, the COMP appears no different to that of a paratactically conjoined clause with an elided S. Here, proof of the infinitive verb form is the inability to express the S in the COMP and retain the knowledge of ability meaning (68). Instead, it would be interpreted as an assertive predicate:

- (67) ki=k^he? [swak] (68) *ki=k^he? [kəhn swak]
3A=know [walk] 3A=know 3S walk
He knew (how) to walk. *He knew (how) to walk.
He knew (that) he had gone.

Only the main clause can be modified in terms of negation, or aspect, but the scope of the modification is over the whole construction.

11.2.2.2 The internal structure of the reduced complement

We will now examine the internal structure of the reduced complement type. The ordering of constituents in the intransitive COMP is V (PP):

- (69) ki=pa?ey swak ?en boh
3A=instruct walk LOC elsewhere
He instructed (the dog) to walk elsewhere.
- (70) ki=?ur la=kaka? [2yot ley bapa? deh]
3A=instruct A=EZ [return TO:up father 3pl]
(His) elder sister instructed (him) to go back up to their father.

The COMP may be a prepositional predicate. This occurs when there is a goal of motion preposition (*laŋ* 'TO:up', *teʔen* 'TO:down', *tet* 'TO' (§8.6), which can be used without the verb of motion as a goal predicate (§10.1):

- (71) ki=?ajak=sən [teʔen ʔate]
3A=invite=SC [TO:down earth]
He invited (her), see, down to the ground.

The ordering of constituents in the transitive COMP is V (O) (PP), identical to the constituent order of the universal clause (§9.2). In (72) the COMP contains an O and a PP. Note that there is no O-marking in COMP clauses (72)–(73).

- (72) ki=?ur [cəŋ] kmyan ?en məŋkə putih
3A=instruct [burn kemayan LOC bowl be.white]
He instructed (them) to burn kemayan in a white bowl.
- (73) ki=?ur [yɔk rapah]
3A=instruct [fetch forest.litter]
She instructed (him) (to) fetch forest litter.

There is a restriction in the COMP against the O preceding the verb:

- (74) ki=?ur *[rapah yɔk]
3A=instruct [forest.litter take]
*She instructed (him), forest litter (to) fetch.

There is however the possibility of the O being fronted to the initial position of the main clause as in (75). The main clause O can never be fronted to this position.

- (75) [tyap ?ikur cə] ki=?ur bək
[each CLF dog] 3A=instructed tether
Each dog, she instructed (them) to tether.

Multiple complements are attested. In (76) *toloŋ* 'to help' is the complement of *cal* 'to tell' and *yɔk* 'to fetch' is in turn the complement of *toloŋ* 'to help'.

- (76) ki=cal [toloŋ [yɔk]]
3A=utter [help [fetch]]
She told (him) (to) help (her) fetch (it).

lən 'to want' as a 'complementiser' A point of interest with the reduced-type COMP is that when the subject in the COMP is co-referential with the main clause subject, the COMP containing an unrealised activity verb is embedded under the verb *lən* 'to desire' (§11.2.3). It remains unclear at this stage as to whether *lən* is functioning as a type of 'complementiser', or if it should be treated as a modal verb (§5.6.1.2).

The clause with *lən* 'to desire' in this function expresses two sequential actions where the first is definite, and the second is unrealised. The subject is co-referential in both clauses. This type of clause occurs where the higher predicate is a verb of knowledge or thought (k^he? 'to know'; da?... sər 'to forget'; t-peker 'to come into one's mind') or feeling (nam 'to sense') relating to the perceiver's own body and expressing self-awareness, or the lack of it, and is not an assertion of fact (§11.2.2.3). Note that here, as elsewhere the verb *lən* 'to desire' expresses an internal or emotion-based desire which cannot be controlled.³

³ *lən* 'to desire' also occurs with a similar function in manner interrogatives concerning possibility/ability (§10.5.2) and in clauses where it is in an expression of probability, or certainty. The presence of *lən* in these contexts may be influenced by Malay, where *mau* 'to want' is used in such circumstances.

lən 'to want' is the initial constituent in the clause. The clause introduced by *lən* 'to want' cannot contain an overt subject, be negated, or be modified by aspect markers. The presence of *lən* in this clause type does not appear to be obligatory.

- (77) *A woman, content with her new life, has forgotten that she wants to return to her own people*

da?	ki=sdər	wɔ?	[lən	?yot]
NEG	3A=remember	longer	[want	return]

She no longer remembered (she wanted) to return.

- (78) *The speaker teases a woman who he has tethered to his house post. She is pleading with him to release her, but he claims he doesn't know how.*

da?	ŋ=k?e?	[lən	th-təh]
NEG	1fA=know	[want	IMPERF-untie]

I don't know (how to) untie.

In the following example the main clause verb is marked for the unrealis. The son has expressed a desire to cure his blind father of his affliction. The mother replies in (79) that she has no knowledge of how to do this. This is more dubitative than (78) and suggests that there is no possibility that she would be able to.

- (79) "da? ma=k?e?" k?leg ʔma? "[lən b<?>la?]"
NEG IRR=know QUOTE mother [want cure<IMPERF>]
"I wouldn't know," said mother. "(how to) cure (him)."

- (80) t-knər lən ca serəh pon
HAPP-to.think want eat betel too
(It) occurred to (him) to chew betel too.

- (81) *The man isn't asleep, but feels like he's going to fall asleep, as a result of chewing the potent betel nuts.*

ki=jam	lən	jtek
3A=feel	want	sleep

He felt (he) wanted to sleep. Or: He felt sleepy.

11.2.2.3 Predicates taking reduced COMPs

We now turn to the various main predicate types that exhibit reduced-clause complements, and examine various idiosyncrasies exhibited by these types.

i. Utterance predicates

tur ʔen [IO]	'to instruct s.o. [IO] to do s.th. [COMP]'
cel ʔen [IO]	'to utter; to declare, pronounce s.th. [COMP] to s.o. [IO]'
smap ʔen [IO]	'to request s.th. [COMP] from s.o. [IO]'
tapa? ʔen [IO]	'to ask s.o. [IO] (s.th.) [COMP]'
ʔajak [O]	'to ask, invite, persuade s.o.[O] to do s.th. [COMP]'
pa?ay	'to call a dog to do s.th. [COMP]'
tawar	'to invite s.o. [O] to do s.th. [COMP]'

It should be noted that the COMP of the utterance predicate only ever expresses indirect discourse, never direct discourse. This is counter to Noonan's claim that all languages have a means of expression of direct discourse in relation to utterance predicates (Noonan 1985: 111). The expression of direct discourse in Semelai is

effected as an independent clause marked off by the quotative *k?leg* 'QUOTE' (§13.1).

In utterance predicates, the agent of the main clause is never co-referential with that of the COMP. The argument deleted from the COMP is co-referential with either the O or the IO of the main clause. In (82), *b?ləu?* 'friend' is the O of the main clause.

- (82) de=ʔajak la=b?ləu? [yvr ple]
3plA=invite A=friend [ascend fruit]
The friends invited (him) to climb for fruit.

The distinguishing feature of this clause type is that core arguments of the main clause can be post-posed following the COMP clause. This ordering reflects the structure of a simple clause (§9.2), and serial constructions (§11.3) where the post-posed agent may (in the former case) and must (in the latter) occur in this position. Post-posing a main constituent to final position results in the apparent embedding of the COMP into the main clause. This could be considered a form of clause union (Noonan 1985: 82). The first example illustrates neutral ordering, the second, the post-posed embedded order:

- (83) ki=ʔur la=kaka? [ʔyot leg bapa? deh]
3A=instruct A=EZ [return TO:up father 3pl]
(His) elder sister instructed (him) to go back up to their father.

- (84) ki=ʔajak [paloh] la=kmpən
3A=invite [flee] A=wife
The wife invited (him) to flee.

In the following example the O of the main clause is placed after the COMP:

- (85) ki=ʔajak [swak] kmpən=hn
3A=invite [go] wife=3POSS
He invited his wife (to) go.

ii. Permissive/non-permissive predicates The only main predicate identified for this construction type is *jon* 'to let, allow'. The agent of the main clause is never co-referential with the COMP. When the main clause is negated, the structure of the COMP is dependent on whether the verb of the lower clause is intransitive or transitive. Intransitive COMPs are shown in the following examples:

- (86) ki=jon la=bapa? [blajar]
3A=allow A=father [study]
The father let (him) study.

- (87) *A group of monkeys want to join a ship, but the captain only wants one of them:*
da? ki=jon [b-TRA-2-ray]
NEG 3A=allow [MID-INTNS-be.many]
He wouldn't let (them) be many.

The transitive verb of non-permissive complements must be marked for the irrealis with the proclitic *ma*=IRR.⁴ This proclitic is associated with deontic modality and is also used to form the negative imperative of transitive clauses (§10.6). Effectively the substitution of a non-permissive main clause by the negative imperative marker *boy* 'NEG:IMP' yields a negative imperative clause. With this predicate type, mood marking in the COMP is an obligatory structural requirement and the irrealis proclitic is never interchangeable with the agent proclitic as it is in the Desiderative type (§11.2.3) below.⁵ As a result, it is treated here as a reduced complement type, despite the presence of the irrealis proclitic.

- (88) *zəc, da?* *ki=jon ma=tɔŋ*
ah NEG 3A=allow IRR=grab.hold
Ah, he wouldn't let (her) grab hold (of) (it).

- (89) *mntuh?=hn dɔs hə? dol, da? ki=jon ma=?ye*
parent.in.law=3POSS arrive AT house NEG 3A=allow IRR=see
(When) the parents-in-law came to the house, he wouldn't let (them) see (her).

The post-posing of the main clause A after the COMP is not attested, differentiating this type from utterance predicates:

- (90) **ki=jon [ca jɔh] la=kaka?*
3A=allow [eat drink] A=EZ
* (His) elder sister let (him) eat (and) drink.

iii. Ability and knowledge predicates This clause type covers a range of predicate types, those which have non-co-referential agents in the main and COMP like the previous types, and those which do have co-referential agents. Both predicate types have similar syntactic structures. This type does not allow the post-posing of the matrix A after the COMP as in utterance predicates, nor can the COMP be marked for mood as in the case of the non-permissives (§11.2.2.4).

- glet* 'to be physically able to perform an activity'
- k^be?* 'to know/intuit (how to V)'
- k^bnal* 'to know (how to V), be familiar (with s.th.)'
- panay* 'to be clever, adept (at V)'
- shjeh* 'to be reluctant (to V)'

The verbs *k^be?* 'to know', *k^bnal* 'to know, be familiar with' and *panay* 'be clever, know how to' are used in this predicate type to express cognitive knowledge of an ability (67), or familiarity with a task or procedure (91)–(92).

- (91) *ki=k^bnal [jɔy k<n>rja?]*
3A=knew [do work<NMZ>]
He knew (how to) do the work.

⁴ The exception is when the subject of the complement is the indefinite pronoun *sma?* 'someone, anyone', in which case the irrealis is not marked on the verb.

⁵ This is the rendering of a negative imperative, a direct statement, into an indirect or reported one: *boy ma=tɔŋ* (NEG:IMP IRR=grab) 'Don't (you) grab (it)!' as opposed to (88). Deontic force is maintained. A similar construction to this is the intransitive predicate *da?* *sot* (NEG permit) 'not allowed (to)'.

- (92) *boy ma=dimpok! deh panay ?yot*
NEG:IMP IRR=follow 3pl be.clever return
Don't follow (them)! They know how (to) return home.

The main clause always has an A co-referential with the lower clause, and co-referential deletion always applies as in (91)–(92) and in the following examples:

- (93) *da?* *ye=glet k<r>mɔŋ*
NEG 1A=able finish<CAUS>
I wasn't able (to) finish (them).

- (94) *?ile?le?=hn ki=shjeh blajar*
then=CONN 3A=be.reluctant study
Eventually he got tired (of) studying.

11.2.2.4 The sense predicate

The sense predicate *pam* 'to sense, feel, taste, smell' covers modes of perception other than visual or aural.

- (95) *A wife has been working hard and feels the consequences:*
kmpon hñc jam jreh ji?
wife THEN feel be.exhausted very
The wife thereupon felt very exhausted.

In transitive clauses, there is potential ambiguity as to whether it is the properties of the object which are felt, or whether it is also the affect on the protagonist. In (96) it is impossible to tell if the betel nut he is chewing tastes potent, or if the man himself is feeling high. There is no ambiguity in (97) where it is clearly the snake's eggs which taste creamy.

- (96) *kl=pam malan*
3A=sense be.high
He felt high. Or: He tasted (the betel nut) which was potent.

- (97) *kl=pam milk*
3A=sense be.creamy
He tasted (them) creamy.

11.2.3 Desiderative and achievement predicates

The desiderative predicate type exhibits paratactic juncture, hence there is no obligatory co-referential deletion and the verb is not infinitival. However, counter to expectation, the scope of negation patterns as for hypotactic juncture, since the negation over the main clause has scope over the COMP.

The subject of the main clause is always co-referential with the subject of the COMP. Main clause predicates which occur in this construction express desire or achievement:

- lən* 'to want, desire, based on inner feeling or emotion'
- moh* 'to want, based on rationale'
- k^bm* 'to get, to manage'

The main clause predicate in this type is *lən* 'to want'. In its function as a lexical verb *lən* 'to desire, want' expresses internal emotion-based physical desire. The

nature of this verb means it cannot be used with non-co-referential agents. In (98) the verb is used in a simple clause and in (99)–(100) the verb has a clausal object.

- (98) da? n=lan hn=kəh
NEG IfA=want O=3
I don't desire him.
- (99) *The blood of royalty, according to Malay tradition, is white.*
ki=lan [krwol mham putih]
3A=want [emerge blood be.white]
He wanted (his) royal blood to emerge.
- (100) sampay ki=moh [2yot leŋ bapa?] so 3A=want [return TO:up father]
So he (was) willing (to) return up to his father.

As a predicate-taking verb, verbs from this group may select between two predicate types dependent on whether the transitive COMP of the desiderative predicate is unrealised (101)–(102), or realised (103)–(104). No such distinction can be made for intransitive COMPs.

- (101) ki=lan [ca] 3A=want [eat]
He wants to eat (it). i.e. He desires something he knows is attainable.
- (102) da? kp=lan=cə? [ca sayor knkən]?
NEG 2fA=want=EM [eat vegetable children]
Don't you want to eat the children's vegetable dish?

In a transitive complement, the verb can host the pronominal proclitics, or the irrealis proclitic, however overt As are not attested in the COMP. Therefore, this clause type is somewhere between a reduced and independent clause and is termed 'semi-reduced'.

- (103) ki=lan [ki=ca] 3A=want [3A=eat]
He wants to eat (it). i.e. He wants to eat the food which is present.

moh 'to want' expresses desire based on reason or cognitive rather than emotion:

- (104) kp=moh [kp=pðɔr-i?] hn=?əp?] kʰləŋ
2fA=want [2A=follow-ITER O=1f] QUOTE
"(Are) you willing to follow me?" (he) asked.

Example (105) illustrates that the clause of this type cannot be interpreted as sequential since the two verbs refer to a single event. If it were sequential, it would be understood as 'he wanted (it) and he ate (it)'. Such a reading is not possible here.

- (105) *A man wants to eat the food his wife has prepared, however a taboo prevents him.*
ki=lan [ki=ca] yd?=hn la=?əmu? ke ləc
3A=want [3A=eat] but=CONN BCS=knowledge that already
He wanted to eat (the food she had prepared), but because (of) that knowledge (he refrained).

- (106) da? ki=kʰm [ki=bukd?] NEG 3A=manage [3A=open]
He didn't manage (to) open (it).

In the following example, as with the previous, the negator has scope over both clauses. If this was a sequential structure, or a main clause with an independent COMP, only the first clause would be negated. In (107) the scope of the negator is obviously over both clauses, which must be interpreted as representing a single proposition.

- (107) de=jpk hə? para? kbe? smə?
3plA=observe AT:above rack torso person
beh, da? de=lan [de=ca]
NO NEG 3plA=want [3plA=eat]
They observed a human body up on the rack, (but) no, they didn't want (to) eat (it).

The COMP clause refers to a realised or currently obtaining state of affairs, as opposed to an unrealised desire which is expressed by a hypotactic construction. The point of issue here is that in order to use this construction, the situation which is desired must actually be realised. Both (108) and (109) illustrate assertions of realised states, in (108) the woman is now married to the man, and in (109) the things are in the ship.

- (108) "da? kp=moh [kp=jɔy knlək d?en]" kʰləŋ
NEG 2fA=want [2fA=make husband past] QUOTE
"You didn't want to make (him) your husband back then," she said.
- (109) ha? kp=ppl ys, baraj ?a=dəŋ ma=lan [ma=ca]
AT ship 1 thing DET=like IRR=desire [IRR=eat]
On my ship (there are) all sorts of things (you) might want to eat.

The following examples show predicates with intransitive COMPs. The agent of the lower clause has zero representation, however examples (110)–(111) show intransitive verbs inflected for the imperfective, indicating that this clause type is not a reduced one.

- (110) da? ki=kʰm [yr-yɔr]
NEG 3A=manage [IMPERF-ascend]
He didn't manage (to) climb.
- (111) ki=kʰm yp-yup, kəh, ki=kʰm st-sət
3A=manage HAVE-cover 3 3A=manage HAVE-wear.sarong
(He) got to have a cover, (and) her, she got to wear a sarong.

11.3 Serial verb constructions

The serial construction consists of two verbs, which combine into a single predicate in a monoclausal structure, forming a single phonological phrase. The serial unit conveys a single proposition. Syntactically, it consists not of two clauses, but of two predicates which share a single set of arguments. Example (114) is not simply the sum of (112) and (113); rather, the second verb is introduced to provide additional

information. Note that in (114) the man is walking, but it is not specified if the child is walking or being carried.

- (112) *ki=goy knon*
3A=bring offspring
He brought the child.

- (113) *knon swak*
offspring walk
The child walked.

- (114) *ki=goy swak knon*
3A=bring walk offspring
He took the child walking.

The defining criteria of the serial verb construction are given in detail in §11.3.1. The function of serialisation is to modify the event depiction in terms of manner or result, where the two verbs effectively pull together encoding a single co-temporal action.

Serial constructions in Semelai are never used as a means of increasing the valence of a clause or expressing oblique roles such as instrument, or goal in a clause. Nor is there any systematic 'verb compound' serialisation of the type encountered, e.g. in Yoruba, *pa run* (hit+crush) 'destroy' (Durie 1997: 322) or Kristang (Malacca Creole Portuguese) *da mpusta* (give borrow) 'to lend' (Kruspe fieldnotes). There is however the occasional idiomatic sequence such as *tlaŋ ʔat* (grab fuck) 'to rape' or *sec dəm* (surreptitiously lie) 'to commit adultery'.

In the majority of cases in the type of serialisation attested in Semelai, one principal verb usually functions to describe the action, and the secondary verb is supplementary, modifying the main verb in terms of manner or orientation (see Jarkey 1991: 81 for Hmong), although in some cases this distinction is not maintained (§11.3.2.4). The relative ordering of the verbs is determined by the scope/type of the supplementary verb, e.g. the supplementary verb describes the activity of the agent, as in manner constructions, and is the initial member in the series:

- (115) *dehn sec paloh*
3S surreptitiously flee
They surreptitiously fled.

- (116) *ʔyot tʰay*
return be.big
(The moon) is waxing.

If the supplementary verb is a resultative modifying the event in terms of the effect on the O, it is the second verb in the series as in (117), although in one rare instance it is initial as in (118).

- (117) *hubi? knon cəŋ*
cassava finish completely
The cassava is completely finished.

- (118) *cə, knon ye=bunuh*
dog exhaustively 1A=kill
The dogs, I killed (them) all.

The ordering of the verbs is essentially iconic, a recurrent feature of serialisation cross-linguistically, (see Durie 1997, Schiller 1990). In Yoruba for instance, verbs which occur as the initial member in the sequence as *V₁*, express the activity of the agent, 'quickly, stealthily, with difficulty', whilst those that occur as the second verb in the sequence as *V₂* pertain to the effect of the action on the patient, e.g. 'completely, off, away' (Durie 1997: 336). This pattern is also attested in Semelai, see §11.3.2.1.

The most significant characteristic of this clause type in Semelai is that one verb functions as the principal verb, and determines the valence of the clause. The valence of the clause is never determined by the composition of the two verbs. In (119), the principal verb *kʰbəs* 'to die' is intransitive, the supplementary verb *cəŋ* 'to pass between' is also intransitive.

- (119) *As a result of evil shamans exercising their malevolent powers, people can die without even manifesting signs of illness.*

- da?* tr-pih. cəŋ kʰbəs ia=pham dehn
NEG HAPP-be.sick immediately die RCS=think 3plS
Without happening to (fall) ill, (people) immediately die because of their (=the evil shaman's) thinking.

When the supplementary verb precedes a transitive principal verb, the supplementary verb hosts the pronominal proclitic. The principal verb *papel* 'to call, summons' is transitive, and *cəŋ* 'to pass between' is again the supplementary verb:

- (120) *de=cəŋ papel nojom*
3plA=immediately summon astrologer
They immediately summoned the astrologers.

A serialised verb clause has only one possible interpretation, as opposed to the ambiguity that is often attested in concatenated clauses (§11.4). The following clause cannot be a serial construction on the basis that neither verb could modify the other in the real world:

- (121) *ki=məc krwəl*
3A=dive.in emerge
He dived in (and then) (re)emerged.

The clause is understood instead as a simple concatenation of two clauses describing two separate events (§11.4.1). Essentially, a serial construction contains two facets, or sub-parts, of the same overall event. In contrast, a case of two temporally consecutive actions – two separate VPs, such as (121) above, or two separate clauses (like those in §11.4.1.1) – conveys two closely related but distinct events, a case of what Matisoff (1973) has labelled 'fortuitous concatenation'.

11.3.1 General features of serial verb constructions

The characteristic cohesiveness of the Serial Verb Construction (SVC) is demonstrated by the following seven criteria, which apply across the board to the various types of serialisation discussed in §11.3.2.

- a) The SVC has the intonation contour of a single clause construction. There is no intonational indication of a clause boundary separating the two predicates. The

introduction of such would change the whole nature of the clause. This is exemplified in b).

b) There is total inseparability of the two verbs. No constituent whatsoever can occur between them, including the pronominal proclitics as shown in (122). Here the transitive principal verb *jon* 'to give' selects the pronominal proclitic, but it attaches not to the principal verb, but instead to the initial verb in the sequence, even if the supplementary verb is intransitive.

- (122) *ki=cəŋ* *jon* *bəs* *rəm* *wəy*
 3A=immediately give sugar.cane WITH knife.
 He immediately gave (him) the sugar cane (along) with the knife.

All NP arguments must follow the sequence of verbs:

- (123) *ki=gonj* *paloh* *la=bapaʔ=hn* *knon=hn*
 3A=take flee A=father=3POSS offspring=3POSS
 The father fled taking his child.

The alternative orderings would be interpreted with a consecutive reading, and not a co-temporal one. Note the commas which indicate the necessary intonation boundary required for the concatenated interpretation:

- (124) *ki=gonj* *la=bapaʔ=hn*, *paloh* *knon=hn*
 3A=take A=father=3POSS flee offspring=3POSS
 His father took (it) (and) (his) child fled.
- (125) *ki=gonj* *la=bapaʔ=hn* *knon=hn*, *paloh*
 3A=take A=father=3POSS offspring=3POSS flee
 His father took his child (and they) fled.

c) The two verbs which form the SVC must share the same set of arguments. There can be potentially only one subject and one object, where applicable, per SVC.

d) Oblique NPs pertain to the SVC as a whole:

- (126) *ki=gonj* *?yot* *he?* *dat*
 3A=bring return AT:above house
 She took (him) back up to the house.

e) Modification in terms of negation, aspect or modality is never ambiguous, since the modifier always has scope over both verbs, as with *gonj* and *masak* in (127).

- (127) *da?* *dapat* *ye=gonj* *masak*
 NEG manage 1A=bring cook
 (It) wasn't possible (that) I engage (myself in) cooking.

f) The maximal number of verbs attested in a serial complex is two. In (128), the third verb belongs to the following concatenated clause, reflected by an intonation break and indicated in the text by the presence of the comma.

- (128) *ki=gonj* *lumpot*, *paloh* *?a=kəh* *no?* *nəy*
 3A=bring leap flee DET=3 this before
 He took (it) leaping, (and) fled/(in order to) flee, the aforementioned him.

g) Only one interpretation of a SVC is possible. There is never any evidence of pragmatic ambiguity unlike that of concatenated clauses, see §11.4.

11.3.2 Types of serial verb construction

The three types of SVCs attested in Semelai are a) manner serialisation (§11.3.2.1), b) motion serialisation (§§11.3.2.2–3) and c) resultative serialisation (§11.3.2.4). The various types are discussed below.

There is a degree of variation in the semantic transparency of the supplementary verbs, consistent with the tendency to grammaticalisation or lexicalisation noted for this construction (Durie 1997: 290). Some of the supplementary verbs assume a radical change in meaning, for example *cəŋ* 'immediately' (§11.3.2.1), while in others the semantics of its function as a main verb are still evident in varying degrees, for instance purposive motion (§11.3.2.3) and resultative serialisation (§11.3.2.4).

11.3.2.1 Manner SVCs

Manner SVCs provide supplementary information regarding the manner in which an event takes place. The supplementary verbs change meaning as demonstrated in Table 11.1. There are two types of supplementary verb, those that precede the principal verb, labelled *V₁*, and those that follow the principal verb, labelled *V₂*.

TABLE 11.1 MANNER SERIAL VERBS

MAIN VERB	SERIAL VERB	
	<i>V₁</i>	<i>V₂</i>
<i>sec</i> 'to steal'	'surreptitiously'	
<i>?yot</i> 'to return'	'revert'	
<i>pay</i> 'to set (food) aside'		'for now'
<i>cəŋ</i> 'to pass through'	'immediately'	'completely'
<i>kmoŋ</i> 'to end, finish'	'exhaustively' (S/O)	'all up (O)' 'totally (S)'

cəŋ 'completely' also functions as a post-verbal modifier in negated clauses with similar semantic consequence (§10.4.1). Neither *sec* 'surreptitiously', *?yot* 'revert', *pay* 'for now', nor *kmoŋ* 'all, totally' function as modifiers in any other capacity.

The various manner SVCs are examined in turn below:

i. *sec* 'surreptitiously' As a main verb *sec* is the transitive verb 'to steal':

- (129) *ki=sec* *dwet*
 3A=steal money
 He stole money.

In a SVC construction, *sec* means 'surreptitiously'. The first example is intransitive (130), the remaining are transitive (131)–(132).

- (130) *dehn* *sec* *paloh*
 3piS surreptitiously flee
 They surreptitiously fled.

- (131) *ki=sec* *bəy* *ha?* *kubur* *?ate*
 3A=surreptitiously dig AT cemetery earth
 He surreptitiously dug the earth in the cemetery.

In the multi-clausal construction in (132) the proclitic is elided.

- (132) dos kaka? deh ha2=sən, sec h2gl? bkɔl=hn
come EZ 3pl AT=SC surreptitiously give provisions=3POSS
Their elder sister there came (and) surreptitiously gave him provisions.

ii. ʔyot 'revert' The intransitive main verb ʔyot means 'to return to one's original point of departure':

- (133) kəhn ʔyot
3S return
He returned/went back.

As a supplementary verb in a SVC ʔyot means 'revert' in terms of returning to an original state, or resuming an activity. It is translated as 'again' or 'to go back to ...'. In conjunction with an intransitive verb, it indicates the return to an original state:

- (134) A woman had stood up to do something:
kɔrɔdor ʔyot dm-dəm
be.female<NMZ> revert IMPERF-lie.down
The woman lay down again.

- (135) A child had been through a period of illness:
ʔyot gmuk
revert be.fat
(He) was fat again.

The following examples are transitive and indicate the resumption of an activity:

- (136) A man hasn't been able to find what he is seeking in the forest. He tries again.
ki=ʔyot reŋ ʔen bri
3A=revert seek LOC forest
He went back to seeking (it) in the forest.

- (137) A man, exhausted from trying to fell a tree, lies down in the incision and falls asleep.
ki=ʔyot kantup la=pat³ir ke
3A=revert close.over A=Sindora coriacea that
That Sindora coriacea tree closed back over (him).

iii. pay 'for now' In a SVC the verb pay 'for now' indicates an action which is carried out, but for the moment has no further consequence; it is translated here as 'for now'. As a main verb, pay means 'to set (food) aside':

- (138) tʃipal syɔk, da? ki=pay ?adi2=hn
leave trace.of.food NEG 3A=set.aside VS=3POSS
Traces of food were left. He hadn't set (any of it) aside (for) his younger sibling.

As a supplementary verb pay always co-occurs with a transitive verb, which it logically follows:

- (139) A were-tiger has killed the mother, but doesn't eat the body straight away.
ki=bunuh pay ʔma? ʔen
3A=kill for.now mother just
He just killed the mother for now.

- (140) ki=cək pay gonroŋ ha? tbəŋ
3A=cast for.now fishing.rod AT river.bank
He cast the fishing rod on the riverbank for now.

- (141) ki=rə?wən pay
3A=abandon for.now
He abandoned (her) for now.

iv. cəŋ 'immediately' As a main verb cəŋ is the intransitive verb 'to pass through; to pass between' (142).

- (142) The speaker uses imagery associated with pushing a boat to land in order to get people on the verandah to move aside for him to pass.
srsar ji hs? tpi?1 ye ga=cəŋ
push.to.land 2 AT:above edge 1 IMM=pass.through
Push yourselves to land! I'm going to pass through.

Example (143) demonstrates the use of cəŋ in conjunction with the negator da? 'NEG', meaning 'not at all' (§10.3.1.1).

- (143) da? de=pdu?i? cəŋ la=kəh
NEG 3pIA=bother completely BCS=3
They weren't bothered at all by her.

In a SVC the verb has the meaning 'immediately, straight away without further ado'. The following examples illustrate its use. In (144) which is intransitive, the third person pronominal is post-verbal.

- (144) cəŋ jtək ʔa=kəh
immediately sleep DET=3
As for him, he immediately slept.

In the transitive clause the supplementary verb hosts the pronominal proclitic:

- (145) ki=cəŋ jaŋ la=cəŋ deg ke
3A=immediately bark.at A=dog like that
The dog immediately barked at (it) like that.

In (146), the consecutive multiclausal construction consists of the serial verb clause, which is concatenated with a preceding clause.

- (146) ʔyot, cəŋ ca
return immediately eat
(He) returned (and) immediately ate.

v. cəŋ 'completely' cəŋ can be a V₂ supplementary verb meaning 'completely', or 'in total' in intransitive clauses as in (147), and 'all O's V-ed at once' in transitive clauses, as in (148).

- (147) There is the implication that there is no more cassava whatsoever; none in the house, none in the swidden.
hubi? kməŋ cəŋ
cassava finish completely
The cassava is completely finished.

- (148) *The protagonist is presented with a number of items which he carries all at once:*

kis=karnik cəj la=miskin
3A=carry.on.hip completely A=pauper
The pauper carried (them) all at once on his hip.

vi. *kmɔŋ 'exhaustively'* *kmɔŋ* is an intransitive verb meaning 'to end, to finish'. The first example is of *kmɔŋ* in its function as a main verb:

- (149) ye ga=kmɔŋ mn-(t)ulis
1 IMM=finish IMPERF-write
I'm going to finish writing.

In a serial verb construction it expresses the 'complete or entire affectedness' of absolute participants under its scope. In the intransitive clause it means 'every S to be V' as in (150), and in the transitive clause it means 'to V every O' as in (151). Alternatively, the two meanings could be conflated as 'to (be) V to X's end'. This SVC is always used in contexts where the referent is a set of individuated entities, and is glossed here as 'exhaustively'. It is unclear whether *kmɔŋ* should be treated as a manner adverbial or an adverbial universal quantifier 'all'. In order to provide a natural translation, it is translated as 'all', or 'every one'.

- (150) ?ma?bapa? kmɔŋ kʰbəs
parents exhaustively be.dead
(Her) parents were (both) dead.

- (151) *We've been eating bread with our afternoon tea. Someone is still hungry, and their request for more is answered with the following. The implication is that while the bread in question has been eaten, there may be more elsewhere, e.g. in the kitchen, or at the shop.*

rɔti? da? da? wo?. kmɔŋ br-ca
bread NEG EXIST longer exhaustively MID-eat
There's no more bread. (It's) all been eaten.

As (151) reveals, this serial construction is distinct from all other serial constructions because although in the transitive clause the 'supplementary' verb is in initial position, it does not host the pronominal proclitic. Compare the example with *kmɔŋ* in (151) with that in (152) where the supplementary verb *?yot 'revert'* hosts the proclitic *ki= '3A='*.

- (152) ki=?yot bʰɔŋ pʰre?
3A=revert be.afraid ghost
She reverted to being scared of the ghosts (of those who have died a violent death).

In transitive clauses, the O tends to occur in pre-verbal position:

- (153) deney, lyaq raʔ-tʰey raʔ-k̥et, kmɔŋ ki=todɔŋ
wall crack COMP-be.big COMP-be.small exhaustively 3A=cover
(As for) the wall, the cracks big (and) small, he covered every (one).

- (154) *Some in-laws provide their destitute nephew with everything he will need to set up camp on his own:*

baba t̥mpɔh lutoš, baraq?aden, kunor, paran, timon, lada?
paddy seven grain whatever gourd gourd cucumber chilli

kmɔŋ de=h?gi? ʔen kəh,
exhaustively 3plA=give LOC 3

cə mə=ʔikur, blahan tuy, lemen tuy
dog one=CLF blow.pipe one spear one
Seven grains of rice (and) whatever, pumpkins, gourds, cucumbers, melons, chilli, they gave everything to him, a dog, a blowpipe, a spear.

In (155) *kmɔŋ* is used in its deverbal form *k<ŋŋ>mol* (*finish<NMZ>*) 'the action of finishing', with what appears to be the same meaning.

- (155) dawen k<ŋŋ>mol kə=pjke?
never.mind finish<NMZ> 2fA=pick
Never mind (if) you pick (them) finishing every (one of them).

vii. *kmɔŋ 'all (O) up; totally S'* *kmɔŋ* 'to finish' is a *V₂* supplementary verb expressing 'the carrying out of an action to full effect, to a point where it can proceed no further'. It is only attested with the two verbs of consumption, *ca* 'to eat' and *jɔh* 'to drink', in 'to eat up/drink until everything is finished', and the predicate *sot* 'to metamorphose (into) s.th., completely'. In this respect, this construction could be considered idiomatic given its low applicability.

Although there is an obvious syntactic difference between this use of *kmɔŋ* and that in the previous section (following rather than preceding the predicate), the meanings partially overlap. Cross-linguistically, 'totally V' is a common extension with singular referents.

In the corpus the object is never represented in the transitive clause:

- (156) ki=ca kmɔŋ la=pədɔŋ
3A=eat totally A=tiger
The tiger ate (it) all up.

11.3.2.2 Associated motion serialisation

The associated motion construction shares the syntactic features of the manner serial verbs of the previous section. The principal verb *goŋ* 'to bring/take; to carry' is a transitive verb of displacement. The supplementary verb in the complex is either an intransitive verb of motion, expressing either manner or orientation of motion, or a process verb.

In (157) *goŋ* 'to bring/take; to carry' is used as a main verb in a simple clause.

- (157) crɛh ki=goŋ teʔen habu?
fish 3A=bring TO:down kitchen.
She brought the fish down to the kitchen.

As the principal verb in an SVC, *goŋ* functions as a transport verb. It is always the *V₁*, and the clause is always transitive because it is the principal verb. The most common usage is with an intransitive verb of motion to derive a Transport SVC. When used in conjunction with an intransitive process verb the SVC expresses engagement in an occupation. The functions are summarised in Table 11.2 and described in turn below.

TABLE 11.2 ASSOCIATED MOTION SERIAL VERB CONSTRUCTIONS

MAIN VERB	SERIAL VERB	
	+ motion verb	+ process verb
gɔŋ 'to take/bring, carry'	a) 'to take s.o./s.th. V-ing' b) 'to take o.self V-ing'	'to engage o.self in V-ing'

It should be noted that SVCs based on verbs of transport such as 'bring/take, carry' are common across SVC languages, see Jarkey (1991) for Hmong, and Durie (1985: 241) for Acehnese. In Acehnese the serial construction *ba* 'to take, carry' followed by an intransitive verb of motion, results in a transitive combination (as for Semelai below) where the undergoer is 'the thing taken'. Durie notes that this construction is unproductive in Acehnese, and cites the only co-occurring verbs as *plueng* 'to run', *phō* 'to fly' and *jak* 'to go'.

i. *gɔŋ* 'to take s.o./s.th. V-ing' When *gɔŋ* 'to bring/take; to carry' co-occurs with a verb of motion it means 'to transport s.th./s.o. in *V₂* manner'. The 'manner' is the method, direction or orientation of movement. The following verbs are attested in my corpus for this construction. The first five are verbs of locomotion, with no trajectory specified: *swak* 'to walk, move', *pər* 'to fly', *lumpot* 'to leap', *paloh* 'to flee' and *c'bom* 'to swim'.

The remaining verbs convey direction: *?yot* 'to return', *masuk* 'to enter' and *?endol* 'to enter a house'. One motion verb which cannot occur in this construction is *dos* 'to come, arrive'.

This is a manner construction, not a causative one, for the essential meaning of *gɔŋ* is 'to displace O'. Both the A and the O are the objects in motion. The following multiclausal construction illustrates the two uses of *gɔŋ*, as an SVC in the first clause, and as a main verb, a verb of displacement, in the second:

- (158) ?psol ga=gɔŋ ?yot lawok kmpən, ki=gɔŋ knət
if IMM=bring return meat wife 3A=bring child
Whenever (he) was going to bring back meat (for) (his) wife, he would take the child.

The O may be animate as in (159) or inanimate as in (160).

- (159) bɔy ma=gɔŋ ?endol ye=?en!
NEG:IMP IRR=bring enter.house 1=AUG
Don't bring (him) inside our (house)!

- (160) ki=gɔŋ ?yot ma=g<n>endol kayen ke
3A=bring return one=length<NMZ> cloth that
He brought back a length of cloth.

The syntactic characteristics of this SVC are the same as the general features outlined in §11.3.1.

ii. *The reflexive use of gɔŋ* Generally the action in a *gɔŋ* SVC is understood to involve a second argument, i.e. an O, but the construction can also be used in clauses with only one actant, the A, in a reflexive construction. There is greater semantic freedom in the choice of the supplementary verb in this construction, compared to that of the associated motion type in i. above. The supplementary verb is

intransitive, either a verb of motion as in (161)–(163) or a process verb in an imperfective form as in (164)–(165) below. The reflexive nature of the clause is not formally marked and in the case of verbs of motion can only be ascertained pragmatically.

- (161) *A wife is fed up with her husband, so she takes herself off to the forest.*

ki=gɔŋ swak kloc bri
3A=take walk inside forest
She took (herself) walking in the forest.

- (162) *A toddler has climbed up onto the railings of the verandah. The speaker alerts her mother:*

mnjipē? mstl? ki=gɔŋ grək, ?mot ?pəs
small.up.high must 3A=take fall mount be.high
Small and up high! She must bring (herself) to fall, getting up high.

- (163) ki=gɔŋ ?m-?əm la=pədɔŋ
3A=bring IMPERF-lower.haunches A=tiger

The tiger lowered (his) haunches.

When the *V₂* is an intransitive process verb the meaning is 'to engage oneself in *V₂*'. The context of these clauses usually implies that there is a specific motivating factor behind the event, as indicated in the following examples.

- (164) *A man is going to test his wife to see if an exorcism was successful. For this he needs a fish.*

ki=gɔŋ mn-jala? la=knlək neŋ
3A=bring USE-cast.net A=husband before
The husband (from) afore took himself fishing with a casting net.

- (165) *A woman relates how in earlier times she would busy herself tapping rubber with the specific intention of getting some money together before the arrival of the itinerant traders.*

ye=gɔŋ mn-(t)ɔrh
1A=bring IMPERF-tap.rubber
I'd take (myself) rubber tapping.

iii. *Cultural activity serialisation with swak 'to go'* A second type of occupation serialisation is found with the verb *swak* 'to go, move' and a verb expressing a typical cultural activity. The second verb is always an intransitive process verb. It is difficult to establish if one verb has primacy over the other, given that both verbs are intransitive.

- (166) kp lpc swak mŋabɔŋ
2f already go IMPERF:cock.fighting
You had already gone cock-fighting.

This construction is only possible with processes which are carried out in various locations over a journey or expedition. One cannot say **swak r<?>mɔŋ* (go reap<IMPERF>) 'to go reaping' because this activity only takes place in one particular prespecified location, and not in a series of locations. In contrast, the collection of rattan, *swak b-dre* (go HAVE-rattan) 'to go rattan collecting', involves

an expedition into the forest, alternating between periods of travelling, seeking the rattan and then the actual collection of it.

- (167) swak b-dre
go HAVE-rattan
(They) went rattan collecting.

- (168) *The tubers sought are non-cultivars which grow in the jungle.*
swak by-bny, bapaʔ=hn 2h-ʔɔh
go IMPERF-dig father=3POSS IMPERF-shoot.blowpipe
(They) were digging tubers, (and/while) their father was hunting with his blowpipe.

Other verbs attested in this SVC are:

- | | | | |
|-----------|--------------------------------|----------|-----------------------|
| brintay | 'to hunt, stalk' | m-rapah | 'to live by foraging' |
| b-buru? | 'to hunt with a dog and spear' | m-rando? | 'to live by foraging' |
| b-pranjop | 'to basket-trap' | | |

- (169) swak ʔh-ʔɔh kloc bri
go IMPERF-blown.pipe inside forest
(They) went blowpiping in the forest.

- (170) *An indolent son has been turned out of his father's house.*
ʔleʔleʔ, swak dom knkon kke,
eventually go AFF child that

swak m-rapa-rapah kloc bktak
go IMPERF-RDP-forest.litter inside undergrowth
Eventually, the child indeed left (that place), (and he) wandered foraging (and
destitute) in the undergrowth.

11.3.2.3 Purposive motion

In the Purposive motion SVC the supplementary verb, which is a verb of motion, is followed by an intransitive or a transitive verb. The second verb is the principal verb given that it determines the overall transitivity of the clause.

The actions are not co-temporal in the sense that they are not necessarily simultaneous; however recalling the characterisation of the SVC in §11.3.1 the construction represents two facets of one event, and not two independent events. The purposive motion SVC involves the expression of directed motion required in order for the subject to be in the right location to carry out the activity of the principal verb.

- (171) *A woman approaches a tree intending to look around it to see what is on the other side.*
ga=k1=dloh lewah
IMM=3A=go.across look.around
(She) was going to go over to look around (it).
- (172) *Malays used to travel upstream to the lake to buy Semelai-made dammar torches.*
gop kwale bra yor rŋ-reg ha? tasik
malay Kuala Bera ascend IMPERF-seek AT lake
Malays (from) Kuala Bera would come up(stream) seeking (them) at the lake.

The purposive SVC is not the same as the often ambiguous paratactic 'purposive' discussed in §11.4.1.2, because here the motion verb is inside the verb complex, demonstrated most clearly by the placement of the subject in initial position.

The transitivity of the principal verb V₂ determines the transitivity of the clause. In (173)–(174), the supplementary verb indicating transverse movement dloh – loh 'to go across' is the same, and it is the principal verb, m n-(t)ugp1 (IMPERF-dibble.rice) 'to dibble rice' in (173) and yok 'to take' in (174), which yields the differing transitivity of the respective clauses.

- (173) kəhn dloh mə-(t)ugp1
3S go.across IMPERF-dibble.rice
He went across to sow rice.

- (174) ki=loh yok knon
3A=go.across fetch offspring
He went across to fetch the child.

Further evidence of the fusion between the two verbs is the inability of a goal PP to intervene. In the following example, the elided pronominal would be an A:

- (175) ʔleʔleʔ loh ʔantat mə=rmoi təʔen jŋ̩ dak
then go:across take REL=be.male TO:down foot water
Then (he) went across to take (the one) who was male down to the lower reaches.

The positioning of the PP between the two would result in a different reading, nonsensical in this instance given the orientation of the verb and the PP do not match:

- (176) *ʔleʔleʔ loh təʔen jŋ̩ dak, ʔantat mə=rmoi
then go:across TO:down foot water take REL=be.male
*Then (he) went across down to the lower reaches of the water, taking (the one)
who was male.

The expression of integral motion is another manifestation of the preponderance in Semelai with the mapping of location and orientation (§8.3).

The participant in the following example heads down to the kitchen, which is located at a lower level in the house:

- (177) "ʔəp Jake", kʰləq ga=sar ca
If then QUOTE IMM=descend eat
"I, in that case," (she) said, "am going down to eat."

Water, both wells and the lake, are always located lower on the vertical axis:

- (178) dehn sar hūm
3pIS descend bathe
They went down to bathe.

The expression sar hūm (descend bathe) 'to go down to bathe' could be considered lexicalised, evidenced by its ability to be affixed in the following way by the circumfix b<>aŋ 'TOG': b<ŋ>saŋ.hūm>aŋ (TOG<descend.bathe>) 'all go down to bathe together'. This appears to be an isolated case; it is not possible for other forms of serialisation where the affix can usually only be marked on the main verb: *b<sec.2end>aŋ (TOG<surreptitiously enter.house>), the correct form is sec

b<?end>an (surreptitiously TOG<enter.house>) 'all together (they) surreptitiously entered the house'.

11.3.2.4 Resultative serialisation

Resultative serialisation expresses a resultant state, which is usually the intended result of the action expressed by the initial verb:

- (179) boy ləh dk̩es k̩!
 NEG:IMP move:lateral be.close 2f
 Don't come close, you! (Lit. Don't move laterally (so as) to be close, you!)

This form differs from previous forms of serialisation discussed above, by virtue of the fact that the second position can only be filled by an adjective; usually either a dimension or a physical property adjective, as in (179)–(180) and (181) respectively, but not a colour term.

- (180) boy ma=ra?wen ?ləp!
 NEG:IMP IRR=abandon be.far
 Don't wander (so as to) be far (from me)!

The serial construction may comprise two intransitive verbs with a shared S, as in (179) above, or the sequence may be O, S, as in (181) where the O of the transitive verb is co-referential with the S of the second verb in the sequence.

- (181) ki=sayor cin, ki=təŋ
 3A=stew.with.vegetable be.cooked 3A=dish.out
 She stewed the vegetable (so it) was cooked, (and) she dished (it) out.

11.3.2.5 Synonymic serialisation

A marginal example of serialisation is 'synonymic serialisation', the pairing of two verbs that are closely related in meaning, generally near synonyms. Given that the verbs are synonyms, they have the same argument structure. In all the examples collected so far, both verbs are transitive.

The clause may have a resultative interpretation (182–183), the result expressed by the second verb.

- (182) ki=sep ca wəc
 3A=suck eat guts
 She sucked (and so) ate the guts (of the fish).

Both verbs in (183) mean 'seek', but the second conveys a sense of fruitlessness. The meaning of *lenjŋɔŋ* 'to seek in vain', prevents a possible reading where *renj* means 'to find':

- (183) ki=renj lenjŋɔŋ cɔ? habu?
 3A=seek seek.in.vain AT:below kitchen
 She sought (and) sought in vain down in the kitchen.

In (184), a man is performing an exorcism on his wife. The two verbs used here have essentially the same meaning, 'to hit with an implement', the difference being a twig in the first case, and a stick in the second.

- (184) ki=r<h>poh tampon rɔm midur
 3A=beat.with.twig<IMPER> beat.with.big.stick wrth plant sp.
 He beat (her) with small twigs (and) he beat (her) with big sticks of *Goniothalamus macrophylla*.

11.4 Concatenated clauses

Clause concatenation is a string of two or more clauses, which a speaker chooses to represent as a string of events in a single unit under a single intonation contour.⁶ This construction is referred to as a sentence. Each clause within the sentence is marked off from the other clauses in the construction by a non-final pause. A comma indicates the clause boundary:

- (185) ki=c7ūm, ki=gɔŋ ?yot
 3A=wrap 3A=bring return
 He wrapped (it) (and) he took (it) home.

Semantically, each clause in this construction is an independent proposition. The concatenated construction can convey the following range of semantic propositions:

- consecutive (§11.4.1.1)
- purposive (§11.4.1.2)
- conditional (§11.4.1.3)
- simultaneous (§11.4.1.4)

The independent clauses are simply juxtaposed in a fortuitous manner, and there is no overt marking of the semantic relationship holding between the clauses. In the absence of any overt markers of the relationship between the clauses, they are ordered in relation to the logical sequence of events, i.e. chronologically – e.g. in the conditional the antecedent must precede the consequent – and are not interchangeable under any circumstances.

Understanding these clauses depends on pragmatic interpretation based on knowledge of the discourse context, or specific cultural knowledge (§11.4.1). Depending on the context of the story, the example in (186) could be interpreted with either of the glosses provided, a sequential reading as in a) or a purposive reading as in b).

- (186) ki=ra?wen, masuk kloc lataj
 3A=abandon enter inside grass
 a) He abandoned (them), (and) entered the long grass.
 b) He abandoned (them), (in order to) enter the long grass.

Muticlaue constructions, which overtly code the semantic relationship between the clauses – conditional, consecutive, cause etc. – with a connective, are discussed in §11.5.

⁶ As stated in the introduction to this chapter, juxtaposed or paratactic constructions are not analysed as serial constructions in this account.

11.4.1 Semantic relationships in concatenated clauses

The four semantic types of concatenated clauses are described in the following sections.

11.4.1.1 Consecutive concatenation

Consecutive concatenation involves two chronologically sequenced clauses which are simply juxtaposed in a paratactic manner, as shown in the following examples. The two clauses must share the same temporal or spatial setting.

- (187) *ki=tibg knon, ki=gog paloh*
3A=grab offspring 3A=bring flee
He grabbed (his) child (and) took (him) fleeing.

- (188) *ki=moc, krwbl rom luluk, rom rapah, rom 2ate*
3A=dive.in emerge WITH mud WITH litter WITH dirt
He dived in, (and) emerged (covered) with mud, litter and dirt.

11.4.1.2 Purposive concatenation

The purposive is slightly idiosyncratic, by virtue of its differing semantics, which in turn is reflected syntactically. The purposive is an event that has not yet come into effect or been realised. The event in the first clause must take place in order to realise the second. Hence, the two actions are necessarily conceptually closer than two separate actions in a chronological sequence of independent events. To obtain a purposive reading the agent must be co-referential in each clause in the construction. There is frequently an absence of marking of person, both by clitic and by free pronoun, in the subsequent clause given that it is understood to be co-referential with the first clause. The purposive can only be expressed in this type of construction.

- (189) *2ep ga=b-layar, ga=reŋ pakayan knon*
If IMM=HAVE-sail IMM=seek clothes offspring
I'm going sailing, (to) seek clothes (for) (our) child.

- (190) *ki=laget, j?oy rakit, ?luc sbo?oŋ krweŋ*
3A=break.off make raft pass.to otherside [name]
He broke (it) off (to) make a raft (to) pass to the otherside (to) Keruing.

- (191) *ye=?en ga=?en tla?ga?, ga=poy hatap ?arch*
1=AUG IMM=TO:down well IMM=make thatch be.new
We are going (to go) down to the well to make a new thatch.

The following example is an imperative. Pronominals are not used in the imperative and this extends to purpose complements adjoined to imperatives (§10.6.1):

- (192) *yok kwen, Jon ha? knon!*
take biscuit give AT offspring
Take a biscuit to give to (your) child! Or: Take a biscuit (and) give (it) to (your) child!

Purposive clauses share the same argument in each clause, whereas consecutive clauses do not have this restriction. This is demonstrated in (193) where a sequential relationship holds between the first two clauses, where S is not co-referential with A,

and a purposive relationship holds between the medial and final clauses where the same actant is understood for both.

- (193) *[i?ec sma? swak], [ki=gog bsi?], [conkil pintu?]*
[already people walk] [3A=bring metal] [pick door]
(When) the people had already gone, he took (a piece of) metal (to) pick/(and) picked the door.

11.4.1.3 Conditional concatenation

Conditional actions are expressed by two chronologically sequenced clauses juxtaposed in a paratactic manner. They are also expressed in a paratactic construction with a lexical connective (§11.5.3.2).

The clause expressing the condition, which is bracketed here, always precedes the consequence reflecting the temporal structure of the conditional. Each example could equally be interpreted as a consecutive clause, indicated by the alternative translations. Only the context determines the translation, although in threats it is common to express conditionals as sequences (194)–(195). As conditionals, these clauses have a predictive reading.

- (194) *[“ko paloh”] k?leŋ “ga=2ep=bunuh”*
[2f flee] QUOTE IMM=1FA=kill
(a) “(If) you flee,” (he) said, “I am going to kill (you).”
(b) “You flee,” (he) said, “(and) I am going to kill (you).”

- (195) *“boy!” k?leŋ kak?2, [de=?yəŋ b?l?u?], de=pren” k?leŋ*
NEG:IMP QUOTE EZ [3plA=hear friend] 3plA=angry QUOTE
(a) “Don’t!” said elder sister, “(if) the friends hear, they will get angry,” (she) said.
(b) “Don’t!” said elder sister, “the friends will hear (and) they (will) get angry,” (she) said.

In (196), the condition is negated.

- (196) *da? ki=?arah la=c?i?gu?*
NEG 3A=banish A=teacher

ga=ki=bunuh la=bapa? dom knon kke
IMM=3A=kill A=father AFF child that
(If) the teacher hadn't banished (him), the father would have killed the child.

The verb in the clause expressing the condition may be cliticised by the irrealis marker *ma=* ‘IRR=’. The presence of the irrealis is associated with a more hypothetical conditional (§10.2.2.5):

- (197) *ma=kali?, swak! da? ma=kali?, p?ot!*
IRR=be.brave walk NEG IRR=be.brave stay
(If) (you) are brave, go! (And) (if) (you) are not brave, stay!

11.4.1.4 Simultaneous concatenation

Simultaneous actions performed by the same subject are expressed by the paratactic juxtaposition of two clauses (see §11.5.3.3 for clauses with different subjects). The simultaneity is not co-temporal, but represents potentially alternating or intermittent action. The verb in the second clause, which expresses the intermittent action, is in the imperfective form in all the following examples (see §5.1.1).

- (198) dm-dəm, ja<cl>yɔk
IMPERF-lie look<IMPERF>
lying down (and) looking about

This construction clearly presents the difficulty in determining the type of structure we are dealing with. The approach here takes the clauses as concatenated. Alternatively, as indicated in the introduction, they could be interpreted as a type of serial construction. The deciding factor against an analysis like that of the purposive motion SVC would be the cliticisation of the pronominal proclitic to the V₁. However this only occurs with the SVC when the V₂ is transitive, which is never the case in this instance.

- (199) mot ye jtek, təŋ yŋ-yŋ
eye 1 sleep ear IMPERF-hear
My eyes were closed, (while) (my) ears were hearing. (Lit. My eyes were asleep ...)

- (200) A mother has asked her son to fetch a water bottle. She watches him return, taking swigs here and there as he approaches. She remarks:

ki=goŋ, j<cl>ʔɔh
3A=bring drink<IMPERF>
He is bringing (it), (taking) swigs.

11.4.2 The syntax of the concatenated clause

Clause concatenation is not open ended and in most cases the construction is limited to two clauses, although up to three may be attested in a construction with a consecutive reading. Clauses entering into a concatenated multiclause construction can be any combination of a non-verbal, simple verb or a serial verb clause. The first example in (201) is of a non-verbal clause followed by a simple verbal clause, see also (191) above.

- (201) A child alerts his father:
[pəðəŋ he? dɔl]NVC, 2ma? kʰbas
[tiger AT:above house] mother be.dead
(A) tiger (is) up at the house, (and) mother is dead.

In (202), a simple verb clause is concatenated with an SVC.

- (202) ki=rus, [goŋ ʔyot]svc
3A=drag [bring return]
a) She dragged (it) (and) took (it) home.
b) She dragged (it) (in order to) take (it) home.

The following points are important given that they form the central distinguishing factors between this form of verb concatenation and the serial verb construction discussed in §11.3. As stated previously, each of the concatenated clauses displays the syntactic features of a simple clause. Each verb has its own set of core arguments and obliques, which stand in direct relation to it; each verb may be marked independently of the other for person, negation, mood and aspect. Examples of these features are provided below, beginning with the representation of core arguments.

11.4.2.1 Core arguments

There is no restriction on which, or how many arguments are shared, and indeed none need be.

In (203) the subject in both clauses is the same, but the object of the first predicate is a knife, and in the second it is the tree which she is about to chop.

- (203) ki=toc hn=wŋy, ga=ki=pøk
3A=pull.out O=knife IMM=3A=chop
She pulled out the knife (and) was going to chop (the tree).

In (204) there are no shared arguments.

- (204) A dim-sighted man has inadvertently cast his fishing net over a woman on the bank, mistaking her for a fish. He removes the net and she consequently goes home.
ki=tarek hn=jalaʔ, krdror hne pon ʔyot
3A=pull O=cast.net woman THEN THEN return
He pulled the net (off) (and) consequently the woman thereupon went home.

The concatenation of clauses does not require any realignment of arguments through syntactic means into concordant roles; in other words there is no syntactic pivot and the clauses are simply concatenated as they stand. The following examples illustrate this point. When an intransitive subject has overt representation in the initial clause, it is post-verbal, as in (205) and (206). Note that in both these clauses the S is co-referential with the A of the second clause.

- (205) krwøl cəŋ, ki=ma<r>suk=sən hn=b2ləʔ
come.out mousedeer 3A=enter<CAUS>=SC O=friend
Mousedeer came out (and) he made his friend enter.

- (206) ʔyot dehn, pael puypŋ=hn
return 3pLS summons shaman=3POSS
They returned (and) summonsed their shaman.

The intransitive subject usually has zero representation in the second clause, irrespective of its status in the initial clause as in (207) and (208).

- (207) ʔyot, cəŋ sar hñm
return immediately descend bathe
(He) returned (and) (he) immediately went down to bathe.

If overt, the S in the second clause is usually pre-verbal, see (204) above.

In (208) the unrepresented S of the second clause is co-referential with the O of the first clause.

- (208) ki=tikam hn=kɔ? ke, kʰbas
3A=stab O=macaque that die
He stabbed that macaque (and) (it) died.

The presence of the proclitic in transitive clauses is necessary at the start of each new set of conjuncts. Each subsequent transitive verb may host a pronominal proclitic, although this is not a necessary condition, see (209)–(211) below. In example (209) the A is co-referential across the clauses, as is the O.

- (209) ki=yɔk, ki=lek, ki=jam mlɪk
 3A=take 3A=lick 3A=taste be.creamy
 She took (it), (and) she licked (it), (and) she tasted (it) creamy.

In concatenations of two transitive clauses where the A is co-referential, the external NP directly follows the first verb in the sequence, never the second:

- (210) ki=tarek la=kmpən, ki=gɔŋ ?yot
 3A=pull A=wife 3A=take return
 The wife pulled (it) (out) (and) she took (it) home.

While the pronominal proclitic is obligatory in the initial clause, it may have zero-representation in the subsequent clause, see also (190) above.

- (211) *A child calls out to his father who looks upward.*
 ki=bɪtɔy la=bapəʔ=hn, jɔk knɔn
 3A=look.upward A=father=3POSS observe child
 Their father looked upwards (and) observed the children.

The only instance where a pronominal proclitic is omitted under anaphora is when the same verb or a semantically similar verb is repeated in a partial repetition of an event. The maximum number of clauses concatenated in this manner is two, as shown in the following adjacent clauses taken from a continuous piece of text:

- (212) *A man assembles the various things necessary for the exorcism of his wife.*
 ki=yɔk gunting, yɔk, ?anu?, kacip
 3A=take scissors take [HES] betel.nut.clippers

 ki=yɔk bunjʌy, yɔk barang ?a=dɛŋ=cəʔ=sən
 3A=fetch aromatic.ginger fetch thing DET=like=EM=SC
 He fetched the scissors (and) fetched, umm, the betel-nut clippers. He fetched aromatic ginger (and) all sorts of things indeed.

In (213) the tiger repeats an action, but it is the orientation which is different in each clause:

- (213) ki=jɔk te ke, jɔk te nɔ?
 3A=observe TO:unspec that observe TO:unspec this
 He looked there (and) looked here.

Lexically represented objects which are shared by the clauses show up in the first clause in the sequence, in post-verbal position. Although both OV and VO are allowable orders in the simple clause (§9.2.1), the preferred pattern in concatenated clauses appears to be VO:

- (214) ki=kɔm crch, ki=gɔŋ ?yot
 3A=get fish 3A=bring return
 He caught fish (and) he took (them) home.

Zero representation of objects is common. In (215), the referent of the unrepresented object of the first two clauses, a knife, is contextually recoverable. In the third clause, the object, hmbacaj 'horse mango', is reintroduced and represented post-verbally.

- (215) ki=roc, ki=gɔŋ masuk, ga=ki=gɔs hmbacaj
 3A=pull.out 3A=take enter IMM=3A=pare horse.mango
 She pulled (the knife) out, she took (it) inside, (and) she was going to pare the horse mango.

The one exception in the corpus where the object is a constituent of the second clause and is in final position is given in (216). The motivation behind this is not clear, although it may be that the speaker wants to represent a more unified action and chooses to manipulate the order in this manner.

- (216) *Some children are frantically securing the door, trying to escape from what they have just realised is a were-tiger masquerading as their mother.*
 de=bək, de=hapti pintu?
 3plA=fasten 3plA=drive.wedge door
 They fastened (and) drove a wedge (under) the door.

11.4.2.2 Peripheral NPs

Each clause can have oblique NPs, which are placed alongside the respective verb independent of the other clauses. In (217) the verb in the first clause has a goal PP, and in (218) the final clause contains an instrumental PP.

- (217) ki=gɔŋ ta?en habu?, ki=bukpɔ?
 3A=bring TO:down kitchen 3A=open
 He took (it) down into the kitchen (and) he opened (it).

- (218) *A woman who was just a head has been killed in an accident. Her sons-in-law prepare the skull for burial as though it were a corpse:*
 de=basuh ktɔŋ, de=pa-yup, de=c?um rɔm kayen putih
 3plA=wash skull 3plA=EQUIP-cover 3plA=wrap WITH cloth be.white
 They washed the skull, they covered (it) (and) they wrapped (it) with white cloth.

11.4.2.3 Aspect, negation and mood

Each clause within the concatenation can be marked independently for aspect, mood and negation. In (219) the second clause is modified by the aspectual adverb lag! 'again'. This is an important distinction between concatenated clauses, and serialised constructions.

- (219) ki=yɔk, lag!, ki=jɔh
 3A=fetch again 3A=drink
 She fetched (water) (and) again he drank.

Where each clause is marked for a particular category, it must be indicated in each instance, e.g. the irrealis mood in the following examples:

- (220) ma=pte? hn=dol wɔy ten priŋka?an, ma=gɔs
 IRR=put O=handle knife LOC vagina IRR=pare
 (She) must have been putting the knife handle in (her) vagina, (while) paring (the mango).

The negator da? 'NEG' likewise only has scope over the clause in which it is marked, as in (197).

11.5 Connective clauses

This section deals with the concatenation of clauses where the interclausal semantic relationship is overtly marked – one clause contains a connective explicitly indicating the causal, temporal or conditional connection between the two clauses. The concatenation is considered a single complex construct, i.e. a sentence. It should be noted that some of these relationships, temporal, simultaneous and conditional, may also be expressed without the use of a lexical morpheme. Lexically unspecified concatenation is discussed in §11.4.

The same set of connectives is also used in interclausal relations as discourse connectives. They provide textual cohesion, marking temporal relations between consecutive independent clauses, for example marking back reference (§11.5.4).

In the following discussion, there is no classification of the connectives in terms of coordination or subordination as there is no evidence which allows the distinction to be convincingly substantiated. For example, the criteria for subordination suggested in Thompson and Longacre (1985) are the presence of: a) a grammatical morpheme, b) a special verb form and c) special word order. Semelai exhibits only the first of these characteristics, the morpheme. What would appear to be reduced verb forms are in fact simply standard cases of discourse-motivated elision as in the first two clauses in (221), and there is no evidence of changes in constituent order in the adverbial clause, as in (221)–(222). (Constituent order in basic clauses is discussed in §9.2.)

- (221) de=yok jalu, yok crch. de=ca.
3plA=take pork take fish 3plA=eat

?luc.tom de=ca, de=?intc
after 3plA=eat 3plA=exit.house

They took pork (and) took fish (and) they ate (it). After they ate (it), they left the house.

- (222) ?luc.tom br-gum, buh lagi? ha? lson
after MID-winnow put again AT mortar
After (it's) winnowed, put (it) in the mortar again.

Thompson and Longacre suggest that preference for ordering the adverbial clause in initial position in some languages is evidence of a special constituent order. This simply reflects a more general ordering in terms of the logical structure of the relationship. The clause containing the connective is dependent, but it is only dependent semantically as the secondary clause. This point is taken up below.

11.5.1 General features of connectives

The clause containing the connective tends to be the initial clause. Concatenated clauses containing connectives display one feature distinguishing them from the concatenated clause type in §11.4 – the ability to interchange the order of clauses where this is semantically plausible.

The connective is always the initial constituent in the clause. It is usually a single word, although there is one phrasal temporal connective ?luc tom (pass from) 'after'. In terms of etymology, only two categories are expressed by Semelai

morphemes: cause and condition; the majority of connectives are borrowed from Malay.

The trend for connectives or conjunctions to be borrowed from the dominant national language is noted for a number of languages in Thompson and Longacre (1985: 204–5). It is not possible to comment why this is so for Semelai, other than to note that in nearly all cases where there is a Malay form, there is no Semelai equivalent. The absence of these overt markers is not remarkable in itself as it is an expected areal feature. In most cases, this type of information is inferable, see §11.4. In fact, many of the terms borrowed from Malay are in turn borrowings from other languages, and many of these relationships can also be expressed in Malay without the use of a conjunction, e.g. a conditional clause may be expressed either paratactically – *ada kereta dia pergi* (be car he go) 'If there is a car he will go', or with a connective *kala* *ada kereta dia pergi* (if be car he go) 'If there is a car he will go'. It is also interesting to note that in the purposive clause (§11.4.1.2) for instance, the Malay conjunctions have not been borrowed into Semelai.

There is not always a match between the form and function of the borrowed term. There are instances where borrowings from Malay have semantic innovations in Semelai, which as far as can be ascertained are not attested in Malay. Consider the temporal *kna?* 'when' (§11.5.2.3). *kena* 'to affect (s.th.)' is a Malay verb, the use of which is often associated with external imposition or adversity toward the thing affected. The usage of *kna?* in (225) 'it happened that when'⁷ as a connective is not found in Malay. Likewise *sampai* 'until', a conjunction in Malay, functions also as a discourse connective *sampay* 'so that; so, at that point' in Semelai (§11.5.3). Both of the Malay usages of these words are attested in Semelai, *kna?* 'to affect (s.th.)' (223), and *sampay* 'until' (224).

- (223) dom ?ma? ke kna? kse?
AFF mother that affect trap
Indeed, the mother (had been) caught (by) the trap.

- (224) ga=b-gantog sampay k^bas
IMM=MID-hang until die
(They) are going to be hung until (they) are dead.

(225) is an example of both lexemes in their respective connective functions.

- (225) sampay kna? t^bey ?us, ki=rej dloj
so when be.big fire 3A=seek stick
So, when (it happened that) the fire was burning well, she sought a stick.

Lists of connectives are set out in each section below.

11.5.2 A note on enclitic =hn 'CONN'

One of the characteristic features of the connectives, both clausal and discourse, is their encliticisation by =hn 'CONN' when they express either temporal or conditional relations. The following connectives may host the enclitic:

⁷ *kna?* could also be construed as a verb 'it happened (that)'. This doesn't convey quite the same nuance.

?ie?le?	'so, eventually'
sampay	'so'
kna?	'it happened that ... when/when'
snih	'whilst'
sta?	'when, after'
?osol	'if, provided that'
yø?	'but, however'

Enclitic =hn 'CONN' appears to be optional in this function. It cannot be considered part of the connective as it only appears sometimes, in each case without any apparent change in meaning, the possible exception being the conditional connectives. The following examples illustrate its presence in (226) and its absence in (227).

- (226) [snih=hn dm-dəm], ki=tikam la=knlək təm lmej
[whilst=CONN IMPERF-lie] 3A=stab A=husband WITH spear
Whilst (the tiger) was lying down, the husband stabbed (it) with a spear.

- (227) [sni-snih by-bøy], ?ma?=hn ki=tłɔj la=podɔj
[RDP-whilst IMPERF-dig] mother-3POSS 3A=grab A=tiger
Right whilst (she) was yam-collecting, the tiger grabbed their mother.

A possible function of the enclitic could be to mark the clause as backgrounded information, against which the main activity is taking place. The clitic =hn does not appear to be functioning anaphorically as an elided pronominal, as is demonstrated in the following examples.⁸ Example (228) is of a temporal conjunction and an intransitive clause, juxtaposed with a transitive clause. Both clauses share the same subject.

- (228) kna?=hn dos he? ?i=miskin, ki=cəŋ ?ɔ?ŋɔr
when=CONN reach AT:above NM=be.poor 3A=immediately offer
When (he) came up to Mr Pauper, he immediately offered (Mr Pauper) (the things).

The enclitic on the connective in (229) is in an independent transitive clause.

- (229) sampay=hn ki=ca me=ha? jɔŋ tipɔ?
so.that=CONN 3A=eat REL=AT foot betel.service
So he ate the one at the foot of the betel service.

11.5.3 The adverbial connective construction

The adverbial clause serves to qualify an adjoined clause, providing context in terms of cause (§11.5.3.1), condition (§11.5.3.2) or temporal sequence (§11.5.3.3). The modifying clause can precede or follow the main clause, though with the exception of the causal clause, the predominant pattern is for the adverbial clause to precede. Variations on this pattern are noted below. The adverbial clause is always a simple declarative clause, apart from the condition clause containing the connective ?osol

⁸ Elsewhere hn functions as an enclitic in the following environments: a) on core-NPs where it functions to mark the third person possessor '3POSS'; b) enclitic on the verb, encoding the elided O (§9.3.1.2); c) in the S-form of third person pronouns, kəhn '3S' and dehn '3plS' (§6.6.1). It is not possible to account for the presence of =hn on connectives, neither when it occurs, nor that it occurs at all, in terms of any of the above functions.

'if, when', which may be in the irrealis (§11.5.3.2). Generally, the construction is limited to a maximum of two clauses.

The connectives are:

Causal:

la	'because, on account of'
----	--------------------------

Conditional:

məhn ~ mən	'if'
mənka ~ məka	'if'
?osol	'when(ever), if, in the event (that ...)' ← Malay <i>asal</i> 'provided (that)'

Temporal:

?luc (tom)	'after' (?luc 'to pass by' təm 'SRC')
sta?	'after, when' probably ← Malay <i>setelah</i> 'after, when it was over'
bila?	'when' ← Malay <i>bila</i> 'when'
kna?	'when' ← Malay <i>kena</i> 'to affect, strike (s.th.)'
sampay	'until' ← Malay <i>sampai</i> 'until'
samel	'while' ← Malay <i>sambil</i> 'while'
siraj	'while'
snih	'while'

11.5.3.1 The causal connective la= 'BCS'

The causal connective introduces either an NP or a clause which provides information about the reason or external causing factor of an event. The clause which describes the antecedent event is introduced by la= 'BCS' 'because (of), on account of, from'. Elsewhere, the morpheme la= is used to encode the post-verbal subject NP in transitive clauses (§9.3.1.3), as in the second occurrence in (232).

In the first two examples the cause is expressed as an NP:

- (230) *A boy has finally learnt how to behave properly in the underworld:*

- cʔen kaka?, bʔlu?, [la=keh]
be.happy EZ friend [BCS=3]
Elder sister (and) the friends were happy with him. (Lit. Elder sister (and) the friends were happy because of him.)

- (231) da? de=ca=hn [la=səc bapa? deh]

- NEG 3plA=eat=O [BCS=meat father 3pl]

- They didn't eat it because (it was) their father's flesh.

The causal clause is a full independent clause in the following examples. The independence of the clause is clear in examples (232)–(233), where conditions for co-referential deletion across the clauses do not apply.

- (232) ki=jłok, kməŋ lawok, [la=de=ca la=deh]

- 3A=observe finish food [BCS=3plA=eat A=3pl]

- He saw the food was finished, (which was) because they had eaten (it).

- (233) da? da? wo?, [la=kəhn swak, kira? b-pn-reŋ]

- NEG EXIST longer BCS=3S go figure HAVE-NMZ-seek

- (She) wasn't there any more, because he had gone, see, seeking things.

The subject of the second clause has been deleted under conditions of co-reference with the S of the initial clause in (234), and the understood O in (235).

- (234) sma? k^bbes, [la=j<r>me?] person die [BCS=drink.alcohol<IMPERF>]
The person died from drinking alcohol.

- (235) A woman is pregnant and craving Salacca fruit. She pesters her husband to take her into the forest in search of them, but he can't be bothered and gets fed up with her.
ki=cce? la=knlək=hn, [la=lan ca klubit?] 3A=resent A=husband=3POSS [BCS=want eat Salacca.conferta]
Her husband resented (her) because (she) wanted to eat Salacca fruit.

The clause containing the consequent generally precedes the reason clause as in the preceding examples. The reverse order where the antecedent occurs in second position, although possible, is infrequent:

- (236) A tiger suddenly realises that a woman's craving for Salacca fruit was due to her pregnancy.
"zəh dom, [la=k^bet=cə?]", k^bləj oh AFF [BCS=pregnant=EM] QUOTE
"kb=lan ca ple klubit?" k^bləj 2fA=want eat fruit Salacca.conferta QUOTE
"Oh indeed, because (you are) pregnant!" he said, "you wanted to eat the Salacca fruit."
(237) [la=zma?ləm yε], ki=b^bɔŋ [BCS=fourth.born.aunt 1] 3A=be.afraid.of
Because of my fourth-born aunt, he was afraid.

The clause containing the cause can stand alone without the consequent expressed. In this context the consequent has already been established in the text as indicated by the affirmative *dom* 'AFF'. In (238) the reason has not been stated explicitly, but the speaker makes the following inference:

- (238) "zah, dom=cə?", k^bləj, "[la=pn-zur kəh]" ah AFF=EM QUOTE BCS=NMZ-instruct 3
"Ah, indeed!" (he) exclaimed, "(it's) because of his instructions."

Example (239) gives a full illustration of where preceding utterances rather than a main clause provide the contrast. The second occurrence of *la*= 'A' cross-references the agent of the causative verb *ni<r>kah* (marry<CAUS>) 'to officiate a marriage'.

- (239) Mousedeer (MD) is up to his usual antics. Caught in a trap and not wanting to be eaten, he calls out to a passing macaque (M), hoping to trick him into swapping places. MD laments to M how he is awaiting his undesired marriage to the princess. MD suggests they change places and that M take advantage of such a prestigious alliance. The name of the princess, *zi=kondō?kandō?*, is derived from *kondō?* an avoidance word for a type of edible gourd. So stupid is M that he doesn't realise that the only 'marriage' he will end up in is with the gourd in the cooking pot!

MD: "zəp" k^bləj=hn "risaw"
If QUOTE=3POSS be.sad

M: "zoh mānde?" k^bləj kɔ? oh why QUOTE macaque

- MD: ["la=ga=ki=ni<r>kah la=twanko? [BCS=IMM=3A=marry<CAUS> A=king
rom ptri? zi=kondō?kandō?] WITH princess NM=gourd]
MD: "I," (he) said, "am sad."
M: "Oh, why?" asked Macaque.
MD: "Because the king is going to marry (me) to Miss Princess Gourd."

If there is more than one antecedent for a given consequent, each clause must be preceded by *la*= 'BCS':

- (240) da? ki=jon dom, wby kəh, [la=da? zilək], [la=?nat] NEG 3A=give AFF knife 3 [BCS=NEG be.good [BCS=bc.ugly]
(He) didn't give his knife indeed, because (it) wasn't good, because (it) (was) ugly.

In (241) the second occurrence of *la*= 'BCS' encodes the source of their fear.

- (241) dehn paloh, [la=de=b^bɔŋ] [la=pran]] 3S flee [BCS=3plA=be.afraid.of [BCS=war]]
They fled because they were afraid because of the war.

II.5.3.2 Conditional clauses

Condition is expressed either by a clause containing a conjunction, or alternatively by a paratactic construction as discussed in §11.4.1.3.

The conditional connectives are:

məhn ~ mən	'if'
mənka ~ məka	'if'
?osol=hn	'if'
?osol	'if, in the event (that ...)/when(ever)' ← Malay <i>asal</i> 'provided (that) ...'

The order of the clauses reflects the chronological sequence of events, so that the condition precedes the consequence (242), although in the case of *mənka* ~ *məka* 'if', the order is reversible (243).

- (242) [mənka yε=k^bom dwet], ?areh ?yot [if 1A=get money] only.then return
If I get money, only then (will I) come back.

- (243) Iməjı da? tahan, [məka da? b-sug?] c?co? [if NEG MID-wash.face be.infront
Teeth won't last if the face isn't washed beforehand.

The clause expressing the condition cannot occur independently.

Two types of conditional are identified, the predictive and the hypothetical. Table 11.3 summarises the distributional possibilities for the four conditional connectives and includes for comparison the possibilities for the paratactic construction discussed in §11.4.1.3. The numbers in the table refer to the relevant examples in the text. The division between realis and irrealis is in terms of morphological category.

TABLE 11.3 SUMMARY OF CONNECTIVE CONDITIONAL CONSTRUCTIONS

CONNECTIVE	MORPHOLOGICAL CATEGORY		
	REALIS		IRREALIS
	PAST	NON-PAST	
məhn 'if'	(248), (250)	(244), (249)	
mənka 'if'		(242), (243)	
?əsɒl(=hn) 'if'		(245)	
?əsɒl 'if'		(246)	(247)

All four of the connectives, as well as the simple paratactic construction, are available to express the conditional in non-past realis clauses. The connective məhn 'if' is the only possibility for the realis past, while mənka ~ məka 'if' is limited to the realis non-past.

Likewise, ?əsɒl(=hn) 'if' is restricted to the non-past realis and ?əsɒl 'if/when(ever)' may apply to the non-past realis or the irrealis. Note in passing that the polarity of either clause is of no significance to the choice of construction.

məhn, ?əsɒl=hn and ?əsɒl in realis clauses all express predictive conditionals, while mənka and ?əsɒl in irrealis clauses express hypothetical conditionals.

The following examples illustrate the similarity between məhn and ?əsɒl=hn. In (244) and (245) the clauses are predictive, based on the likely outcome. The first uses məhn, the second ?əsɒl, encliticised with =hn 'CONN'.

- (244) *The speaker is stating what will happen if a person's craving is not satisfied and a particular taboo transgressed.*

məhn da? ki=k^bom, ki=halaw kley
if NEG 3A=get 3A=chase millipede
If she doesn't succeed, a millipede will chase (her).

- (245) ?əsɒl=hn da? ki=k^bom, ki=halaw klay
if=CONN NEG 3A=get 3A=chase millipede
If she doesn't succeed, a millipede will chase (her).

In (246) ?əsɒl does not host the enclitic. Here it is used to express regularity or habitual activity 'when(ever), in the event (that) ...'.

- (246) [?əsɒl ga=gɔŋ ?yot lawok kmpan], ki=gɔŋ knon
if IMM=bring return game wife 3A=bring offspring
If/whenever (he) was going to bring game home (for) (his) wife, he took the child (with him).

Note that ?əsɒl 'if/when(ever)' may co-occur with the irrealis, but ?əsɒl=hn (if=CONN) only occurs in the declarative. The motivation for this is discussed in §10.2.2.5.

- (247) ?əsɒl deh ma=t-genen. wc-woc deh
if 3pl IRR=HAPP-disturb IMPERF-curl.up 3pl
If/whenever they happen to be disturbed, they curl up.

The following examples serve to further illustrate the distinctions set out in Table 11.3 above. The first example is of an event in the past.

- (248) *A woman has died giving birth and metamorphosed into a birth-vampire. Her husband retrieves his child by means of deception, but as the narrator states, he had no other choice.*

məhn da? ki=j^bɔy den ke, da? ki=k^bom
if NEG 3A=do like that, NEG 3A=get
If he hadn't done (it) like that, he wouldn't have got (him).

In (249), a non-past usage of məhn is illustrated in which the speaker makes a prediction based on past experience.

- (249) *An estranged wife is afraid to return to her village lest she encounter her husband who has previously abused her.*

"məhn ?yot, k^bas ?ap" k^bɔŋ
[if return] die If QUOTE
"If (I) return, I will die," (she) said.

Example (250) illustrates two conditional clauses. The first, which refers to a factual past event, contains the connective məhn. The second expresses conditionality with a paratactic construction, the conditional clause marked with the irrealis infers a degree of uncertainty, the hypothetical nature of the assertion.

- (250) *The communist insurgents were ruthless. The Semelai knew that if they didn't give them what they wanted they would be killed. They hoped that compliance would spare them.*

məhn da? de=jon baran, ga=de=bunuh, ma=jon, beh
if IRR 3plA=give thing IMM=3plA=kill IRR=give NO
If they hadn't given (them) something, they would have killed (them). If (they) gave, they wouldn't.

11.5.3.3 Temporal clauses

The following are some of the temporal connectives used in Semelai. The majority are borrowed from Malay, with the exception of ?luc tom (pass from), ?arch (to be new) 'at last, only then' and snih and siraj which mean 'while'.

The temporal connectives are:

?luc (tom)	'after' (< ?luc 'to pass by'; tom 'SRC')
?arch	'only then'
sta?	'after, when'
bila?	'when'
kna?	'when; (it) happened that ...'
siraj	'while'
samel	'while'
snih	'while'
sampay	'until'

The order of the clauses always reflects the chronological sequence of the events, one clause being subsequent to the other. Accordingly, in the first two examples the second clause, which contains the connective, is subsequent to the initial one:

- (251) ga=b-gantɔŋ, [sampay k^bas]
IMM=MID-hang [until die]
(They) are going to be hung until so that (they) die.

- (252) [bilə? de=trmpa? b?lu? ramay], de=lən la=b?lə?
 [when 3plA=encounter friends be.many] 3plA=desire A=friend
 Whenever (they) happened to encounter (his) many friends, his friends desired (her).

- (253) [sta? ki=k?om jampi?], kehn pon ?yot
 [after 3A=obtain incantation] 3S THEN return
 When he obtained the incantations, he thereupon returned home.

In the case of simultaneous or intersecting actions, the ordering of the clauses is flexible:

- (254) [siraj kehn jtak], knən nə? neŋ kuruj pintu?
 [while 3S sleep] offspring this earlier shut door
 While she was asleep, the aforementioned children shut the door.

- (255) ki=?ur=ca? jaga? ?en ?ma?, [siraj sma? ?yot]
 3A=instruct=EM watch.over LOC mother [while person return]
 He instructed the mother to watch over (it), while the people returned (to work).

The temporal clause may be nominalised:

- (256) lag? ki=ca, [sampay k<nj>moŋ gorom mə=e<pn>?dm]
 still 3A=eat [until finish=NMZ> salt one=package<NMZ>]
 Still she ate, until finishing one package of salt.

The verb *kna?* 'to affect s.th.' functions as a temporal conjunction indicating an achieved state still current at the time of the event described in the second clause (see also (225) above):

- (257) *A man who abandoned his wife to go cock-fighting returns home after a number of years only to discover that she has returned to her parents' place of residence and remarried.*
kna? dos he? ke, loc b-knək kmpan
when arrive AT:above there already HAVE-husband wife
(Alas), when (he) arrived up there, (his) wife had already (re)married.

The following clauses illustrate the differences between *siraj*, *samel*, and *snih* which all mean 'while'.

siraj 'while' is used when different people perform different but simultaneous actions 'A does V_X at the same time as B does V_Y', see examples (254)–(255) above.

samel 'while' is used when simultaneous actions are carried out by the same person, 'A does V_X at the same time as V_Y'. This is the same usage as *sambil*, the corresponding Malay term. It is more usual to express the second of these with a sequence of two verbs, the second in the imperfective, *ca gg-goŋ* (eat IMPERF-carry) 'to eat while carrying' (§11.4.1.4).

- (258) ca samel swak
 eat while walk
 (He) is eating while walking.

snih 'while' is like *siraj*, used where simultaneous independent actions are performed by different individuals, 'B does V_X to A at the same time as A does V_Y'. *snih* may host the enclitic =hn 'CONN'. See also (226)–(227) above.

- (259) [snih=hn dm-dəm], ki=tikam la=knək rəm lmeŋ
 [while=CONN IMPERF-lie.down] 3A=stab A=husband WITH spear
 While (the tiger) was lying down, the husband stabbed (it) with a spear.

?areh is glossed as 'only then' or 'now at last' (and is similar to the Malay use of *baru* (to be new) as 'now at last'). The presence of ?areh profiles a new state of affairs. The conditioning event is usually presented in the previous clause, never the following one. It is understood that only from that point in time as expressed in the previous clause does the predicate hold true, and not before.⁹

- (260) ye bnapi?, ?areh ki=jtək
 I sing only.then 3A=sleep
 I sang (and) only then (did) he fall asleep.

In (261) it marks a new activity in a sequence of events:

- (261) tonoy sake? jmur, ?areh tan r?r?
 wait pandanus be.dry only.then weave basket
 (One) waits for the pandanus to dry (and) only then (does one) weave the basket.

In (262) ?areh is present in both clauses. The best translation here is 'as soon as'.

- (262) ?areh ye=bl?, ?areh ye=pakay
 only.then 1A=buy only.then 1A=use
 As soon as I buy (it), I use (it).

In the preceding examples, ?areh is a clausal connective. It may also function as a discourse connective (263), with the same meaning (§11.5.4).

The following passage of text illustrates the context in which ?areh 'only then' is used.

- (263) *A successful exorcism is performed:*
?yot dəs he? dol ?ma?, da? ki=sdar.
return come AT:above house mother NEG 3A=remember
pe? ?are? pe? ptom, la=b?b?ŋ,
three day three night BCS=fear
de=p-bliyan la=?ma?. ?areh ki=sdar
3plA=CAUS-exorcise BCS=mother just.then 3A=remember
(She) came back (home) up to (her) mother's house, (but) she (could) not remember (it). (For) three days (and) three nights, because of fear, the mother made (them) exorcise (the daughter). At last she remembered.

?luc.tom 'after' is possibly influenced by the Malay *lalu* 'to pass by', which is used as a conjunction meaning 'then; afterwards', or *lepas* 'to pass by', used as a connective (*se*)*lepas itu* 'after that' respectively. *lalu* has been borrowed into Semelai as *lalu?* a discourse connective, not as a clause connective (§11.5.4).

- (264) de=yok jalū, yok crth, de=ca.
 3plA=take pork take fish 3plA=eat

⁹ This is like the Malay use of *baru* 'new', which also functions as a temporal expression with this meaning (Mintz 1994: 306).

- ?luc.tom de=ca, de=?intc
after 3plA=eat 3plA=exit.house
They took pork (and) took fish (and) they ate (it). After they ate (it), they left the house.

There is an absence in the list of connectives of a monomorphemic term to express 'before' in a sequence of events. Given that 'before' describes an event that has not happened, the usual manner of expressing this is by using an adverbial which expresses continuity, *lagi?* 'still' (265). When *lagi?* precedes the negated clause as in (266), *lagi? da?* (still not) 'not yet' expresses an event or state which has still not happened.

- (265) *lagi? ki=jɔh*
still 3A=drink
Still he drank.

- (266) *lagi? da? dos*
still NEG come
He still (has) not come.

The construction exemplified in (266) can also be understood as a temporal clause connective 'before ...', to express a prior existing state.

- (267) *lagi? da? b-knɔn, ye m-rok*
still NEG HAVE.offspring I PERFM-smoke
Before (I) had children, I smoked.

11.5.3.4 *yo?* 'but, however'

The discussion now turns to the connective *yo?* 'but, however' which like the causal connective may also occur in an independent clause as a discourse connective.

The clause containing *yo?* 'but' is usually part of a concatenated structure, the two clauses expressing paired contrast or opposition.

- (268) *The food isn't cooked, but they eat it anyway:*
?eh da? kmɔŋ, yo? ye=ca
eh NEG finish but 1A=eat
Eh, (it's) not finished, but we'll eat (it anyway).

- (269) *A man is trying to figure out if his wife remembers her time spent as a birth-vampire.*
dom kb, yo? kb=sda?
AFF 2f but 2fA=remember
Indeed, (it's) you, but do you remember?

The two clauses are interchangeable, although the favoured order is for the one containing the connective to be initial. However, provided the opposition is contextually recoverable, it may stand as an independent clause. In this function, it resembles a discourse connective (§11.5.4).

The distribution of *yo?* differs from other connectives. It may be used in clause-final position. In this case, a more suitable translation would be 'however', though it is also like the non-standard Australian English usage of 'but' as shown in the alternative translation:

- (270) *The speaker protests about the suggestion that he doesn't know, but tries to cover for himself.*
"k^be?", do? ?əp=k^be?, [hilaj yo?] k^bleŋ
know of.course IfA=know [disappear but] QUOTE
"Know? Of course I know. (It's) slipped (my mind) however," (he) said.
Or "Know? Of course I know. (It's) slipped (my mind) but," (he) said.

- (271) "eh," k^bleŋ=hn, "da? k^bbəs yo? hne," k^bleŋ
eh QUOTh=3POSS NEG die but THEN QUOTE
"Eh," he said (surprised), "so (it) didn't die but," (he) said.

The clause containing the connective can be used in isolation to respond to a previous statement made by someone else, usually a proposition of some sort.

- (272) *The Pauper wants to flee, but the king counters this by reminding him of his destiny.*
"yo?" k^bleŋ=hn "nasip kb rɔm knɔn ke" k^bleŋ
but QUOTh=3POSS destiny 2f WITH child that QUOTE
"But," he said, "your destiny is with that child," (he) said.

yo? can be encliticised by =hn 'CONN' as noted in §11.5.2. The presence of =hn 'CONN' is only attested where the connective is in initial position in the second clause, suggesting that it is anaphorically motivated. In (273) the antecedent is an NP.

- (273) *The speaker is describing a type of wild tuber.*
jmo?, [yo?=hn ple ʔen ʔats]
vine [but=CONN fruit LOC earth]
(It's) a vine, but its fruit (is) in the ground.

- (274) *The speaker is describing a gul 'a small round pool of open water'.*
bu:bt mɔ:com lubuk. [yo?=hn ra?k̚et tom lubuk]
be.round like open.water [but=CONN COMP-be.small SRC open.water]
(It's) round like a stretch of open water. But it's smaller than a stretch of open water.

In (275) the clause containing the connective is verbal.

- (275) *A tiger is taken aback and protests at a woman's suggestion that he hold her baby. He retorts:*
t^b!, ?əp da?! [yo?=hn t<ŋ>lɔŋ da? ɻ=k^be?]
hand If EXIST [but=CONN hold<IMPERF> NEG IfA=know]
Hands, I have! but I don't know (how to) hold.

11.5.4 Discourse connectives

Discourse connectives establish cohesiveness in a discourse, providing temporal connections between independent clauses, and in some cases even larger tracts of text. They tend to frame events or units of events within the text. This section will not deal with them in depth, but merely flag their presence and distinguish them from intra-clausal connectives.

The discourse connectives are indistinguishable from the clausal connectives, apart from the fact that they are only used in single clause constructions, i.e. clauses which are separated at the point of juncture by a sentence-final rather than an

internal pause, marking inter-clausal relations. Like the clausal connectives, most discourse connectives are also borrowed from Malay, again in form although not necessarily in function. The connectives occur clause initially, and they may be encriticised by =hn 'CONN', with the exception of jake 'in that case'.

The most frequently occurring connectives are listed here:

?arch	'(now) 'at last, only then'
lalu?	'then' ← Malay <i>lalu</i> 'then, after that'
sampay	'so' ← Malay <i>sampai</i> 'until, up to a particular point'
jadi?	'consequently, so' ← Malay <i>jadi</i> 'thus, so, consequently, therefore'
jake	'then, in that case'
cəŋ=cə?	'from that point on' (pass through=EM)
yb?	'but, however'
beh=hn	'not that ...'
da? beh	'it isn't because ...'
?le?le?	'eventually, then, so' ← ?le? 'be long' ← a calque of Malay <i>lama-lama</i> 'finally'

The clauses can be characterised as expressing the consequence of a previous event, or indicating a new turn of events. A few notes will be made on the distribution of the connectives, followed by examples of individual cases.

The following examples are the opening lines in a story, so there is no option but to treat ?le?le? and sampay as discourse connectives and not as clausal connectives, recalling that some discourse connectives, like ?arch 'only then, at last', also function as clausal connectives (§11.5.3.3):

- (276) ?le?le?=hn, knlək=hn ke swak. p?st kah, b-knon
so=CONN husband=3POSS that go stay 3 HAVE-offspring
So, her husband went. She stayed (and) gave birth.
- (277) sampay swak b-prəŋkəp cŋj, dom raja?
so walk HAVE-basket trap mousedeer AFF prince
So, the crown prince, yes, went trapping (for) mousedeer.

In the following clause, sampay marks the completion of a chain of events, a scene of a long tale, providing a bridging context in the discourse:

- (278) sampay=hn ki=cə mə=ha? joŋ tipə?
so=CONN 3A=eat REL=AT foot betel.service
So he ate the one at the foot of the betel service.
- (279) sampay nə? ki=cəl ni<r>kah
so now 3A=pronounce marry<CAUS>
So now he pronounced (them) married.

Discourse connectives always precede clausal connectives:

- (280) Now that a woman has lured her husband up to the sleeping area, she's making sure he stays.
sampay, siraq jtek, ki=bək la=kmpən hə? ke
so while sleep 3A=tether A=wife AT:above there
So, while (he was) sleeping, the wife tethered (the husband) up there.

Two connectives can occur adjacent to each other. In (281) there are two synonymous connectives expressing consequence, jake 'in that case' and jadi? 'then, consequently'.

- (281) A woman starts to comprehend her husband's wishes:

?ec, ki=?yəŋ la=kmpən deŋ ke
ah 3A=hear A=wife like that
"ʔeh, jake, jadi?" kʰləŋ=hn "ga=?əŋ=cəŋj"
ah in that case consequently QUOTE=3POSS IMM=1fA=burn
Oh no, (when) the wife heard (it) like that, "Ah, in that case then," she said "I'm going to burn (it)."

All of the connectives, other than jake 'in that case' may be encriticised by =hn 'CONN'. The motivation for this is not understood, as discussed earlier in §11.5.1.1.

- (282) ?le?le?=hn jtek, mə=ptəm, dwa? ptəm.
so=CONN sleep one=night two night

?le?le?, bula-bulan, ?ah
so RDP-month ah

So then (he took to) sleeping one or two nights (away). Then (it became) months (at a time).

In the two examples which follow, =hn 'CONN' is encriticised to the connective, an intransitive clause in (283), and a transitive clause in (284).

- (283) lalu?=hn ryoh=cə? reŋkəŋ cə? paya?
then=CONN make.noise=EM toud.sp. AT:below stream
Then a toad down at the river made a noise.

- (284) lalu?=hn, ki=pantaŋ la=cəŋ, "ʔəh jə kə?", kʰləŋ
then=CONN 3A=call.out A=mousedeer oh 2 macaque QUOTE
Then the mousedeer called out, "Oh you Macaque," (he) called.

?arch can host the emphatic enclitic =cə? 'EM' (§11.4.1):

- (285) A boy has been impatiently awaiting his circumcision.
ki=sunət dom la=sma? gdo kke. ?arch=cə? sunət
3A=circumcise AFF BCS=person be.old that only.then=EM circumcise.
The old man indeed circumcised (him). Now at last, (he) was circumcised.

lalu? 'then' can also express the next event in a series:

- (286) lalu? de=kabar ?en gdo
then 3plA=tell.news LOC old
Then they told the news to the old (people).

- (287) lalu? ki=capay hc? kʰoy tipə?, ki=cə
then 3A=grab.hold AT:above head betel.service 3A=eat
Then he grabbed hold (of one) at the head of the service, (and) he ate (it).

The following example shows the context of ?le?le? 'eventually' the most frequent of the connectives. This is a reduplication of the Semelai word ?le? 'be long, (of time)'. Used here in its reduplicated form it is 'so, then'. In (288) it is ambiguous whether it should be interpreted as a connective or an adverb.

- (288) solor j̥m>?. ʔieʔleʔ sot dloŋ
first vine eventually metamorphose.into tree
(At) first (it grows as) a vine. Then (it) metamorphoses (into) a tree.

In (289) the 'connective' is not encliticised. This clause does not contain backgrounded information and the function of ʔieʔleʔ is adverbial.

- (289) ʔieʔleʔ c<cl>rak k<r>dor ke
RDP-be.long.time scream<IMPERF> be.female<NMZ> that
Then the woman was screaming for a long time.

The connective can mark the consequence as resulting directly from a previous event.

- (290) A pauper has been selected to marry the king's daughter, and the officials, incensed, turn their backs on him.

ki=ʔireŋ la=twanko?, ʔyot teʔen dol.
3A=lead A=king return TO:down house
jadl?=?son, dom kraba-krabat de=c<r>lon
consequently=SC AFF RDP-official 3plA=back<EQUIP>
The king led (him) (and) (they) went back down to the house. Consequently, yes, (every) official turned (his) back (to him).

- (291) ... ki=gɔŋ ʔyot tet k<r>dor ...
... 3A=bring return TO:spec be.female<NMZ>

jadl?, knon ha? no?, ki=kʰom yip-yup,
consequently offspring AT here 3A=get HAVE-sleeping.cover
keh, ki=kʰom st-sot
3 3A=get HAVE-sarong
He brought (the cloth) back to the woman. Consequently, the child here got to have a sleeping cover, (and) she got to have a sarong.

jake functions to introduce a response or retort, framing it as a consequence of the previous discourse. The speaker announces his/her intention as in (292) or opinion as in (293).

- (292) A: "beh," kʰlaŋ "yeʔ-en, yeʔ=yam ʔma? yeʔ-en"
NO QUOTE 1=AUG 1A=cry mother 1=AUG

B: "jake ht=yok"
in.that.case 1&2A=take
A: "No", (they) said, "we are crying for our mother."
B: "Then in that case we'll fetch (her)."

- (293) "jake" kʰlaŋ ʔma? "ga=kniłek ko sarek"
then QUOTE mother IMM=husband 2f future
"Then in that case," said the mother "(he) will be your husband (in the) future."

cəŋ=cəʔ 'from that point on', is usually used to mark a very important turn of events in a text, and flags a new scene in a story. It is generally accompanied by a lack of reference to events or circumstances in the past. In the following story it

marks the point where the siblings commence their life of incest, turning their backs on their parents and instigating the ensuing chain of events:

- (294) cəŋ=cəʔ tɔm ha? kke, deh jtek muy, swak muy
From that point=EM SRC AT there 3pl sleep one walk one
From that point on, they slept (as) one, went (about) (as) one.

12 Expressives

Expressives are a minor word class in Semelai. This word class can be considered iconic, consisting of non-arbitrary signs which invoke a sensate response (§4.3.2). Cross-linguistically 'expressives' are more commonly termed 'ideophones' (Alpher 1991, Hinton et al. 1994); 'expressive', as the term commonly used within Austroasiatic (Diffloth 1972, 1976b, 1976c), is adopted.¹ Expressives are often cited as a distinctive shared feature of the Austroasiatic language family (Diffloth and Zide 1992; Osada 1992 (Mundari); Svantesson 1983 (Kammu)). Some examples are given in a) below from the Aslian languages Semelai, Semai (Diffloth 1976b), Jah Hut (Diffloth 1976c) and Temiar (Benjamin 1976b). Expressives are a common typological feature of the Malay Peninsula, not restricted to Aslian languages, but attested also in non-standard varieties of Malay. Data in b) below is from two varieties of Malay, the first a dialect spoken by Jakun at Kampong Landai, Pahang (Collins 1985); the second, a variety of Malay spoken on Langkawi Island, Kedah (Collins 1979).²

a) Semelai:

- c^bŋis 'the smell of ammonia, urine'
- crālāp 'the sound of s.o./s.th. entering the undergrowth'
- grw^bŋi 'the appearance of eyes, sunken and dark from illness'

Semai:

- blb^bəl 'the feeling of painful embarrassment'
- cnayur 'the appearance of bushiness in several places'
- grē:p 'the noise of chewing large, somewhat soft things'

¹ I do not make a distinction between expressives and ideophones. Diffloth (1976b: 264) proposes a taxonomy consisting of expressive, ideophone and onomatopoeic words. I distinguish expressives from onomatopoeic forms, although the two probably overlap. I do not have any clear data on such issues as functional differences, and it is therefore difficult to provide a rigorous distinction. I tentatively define onomatopoeic items as instances of acoustic replication, rather than acoustic symbolism (expressives), i.e. words which are used simply to imitate or mimic sounds, like the replication of animal calls, e.g. gwuk 'the frog's croak'. These are not specifically mentioned here. Some nouns in the language, particularly the names of certain fauna are onomatopoeic, mimicking the noise made by a particular animal, e.g. example ʔk^b?ok 'crow', paypay 'an insect that hums at dusk', kɔŋkɔlb 'the nightjar bird'.

² As Collins notes, the presence of expressives in Malay is usually overlooked, due to the fact that studies in Malay are based on formal varieties of the language. Indeed, a quick glance through a Malay dictionary, for example Wilkinson's Malay-English dictionary, or Coope's MacMillan's student Malay dictionary, reveals many 'onomatopoeic' or expressive-type forms, including both aural and visual concepts. Wilkinson (1927): *legum* 'the patter of feet on wooden flooring'; *legup-legap* 'the sound of blows', also *gup-gap*, *degup-degap*, *chekup-chekap*, *dup-dap*, *chela-chelas* 'passing in and out familiarly; taking liberties; making oneself at home; swishing in and out'. Coope (1978): *geregok-geregau* 'clawing at, grabbing at, as when searching for food'; *gentang-genit* 'wavering and tapering to a point (as a kris)'.

Jah Hut:

- | | |
|-------|--|
| lpuij | 'the sound of heavy fruit falling on the ground' |
| glhuŋ | 'caved in, of eyes' |

Temiar:

- | | |
|-------------|--------------------------------|
| kərləg log | 'the sound of heavy footsteps' |
| cəra?uk ?uk | 'stomach queeziness' |
| gengerlat | 'spindly-ness' |

b) Jakun (Aboriginal Malay) dialect spoken by Jakun, Kampong Landai, Pahang:

- | | |
|--------|--|
| pru?up | 'pitch black' (<i>hitam legam</i>) |
| pū? | 'the sound of something small falling into water' (<i>bunyi (benda kecil)</i>
<i>jatuh dalam air</i>) |
| həjc̩t | 'the sound of mice' (<i>bunyi tikus</i>) |

Kedah Malay, Langkawi Island:

- | | |
|-----------|--|
| ku?̩ ku?̩ | 'of writhing movements (snakes in water, fish in mud)' |
| c̩t | 'of a mosquito bite, medical injection' |
| klen kloŋ | 'of Thai being spoken' |

12.1 The semantics of expressives

Expressives provide an iconic summary of an event, or an aspect of an event or entity, in the form of a single lexical item. They encompass all aspects of sensory perception: aural, visual, tactile, smell, taste and internal feeling. This is reflected in the speakers' metalinguistic observations of expressives in terms of the mode of perception: aural – *swara?* 'voice, sound' (?ānṛāk 'the sound of many people, e.g. in a working party'); visual – *cara?* 'appearance' (rmoŋkōp 'the appearance of antlers'); k^bləm 'smell' (c^bŋis 'the smell of snakes, menstrual blood'); and *pjam* 'feeling, internal' (?pac 'bewildering appearance') or 'tactile (p^btac 'feeling of a cord snapping') and taste'. The classifications are loose, and the variation in classifications of individual lexical items is unremarkable as we shall see below.

(1)	c̩p,	c̩p,	
	<i>winnow.rice.inexperienced!</i>	<i>winnow.rice.inexperienced!</i>	
	c̩p.	swara?	bras
	<i>winnow.rice.inexperienced!</i>	sound	husked.rice
	c̩p c̩p c̩p (is) the sound (of) husked rice (being) winnowed.		MID-winnow

Expressives are a linguistic expression of a speaker's conceptualisation of a sensate response. Providing precise definitions for expressives is problematic due to the idiosyncrasies of each speaker. Expressives are characterised as having loose 'conceptual' definitions as reflected in their classification, discussed earlier.

Expressives are not necessarily sense specific, what one person may consider a visual expressive, another may consider aural, to someone else it may be tactile/feeling. For example, some speakers defined p^bur/p^bar 'the sound of someone jumping down or leaping from a height' to be aural, while to others it was a visual-based expression of the same act. The essential point here is that the event or state was not a matter of contention, for instance speakers were unanimous in the

opinion that these terms represented movement through space associated with someone jumping down or leaping from a height. A similar situation is found with *kräsäk*, which can be interpreted agentively as in 'the sound of fish biting, taking bait', or as in 'the sound of a trap closing on prey'. This overlap and lack of definition does not appear to be problematic, but just part of the nature of this type of description. Diffloth (1976b: 257) suggests that expressives be defined "as a cluster of elementary sensations".³ This scope for the interpretation (or reinterpretation) of the meaning is perhaps natural given that we are dealing with "a person's reaction or interpretation of an event" (Collins 1979: 389). Collins provides an interesting commentary on 'expressives and expressers' in his account of these lexemes in Kedah Malay. An informant commented to him that "the sounds of the expressive are relevant at the moment of reaction and that each person's reactions are different" (1979: 395). This astute observation is relevant also to the use of expressives in Semelai.

With lexical items like nouns or verbs, a change in the phonological shape of the form either changes the meaning, or renders the form meaningless. In contrast, a change in the phonological shape of an expressive may introduce a new aspect of meaning, but retain the basic core semantics of the original form. This does not apply systematically to all forms. Some relevant examples are provided below. In the first pair of examples, there is a change in the final consonant of each syllable:

- rɔpräp 'something large walking through twigs'
rɔkräk 'something large walking through pandanus'

In the second set there is a change in vowel height:

- grēŋ 'the appearance of a little hair standing up straight'
grāŋ 'the appearance of a lot of hair standing up straight'

Certain expressives appear to be related not only conceptually, but also acoustically/phonetically to particular lexical items from the class of nouns or verbs. Phonological changes are achieved by a modification of the vowel quality (a change in height or nasality) or the quality of a consonant, as in the final liquid in the example: pʰu! 'to descend down s.th., to start blowing (of wind)' has seemingly related expressives: pʰur/pʰar 'the sound/appearance of someone jumping down or bounding along or leaping'. Likewise the verb pläc 'to emerge' is similar to pläc/pläw 'the sound, appearance of s.o./s.th. emerging'. The nominal plat 'the exit point of fish from reeds' resembles plät 'the sudden escape of s.th. and the accompanied surprise of the captor'. The speaker provided the scenario of opening a box and a mouse jumping out unexpectedly.

Relationships between expressives and other word classes are also attested in Semai (Diffloth 1976b: 261) and Temiar (Benjamin 1976b: 177-8). In these languages expressives feed morphological processes: *gengerlat* 'spindliness' ← *gerlat* 'long and thin, of small objects'. Some examples of this in Semelai are provided here. The noun *kdot* 'fontanelle' appears to be related to the expressive

³ This is reflected in my approach to glossing expressives. I have attempted to overcome the problem by listing the basic associated concepts in apposition, separated off from each other by a point. The gloss are in italics, followed by an exclamation mark to avoid confusion with other word classes. Translating them in context has proved equally difficult.

kidot 'the appearance of a baby's hair standing up on end', derived by the infixation of <|> and a copy of the final consonant <p> (§3.2.4.2). The expressive *kipet*, meaning 'small and motionless (of an entity)' appears to be derived by the infixation of <|p> into the root *köt* 'to be small'.

In some cases, the association is more tenuous. For example *kilmuk* 'the appearance of someone healthy with fat cheeks', is similar to *gmuk* 'to be fat'; *jep* is 'to winnow yams vertically on a tray', *cép* is 'the sound of someone inexperienced winnowing rice on a tray', and *snekép* 'the sound of someone experienced winnowing rice on a tray'. Some further notes on the phonology and morphology of expressives are made in §12.3.

One other common feature of this word class is the co-occurrence of certain expressives with particular verbs, although they are not specifically associated with that verb, (see Alpher 1991: 7 for similar observations in relation to Australian languages). Some examples from Semelai are given here: *pläc* 'the sound, appearance of s.o./s.th. emerging' is often paired with *dos* 'to arrive'; *gom* 'startled movement, either jumping to a standing position or falling flat' and also 'startled and running off' usually occurs with *brabluh* 'to run', or *paloh* 'to flee' as in (2), and 'falling flat' with *t-pa-dəm* (HAPP-CAUS-lie) 'to fall to a lying position'.

In (2), the notion of movement on being startled is introduced by the expressive *gom* 'startled.movement!'

- (2) *gom* means to experience a fright-flight response to something, and denotes the point of inception of that flight.

gom, paloh, de=cəŋ halaw la=co
startled.movement! flee 3pIA=immediately chase A=dog
Started movement! (he) fled, (and) the dogs immediately chased (him).

gräbäc is only associated with *tłɔj* 'to hold, to grab', expressing the element of surprise, or being startled from the point of view of the captee:

- (3) gräbäc, ki=tłɔj knən, ki=goy paloh
grab.startle! 3A=grab offspring 3A=bring flee
From behind, starting (the captee), he grabbed the child (and) fled (with it).

12.2 The syntax of expressives

Expressives function in an adjunct-like manner, syntactically independent of the associated constituent (whether clause or phrase), hence their treatment here rather than in previous chapters on clausal syntax. Expressives are defined syntactically by the following features:

- They only occur in relation to positive declarative clauses.⁴
- They cannot be negated, quantified, or modified. This distinguishes them from manner adverbials with which they could otherwise be confused.
- Expressives are usually, though not always, distinguished by an intonation break separating them from the constituent which precedes or follows. This is

⁴ The inadmissibility of expressives in negated clauses is attested cross-linguistically in genetically and areally unrelated languages such as the African language Chadic and the Austronesian language Sundanese. In Korean, expressives can only occur with metalinguistic negation indicating the inappropriate use of an expressive (Diffloth 1972: 446).

represented in the examples by the comma and in the gloss by the exclamation mark.

Expressives are never syntactically integrated, i.e. they never occur in any position which is inside a phrase boundary. In this respect, they bear some resemblance to the relationship between the quotative marker *k^bləŋ* 'QUOTE' (§13.1) and its accompanying direct speech. In fact, aural expressives may be followed by *k^bləŋ* 'QUOTE':

- (4) "prāc," k^bləŋ kris, lalu?, "pār" k^bləŋ bsurak
draw.weapon! QUOTE dagger then *sound:shout!* QUOTE shout
 "prāc" went the dagger as it was drawn, then, "pār" went sounds of shouting.

The expressive either immediately precedes or follows the construction with which it is apposed. Although expressives are not syntactically integrated, their distribution can be stated in terms of their commonest type of frame, characterised here as i) perception and ii) generalisation of a statement.

i. Perception When expressives occur in relation to a perception predicate (*?ye* 'to see', *jŋsk* 'to look at', *?yəŋ* 'to hear') it follows the predicate:

- (5) ki=?yəŋ, bldbk
 3A=hear lips.smacking.jowls.moving!
 He listened to (it), (something) chewing noisily, lips smacking.

In (6) the expressive precedes the dependent clause.

- (6) sampay, ki=jnɔk, kltput=sdn.
 so 3A=look.at bob.up.and.down!=SC
 ga=gəl, la=ki=tut la=rɪbut
 IMM=float.away BCS=3A=blow A=wind
 So, he watched (it) bobbing up and down. (The boat) was going to float away, because the wind was blowing (it).

In (7) the expressive is between the verb and the post-posed subject. *?ɔlok.zelek* refers to something 'upright swaying from side to side', or something 'lying and turning, facing up, then facing down again in a continuous motion'.

- (7) sampay, ?le?le? ki=?ye b-bway, ?ɔlok.zelek,
 so eventually 3A=see MID-swing oscillating.to and fro!
 mə=?ikur puru? ?en karom
 one=CLF toad LOC underneath
 Then eventually he saw swinging to and fro, one toad in the space underneath (the house).

In (8), the expressive is simply apposed with the organ of perception.

- (8) təŋ, ?yʒm
 ear sound of wind blowing across a small opening!
 (In his) ears, the sound made by wind blowing past, like it does across a small opening.

ii. Generalisation of a statement When an expressive provides a summary of a statement, the expressive generally directly precedes the clause it summarises:

- (9) rēncērencēc, mm-(p)antaw, ki=?ajak=sdn tə?en ?ate
 sound.shout! IMPERF-call.out 3A=invite=SC TO:down ground
 (With) the sound of shouting, (he) was calling out, (and) invited (them) to (come) down to the ground.
- (10) p^bur, ki=gooŋ lumpat
jumping down! 3A=bring leap
 Jumping down, he carried (it) leaping.
- (11) *The sight of an unwed daughter's pregnant belly alerts her father to her condition and causes a sense of grief.*
 ?pac k^bet, ?eh, ki=risaw-17 la=bapa?
 bewildering.appearance! be.pregnant ah 3A=have.pity-APPL A=father
 The bewildering sight of (her) pregnant, ah, the father grieved (over her).
- (12) *The character snggøe bmbau has taken a bath in a pool of dirty water:*
 loc hūm, drek.drak, kbə? j?ji?
 already bathe dirty! body be.dirty
 After bathing, dirty, (his) body was dirty.
- (13) de=?ye tom ha? d?oh, masuk bri,
 3plA=see SRC AT swidden enter jungle
 gwenwët, tran lmeŋ=sdn
 moving in and out of view! shaft spear=SC
 They saw (him) from the swidden enter the forest, the shaft of his spear moving in and out of view.

In a clause containing a connective, the expressive naturally follows it. The expressive always precedes the quotative marker *k^bləŋ* 'QUOTE':

- (14) lalu?, "cyāp" k^bləŋ=hn
 then sound.snapping shut! QUOTE=3POSS
 Then, the sound of something snapping shut.

12.3 Notes on the phonology and morphology of expressives

As many of the preceding examples show, expressives usually consist of more than one syllable:

crop	'the sound of something entering the water'
jlŋaw	'the appearance of someone thin from illness'
gmarasak	'the sound of walking through forest litter'

It is not uncommon to have phonemic vowels and nasalised vowels in non-final syllable position, the presence of nasalised vowels being at variance with the phonotactics of other word classes (§2.3.3.1). There is generally a higher occurrence of nasal vowels in expressives. Other than this, they are phonotactically unremarkable.

The lexeme may or may not exhibit morphological complexity. The most frequently attested method of word formation is the compounding of a complete copy of the base, but in which the final vowel of each base differs from the other: *tuj.taj* (**v**tuj*) 'the sound of something descending, like fruit falling or being thrown to the ground, or the sound of a person or monkey descending a tree'.

This differs from reduplication elsewhere in the language in that the whole root, whether mono- (as above) or disyllabic, is reduplicated: *grōbōk.grābāk* (**v**grōbōk*) 'the sound of something slapping, like rubber thongs which are too big'. Elsewhere reduplicative processes are limited to the reduplication of a single consonant (see coda copy §3.2.2.2) or light syllable reduplication (§3.2.4). The direction of reduplication appears to be from left to right, on the basis of the predictability of the vowel /a/ in the right hand base. The usual direction of reduplication is from right to left: *pros.pras* (**v**pros*) 'the sound something descending, like fruit falling or being thrown to the ground; or the sound of a person or monkey descending a tree'.

Semantically, reduplication appears to be associated with the repetition of an action, or its continuation over time.

Some expressives appear to be derived from other expressives, e.g. *craw* 'the sound of the movement of water, like someone bathing; a boat moving through water' appears to be the root of *cmaraw* 'the sound of something being poured, like water, sugar or rice', which appears to be derived by the infixation of <*ma*>. However, it is harder to motivate the following forms, which appear also to be related: *praw* 'the sound of something light hitting the water, like leaves', *cbaw* 'the sound of water as someone steps into it'. This hints at the possibility that the individual segments are phonaesthemes, examples of sound symbolism. Unfortunately, the examples of attested similarities collected to date are too few and too disparate to be able to decompose the words into possible smaller meaningful units (cf. Diffloth 1976b for Semai, Collins 1979 for Kedah Malay, and McGregor 1996 for the Australian language Gooniyandi).

13 The quotative marker, interjections and discourse clitics

In this chapter, three topics which fall outside the scope of the preceding chapters are examined. The quotative marker *k^hləŋ* 'QUOTE' which is used to frame speech, thought or aural expressives is discussed in §13.1. Like the expressives in the previous chapter, this lexeme is also syntactically independent. Interjections are discussed in §13.2, and finally, discourse clitics in §13.3.

13.1 The quotative marker

k^hləŋ 'QUOTE' marks reported speech. In Semelai there is no means of reporting speech other than as an actual direct quotation. *k^hləŋ* 'QUOTE' functions to frame the speech as being other than that of the speaker. The quote is indexed as the actual words of the person to whom it is attributed, removing the viewpoint of the narrator. In addition, it is used to frame expressives, interjections, exclamations and non-communicative aural events.

In Semelai utterance verbs are never used to frame the reported speech. The presence of such a verb is only to convey the manner in which the transfer of information takes place and to identify the participants.¹ Although *k^hləŋ* is glossed as 'QUOTE', it is translated with a range of meanings depending on the context, for example 'to say', 'to ask', 'to reply', 'to exclaim', 'to say to oneself, to think'.

- (1) "tɔh bborɔŋ-kadeh" k^hləŋ. cəŋ cəl ʔen kɔʔ.
oh thing-who QUOTE mousedeer speak LOC macaque
"ji=ien tɔh co? hn=yɛ?"
2A-want untie AT:below ABS=1
"Oh, whoever," (he) said. Mousedeer is speaking to Macaque. "Do you want to untie me down here?"
- (2) "ah k^hom creh tʰay=cəʔ?"! k^hləŋ snaʔ rabon
ah get fish be.big=EM QUOTE person be.short.sighted
"Ah, (I've) got a big fish!" exclaimed the short-sighted man.

In (3) the first instance of *k^hləŋ* is translated as 'ask' and the second as 'reply':

¹ With the exception of *cəl* 'to speak, pronounce', the remainder are all borrowings from Malay: *cakap* 'to speak, talk' ← *cakap* 'to say, speak'; *tapaʔ* 'to ask' ← *tanya* 'to enquire'; *sahut* 'to answer'; *bbwəl* 'to chat' ← *bual* 'to gossip'; *perber* 'to think' ← *fikir* 'to think'.

- (3) A shaman is trying to establish if an exorcised person recalls the period when he was possessed.

A: "k^bləŋ sma? ha? h^bñ?" k^bləŋ puyŋ
2f person AT where QUOTE shaman

B: "beh, da? ñ=k^bε?" k^bləŋ
NO NEG IfA=know QUOTE
"You (are) a person from where?" asked the shaman.
"No, I don't know," (he) replied.

Speakers were unable to assign a referential meaning to k^bləŋ 'QUOTE'. k^bləŋ is not a verb of speech, as it does not fit into the verbal part of speech category – it cannot take arguments, be negated, be modified aspectually or derived. Etymologically, k^bləŋ 'QUOTE' was probably a verb; indeed in the Aslian language Jah Hut (Central), the verb 'to speak' is k^bləŋ (Diffloth 1976c: 96) and in the Northern language Cheq Wong kluŋ is the verb 'to say' (Kruspe fieldnotes); in Mah Meri (Southern) there is klək 'to say or think to oneself' (Kruspe fieldnotes).² Cross-linguistically, most quotative or reportative markers are derived from a verb 'to say' (Heine and Kuteva 2002: 267–8).

k^bləŋ is also used to flag exclamations, interjections and other indexical expressions of human interaction.

- (4) A man has had enough of his wife's pestering.

"has!" k^bləŋ knlək
enough! QUOTE husband
"Enough!" said the husband.

- (5) A woman is agreeing to comply with her son's instructions.

"hmm," k^bləŋ ?ma?
hmm QUOTE mother
"Hmm," agreed the mother.

The interpretation of the quotative is often coloured by accompanying interjections such as: bah 'NO!' 'strong disagreement', ?ac 'displeasure, disgust', ?ec 'negative surprise'; so the translation is composed of k^bləŋ plus the interjection, e.g. hmm ... k^bləŋ 'agreed' as in (5) above, and ?ac ... k^bləŋ 'complained'. Interjections are discussed in §13.2.

In the same manner, the intonation of the utterance will also influence the interpretation of the quotative marker. Interrogative intonation will be translated as 'ask, enquire', (11) below, and imperative intonation, (or indeed the actual presence of the negative imperative marker) will also affect the translation.

A third function of k^bləŋ is to mark non-communicative aural events, in particular the sounds associated with aural based expressives, i.e. 'to make X noise'. In the following examples, the expressives convey the sound of the protagonists' passage through space:

- (6) k^bləŋcompok hn=tilam c7cɔ? ha? malanj dol. "p'ur"
3A=throw O=mattress be.infront AT side house jump.down!

² It is a common reflex for Jah Hut /u/ to be realised as /ə/ in Semelai: 'to bark' JH /ju/ is SML /jət/, in Mah Meri final nasals become voiceless stops preceded by a second register vowel.

k^bləŋ kəh he? raboŋ, g̬atm paloh
QUOTE 3 AT:above roof.ridge startled.movement! flee

He threw (down) the mattress first at the side of the house. There was the sound (of) him jumping down through the air (from) up on the roof, (and with) the appearance of some one taking flight, (he) fled.

The identification of the speaker or source of the sound is achieved by placing a nominal or a pronominal after the quotative marker, see examples (2)–(4) and (7).

- (7) beh naŋ beh, "pläc" k^bləŋ ks?
NO AFF NO emerge! QUOTE macaque
Indeed "pläc!", the sound of the macaque emerging.

k^bləŋ may be encliticised by =hn '3POSS', the enclitic never co-occurring with an overt mention of the person to whom the quote (16, line 7), or noise, is attributed (7). In light of the presence of the possessive enclitic, a more representative translation of (2) and (7) would be 'the mother's quote/utterance' and 'the macaque's sound' respectively, rather than the more literal translation given.

Usually the speaker is not encoded and must be contextually determined as in (8) and (9) below.

13.1.1 The syntactic distribution of k^bləŋ 'QUOTE'

The distribution of k^bləŋ 'QUOTE' is extremely flexible, occurring at either phrase or clause boundaries. As such, it is significant as a syntactic test for constituency. It may occur more than once in relation to the quoted speech. There is always a more marked pause after the quotative marker rather than before it, represented in the text by a comma.

When k^bləŋ is inserted at a clause boundary, it usually occurs in clause-final position:

- (8) "bøy ma=jəl ?əj!" k^bləŋ
NEG:IMP IRR=bark.at If QUOTE
"Don't bark at me!" (he) warned.

In clause-initial position it may be preceded by a discourse connective, see (14) and (16, line 4).

Within the clause, k^bləŋ is attested at the following phrasal boundaries:

- a) Between the subject of an NVC and its predicate as in (9) and (13) below.

- (9) "?ma? ye=?en" k^bləŋ, "k^boy nehneh"
mother 1=AUG QUOTE head only
"Our mother," (they) said, "(is) just a head."

- b) Usually k^bləŋ follows the verb phrase, occurring immediately adjacent to it as in (10) where the quotative marker is between the verb and the post-posed subject.

- (10) "beh, lən j<hb>?əh" k^bləŋ, "?əp."
NO want drink<IMPERF> QUOTE If
"No, I want to drink," (she) said.

In (11) it is between the verb and its object, (and then again following the object in clause-final position).

- (11) "sak tah" k^bləŋ, "bos? no?no?" k^bləŋ
 peel Q QUOTE sugar.cane this QUOTE
 "Aren't you going to peel," (he) asked, "this sugar cane?" (he) asked.

Here k^bləŋ is between a verb and an oblique:

- (12) "ko? ləc? ?ie? po?" k^bləŋ, "ha? no?"
 2f already belong very QUOTE AT here
 "You (have) already (been) here a long time," (she) said.

c) It may also occur between the pre-verbal ignorative and the verb as in the first two lines in this example:

- (13) "hōn?" k^bləŋ, "lən b-pri?to??" k^bləŋ.
 where QUOTE want MID-prepare.chicken QUOTE
 "ha? kadeh" k^bləŋ, "ko=?ur pri?to??"
 AT who QUOTE 2fA-instruct prepare.chicken
 kmpən k^bləŋ, "beh, da? da? wo?"
 wife 2f QUOTE NO NEG EXIST more
 "How," (she) asked, "would (it) be prepared?" (she) asked.
 "Whom," (she) asked, "did you instruct to prepare (it)?
 Your wife," (she) said, "no, (she) isn't (here) anymore."

The quotative marker may occur at several boundaries, both clausal and phrasal, in the one utterance. In (14) it follows the initial connective in the first instance, then occurs between fronted object and the verb, and in the third instance in clause-final position following the verb.

- (14) ?ie?ie? k^bləŋ=hn, "əj? no?" k^bləŋ,
 eventually QUOTE=3POSS If this QUOTE
 "da? ga=k^bp-r-luc?" k^bləŋ=hn
 NEG IMM=2fA=CAUS-pass QUOTE=3POSS
 Then, she said, "me here," (she) said, "aren't you going to release me?" she asked.

In complex constructions, k^bləŋ may occur at the juncture of two paratactically concatenated clauses, attached phonologically to the initial clause. In the first example it occurs twice in the antecedent clause, first at the juncture of a paratactic complement structure, and in the second case at the juncture of the two clauses of the larger structure:

- (15) [("məhn ga=k^bp=məh") k^bləŋ, ("pdər-i? hn=?əj,")]
 [if IMM=2fA=want] QUOTE [follow-ITER O=1f]
 k^bləŋ, ["n=jon"]
 QUOTE [1fA=give]
 "If you are going to agree," (he) said, "to accompany me," (he) said, "I will give (them to you)."

The final example contains the full context of some examples referred to above:

- (16) swak, t-c<n>roh rɔm baseŋ.
 walk HAPP-meet<CAUS> WITH leaf.monkey
 ?ah=sən dom kəh, kira? de=b^bŋ=sən
 ah=CL AFF 3 figure 3plA=be.afraid.of=SC
 ma=p-ijib-ijib=sən pacə?, bndə? da? ma=?ye
 IRR=CAUS-INTNS-peep=SC likewise thing NEG IRR=see
 lalu? k^bləŋ, "mande ??" k^bləŋ, "mande ma=b^bŋ-i? ye?
 then QUOTE why oh QUOTE why IRR=be.afraid-APPL 1
 da? da? hal ma=b^bŋ-i? ye.
 NEG EXIST reason IRR=be.afraid-APPL 1
 kira?, da? bndə? me=ma=lən" k^bləŋ,
 figure EXIST thing REL=IRR=want QUOTE
 "p^bəl=ca?!" k^bləŋ=hn, "co? no?" k^bləŋ,
 go.down=IMP, QUOTE=3POSS AT:below this QUOTE
 "tray" k^bləŋ, "bndə?" k^bləŋ, "ha? kapəl ye=?en.
 be.many QUOTE thing QUOTE AT ship I=AUG
 bərəŋ.2a=dəŋ ma=lən ma=ca"
 thing.DET=like IRR=want IRR=eat
 (While) walking (he) happened to meet some leaf monkeys. Ah, as it turned out, they were afraid of him. (They) peered (through their fingers) likewise, at something they had never seen. Then (he) asked, "Why?" he asked, "are you afraid of me? There's no reason to be afraid of me. You see, there are things that you would want," he said. "Come down!" (he) said, "down here," (he) said, "(There are) many," (he) said, "things," (he) said, "in our boat. All sorts of things you would want to eat."

13.2 Interjections

This section provides a brief account of some of the interjections used in Semelai. Interjections, including exclamations, are a closed class of words, which may stand alone as complete utterances.

13.2.1 Expressions of negative attitude

There are three markers of negative attitude: bah, ?ec and ?ac. They always co-occur with negated propositions. Each is described in turn below.

bah 'NO!' is an emphatic variation on beh 'NO' and expresses strong opposition to, or the rejection of, the asserted proposition. It may be used by the speaker, as in (17), or by the narrator with similar effect, as in (18).

- (17) do? ?əj?, bah, da? da? hal rɔm ?əj?
 as.for If no! NEG EXIST concern WITH If
 As for me, no! (you) don't have any concerns with me.

- (18) ?ah bah! da? ki=pakay
EXCL no! NEG 3A=use
Ah no! He wouldn't accept (it).

?ec and ?ac, both meaning 'no!', are used in conjunction with the external negator beh 'NEG'. There are no examples in the corpus of them co-occurring with b?en 'NOT'. ?ac is a stronger protest than ?ec; ?ac is used to express disgust toward a statement; ?ec to express mild opposition. Note the similarity of ?ec to ?es which is used as a gentle alternative to utterances containing b?oy 'NEG:IMP', the imperative negator (§10.6.3), especially when addressing a child.

- (19) A woman expresses fear that a tiger has eaten her newborn child's placenta and not buried it as she requested. The tiger, disgusted at her lack of faith in him, protests.

A:	t?oŋ	k?o=ca	hns,	beh,"	k?leŋ,
B:	fear	2fA=eat	THEN	NO	QUOTE
B:	"?ac	beh!"	k?leŋ=hn,	"da? n=ca"	
	ah	NO	QUOTE=3POSS	NEG	IfA=eat

A: "(I) fear you ate (it) then, no," (she) said.

B: "No way!" he protested, "I didn't eat (it)."

- (20) The speaker protests mildly about a proposition.

?ec	beh!"	k?leŋ	"da? moh"
EXCL	NO	QUOTE	NEG want

"Oh no!" (he) protested, "(I) don't want to."

13.2.2 Interjections with imperative force

The following interjections express commands:

h?2	'see! look out!'	?2a?	'animal!'
?oy	'you there!'	gi?	'give!'
has	'enough! forget it! quiet!'	nyanj	'here! (to dogs)'
sot=cə?	'stop it!'	cuh	'shoo! (chickens)'
?es	'don't!'	ce?	'get! (to dogs)'
bah hch	'no way!'	pes	'get! (to cats)'

h?2 'see! look out!' is used to draw someone's attention to something. It is used in two contexts: a) it is uttered on presenting or passing something to someone, or b) as an interjection 'look out!' to alert someone to something, e.g. if someone is about to inadvertently step on something. Elsewhere it is an enclitic on third person pronouns (§6.1.3).

- (21) "h?2!" k?leŋ k<r>dor ke hns, "ki=tloŋ knɔn"
look.out QUOTE be.female<NMZ> that THEN 3A=grab offspring
"Look out!" warned the woman then, (and) she grabbed (her) child.

?oy 'hey!' is a vocative used to draw someone's attention in order to address them. One elderly couple has their own variant, c?oy 'hey!', which they use exclusively between each other.

- (22) ?oy ji! ?ma? ?ame, ji=lən?
hey 2 mother Amelia 2A=want
Hey you there! Amelia's mother, do you want (it)?

has 'enough, forget it quiet!', expresses annoyance or a lack of interest in someone else's endeavours to communicate with them, see (4) above.

sot=cə? (must=IMP_i) 'stop it!' is an imperative used as a request to stop someone's annoying behaviour:

- (23) "sot=cə?!" ?pih ?əp" k?leŋ
must=IMP_i be.in.pain If QUOTE
"Stop it! I'm in pain," (she) cried.

?es 'don't!' is a gentle negative imperative. An alternative to b?oy 'NEG:IMP', it is used when addressing young children, or to distract their attention from something.

?2a? is babyspeak for 'animal' ← p?ia? 'animal'. This is uttered with imperative type force to instil fear in a child wandering ahead of, or away from, parents, or touching something she should not.

gi? 'give!' is the final syllable of h?gi? 'to give', and is used to coax a child in order to gently wield something out of their possession.

The following interjections are used as commands to dogs. nyan 'here!' is used to call a dog. The owner calls it repeatedly, in a continuous stream at an ever higher pitch. ce? 'get!' is the verb 'to be lost', used here to drive away dogs. The equivalent to drive away cats is pes!

13.2.3 Other interjections

Some other interjections are provided below, followed by a brief summary.

dom	'AFF'
k?e?	'who knows'
?anu?	[HES]

The following rough translations are given as a loose guide:

?o?	~ ?2o?	'I see, I'm following'
hmm		'agreement'
hnən		'agreement', 'so it is'
?ah		'ah!'
zeh, ze?		like English 'eh'

dom 'AFF' is sometimes used to provide an affirmative response to a question. It is usually uttered in agreement or affirmation of an accompanying statement, made either by the speaker or the addressee, as in (25) below. (See §10.5.1.3 for responses to polar interrogatives.)

k?e? 'who knows?', is used to express a lack of interest or even annoyance towards someone's question. It is the verb 'to know', uttered in a curt manner, with an emphasis on the aspiration of the initial segment, and accompanied by rising intonation.

?anu? [HES] 'umm' ← Malay *anu* is a hesitation marker which is interpolated into what is being said:

- (24) blaħan. pn-żoh dom, żanu?, p?la?, cim
blowpipe NMZ-shoot,blowpipe AFF [HES] animal bird
(It was) a blowpipe, (for) shooting indeed, umm, animals, birds.

żo? ~ żożo? is an expression of interest, interpolated by a hearer into what the speaker is saying.

The interpretation of the following depends on the intonation and the context in which they are uttered.

żah normally expresses resignation, or agreement.

- (25) "żah," k'ləj "nasip żəj=cə? mdom"
ah QUOTE fate If=EM actually
"Ah," (she) said, "(it's) my fate actually."

In other contexts it may express delight, or the realisation of something:

- (26) "żah! hə? ke," k'ləj
ah, AT:above that QUOTE
"Ah! up there," they exclaimed.

żeh expresses surprise, but not delight, sometimes accompanied by a sense of bewilderment. It is especially used when the speaker is caught off guard as in (27). In (28) it is used to introduce an admonishment.

- (27) "żeh m?ap=cə? żəj!" k'ləj
eh excuse=EM If QUOTE
"Eh, excuse me!" (he) said.

- (28) żeh, ma=pnɔn!
eh IRR=tiger.taboo
Eh, (don't do that) (it's) the tiger taboo!

13.3 Discourse clitics

The discourse clitics are difficult to characterise in terms of either semantic or pragmatic functions. Speakers were unable to provide referential meanings for any of these morphemes. In effect, the clitics are concerned with textual cohesion. They may index plot-level relations within a text where they highlight events or states which are central to the unfolding narrative, or bring events to a conclusion; others express speaker/narrator stance in terms of event verification, deduction, and so forth.

Syntactically and phonologically, the clitics enter into phrase level constructions; however semantically and pragmatically they can modify the whole clause. For the most part the clitics function at the discourse level, beyond that of the clause. They are not a necessary constituent of the clause, as is apparent from their lack of appearance in elicited data. In general, the clitics tend to occur in declarative realis clauses, apart from =sən which may also appear in irrealis clauses (§13.3.2), and =cə? IMP₁ and =ja IMP₂ which are used in the imperative (see §10.6).

This section confines itself to flagging the existence of these morphemes and provides some rudimentary notes on the possible range of functions and syntactic distribution.

The four discourse clitics are all phrase final enclitics. They are =cə? 'EM(phatic)', =sən 'SC', =ja 'CL(itic)' and pa= 'FACT(ual)'. The first two, =cə? 'EM' and =sən 'SC', exhibit the widest range of distribution which is matched by a variety of situational uses, and hence implications. These two clitics may co-occur. In contrast, =ja 'CL' and pa= 'FACT' have a narrow range of distribution reflected in a lower range of uses.

13.3.1 Emphatic =cə? 'EM'

The enclitic =cə? 'EM' is the most frequently occurring of the discourse clitics. Its main function appears to be to mark plot-level relations highlighting central events in a narrative. In this function =cə? 'EM' is only attested in realis declarative clauses.

=cə? was encountered in §10.6 as a constituent of the imperative clause where it encliticises to the verb in transitive imperatives, and to the imperative negator bɔy 'NEG:IMP' in negative transitive imperatives when the verb is elided. The clitic itself does not convey the imperative.

The fact that =cə? 'EM' is only attested in the realis mode suggests that it indexes actuality, or verification of an event, and even perhaps the prominence of a state of affairs or participant to the narrative. Evidence of this is found with respect to its co-occurrence with the two discourse connectives which indicate achievement or completion: żareħ 'at last, only then' and cəj 'from then on'. I have glossed it '=EM' to express its discourse emphatic possibilities. In the light of this, I translate it as 'and so' indicating 'as previously mentioned or implied, this resulted in X'.

The positional options are discussed in (§13.3.1.1), followed by an examination of the features of this clitic in (§13.3.1.2).

13.3.1.1 Syntactic distribution of =cə? 'EM'

=cə? 'EM' can be hosted by: a) predators; b) NPs; c) attitudinals and d) the temporal discourse connectives żareħ 'at last, only then' and cəj 'from then on'. Interjections and discourse connectives, other than those specified in d), are excluded as hosts. For examples of temporal connective hosts see (42) and (43). The host is the first available constituent in the clause (29)–(34).

- (29) A man has been longing for a child:
sampay nɔ?, dom=cə? k'et
so this AFF=EM be.pregnant
So now, so indeed, (she) was pregnant.

=cə? EM may itself host the enclitic =sən 'SC' (§13.3.2), as in (30) and (48).

- (30) ki=yok buŋlay, yok barang ɬa=dəg=cə?=sən, midur
3A=take Zingiber sp. take thing DET=like=EM=SC tree sp.
He fetched the aromatic ginger, fetched all kinds of things, midur.

Some examples and further comments on potential hosts are presented below.

a) =cə? 'EM' may be enclitic to any predictor, verbal (31)–(32) or non-verbal as in (36).

- (31) *A tiger has a human wife. Even though he is a tiger, he knows how to behave like a dutiful human husband.*

p?ot=ca? podoŋ ha? neŋ
stay=EM tiger AT before
And so (he) stayed, the tiger who was here before.

- (32) xi=thpoh=ca? me=bje? ple hmbacar
3A=pick.up=EM one=CLF fruit mangga.foetida
And so she picked up a horse mango.

In clauses which have a lexically represented object which is non-referential, the enclitic is hosted not by the verb, but the final constituent, in this case an NP:

- (33) *Some people are talking about plans to find the unidentified father of a bastard child.*

dom, ga=reŋ bapa?=ca?
AFF IMM=seek father=EM
Indeed, (they) were going father-seeking.

In complex or multi-clause constructions, the clitic can only be hosted by the main verb (34), or the verb in the initial clause (35). It is never hosted by dependent verbs.

- (34) ki=tawar=ca? ʔendɔl dom, la=sma? kke
3A=invite=EM enter.house AFF A=person that
And so indeed, that person invited (him) to come inside.

- (35) yp?=hn, ʔah ki=korek=ca?, pte? ʔen puteŋ wpy
but=CONN ah 3A=bore.out=EM put LOC tang knife
However, ah and so he bored (it) out (and) put (it) on the tang of the knife.

In the non-verbal clause in (36), =ca? 'EM' is enclitic to the negated existential predicate da? 'EXIST'.

- (36) *The protagonist has endured a few rough years wandering through the forest. Then his luck changes and he settles down.*

ʔah kira?, da? da?=ca? mandeh ?a=ksh,
ah see NEG EXIST=EM what DET=3

kde? dwa? pe? tahan
stay two three year

Ah see, and so there was nothing (more) (for him), (and) he stayed two or three years.

b) =ca? 'EM' may also be enclitic to NPs, like the free pronoun in (37), the manner NP in (38) and to the predicative NP in (39).

- (37) ʔep=ca?, ga=tar-mot ʔus
1=EM IMM=MCAUS-mount fire
So me, (I) am going to build up the fire.

- (38) ʔe?le?, "deŋ ke=ca?" kʰleg bapa?. "da?
so like that=EM QUOTE father NEG
de=pakay cokdop ye?en"
3plA=heed talk 1=AUG
So, "and so (it's) like that," said the father. "They don't heed our advice."

- (39) pokok b?iu?=ca? deh
choice friend=EM 3pl
And so they (were) at (their) friends' mercy.

c) it may be enclitic to attitudinals as in (40), the metalinguistics negator beh 'NEG' in (41) and affirmative lexemes in (29).

- (40) "ʔah dawen=ca?, lan tray" kʰleg
ah never.mind=EM want try QUOTE
"Ah, and so never mind, (I) want (to) try," (he) said.

- (41) beh=ca?, da? de=pdull? caŋ la=ksh
NO=EM NEG 3plA=concern at.all BCS=3
And so on the contrary, they weren't concerned at all by her.

13.3.1.2 Functions of =ca? 'EM'

In all of the functions listed here, the common feature is the relevance of the information to the storyline of the unfolding narrative. In the case of direct speech, it indicates relevance to the preceding discourse, marking either a change in state or making a report on progress so far. The most transparent example of this function is where the enclitic is used with a temporal discourse connective (§11.5.4), clearly indicating the relevance of the event to the storyline:

- (42) *A boy was impatient to be circumcised, to become an adult. At last it happened.*
ʔarch=ca?, sunst
only.then=EM circumcise
And so, at last (he was) circumcised (which he hadn't been before).

- (43) caŋ=ca?, xi=təŋoy ha? ke brapa-brapa? ʔare?
from.then.on=EM 3A=wait.for AT there RDP-how.many day
And so from that point on he waited there (for) several days.

The presence of the clitic may highlight the point where a new state of affairs actually obtains, as in (29) above.

In the extract presented in (44), the clitic marks an event as central to the unfolding story. The husband sets sail, even though his wife is pregnant. Later on, she dies alone whilst giving birth, the event which leads to her transfiguration into a birth-vampire.

- (44) dos he? ?ma? bapa? he?, "bapa?" kʰleg, "ʔma?",
arrive AT:above parents AT:above father QUOTE mother
ʔulŋ kmpen!" kʰleg, "kmpen," kʰleg "ra?wen,
go.back.and.forth wife QUOTE wife QUOTE leave.behind
ʔep ga=b-layar, ga=reŋ pakayan knon." b-layar=ca? ksh
If USE-sail IMM=seek clothing offspring MID-sail=EM 3S
(He) arrived up at his parents. "Father," (he) said. "Mother, go back and forth (to my wife!)" (he) said. "(My) wife," (he) said. "(I'm) leaving (her) behind. I'm going sailing, going to seek clothing (for) the child." And so he sailed off.

A second function of =ca? 'EM' indicates a protagonist's resignation to a state of affairs. Again this is underscored by the significance of the event to the storyline:

- (45) A pauper has been handed a knife and sugar cane, indicating that he is the father of a noble child. He does not want to face it, but cannot flee.

ya ta? ya, ki=spl=cə? hn=bos
yes no yes 3A=peel=EM O=sugarcane
And so, with that, he peeled the sugar cane.

- (46) A orphaned girl who lives alone in the forest is being courted by a ghost. He remarks on her apparent fearlessness and she replies, her resignation inferred:

pokok kp=cə?, len kp=cə 2əj
choice 2f=EM want 2fA=eat If
Well, (it's) your decision, (if) you want to eat me.

- (47) Against everyone's advice, a father resolves to try:

sampay dom=cə? ki=duga?
so AFF=EM 3A=try
So, indeed he tried.

In direct speech, this encompasses emphasis and conviction:

- (48) "ʔəh bah=cə?=son" kʰəŋ
oh NO!=EM=SC QUOTE
"Oh, no way!" (he) said.

13.3.2 Speaker conclusion =son 'SC'

As with =cə? 'EM', it is difficult to assign a function to =son 'SC', although a characterisation of the clitic would suggest that it is used primarily in relation to background information, and the deduction or drawing to a conclusion in terms of the narrator's or speaker's belief or opinion: 'due to the event or state X, I draw the conclusion Y'. It is glossed 'SC' 'speaker conclusion' and is translated as 'evidently' or 'naturally'.

=son 'SC' was never observed in materials obtained through direct elicitation. It occurs infrequently; in a rough count in a selection of two traditional narrative texts from one speaker the following rates of occurrences were recorded: 13 out of 112 clauses in the first text and 1 from 130 clauses in the second. A general description of the distribution and functions of =son 'SC' is set out below in §§13.3.2.1–2.

13.3.2.1 Syntactic distribution of =son 'SC'

=son 'SC' is used in both realis and irrealis clauses; it is the only enclitic compatible with the irrealis mood.

=son 'SC' may be hosted by one of the following constituents of the initial constituent in the clause, other than the discourse connectives c) below. Possible hosts are: a) a predicate; b) an NP; c) the discourse connectives: ?le?le? 'then' and sampay 'so'; d) negative and affirmative lexemes, or e) the temporal clause connective kna? 'it happened that when ...' and jadi? 'when'.

Some examples and further comments on the hosts listed above are presented below.

- a) The enclitic only attaches to main verbs; it is never found in dependent clauses. The verb may host the irrealis proclitic (51).

- (49) da? ki=kali=son, b-knlek
NEG 3A=be.brave=SC HAVE-husband
She wasn't brave (to) marry.

- (50) risaw=son, gres
be.sad=SC liver
(He) evidently felt sad. (Lit. (His) liver was sad.)

- (51) ma=gos=son nm-k'om deŋ ke
IRR=slice=SC NMZ-sit like that
(She) evidently spliced (it) sitting like that.

b) It can be enclitic to an NP: a noun as in (52), a demonstrative as in (53) or the subject of an NVC as in (54).

- (52) kira? knon, dom twanptri?=son,
see offspring AFF princess=SC
ma=gos dom ple hmbakan ke
IRR=peel AFF fruit horse,mango that
See the child, indeed evidently the princess, (she) peeled the mango.

In (53) the demonstrative hosting the clitic is co-referential with the NP_O of the following imperative. The clitic appears to be linking the independent constituent.

- (53) "dom no?=son," kʰəŋ, "no? cəŋ, masuk kloc rban!"
AFF this=SC QUOTE this mousedeer enter into pen
"Indeed this," he said, "this mousedeer, put (it) in the pen!"

- (54) kira?, me=ki=bunuh dom rəm sma? gdo=son, dom ʔadi?
see REL=3A=kill AFF WITH person be.old=SC AFF YS
See, indeed the one he killed accompanied by the old man, indeed (he was) the younger sibling.

c) =son 'SC' can be enclitic to the following discourse connectives: ?le?le? 'then', sampay 'so' and jadi? 'happen, when'. It never occurs on a discourse connective if it is already encliticised by =hn 'CONN'. In (55) it is enclitic to the second of the connectives:

- (55) sampay, ?le?le?=son, cin, kira? ninca
so then=SC be.cooked see food
So then evidently, the food see, was cooked.

d) Enclitic to negative lexemes as in (48) above, affirmative lexemes (56) and interjections (57):

- (56) dom=son, kira? bʔla? br-tsh
AFF=SC see friend MID-untie
Indeed evidently, see, the friend (was) untied.

- (57) ?ah=son, dom kəh, kira? de=bt'ɔŋ=son
ah=SC AFF 3 figure 3plA=be.afraid.of=SC
Ah evidently, indeed, him, see they were evidently afraid of (him).

e) When =son 'SC' co-occurs with the clausal connective kna? 'it happened that when ...', it is enclitic to the final constituent of the clause:

- (58) A father awaits the birth of his child. As soon as he is born, he goes wandering off again, abandoning his wife and child.
 kna? cəh knon=son, swak
 happen.when be.born child=SC go
 It happened that when the child (was) born, (the father) went.

13.3.2.2 Functions of =son 'SC'

Two contexts are identified in which =son 'SC' is used: a) to draw a conclusion based on either evidence from the narrative, or the speaker's personal belief, and b) to express a natural consequence. The second function is usually confined to backgrounded clauses.

a) Speaker belief or the drawing of a conclusion. The following is a good example of the expression of speaker conclusion, embedded in a quotation.

- (59) A mother's death is attributed to the fact that her child is ill-omened.
 k'bəs ʔmaʔ, bən de=cceʔ? la=?c? kakaʔ=hn
 die mother NEG 3plA=hate A=BB EZ=3POSS
 "knkən clakaʔ=son" kʰlaŋ
 child ill-omened=SC QUOTE
 The mother died, (and) no, his elder brothers and sisters hated (him). "Evidently, an ill-omened child," (they) concluded.

- (60) Q: do kadeh do?
 of who of
 A: do?, dde=yə=son
 AFF OF=I=SC
 Q: Whose is (it)?
 A: Indeed, I'd say (it's) mine.

The clitic is also used in clauses in which the speaker makes an assertion about a third person's internal state (61), and (49)–(50) above.

- (61) sampay pəŋel tɔʔ. da? de=məh=son
 so call AT:across NEG 3plA=want=SC
 So they called across. Evidently, they didn't want to.

=son 'SC' is used to mark the closure of a narrative (62). It may not be an actual endpoint, but it is where the speaker chooses to finish the retelling.

- (62) kməŋ crite=son
 finish story=SC
 (Evidently) the story is finished.

b) **Consequence.** The clitic is also used to highlight an event as relevant to determining what will happen next. In this respect it differs from =ce? 'EM' which marks an event as relevant in terms of progress so far. Again, the clauses in which it co-occurs with the discourse connectives provide evidence of its function. It is particularly frequent with the conjunction kna? 'when', indicating a background state required to set the conditions in order for the next event to take place:

- (63) kna? tʰey, jlaŋ=son, kəh len sunot
 happen.when be.big be.long=SC 3 want circumcise
 Naturally, when he (grew) big (and) tall, he wanted (to be) circumcised.

In (64), a woman lies down and goes to sleep, and out of this a new chain of events develops:

- (64) ʔluc.təm masak, t-pekbər len jtək war
 after cook HAPP-think want sleep instead
 ki=ʔuray hn=suk, məcom dde=ʔəp deŋ ke
 3A=let.down O=hair like OF=1 manner that
 dm=dəm ha? bnul ponog kəh, jtək=son.
 IMPERF-lie.down AT beam hut 3 sleep=SC
 sampay deh 2yot, ki=ʔintay
 so 3pl return 3A=stalk
 After cooking, (she) happened to feel like sleeping instead. She let down her hair, like mine, in this manner, lay down on the beam of the hut, (and) evidently (she) fell asleep. So they returned (and) he crept up (on her).

- (65) A mousedeer has been trying to entice a macaque into taking his place in the enclosure. The macaque finally agrees.
 krwpl cəŋ, ki=ma<r>suk=son hn=bʔia? ke
 come.out mousedeer 3A-enter<CAUS>=SC O=friend that
 Mousedeer came out, (and) naturally he got the friend to go in.

The events described in the irrealis mood are actual within the context of the narrative and not associated with non-actuality:

- (66) ma=cokel=son rəm dlog ma=center
 IRR=pierce=SC WITH stick IRR=carry.by.side
 Evidently (it) was pierced with a stick (and) (it) was carried.

13.3.3 Clitic =ja 'CL'

=ja 'CL' is attested infrequently. Formally, it would appear to be a borrowing from the Malay *nya* 'third person possessive pronoun', however functionally the resemblance is tenuous.

=ja 'IMP₂' encliticises to the verb in transitive imperative clauses (§10.6.2) and to the negative imperative marker boy 'NEG:IMP' when the verb is elided (§10.6.3). It is interchangeable with the enclitic =ce? 'IMP₁' (§13.3.2). In this function the enclitic bears the clause-final stress usually associated with the verb (67), unlike elsewhere where it is not stressed (68). This observation also holds for =ce? 'IMP₁'.

- (67) "duga?=ja ?ma?" kʰlaŋ
 try=IMP₂ mother QUOTE
 "Try mother!" (he) said.

- (68) kəhn pon 'sar=pa ʔintc
 3S TIJEN descend=CL exit.house
 She thereupon went down outside.

Elsewhere =pa is used in transitive and intransitive declarative clauses, enclitic either to the verb (69), or an NP (70).

- (69) yok sarek dom dos=pa ptom
take tomorrow AFF arrive=CL night
The next day, indeed (they) arrived (at) night.

In the following clause =pa is enclitic to the NP complement of the intransitive verb:

- (70) dm-dəm kəh, "ga=jtek mot=pa" kʰləŋ
IMPERF-lie.down 3 IMM-sleep eye=CL QUOTE
She lay down, "(My) eyes are going to sleep (on me)," (she) said.

In transitive clauses =pa 'CL' appears to function anaphorically, linking events or participants. In (71), a plate of fruit has been set down on the floor, and those present invited to partake. The following was given as a polite form of acceptance, accompanying the speaker's action of taking the fruit:

- (71) ye=yok=pa muy
1A=take=CL one
I'm taking one (of them).

In (72) the enclitic would appear to be anaphoric with tulis 'writing'.

- (72) da? məcəm tulls, de=kʰe?=pa, ma=crdek
EXIST like write 3plA=know=CL IRR=be.clever
(There) was writing, they knew (it), (those) (who) were clever.

In (73) it would appear to be the proposition of discarding the new husband, the unexpressed COMP of the final main verb kali 'be brave (to)' that is being encoded, rather than just the NP.

- (73) "ʔap" kʰləŋ-hn "da? wen kniék ma=jaman,
If QUOTE=3POSS NEG discard husband REL=be.former
ma=?arch pon" kʰləŋ "da? ʔap=kali=pa"
REL=be.old THEN QUOTE NEG IfA=be.brave=CL
"I," she said, "didn't discard my former husband, (and) as for the new one," she
said, "I'm not brave (enough) to (either)."

In the final example it is not enclitic to a verb, but =pa 'CL' is hosted by the deictic manner expression and refers forward to the number of visits:

- (74) A ghost courts a woman:
dom deg ke=pa, b-biləŋ ptom ki=?uləŋ-i?
AFF like that=CL MID-count night 3A=go.back.and.forth-JTER
Indeed, (it was) like that, count the nights he kept calling by (on her).

In intransitive clauses, the function of pa 'CL' is not obvious, and no conclusions are drawn. An example of its use in an intransitive clause was given in (70) above.

13.3.4 Factual =pa 'FACT'

=pa 'FACT' is associated with assertions expressing a firm conviction of fact on the part of the speaker, and is simply glossed 'FACT'. It is attested enclitic to verbs (75), pronominals (76) and ignoratives (77). It occurs infrequently.

- (75) hūm=pa deh. shnəh ga=dmpək
bathe=FACT 3pl be.reluctant IMM=follow
The fact is, they are bathing. (I'm) not up (to) following (them).
- (76) ns-gəs də gəp nə?=pa
NMZ-peel OF malay this=FACT
The fact is this is the Malay way of peeling.
- (77) "mandeh=pa hne lən b-c?ŋ?" kʰləŋ "kníék"
why=FACT THEN want MID-burn QUOTE husband
"(For) what reason does your husband want (it) burnt?" (he) wondered.

14 Texts

The text collection presented here provides a sample of Semelai oral literature. Section 14.1 is a selection of *ttikyan* 'riddles'; §14.2 consists of excerpts from an elderly woman's autobiographical narrative; in §14.3 a man talks about the differences between Semelai practices of the present day, and those of previous generations, and in §14.4 a collection of four *cnnman* 'traditional narratives' is presented.

The texts are presented in the same format as the examples. Note that where the text is bracketed [] an exchange takes place between a member of the audience and the principal narrator.

14.1 *ttikyan* 'Riddles'

To 'tell riddles' is *b-ttikyan*. The term comes from the Malay *teka teki* 'riddles, brain-teasers'. If one doesn't know the answer to a riddle, one replies with *?apacom ?ahem*. Speakers were unable to gloss this construction, only transliterate it as *da? ye=k?e?* 'I don't know (it)'.

The following riddles were collected in a group situation from an extended family, ranging from an elderly couple to their young grandchildren. The session began spontaneously when we were sitting around one evening and someone happened to ask me one of these riddles. Everyone then proceeded to join in, taking great delight in my attempts to come up with the correct answers.

Each clue is presented here together with the answer (A:). The riddles are not only entertaining, but provide interesting insights into the Semelai perception of their environment. Explanatory notes are provided where necessary.

1 *dr? kp-k^bop,* *k^beh pa<r>jgar*
 Iself IMPERF-lie.face.down 3 lie.face.up<CAUS>
 We lie face down, he is made to lie face up.

A: *kaltog*
 knee
 Knee

2 *ma=ptam co?* *pamah, pdo? ht? c^bng*
 REL=plant AT:below flat.land sprout AT:above hill
 The one planted down on the plain, sprouts up on the hill.

A: *sigeret*
 cigarette
 A cigarette.
The cigarette is placed in the mouth, the smoke emerges from the nose.

- 3 *tihag dwa? d<p>loj;* *hatap=hn dwa?*
 post two tree<UNIT> thatch=3POSS two
 Two house-posts; it's thatch, (consists of) two (sheets).
- A: *hayam*
 chicken
 A chicken.
Think of a Semelai house up on its posts, these are the chicken's legs. The two sheets of thatch are the wings folded over the chicken's head.
- 4 *co? 7ate, largemilar?; he? dol*
 AT:below ground writhing?; AT:above house
 sot papan
 metamorphose.into board
 Down on the ground, snake-like (it) writhes; up in the house it becomes a board.
- A: *tikar*
 mat
 A mat.
*The speaker defined *largemilar* as '*b-bway, ki=tut ribut, mocom tijo*' (MID-sway 3A=blow wind like snake) '(When) the wind blows (it), (it is) like a snake (writhing)'.*
- 5 *troj dwa?; mə=?ikur sma? mi=tən sok*
 path two one=CLF person one=instance along
 (There are) two paths; one person (goes) along (both) at once.
- A: *srwol*
 trousers
 A pair of trousers.
- 6 *swara? t^bay; kba? k^bt*
 voice be.big body be.small
 A big voice; a small body.
- A: *bdil*
 gun
 A gun.
- 7 *co, kuceŋ, hayam, ki=jon muka?; raja?,*
 dog cat chicken 3A=give face sultan
 pjhulu?, ki=c<ar>loŋ
 headman 3A=back<EQUIP>
 (To) dogs, cats (and) chickens, he shows his face; (to) sultans (and) headmen, he turns (his) back.
- A: *dij*
 bamboo
 Bamboo (flooring).
This is the split bamboo used for flooring. To the animals on the ground, the inside or 'face' of the split bamboo faces down toward them. In the house, the occupants see the exterior, or 'back' of the bamboo.

- | | | |
|--|---------------------------------------|-------|
| 8 | podɔŋ̩ mrwɔŋ̩, krbaŋ̩ b-tali? | hmpon |
| | tiger growl buffalo HAVE-mooring four | |
| The tiger growls, (but) the buffalo (has) four moorings. | | |

- A: smoj rom klambu?
mosquito WITH mosquito.net
Mosquitos (outside) a mosquito net.

- | | | | | | |
|--|------|-----------------|-------|--------|-------|
| 9 | baba | me=t<n>arkay | sbon | kloc | dɔl |
| | padi | one=stalk<UNIT> | store | inside | house |
| One stalk of padi stored in the house. | | | | | |

There are two possible answers to this riddle:

- | | |
|---------------|------------------|
| a) <i>Tus</i> | b) <i>plita?</i> |
| <i>Fire</i> | <i>lamp</i> |
| <i>Fire.</i> | <i>A lamp.</i> |

- 10 kwarga? ma=7ye, ma=goy b-cokop, da? de=sahot
 family IRR=sec IRR=bring MID-speak NEG 3plA=reply
 The family is visible; (but) if (one were to) speak (to them), they wouldn't reply.

- A: gambar
photo
A photograph or picture.

- II ma=c^ək ʔəp; ma=tarek dk^əes
 JRR=throw.with.aim be.far IRR=pull be.close
 (They) can be thrown far; (they) can be drawn close.

- A: mot
eye
Eyes.

- 12 *ma=tul^k* *he?* *k'hoj;* *ma=pawot* *co?* *pigan*
 IRR=set.off AT:above head IRR=lay.across AT:below waist
 Pushed away at (its) head, laid across at (its) waist.

- A: p̄jauh
paddle
A paddle.
pawot is a term used to describe the laying across of the fibres in mat-making.

- 13 swara? ma=?yəŋ; ma=?ye, beh
 sound JRR-hear IRR=see NO
 The sound can be heard; on the contrary, (it) can't be seen.

- A: guruh
thunder
Thunder.

- 14 *təŋ da?*; *yŋ-yəŋ* beh
 ear EXIST IMPERF-hear NO
 (it) has ears; (it) hears not.

- A: kwall?
wok
A wok.
The word təŋ 'ear' has a secondary meaning 'handle'.

- 15 bubuh tkoh; crsh beh
fish.trap be.wet fish NO
The fish trap is wet; (but) not the fish (inside).

- A: do
 house
 A house.
 The

- 16 p_gor b_{lu?} ma=ye; p_gor dri?, beh
 fence friend IRR=see fence lself NO
 (Our) friend's fence can be seen; our own fence (can)not.

- A: *Imp*
tooth
Teeth.

- 17 bnum b?lu? ye=?ye; bnum ye, beh
 mountain friend 1A=see mountain 1 NO
 (Our) friend's mountain, we see; our mountain, (we) don't

- A: knee
forehead
Forehead.

- 18 b-sisik, b?en da? naga?; b-jambol, b?en da? syol
 HAVE-scale NOT EXIST cobra HAVE-tuft NOT EXIST crest
 It has scales, but (it's) not a cobra; it has a tuft, but (it's) not a crest.

- A: nas
pineapple
A pineapple.
The teacher

- | | | | | | |
|----|--|----------------|-------------------|--------------|------------|
| 19 | dlong | ma=d<ŋ>lɔŋ, | kɪpm | ma=ribu? | sma?, |
| | tree | one=tree<UNIT> | carry.on.shoulder | one=thousand | person |
| | da? | b-glet. | ʔarəch | b-baʔi, | ki=glet |
| | NEG | MID-be.able | be.new | MID-be.alone | 3A=be.able |
| | One (felled and cleaned) tree, (if) one thousand men (were) to carry it on (their) | | | | |
| | shoulders, (they wouldn't) be able, (but) alone, it is able (to carry them). | | | | |

- A: troj
path
A path.

- 20 *7ac* *mande* *mə=də?* *sɪt?*
excreta *what,sort* *REL=NEG* *stink*
 What sort of excreta doesn't stink?

- A: ?əc dər
excreta rattan
The by-product from (preparing) rattan.

14.2 Recollections

The following excerpts are from a long autobiographical narrative, spoken by a woman, kɔp, aged approximately 70 years.

14.2.1 Establishing the identity of ga=tɔʔɛ?

The narrator has been talking about the period of the Communist Emergency (a period of civil unrest from the 1940s until the early 1960s, see §1), and in particular about an elder male 'sibling', actually a first cousin, who had assisted the Communist insurgents.

The Communist insurgents, who were Chinese, sought help from the Semelai in order to negotiate their way through and survive in the jungle. The British rulers had put a price on the head of anyone who aided and abetted the enemy. The 'Europeans' mentioned in this text are the British forces.

- 1 da? ʔiʔɛ?, dom sdara? mə=br-dɔl hɛ? kke d2en
EXIST EB AFF relation REL=HAVE-house AT:above that long.ago
(I) had an elder male sibling, yes a relation who had a house up there back then.
- 2 kʰbəs
die
(He) is dead (now).
- 3 ʔiʔɛ? yɛ=?en=cəʔ=sən, dom ??ɛ? gdo tuli? ke
EB 1=AUG=EM=SC AFF EB be.old be.deaf that
(He was) an elder brother (to) us all, see; indeed, (he was) the elder brother (of) that old deaf (one).
- 4 ji=kʰɛ? gato?ɛ? ke
2A=know [name] that
You know that gato?ɛ?.
- 5 ??ɛ? dom kəh, ʔabəŋ
EB AFF 3 EB
(He was) her elder brother. Elder brother.¹
- 6 [kampong bapak, mə=ha? dɔl baten]
[settlement [top] REL=AT house batin]
[(You mean at) Kampong Bapak, the one at the headman's house.]
- 7 dom kke=cəʔ
AFF that=EM
Yes, that one.

¹ The narrator adds the term *abang*, the Malay term for 'elder male sibling' and the equivalent of ??ɛ?, making sure I understand the relationship.

- 8 ʔiʔɛ? sprot, kira? sə=ma? sə=bapa?
EB one.womb figure one=person one=father
(He was her) elder brother (from) the same womb, (you) know, the same mother and father.

14.2.2 gato?ɛ? and her elder brother

In the following excerpt, the narrator continues talking about gato?ɛ?'s behaviour toward her elder brother, and she makes her own feelings known. In this excerpt the speaker refers to gato?ɛ? by her kin term, ?ma?itam 'fourth-born sibling, female'.

- 1 ki=lən heler
3A=want travel.downstream
(He) wanted to travel downstream.
- 2 dom ma?itam he? ke ki=bt̥ɔŋ
AFF fourth.born.female AT:above that 3A=fear
Indeed, his fourth-born sibling up there, he was afraid of (her).
- 3 bt̥ɔŋ
be.afraid
(He) was afraid.
- 4 ki=bt̥ɔŋ la=molot, molot da? ʔilɔk
3A=fear BCS=mouth mouth NEG be.good
He was afraid of (her) mouth, (her) bad mouth.
- 5 ki=par-yəŋ-iʔ ʔen sma? putih, ʔen cina?
3A=CAUS-hear-ITER LOC person be.white LOC chinese
She repeatedly informed to the Europeans, (and) to the Chinese.
- 6 la=?ma?itam yɛ, ki=bt̥ɔŋ
CAUS=fourth.born.female 1 3A=fear
(It was) because of our fourth-born sibling, (that) he was afraid.
- 7 sdom nɔʔnɔ? ki=sayag
only this 3A=be.fond.of
Only this one, she was fond of (him).
- 8 [dom ?ma? ʔayəl neŋ]
[AFF mother [name] before]
[Actually, the mother of ʔayəl (who was) here just now.]
- 9 la=rəm sma? putih, da? ki=par-yəŋ=sən,
BCS=WITH person be.white NEG 3A=CAUS-hear=SC
komunis, da? ki=par-yəŋ
communist NEG 3A=CAUS-hear
Because the Europeans, she didn't tell (them), see. The communists, she didn't inform (them either).
- 10 ?ma?itam he? ke ki=par-yəŋ ʔeh dom,
fourth.born.female AT:above that 3A=CAUS-hear ah AFF
məcəm da? ʔadi?, ʔiʔɛ? kəh mə=gdo
like NEG YS EB 3 REL=be.old
Fourth-born sibling up there, she informed; ah indeed, like (she was) not her elder brother's younger sibling.

- 11 ??c7 yc, msti2 ye=sayan
EB 1 must 1A=be.fond.of
(If he were) my elder brother, (I) would have to be fond of (him).
- 12 dom
AFF
Yes.
- 13 da? ji=k^be?, ? pspl ye=?y^bg grgor bom ha? ke, ye ke
NEG 2A=know if 1A=hear rumble bomb AT there I that
sampay da? moh ca, ia=t-peker hn=kehn ke
so.that NEG want eat BCS=HAPP-think O=3S that
You don't know (this), (but) whenever I heard the rumbling of the bombs there, I'd
get so that I couldn't eat, on account of thinking of him.
- 14 yo? beh, nasip ?iick tom ye
but NO fate be.good SRC 1
But no, in my opinion (his) destiny was good (and I need not have worried).
- 15 k^bbas kmuc ?a=ksh=h3?
die ghost DET=3=ATTN
He died a natural death.²

14.3 Then and now

The narrator, a middle-aged male talks about past customs relating to the disposal of corpses. He starts out by setting the scene. In earlier times, the Semelai didn't live in settlements, but in their swiddens in the manner explained below. This pattern of settlement was forcibly disrupted by the British (referred to as the *sma?* *putih* 'Europeans'), during the Communist Emergency in the 1950s-1960s. The Semelai were made to live in settlements (*kampot*). The majority of the community has remained in these until this day.

- 1 beh mocom no?no?
NO like this
(It) wasn't like this.
- 2 me=c7co? muy, muy, d7oh
REL=in.front one one house
(For our) forebears (it was) one (family), one swidden.
- 3 muy d7oh, muy dol, dloq ddes
one swidden one house tree be.close
One swidden, one house (and) trees nearby.
- 4 b2en mocom no?no?
not like this
(It) wasn't like this.

² k^bbas kmuc (die ghost) 'means 'to die naturally'. See Text 3 §14.4. See §14.4.3 for a commentary on ghosts.

- 5 no?no?, dloq brapa? ?ap, susah
this tree how.many be.far be.difficult
Now, the trees are so far, (it's) difficult.
- 6 tyap tyap d7oh
every every swidden
(Before) every (one) had a swidden.
- 7 masa? kampot, kampot sma? putih d2en
era settlement settlement person be.white back.then
(Now) it's the era of the (village), the villages of the Europeans from (that time)
back then.
- 8 ?yot=c? ttap
return=EM be.firm
(It's) fixed (residence).
- 9 solon=hn beh, kira?=c? ?asej sdek?et
beginning=CONN NO figure=EM be.different little
In the beginning, no (it wasn't like that), see, (it) was a bit different.
- 10 tu?aki-tu?aki? ye=?en mocom tuyob-muyob mocom jace?
RDP-grandfather 1=AUG like RDP-great.grandparents like likewise
(For our) grandfathers (it was) the same as for their great-grandparents (and) so
forth. ((It was) different back through the generations.)
- Note in the next excerpt from the same text how the narrator draws a comparison between past customs and the preservation of the knowledge of these customs in the c<nn>man (cman<NMZ>) 'traditional narratives'. When someone died, as he explains, the corpse was smoked in the house. The residents would abandon the house and swidden and stay with another family in a different location. The narrator doesn't actually state what the ghost's bad behaviour involved.
- 11 bila? k^bbas, de=?ur kem
when die 3plA=instruc bury
When (someone) died, they instructed (them) to bury (them).
- 12 ha? ke mocom 2en c<nn>man
AT that like LOC retell.myth<NMZ>
In (earlier times) there (it was) like in a traditional narrative.
- 13 lbeh kuraj, crte c<nn>man=c?
more less story retell.myth<NMZ>=EM
More or less (like) a story, a traditional narrative.
- 14 dom bagay no?no?
AFF like this
Yes, like this.
- 15 dom no?, ia=b-bt^boy, dom la=kmuc jadi? wer
AFF this BCS=MID-fear AFF BCS=ghost happen instead
pakay br-kem
practise MID-bury
Indeed (it was) this, on account of being afraid, indeed (because of their fear) of the
ghosts, so instead (they) practised burying (them).

- 16 de=ur br-kem
3plA=instruct MID-bury
They instructed (that they) be buried.
- 17 soloj b-ptet? ten pran mpcem no?no?, b-salay
first MID-put LOC mat.rack like this MID-smoke
In the beginning, (they) were put in the mat racks, like this, (and) smoked.
- 18 b-salay tom co? ?ate ?us t'ey, sampay krey
MID-smoke SRC AT:below earth fire be.big until be.dry
Smoked from down on the ground, (with) a big fire, until (they) were dry.
- 19 sampay krey sma?; dom
until be.dry person AFF
Until the person was dry; yes.
- 20 [dol b-ra?wen?]
[house MID-abandon]
[Was the house abandoned?]
- 21 dol b-ra?wen
house MID-abandon
The house was abandoned.
- 22 y?=?hn, b-salay rom ?us t'ey co?
but=CONN MID-smoke WITH fire be.big AT:below underneath
In contrast (to nowadays), they were smoked by a big fire down in the space underneath the house.
- 23 y?=?hn dom lan j?oy jahat kmuc, da? de=jon
but=CONN AFF want do be.bad ghost NEG 3plA=give
w? ma=j?oy dej ke
longer IRR=do manner that
However, indeed, the ghosts wanted to do bad, (and so) (the elders) they no longer permitted (that) they would do (it) like that.
- 24 de=ur kem
3plA=instruct bury
They instructed (that they) bury (the dead).
- 25 jadi? dom no?=hn, br-kem=co?; jam no?=co?, pakay br-kam
then AFF this=CONN MID=bury=EM time this=EM use MID-bury
So indeed nowadays, see (they) are buried, (up until) the present day, burial is practised.
- 26 ?asol br-kem me=c?co? da? bagay no?no?
if MID-bury REL=be.in.front NEG like this
When (they) were buried, our forebears, (it was) not like this.
- 27 no?, pakay b-simin
this practise MID-cement
Now cementing (is) practised.
- 28 c?co? beh, b-ptet? hn=dloj.tros
be.in.front NO MID-place O=heartwood
(In) earlier (times), in contrast, hardwood was placed (on the grave).

- 29 b-rej dloj tha-than mpcem dloj cnay; ha? me=than
MID-seek wood INTNS-endure like wood [name] AT REL=endure
Really lasting wood was sought, like cnay wood; that (is) one which lasts.
- 30 me=c?co?=s?n beh; pakay mpcem no? ha? c<nn>man
REL=be.in.front=SC NO use like this AT retell.story<NMZ>
Our forebears see, no; they did (it) like this (as if) in a traditional narrative.
- 31 ?asen, pakay b-salay, lag? me=c?co?
be.different practise MID-smoke still REL=be.in.front
w? pakay dom dloj papan=ca?
instead use AFF wood plank=EM
(It) was different, (they) still used smoking, further, (our) forebears (then) in their stead (they) used planks (of) wood see.
- 32 kira? me=rajin, me=j?oy pakay dloj mpcem papan kke
figure REL=be.industrious REL=do use wood like plank that
(I) reckon, (they) were the industrious ones, the ones who did (it) using wood like the planks there.
- 33 b-sapu? rom simin do? no?, dej ke=c?o?
MID-wipe WITH cement AFF this manner that=EM
Indeed, (they) are smeared with cement now, (it's) like that, see.
- 34 ?asol me=c?co? ?asen; no?no? ?asen
if REL=be.in.front be.different this be.different
As for (our) forebears, (for them it) was different; now (for us it) is different.

14.4 Traditional narratives

Four short traditional narratives (*c<nn>man* (retell.narrative<NMZ>)) are presented below. The first three narratives were retold by Normah Het, a 30-year-old woman. The fourth narrative is retold by ?amay (Amrita Isa), an elderly woman, renowned for her passion for story-telling.

14.4.1 *c<nn>man sma? ?adi?bradi? rom romgasi* 'The narrative of the siblings and romgasi'

In this first text, *c<nn>man sma? ?adi?bradi? rom romgasi*³ (retell.narrative<NMZ> person sibling with romgasi), the siblings encounter the ogre romgasi. mijan is both the fine hairs inside bamboo and the smarting itch caused by them. Note how the breath ?ahom, which is life itself, is kept in a container.

- I swak kloc bri ?h-?h,
walk inside jungle IMPERF-shoot.with.blowpipe
t-harog rom dol sma? gdo
HAPP-encounter WITH house person be.old
They were walking in the jungle blowpiping (when) they happened to come across the house of an old person.

³ The ogre is also known as rengasi, a name used by Malays for 'giant'.

- 2 ʔendol ha? dol
enter.house AT house
(They) went into the house.
- 3 lpac deh sleh
stomach 3pl be.hungry
They were hungry.
- 4 de=jpk co? habu? thi sma? kloc takar
3plA=observe AT:below kitchen hand person inside jar
Down in the kitchen, they saw a human hand in a jar.
- 5 de=jpk he? para? kba? sma?
3plA=observe AT:above rack body person
They saw a human body up on the rack.
- 6 beh da? de=lən de=ca
NO NEG 3plA=want 3plA=cat
No, they didn't want to eat (that).
- 7 de=yok jalu yok crəh, de=ca
3plA=take pig take fish 3plA=eat
They took pork (and) took fish, (and) they ate (that).
- 8 ?luc.təm de=ca, de=?int̪, ?j-?ɔj karom pontɔŋ
after 3plA=eat 3plA=exit.house IMPERF-hide underneath firewood
After they had eaten, they went down outside (and) hid themselves underneath the firewood.
- 9 dom sma? gdo ʔyot ha? dol
AFF person be.old return AT house
Indeed, the old man came home.
- 10 ki=pam kʰiem deh
3A=sense odour 3pl
He smelt their odour.
- 11 ki=reŋ la=lən ki=ca
3A=seek A=want 3A=eat
He sought (them), because (he) wanted to eat (them).
- 12 beh, da? ki=croh
NO NEG 3A=meet
But no, he didn't find (them).
- 13 sar co? habu?, ki=jpk lawok kmɔŋ
descend AT:below kitchen 3A=observe game finish
(When he) went down to the kitchen, he saw the food was finished.
- 14 dom de=ca ia=dəh
AFF 3plA=eat A=3pl
Indeed, they had eaten (it).
- 15 da? de=kʰt̪? lən mmbunuh la=sma? ?adi?bradi?
NEG 3plA=know want IMPERF:kill A=person sibling
The siblings hadn't known (that) (he) would want to kill (them).

- 16 t-masuk ha? gres ?adi?
HAPP-enter AT liver YS
Then (the) younger sibling had (an idea) come (to him).
- 17 ki=?ajak ?t̪e?
3A=invite EB
He invited the elder brother (to join him).
- 18 ki=reŋ mipaq rboŋ din ki=tabor ha? dol, he?
3A=seek fine.hair shoot bamboo 3A=sprinkle AT house AT:above
j<n>tek, co? habu?,
sleep<NMZ> AT:below kitchen
co? ʔate?, sok troŋ tlaga?
AT:below ground along path well
He sought the fine hairs of the bamboo shoot and he sprinkled (them) in the house, up in the sleeping area, down in the kitchen, down onto the ground, (and) along the path (to the) well.
- 19 kehn ʔyot tom bburu?
3S return SRC hunt.with.dog.and.spear.
He (the old man) returned from hunting with his dog and spear.
- 20 ʔendol he? dol tar-koh baju?
enter.house AT:above house CAUS-undo clothes
(He) went up into the house (and) took off his clothes.
- 21 ki=pam mipaq kba?
3A=sense itchiness body
He felt the itch (on his) body.
- 22 yar he? j<n>tek, pon mipaq
ascend AT:above sleep<NMZ> too itchiness
(He) went up to the sleeping area, (and) he was itchy (up there) too.
- 23 sar jɔh co? habu?, lagi? kwat mipaq
descend drink AT:below kitchen still be.strong itchiness
(He) went down to drink in the kitchen and the itchiness was stronger.
- 24 sar te?en tlaga?, mipaq yar=co?
descend TO:down well itchiness increased=EM
(He) went down to the well (and) the itchiness increased.
- 25 dom, sma? ?adi?bradi? yok taboŋ ?ahom
indeed person sibling take vessel breath
Indeed, the siblings took the vessel (containing) (his) breath.
- 26 ki=?ye deh, "ha?ha? lawok ?əp,"
3A=see 3pl DEM-AT game If
yə?hn ki=jpk taboŋ ?ahom ha?hn deh
however=CONN 3A=observe vessel breath AT=ABS 3pl
He saw them, "these ones will be my dinner." (he thought), however he saw they had the vessel (containing) his breath.

- 27 de=yok rmp^{ha}l co? ?ate
3plA=take throw.down AT:below ground
They took (it) (and) threw (it) down on the ground.
- 28 de=tam, ?areh kəhn k^hbəs
3plA=stamp just.then 3S= die
They stamped (on it) (and) then he died.
- 29 kməŋ
end
The end.

14.4.2 sma? ki=ca podɔŋ 'The person the tiger ate'⁴

- 1 knək jtek kloc dol,
husband sleep inside house

kmpən tjp-tap co? balay
wife IMPERF-weave AT:below verandah
Her husband was asleep up in the house, (and his) wife was weaving down on the verandah.
- 2 ki=kunci? pintu? he? iwar, kbran
3A=lock door AT:above outside be.apprehensive.lest

ki=k^he? knək
3A=know husband
She had locked the door from outside, apprehensive lest (her) husband find out.
- 3 ki=jam sake?, ko=tarek, kəh
3A=sense pandanus 3UA=pull 3

pon ?inte, pjam sniymin dɔs co? ?ate
THEN go.down.ground feeling lover arrive AT:below ground
(When) she felt someone pull at the pandanus, she went down thinking (her) lover had arrived down on the ground.
- 4 ki=tloŋ la=podɔŋ
3A=grab A=tiger
A tiger grabbed (her).
- 5 ki=goŋ la=podɔŋ sok troy, b-croh rom sniymin
3A=bring A=tiger along path MID-meet WITH lover
The tiger took (her) along the path (and they) ran into (her) lover.
- 6 ki=tikam podɔŋ rom tomok
3A=stab tiger WITH dagger
He stabbed the tiger with his dagger.
- 7 k^hbəs podɔŋ
die tiger
The tiger died.

⁴ sma? ki=ca podɔŋ (person 3A=eat tiger) 'The person the tiger ate'.

- 8 ki=goŋ ?yot k<r> dor tet dol knək
3A=bring return be.female<NMZ> TO:spec house husband
He took the woman back to her husband's house.
- 9 ki=tapa? la=knək ha? h̄n b-croh
3A=ask A=husband AT where MID-meet
The husband asked where they had met.
- 10 he? tŋah trop
AT:above middle path
In the middle of the path.
- 11 sampay knək=hn yok tomok r<l>məl
so husband=3POSS take dagger be.male<NMZ>
So, her husband took the man's dagger.
- 12 ki=tikam ha? r<l>məl, kəhn k^hbəs
3A=stab AT be.male<NMZ> 3S die
He stabbed the man (and) he died.

14.4.3 c<nn>man sma? pe? b-knon 'The narrative of the person who had three children'

A child's elder sister has died from an illness, and the child travels down to the underworld, the bri kmuc (jungle ghost) 'jungle (of) the ghosts', seeking the elder sister. The child refuses to leave, and so the story unfolds.

People who die either a natural death become ghosts called kmuc. The Semelai recognise numerous types of ghosts, depending on the manner and age at which a person passes away. kmuc, like the elder sister and her friends in the story, are harmless ghosts, identifiable by their red eyes. These ghosts are not malevolent, just rather annoying because they give people a fright. This kmuc see the world in an opposite way to humans. For example, to a ghost night is day, the moon is the sun and the thorns which people place above their doorways to keep such ghosts out of the house are perceived by the kmuc as kmtk^het 'stinging insects'. As we will see in the story, this causes some confusion for the child during her stay in the underworld, until things change ...

Note that the gender of the younger sibling is never specified, nor is the child named. The narrator could not decide if the child was male or female. This is typical of Semelai narratives. I have glossed the third person references as female for the purpose of a more natural translation.

Some terms used in the text which require explanatory notes are:

- b<?>la? 'an unrelated companion, or friend'
- k^hru? (fish.with.poison<IMPERF>) is to fish with an ichthycide in a pool of water. The stunned fish are then speared as they float up to the surface, or simply gathered and placed in creels.
- seŋ is a Derris sp. shrub, the sap of which is used as an ichthycide. The root is pounded at the fishing site to release the sap, which is then mixed through the water.
- rapah is jungle litter, fallen leaves, twigs and the like.

- c<n>aro? is a vulgar term for food, derived by nominalisation from the verb caro? 'to eat, (vulg.)'. It is a serious offence to curse or speak deprecatively of food or potential sources of food.
- tampoy is a wild mangosteen (*Baccaurea* sp.) ← Malay *tampoi*. The fruit is either eaten, or brewed into an alcoholic beverage. The Semelai, the only Orang Asli who brew alcohol, also make alcoholic beverages from sugar cane (bos) and bananas (tyək).

- 1 k^bbes kaka?; p^bakit sdom
die EZ illness only
The elder sister died; only from illness.
- 2 ?adi?=hn ?yam, la=saya_j kaka?, sampay risaw
YS=3POSS cry BCS=be.fond.of EZ until grieve
The younger sibling cried, because (she) had been fond of (her) elder sister, until (she) grieved.
- 3 lalu?=hn p?ot ha? jŋ taja?, ki=gap la=lsbs
then=CONN stay AT foot stair 3A=bite A=flying.ant
Then (one day) (she) was sitting at the foot of the stairs, (and) a flying ant bit (her).
- 4 "7əj=ən gren kd"
IfA=want be.angry 2f
"I'll get angry, you!"
- 5 "hmoh mə=kø ?yam?"
what REL=2f cry
"What (is it) that are you crying about?" (it asked).
- 6 "7əj=risaw hn=kaka? k^bbes"
1Af=grieve O=EZ be.dead
"I'm grieving for (my) elder sister who has died."
- 7 "lah pdor hn=?ap, dom," k^bləŋ lsbs, "te?en bri.kmuc"
HORT₂ follow O=1f AFF QUOTE flying.ant to:below jungle.ghost
"Come! Come with me, yes," said the flying ant, "down to the underworld."
- 8 dom knkon pdor hn=lsbs
AFF child follow O=flying.ant
(And) indeed, the child followed the flying ant.
- 9 dos co? kaka?, k^bləŋ kaka?, "hmoh mə=kø=?"
arrive AT:below EZ QUOTE EZ what REL=2fA=seek
(The child) arrived down at the elder sister, (and) elder sister asked, "What (is it) that you are seeking?"
- 10 "7əj=te?en hn=kø, s?rɔ?"
1fA=TO:down O=2f miss
"I (came) down to you. (I) miss (you)," (she replied).
- 11 ki=jon la=kaka? ca j?oh
3A=give A=EZ eat drink
The elder sister let (the child) eat (and) drink.

- 12 2luc.tom j?oh, ki=ur ?yot
after drink 3A=instruct return
After (the child) had drunk, she instructed (her) to go back (home).
- 13 da? ki=moh
NEG 3A=want
She didn't want to.
- 14 sampay tr-ye hn=b?lu? ga=k<?>ru?
so HAPP-see O=friend IMM=fish.with.poison<IMPERF>
kəh len pdor
3 want follow
So, (when the child) happened to notice (her sister's) friends going fish-poisoning, she wanted to go with (them).
- 15 ki=?ur la=kaka? pdor, yd?=hn cokbp kaka? kəh:
3A=instruct A=EZ follow but=CONN speak EZ 3
da? ki=jon tna=yøk crch, ki=?ur yøk rapah
NEG 3A=permit IRR=take fish 3A=instruct take litter
The elder sister told (the child) to follow, however, her elder sister advised (thus): She forbade (the child) to collect fish. She instructed (her) to collect jungle litter.
- 16 dos he? tmpat k<?>ru?, køŋ scŋ,
arrive AT:above place fish.with.poison<IMPERF> beat ichthyocide
de=haro?
3plA=distribute.in.water
(They) arrived at the fishing spot, beat the ichthyocide (and) they distributed (it) through the water.
- 17 kəh, ki=?ye crch, ki=yøk
3 3A=see fish 3A=take
Her, she saw the fish, and (she) collected (them).
- 18 b?lu?=sbn, yøk rapah, tŋkel dloŋ, de=buh kloc rentəŋ
friend=SC take litter length stick 3plA=put inside creel
The friends see, collected litter (and) lengths (of) stick, (and) they put (them) in (their) creels.
- 19 kna? kmøŋ crch deh ha?, ?are? .doy, dehn ?yot
happen finish fish 3pl AT day.afternoon 3pls return
leŋ dol
TO:up house
It happened that (by the time) their fish thereabouts were finished, (it was) dusk, (and so) they went back up home.
- 20 dos he? dol, ki=jon crch ha? kaka?
arrive AT:above house 3A=give fish AT EZ
Arriving up at the house, (the child) gave the fish to the elder sister.
- 21 ?arrh ki=?urah la=kaka?, sot rapah,
just.then 3A=tip.out A=EZ metamorphose.into litter

- 15 sot dloj
metamorphose.into stick
Just when the elder sister tipped (them) out, (the fish) transformed into litter, (and) transformed into sticks.
- 22 dom b?lu?, de=curah tom ha? renten, sot crch
AFF friend 3plA=tip.out SRC AT creel metamorphose.into fish
Indeed, the friends, they tipped out (their's) from in the creels (and) (they) transformed into fish.
- 23 gren k?h
be.angry 3
She was angry.
- 24 "k? lpc?" k?l?j kaka?, "da? ?ap=jon ma=yok crch, bnda?=hn
2f already QUOTE EZ NEG !A=give IRR=take fish thing=3POSS
sot rapah" k?l?j "m?=lpc dos ha? dol
metamorphose.into litter QUOTE REL=already arrive AT house
sot crch ?ap=jur yok rapah" k?l?j
metamorphose.into fish !A=instruct take litter QUOTE
"What have you done?" exclaimed the elder sister, "I didn't permit (you) to collect fish. The things have become litter," (she) said. "See, the ones (of the friends) who have come home are already fish. I instructed (you) to collect litter," (she) said.
- 25 ki=tar-ypr ninca da? ?ilok
3A=CAUS-ascend food NEG be.good
She put the bad food away.
- 26 "c<n>aro? kmuc!" k?l?j
eat<NMZ> ghost QUOTE
"Ghost swill!" (the child) swore.
- 27 "boj!" k?l?j kaka?=hn, "de=gren b?lu?"
NEG:IMP QUOTE EZ=3POSS 3plA=be.angry friend
"Don't!" warned her elder sister, "(my) friends will get angry."
- 28 sampay=hn yok sarek, tr-ye hn=b?lu? ga=rej ple
so=CONN take tomorrow HAPP-see O=friend IMM=seek fruit
kloc bri: ple tampoy, duryan, rambutan.bri
inside jungle fruit wild.mangosteen durian rambutan.jungle
So the next day, (the child) happened to see the friends going to look for fruit in the jungle: wild mangosteen, durian, wild rambutan.
- 29 swak deh, pdor k?h
go 3pl follow 3
They went, (and) (she) followed.
- 30 dos kloc bri deh, bri jruh, tr-ye ple deh
arrive inside jungle 3pl jungle be.deep HAPP-see fruit 3pl
(When) they arrived in the jungle, the deep jungle, they happened to spot (some) fruit.

- 31 de=?ajak la=b?lu? yor ple
3plA=invite A=friend ascend fruit
The friends invited (her) to climb (the trees) for fruit.
- 32 da? ki=moh
NEG 3A=want
She didn't want to.
- 33 de=?ur k?h ps-pos, dom pie
3plA=instruct 3 IMPERF-gather AFF fruit
They instructed her to gather (fallen) fruit, yes.
- 34 ki=pos, yok m?=?nom, ki=buh kloc ?amoy
3A=gather take REL=be.ripe 3A=put inside back.basket
She gathered fruit, collecting the ripe (ones) (and she) put (them) in (her) back-basket.
- 35 dom b?lu?, de=yok m?=?nom, de=buh ha? ?amoy
AFF friend 3plA=take REL=be.unripe 3plA=put AT back.basket
Indeed the friends, they collected unripe (ones) (and) they put (them) in (their) back-baskets.
- 36 ?are? doy, dehn ?yot
day afternoon 3pl return
(At) dusk they returned home.
- 37 dos deh he? dol, ki=curah
arrive 3pl AT:up house 3A=tip.out
They arrived up at the house, (and) (she) tipped (the fruit) out.
- 38 dom pie sot muda?, do b?lu?, ?nom
AFF fruit metamorphose.into be.unripe OF friend be.ripe
Indeed (her) fruit was unripe; (that) of the friends was ripe.
- 39 gren lagij? k?h
be.angry again 3
(She) was angry again.
- 40 de=h?gi? la=b?lu?, dom ple
3plA=give A=friend AFF fruit
The friends gave (her some of their) fruit.
- 41 jtek pt?m, srder som, sat bsugi?
sleep night awake morning descend perform.ablations
co? tla?ga?
AT:below well
(That) night (the child) slept; (she) awoke (next) morning and went to perform her ablutions down at the well.
- 42 ki=p?ul kubag
3A=go.down pool
She went down to a pool of water.
- 43 ki=yok kawot rapah, ki=ta<k>rek t?kel dloj tom
3A=take scoop.up litter 3A=pull<IMPERF> length tree SRC

- kloc ?alor dak
in channel water
She collected litter, scooping (it) up, and repeatedly pulled sticks out of the channel of water.
- 44 ki=goj ?yot leg dol
3A=bring return TO:up house
She took (them) back up to the house.
- 45 c?en kaka?, b?lu? la=kah, panay reg creh, m?-(s)lay
be.happy EZ friend A=3 be.clever seek fish PERFM-smoke
The elder sister (and her) friends were happy with her, (she) was skilful at finding fish (and) smoking (them).
- 46 "t?o?" k?leg kah, "dej no?, knet, c<n>aro? kmuc, kmog"
oh QUOTE 3 manner this before eat<NMZ> ghost finish
"Oh (I see)," (the child) said, "(it's done) (in) this manner. Before, (the) ghost's swill (episode), (that) is finished."⁵
- 47 "boy!" k?leg kaka?, "de=?y?n b?lu?, de=?gren"
NEG:IMP QUOTE EZ 3plA=hear friend 3plA=be.angry
"Don't!" warned the elder sister. "(If my) friends hear (you), they will get angry."
- 48 "bah!" k?leg, "?ep da? n=k?e? ninca kmuc"
NO! QUOTE 1 NEG 1A=know food ghost
"No way!" (she) protested. "Me, I don't understand ghost food."
- 49 ki=?ur la=kaka?, ?yot leg bapa? deh
3A=instruct A=EZ return TO:up father 3pl
The elder sister instructed (her) to go back up to (their) father.
- 50 "beh da? moh," k?leg, da? ki=moh
NO NEG want QUOTE NEG 3A=want
"No, (I) don't want to," (she) said. She didn't want to.
- 51 "boy!" k?leg kaka?, "kd loc 2le? j?o?" k?leg,
NEG:IMP QUOTE EZ 2f already be.long very QUOTE
"ha? no?, t?m?oh ?are?. ?yot leg bapa?, k?leg
AT this seven day return TO:up father QUOTE
"Don't!" said the elder sister. "You've already been here a very long time," (she) said, "(It's been) seven days. Go back up to father," (she) urged.
- 52 "dos kd he? ke, t?m?oh ?are? m?o?m r?m ?ep"
arrive 2f AT:above that seven day like WITH If
"(After) you arrive up there, on the seventh day (you) (will be) like me."
- 53 sampay ki=moh ?yot leg bapa?
so 3A=want return TO:up father
So then, she was willing to return up to her father.
- 54 dos he? bapa?, "te h?n kd swak?
arrive AT:above father TO:unspec where 2f go

⁵ The child is now a ghost, and here she refers back to the fish and ripe fruit which she collected previously as being 'ghost food'.

jnuh ?ep=re?i?, da? j?ye"
exhausted 1A=seek-ITER NEG 1A=see
When she arrived, "Where have you been? I kept looking (for you) (until I was) exhausted, (but) I didn't see (you)," (he said).

- 55 "t?ep=t?en kaka?" k?leg,
1fA=TO:down EZ QUOTE
"I (went) down to elder sister," (she) replied.
- 56 "t?o?" k?leg bapa?, "t?en bri.kmuc" k?leg
oh QUOTE father TO:down jungle.ghost QUOTE
"Oh?" said (her) father (taken aback), "down to the underworld?" (he) asked.
- 57 "h?n?" k?leg kah
hnnn QUOTE 3
"Yes" she replied.
- 58 dos t?m?oh ?are?, kahn k?bes
arrive seven day 3S die
The seventh day came, (and) she died.

14.4.4 c<nn>man sma? k?oy 'The narrative of the person who was a head'

- 1 kahn ke me=bje? k?oy, y? knon pe? ?ikur
3S that one=CLF head but offspring three CLF
She there, was a head, but (she had) three children.
- 2 k?oy b-knon
head HAVE-offspring
The head had given birth.
- 3 pe? ?ikur knon=h?n ke=c?i, p?st
three CLF offspring=3POSS that=EM stay
(She had) her three children there (and then she) had no more.
- 4 ?ospi swak kah, 2ma?hn ke goj ?yot creh hne
whenever go 3 mother=3POSS that bring return fish THEN
Whenever she went out, the (children's) mother, she brought back fish (for them).
- 5 ki=goj c<n>aro? creh rom mol?, rom lmop,
3A=bit skewer<NMZ> fish WITH mouth WITH tooth
ki=rus, gon ?yot
3A=drag bring return
She bit the skewer (of) fish with (her) mouth, with (her) teeth, (and) she dragged (it) back home.
- 6 dos he? dol hne
arrive AT:above house THEN
Arriving back at the house then ...
- 7 ?'sma? mande?"
{person what.sort
["What sort of person (was she)?"]

- 8 "sma? k^boy m^e=bj^e?, kb^a? da? da?, k^boy nehneh"] person head one=CLF body NEG EXIST head only] A 'head' person, (she) didn't have a body, just a head.]
- 9 ?le?le?, p?ot=ce? deh ha? ncj, swak baranj-h^on, eventually stay=EM 3pl AT before go where.ever sjnu-jnuh ?yot, gon ?yot crsh INTNS-be.exhausted return bring return fish So, they stayed there, the aforementioned (children), (while) the mother went wherever, (and) (when she had) had enough, (she) returned, bringing home fish.
- 10 mas^a?, bkira? knon=h^a, ca cook cook offspring=3POSS eat Her children cooked, cooked (and) (she) ate.
- 11 de=t^aj ?ma?=hn, ca=ce? 3plA=dish.out mother=3POSS eat=EM They dished out (for) their mother (and) so she ate.
- 12 ?ma?=hn ca rom mol^at mother=3POSS eat WITH mouth Their mother ate with her mouth.
- 13 creh brcnj, wen j?aj fish bony throw.away bone (If) the fish was bony, (the children) discarded the bones.
- 14 ["?oh ?ma? da? br-k^boy!" [EXCL mother NEG HAVE-head ["Oh the mother didn't have a body!"]
- 15 "?oh k^boy nehneh" oh head only "Oh, (she) was just a head"
- 16 "knon=h^a?" offspring=3POSS "(And) her children?"
- 17 "knon cukup ... k^boy, j^bnj, t^bl" offspring enough head foot hand "The children were complete ... head, feet, hands."
- 18 "h^on.mandeh b-knon?" where.what.sort HAVE-offspring "(So) how did she give birth?"
- 19 "k^be?" know] "(How should I know?")
- 20 ?le?le? tah mande-mandeh ... eventually Q RDP-what.sort So, anyway ...

- 21 de=t-haronj la=raja? bburu? 3plA=HAPP-encounter A=sultan hunt.with.dog.and.spear (One day) some princes out hunting with their dogs and spears happened to come across her children.
- 22 knon=h^a de=t^aj, de=g^aj tara? m^e=?ikur deh: offspring=3POSS 3plA=grab 3plA=bring each one=CLF 3pl raja?, raja?.muda?, raja?.naga?, prince crown.prince prince.underling de=g^aj ?yot j?ay kmpen ... 3plA=bring return make wife They grabbed (the siblings), (and) each one took one (of) them: the prince, the crown prince (and) the underling prince, they took (them) back (and) made them (their) wives ...
- 23 ?yam deh la=?ma? deh cry 3pl BCS=mother 3pl They cried, on account of their mother.
- 24 ?are?doy ?yot ?ma? tom k<c>bac day.afternoon return mother SRC fish.with.pole<IMPERF> (At) dusk, their mother returned from fishing.
- 25 ki=jpk, beh, da? da? knon 3A=observe NO NEG EXIST offspring She saw her children weren't there.
- 26 ?yam kah no? neg cry 3 this before The aforementioned cried.
- 27 ki=pantaw, t-br-k^be-br-k^be? 3A=call.out HAPP-MID-know-MID-know She called out, unaware (of what had happened).
- 28 knon pon ?yam=j^a he? dol de=g^aj raja? offspring too cry=CL AT:above house 3plA=bring prince (Her) children were crying too, up at the house where the princes had brought (them).
- 29 "mandehmoh=ce?" k^blay, "m^e=je=?en ?yam?" what=EM QUOTE REL=2=AUG cry "What (is it)?" (they) asked, "that you are crying for?"
- 30 "da? da? ... len den no?, snaj len ca, len pakay" NEG EXIST waat manner this be.pleasant want eat want wear "There's nothing ... (you) would want for. Like this, (it's) easy to eat, dress ..."
- 31 "beh!" k^blay "ye=?en ye=?yam ?ma? ye=?en NO QUOTE I=AUG 1A=cry mother I=AUG "No!" (they) cried, "we are crying (for) our mother."

- 32 "jake he=yɔ:k"
in.that.case 1&2A=fetch
"In that case we'll fetch (her)."
- 33 "beh!" k^bləŋ, ?ma? ye=?en, ?aseŋ sma?"
NO QUOTE mother 1=AUG be.different person
"No!" (they) protested, "our mother is different (to) other people."
- 34 ?ma? sma? ?aseŋ."
mother person be.different
"Our mother is an unusual person."
- 35 "sma? ?aseŋ mandehmoh?"
person be.different what
"What (sort of) unusual person?"
- 36 ?ma? ye=?en" k^bləŋ "k^boy nehneh, da? da? kbə?"
mother 1=AUG QUOTE head only NEG EXIST body
"Our mother," (they) replied, "is just a head. (She) doesn't have a body."
- 37 "sabar" k^bləŋ raja? "la=da? da? kbə?"
patient QUOTE prince BCS=NEG EXIST body
"Poor thing," remarked the princes, "(if) (she) doesn't have a body."
- 38 "ga=k^boy nehneh h^bn.mande lən ?ris?"
IMM=head only how want be.alive
"(If she) is just a head, how (then) can she be alive?"
- 39 "?oh ?ris," k^bləŋ=hn
oh alive QUOTE=3POSS
"Oh, (she) is alive," they replied.
- 40 de=ləh yɔ:k la=knək=hn
3plA=go.across fetch A=husband=3POSS
Their husbands went to fetch (her).
- 41 de=ləh paneŋ
3plA=go.across summon
They went calling out (to her).
- 42 de=yɔ:k, de=taten, de=pte? de=kanar k^boy la=deh
3plA=fetch 3plA=remove 3plA=place 3plA=carry head A=3pl
They fetched (her), they took (her there) and they carried (her).
- 43 de=gɔŋ=hn
3plA=bring=O
They brought her (back).
- 44 dɔ:s jŋɔ:k dom, dom ki=cən la=knən=hn
arrive observe AFF AFF 3A=be.happy A=offspring=3POSS
When they arrived (and) saw indeed, (her) children were happy.
- 45 de=pan-ca la=knən=hn, ca=cə? kəh
3plA=CAUS-eat A=offspring=3POSS eat=EM 3pl
Her children fed her (and) she ate.

- 46 mulb? kde? rom knən=hn soloŋ lpəc
beginning dwell WITH offspring=3POSS first stomach
First she stayed with her first-born child.
- 47 kde? rom knən tom ke pnŋahan
dwell WITH offspring SRC there middle
From there (she went and) stayed with the middle child.
- 48 zluc tom ke kde? rom knən mə=bunjs?
after that dwell WITH offspring REL=youngest.born
After that, she stayed with the youngest child.
- 49 2are? muy, sjnu-jnuh beh kəhn ke
day one INTNS-exhausted NO 3S that
One day, she had had enough there.
- 50 da? than kde? 2endəl nehneh, kde? kloc ro?
NEG endure dwell inside.house just dwell inside basket
(She) couldn't bear to just stay inside, stay in a basket.
- 51 de=pan-ro?
3plA=EQUIP-basket LOC kitchen
They had housed her in a basket in the kitchen.
- 52 sar knppup tə?en habu?
descend offspring-in-law TO:down kitchen
(Her) son-in-law came down into the kitchen.
- 53 ki=jpk dom k^boy ke, ?ma?, ns?" k^bləŋ=hn,
3A=observe AFF head that mother this QUOTE=3POSS
"beh," k^bləŋ=hn, knppup, "kde? 2en habu?,
NO QUOTE=3POSS offspring-in-law dwell LOC kitchen
kde? he? bilik"
dwell AT:above room
He looked at the head, "mother," (he) said, "no," he said, the son-in-law, "don't live
in the kitchen, live up in the room."
- 54 "beh," k^bləŋ, "da? moh. ?əŋ nɔ? byasa?," k^bləŋ
NO QUOTE NEG want I this be.accustomed QUOTE
"No," (she) replied, "I don't want to. I'm used to this," (she) said.
- 55 "?oh boy!" k^bləŋ, "kde? ha? habu?, ki=ca zus" k^bləŋ
oh NEG:IMP QUOTE dwell AT kitchen 3A=consume fire QUOTE
"ki=t'et sma?, ?ənū?" k^bləŋ, "sma? bajə?," k^bləŋ "grək"
3A=step.on person [HES] QUOTE person be.many QUOTE fall
"Oh," (he) pleaded, "don't stay in the kitchen, (you'll) get burnt; (or) someone
(will) step (on you), umm," (he) said, "(there's) lots of people, (or you'll) fall," (he)
said.
- 56 "?ah, dawen!" k^bləŋ
oh never.mind QUOTE
"Oh never mind!" (she) replied.

- 57 "boy!" k^bleg knjjpup
NEG:IMP QUOTE offspring-in-law
"Don't!" begged the son-in-law.
- 58 ?le?le? ?yot kde? ?en knon mə=gdo dom dej ke
so return dwell LOC offspring REL=be.old AFF manner that
Eventually (she) returned to live with her eldest child, yes, in that manner (as before).
- 59 ?le?le? ?are? muy ga=svak bburu?
so day one IMM=go hunt.with.dog.and.spear
So, one day she was going hunting with (her) dogs and spear.
- 60 kahn gon co, co, ?anu? ...
3S= bring dog dog [HES]
She brought dogs, dogs, umm ...
- 61 k^boy, gurut grit, gurut grit, k^boy ke
head rolling.bouncing.along.rough.ground! head that
The head, went rolling along, bouncing over the terrain, the head there.
- 62 de=pdr=c? la=co
3plA=follow=EM A=dog
The dogs went with (her).
- 63 dos he? ke, jat co no? nej
arrive AT:above that bark dog this before
(When they) got up there, the aforementioned dogs barked.
- 64 kah hne k^boy hnt p^bkokn thng, pintu? thng
3 THEN head THEN supported hole door hole
She then, the head, helped at the hole, at the entrance (to) the hole.
- 65 co masuk krja?-i? jkos
dog enter work-ITER porcupine
The dogs went in and harried the porcupine.
- 66 ?le?le? paioh jkos, krw^bl pdor halaw la=co
so flee porcupine emerge follow chase A=dog
Eventually the porcupine fled, the dogs emerging in pursuit.
- 67 pnam dehn, co, pon da? sabur de=tkan dom k^boy
feeling 3pl dog THBN EXIST melee 3plA=hold.down AFF head
(As for) the feelings (of) the dogs (it was a porcupine), then (in) the melee, indeed they held down the head.
- 68 de=tkan bramatamey
3plA=hold.down all.in!
They snatched (it), all together in the confused melee.
- 69 k^bbəs ?a=kah=h5?
die DET=3=ATTN
(And so) she died.

- 70 sta? k^bbəs, b-krntah ?are?, ?anu? ...
after die MID-lightning day [HES]
After (she) died, (there was) lightning, umm ...
- 71 "eh beh," k^bleg knon, "ga=te h5n ?ma?"
eh NO QUOTE offspring IMM=TO:unspec where mother
k^bleg=hn
QUOTE=3POSS
"Oh no," said the children, "where did mother go?" they asked.
- 72 "ren!" k^bleg, "2ma?=hn" k^bleg=hn
seek QUOTE mother=3POSS QUOTE=3POSS
"Seek!" (they) cried, "our mother," they said.
- 73 ki=phor knlek deh p?=pe?, ren mntuhb? deh
3A=chase.out husband=3POSS
They chased out their husbands.
- 74 swak=c? knlek deh p?=pe?, ren mntuhb? deh
go=EM husband 3pl BE=three seek parent.in.law 3pl
The three husbands went off to find their mother-in-law.
- 75 jnok co nej, rbut k^boy mntuhb?=hn de=ca
observe dog before snatch.off head parent.in.law=3POSS 3plA=eat
(They) saw the dogs from before, (which) had snatched the mother's head (and) eaten (it).
- 76 "eh!" k^bleg, "pkima?," k^bleg, "co na? ha?"
eh QUOTE ill-fated QUOTE dog DEM-AT
"Oh!" (they) exclaimed, "ill-fated," (they) said, "these dogs here."
- 77 de=tampor co,
3plA=hit dog
They beat the dogs.
- 78 k^bbəs co ki=ccyet k^boy mntuhb?
die dog 3A=hold.by head parent.in.law
The dogs who held the mother-in-law's head died.
- 79 mande, da? da? wo? kulit k^boy
what NEG EXIST longer skin head
Why! there was no skin left (on her) head.
- 80 tinal ktog k^boy
remain skull head
(Only) the skull remained.
- 81 de=ca la=co
3plA=eat A=dog
The dogs had eaten (her).
- 82 ki=gon ?yot, ki=kabar ?en kmpən
3A=bring return 3A=tell.news LOC wife
They took it home (and) told (their) wives the news.

- 83 bila? ga=da? kmpen
when IMM=NEG wife
When (they did), (it was as if they) weren't (their) wives.
- 84 ?yam, beh da? ga=?ojo?
cry NO NEG IMM=joke
(They) cried, no (it was) no joke.
- 85 knon me=buru? ha?, ga=wen dri?=san.
offspring REL=youngest.born AT IMM=discard 1self=SC
ga=bunuh wen dri?
IMM-kill discard 1self
The youngest child there was going to commit suicide, was going to kill herself.
- 86 ki=tloj la=knlèk=hn
3A=grab A=husband=3POSS
Her husband restrained (her).
- 87 de=pojo-pojo? kmpen deh
3plA=INTNS-cajole wife 3pl
They cajoled their wives.
- 88 de=basuh ktog. k'oy de=pa-yup, de=c'um
3plA=wash skull head 3plA=EQUIP-sleeping.cover 3plA=wrap
r'm kayen putih
WITH cloth be.white
They bathed the skull. The head, they covered (it), wrapped (it) in white cloth.
- 89 ki=pte? la=raja? kloc krnda?
3A=put A=prince inside bier
The princes placed (it) in a coffin.
- 90 de=smbaya?, de=smbaya?=hn, de=kem ?en krnda?
3plA=pray 3plA=pray=O 3plA=bury LOC bier
They prayed, they prayed (over) it, (and then) they buried (it) on the bier.
- 91 ?yam knon=hn nej
cry offspring=3POSS before
Her children were crying from before.
- 92 "j'o," k'lèj, "je=saya? co"
very QUOTE 2A=be.fond.of dog
(She) cried, "you were so fond of the dogs."
- 93 sayan tom hñ? k'lèj knlèk
be.fond.of SRC where QUOTE husband
"Fond of what?" asked the husband.
- 94 "co kmøj ye=bunuh," k'lèj. "kmøj
dog exhaustively 1A=kill QUOTE exhaustively
k'bas," k'lèj=hn
die QUOTE=3POSS
"The dogs, we killed them all," (they) said. "They're all dead," they said.

- 95 "kah? ben," k'lèj, "da? samo?" k'lèj "co
Q NO QUOTE NEG be.equal QUOTE dog
r'm sma?" k'lèj
WITH person QUOTE
"What! No," (they) said, "(they're) not the same," (they) protested, "dogs and people," (they) said.
- 96 pwa-pwas ?yam. ?le?ie? brnti? ?a=kéh
INTNS-content cry eventually stop DET=3
(They) cried to their hearts' content. Eventually, as for them, (they) stopped.
- 97 de=pojo-pojo? la=knlèk=hn
3plA=INTNS-cajole A=husband=3POSS
The husbands cajoled (them).
- 98 cej=ca?, ?anu? ...
from.that.point.on [HES]
From then on, umm ...
- 99 pc? ?are?, mi=tan ki=dos smbaya?-i? la=knpipunj
three day one=time 3A=arrive pray-ITER A=offspring.in.law
ha? kubur
AT cemetery
(On) the third day, once again the sons-in-law came to pray at the graveyard.
- 100 dos tñph? ?are?, mi=tan de=dos smbaya?
arrive seven day one=time 3plA=arrive pray
knduri? bagay ?en
feast like just
(On) the seventh day, once again they came to pray. (They held) a feast, (and) all of that.
- 101 de=j?oy mocom b-cuku-cukup t?i, j?y, k?e?
3plA=behave like MID-be.complete hand foot body
They behaved as (if) (her) hands, feet and body had been complete.
- 102 tom ha? kke beh da? de=?yam w?; da? t=?ya=?yam lpc
SRC AT there NO NEG 3plA=cry longer NEG HAPP-INTNS-cry already
From then on, no, they didn't cry over it anymore; didn't cry inconsolably.
- 103 k?møj d? ke
finish OF there
(The story) finishes there.

Vocabulary

The Semelai–English vocabulary presented here lists the 1,318 lexical entries which appear in the preceding grammatical description. They are drawn from a current lexicon of 4,128 entries.

Entries are given in the citation form, which in most cases is a monomorphemic lexeme. In a few instances, a derived form is listed as an entry, because a) it is the citation form, or b) the derivation is not predictable or synchronically motivated, but there is an obvious relationship, both morphological and semantic, to a synchronic root elsewhere in the lexicon.

Each entry consists of a headword, word class and English definition. Multiple senses are indexed numerically. An asterisk denotes a lexeme which occurs in an idiom and has no independent meaning. Avoidance speech style terms are identified as (av.). Scientific names are provided where they have been identified. A full list of abbreviations is provided below.

Arrangement of entries The arrangement of the entries is based on the place of articulation of the onset segment, beginning with glottal and progressing through to bilabial. Entries for each place of articulation are then organised according to mode of articulation – stop, fricative, rhotic, lateral, glide. For oral stops, the ordering is voiceless unaspirated, voiceless aspirated and voiced. For nasals it is voiced, glottalised, then voiceless. The ordering for consonants is: ?, h, k, k^h, g, g^h, hy, c, c^h, j, J, ?j, s, y, ?y, t, t^h, d, n, ?n, hn, l, ?, r, ?r, p, p^h, b, m, ?m, hm, w.

Vowels are ordered from front to back (or rounded to unrounded), commencing with the high front vowel i and ending with the low back vowel o. Oral vowels are listed before their nasal counterparts. The ordering is: i, ī, e, ē, ε, ē̄, a, ā, u, ū, ə, ī̄, u, ū̄, o, ī̄, ə̄, d, ə̄̄.

List of word class abbreviations adj. adjective; adv. adverb; aff. affirmative; asp. aspectual; av. avoidance speech style; clf. classifier; clt. clitic; conn. connective; dem. demonstrative; exp. expressive; hort. hortative; ign. ignorative; imp. imperative; interj. interjection; modal modal verb; n. noun; neg. negator; num. numeral; onom. onomatopoeia; part. interrogative particle; pred. predictor; prep. preposition; pro. pronoun; proclt. pronominal clitic; quant. quantifier; quot. quotative marker; top. toponym; v.i. intransitive verb; v.t. transitive verb; v.supp. supplemental (serial) verb; voc. vocative.

?i= <i>clt.</i> title or name-forming clitic	?amɔn <i>n.</i> pet, plaything
?i?e? <i>n.</i> classificatory elder brother	?aŋ <i>v.t.</i> carry with a head strap
?ikur <i>n.</i> tail	?anrak <i>exp.</i> sound of many people gathered
— <i>clf.</i> classifier for animals	working together
?iŋat <i>v.t.</i> think, hold an opinion	?ac <i>n.</i> faeces, excrement, exudate
?i=cm <i>n. (av.)</i> cassava	— <i>v.t.</i> shit
?iŋy <i>n.</i> millet	?əŋi <i>pro.</i> first person singular familiar
?isi? <i>n.</i> contents	?əŋi= <i>pro.</i> first person singular familiar
?i=tog <i>name</i> ancestor giant	proclitic form
?inci? <i>n.</i> inch	?ənam <i>num.</i> six
?inte <i>v.t.</i> exit raised house, go down to the	?əm <i>v.t.</i> lie prone
ground	?uŋuŋ <i>n.</i> smouldering log
?ilɔk <i>adj.</i> good	?us <i>n.</i> 1) fire, flame. 2) electricity
?ipar <i>n.</i> sibling-in-law	?uda? <i>n.</i> fifth born aunt
?ec <i>interj.</i> expression of negative surprise	?ulor <i>v.t.</i> lower (something) down into
?es <i>neg. imp.</i> don't! (Directed at small	(something)
children.)	?ulŋj <i>v.i.</i> come and go repeatedly
?en ₁ <i>prep.</i> in; on	?ur <i>v.t.</i> instruct
?en ₂ <i>adv.</i> just	?urat <i>n.</i> vein, sinew, muscle
=?en ₃ <i>clt.</i> marks plural number in first and	?ubot <i>n.</i> 1) medicine. 2) battery
second person pronouns	?ŋ?por <i>v.t.</i> offer
?endol <i>v.i.</i> enter a house	?oh <i>v.t.</i> 1) blow (into something); breathe life
?a= <i>clt.</i> clitic to third person pronouns	(into something); to come/ejaculate in.
?ahom <i>n.</i> breath	2) shoot with a blowpipe
?ankot <i>v.t.</i> bring (some) at a time	?ɔŋ <i>v.i.</i> howl
?ac <i>interj.</i> expression of disgust, displeasure	?ɔŋ <i>v.t.</i> hide
?ajak <i>v.t.</i> invite	?ɔsok <i>n.</i> placenta
?aseg <i>v.i.</i> different; unusual, peculiar	?ɔy <i>voc.</i> hey you!
?asd ₁ <i>conn.</i> when(ever)	?ɔlɔ? <i>n.</i> joke
?ayer hitam <i>top.</i> Ayer Hitam	?ɔlɔk?elek <i>exp.</i> appearance of something
?ayon <i>v.i.</i> be slow	oscillating
?ate <i>n.</i> ground, dirt	?ɔw <i>onom.</i> call of the Malay Peacock-
?adi? <i>n.</i> younger sibling	peasant [kawɔŋ]
?anu? <i>interj.</i> hesitation marker	?ɔŋlɔŋ <i>n.</i> catwalk
?antat <i>v.t.</i> bring, take	?ɔs <i>v.i.</i> swollen
?alus <i>adj.</i> fine, subtle, refined	??e? <i>n.</i> elder brother
?alor <i>n.</i> channel, hollowed-out bed of a	??a? <i>n.</i> animal (child speak)
stream	?k?pk <i>n.</i> crow
?are? <i>n.</i> day	?ŋ?ŋj <i>n.</i> termite
?are? <i>n.</i> rain	?sta? <i>n.</i> ell
?areh <i>adj.</i> new	?ləmu? <i>n.</i> knowledge, traditional
— <i>asp. adv.</i> recently, just now	?mpat <i>num.</i> four
?arah <i>v.i.</i> head, go in a direction	?w?aw <i>v.i.</i> stand
?amog <i>n.</i> back-basket	

hijaw *adj.* green, light blue
hitam *adj.* black, dark

hindi? *n.* implement for pounding rice
hir *v.t.* growl (at), of dogs

hej *n.* way up yonder
heler *v.t.* go downstream
he *pro.* first person inclusive
he= *pro.* first person inclusive proclitic
he? *prep.* up at
hen *n.* up yonder
ha? *prep.* at
has *interj.* expression of annoyance
hayir *n.* smell of burning flesh or bones, or postnatal blood
hayam *n.* chicken
hatap *n.* roof
hata? *hort.* Let us go now!
hadap *v.i.* present one's face, to face first (something)
halaman *n.* compound, area around a house
halaw *v.t.* chase
haroŋ *v.t.* pass by
haro? *v.t.* distribute in water
haroy *v.t.* sprinkle (something) into the wind

hapit *v.t.* pinch between (two things)
habu? *n.* hearth; kitchen; ash; dust
hamponj *adj.* light
huc *n.* rice, cooked
huris *v.t.* slice, slit
hubi? *n.* cassava. *Manihot esculenta Crantz*
huban *n.* grey hair
hüm *v.t.* bathe
hoŋ *n.* hornet
=hɔ? *clit.* enclitic to third person pronominal forms; preservative
hɔn *ign.* where; wherever
hpp *v.t.* place (food) in (one's) mouth
h?gi? *v.t.* give
hyhoy *v.i.* yawn
hlay *clf.* 1) strand, of hair. 2) sheet
hmpe? *num.* three
hmpoŋ *num.* four
hmbacaj *n.* horse mango. *Mangga foetida*

k

ki= *pro.* third person singular proclitic form
kis *v.t.* scrape hair, scales off a carcass; run a knife down hair to remove nits
kirim *v.t.* send (something) with (someone); send (someone) on an errand
kira? *v.t.* 1) reckon, count. 2) act stingy
kit *adj.* tiny (see kɛt)
ke *dem.* that
kem *n.* military camp
keti *n.* measure, weight
kĕt *adj.* small
ka?ip *n.* centipede
kah *part.* interrogative particle expressing alternative choice
kaka? *n.* elder sister
kacip *n.* betel nut clippers
kacaw *v.t.* annoy (someone)
kajan *n.* pandanus thatch
kasar *adj.* coarse; vulgar
kasaw *n.* rafter, vertical
kasut *n.* shoe
kayen *n.* cloth
kayen sot *n.* sarong
kayen suy *n.* loincloth
kayen yup *n.* sleeping cover
kayuh *v.t.* paddle; peddle
kadeh *ign.* who; someone; whoever
kadap *adv.* sometimes

hapit *v.t.* pinch between (two things)
habu? *n.* confluence
kucinj *n.* cat
kuto? *n.* box
kudes *n.* skin infection
kunig *adj.* yellow, orange, gold
kunor *n.* gourd
kunci? *n.* key
— *v.t.* lock
kul *n.* skin disease
kuraj *v.i.* be less
kurug *v.t.* close; shut
kurus *adj.* thin
kubur *n.* Red Giant Flying Squirrel. *Petaurista petaurista* (*Pallas*)
kubur *n.* 1) grave. 2) graveyard
kubog *n.* pool of water
kuman *n.* germ
kumpet *n.* great-great-grandchild
ko= *pro.* third person unknown agent proclitic
korek *v.t.* bore out
kopi? *n.* small tin can
ko? *n.* Pig-tailed Macaque. *Macaca nemestrina* (*Linnæus*)
kɔŋ *v.t.* hit (with an implement)
kɔŋkɔlōh *n.* nightjar bird
kɔleh *n.* cup, mug
komyunis *n.* Communist
kɔʃʃt *v.t.* strangle, garotte
ko *pro.* second person singular familiar
ko= *pro.* second person singular familiar proclitic
kv̥ *v.i.* come undone, fall out
khew *n.* Crested Wood Partridge. *Rollulus roulei*
khkuh *n.* termite mound
kke *dem.* that
kklon *n.* slater bug
kgdo *n.* (av.) elephant
kcbɔŋ *n.* insect sp.
kjah *adj.* heavy
ksep *v.t.* eat a little
kse? *n.* trap
ksum *n.* cocoon, beeswax, sac
ksnduduk *n.* melastoma sp. shrub
ktam *n.* crab
kton *n.* skull, shell
kde? *v.t.* dwell
kdeg *n.* rice cooking pot
kdah *n.* palm sp. *Licuala* sp.
kday *n.* shop
kdor *adj.* female

kdop *n.* fontanelle
knej *n.* forehead
knen *adv.* before
kna? *v.t.* incur
knɔn *n.* 1) offspring, 2) smaller component part
knkɔn *n.* child
knŋpuŋi *n.* child-in-law, sibling's child-in-law
knṭe? *n.* body lice
knṭoŋ *v.i.* fall off
knduri? *n.* feast
knlək *n.* husband
klan *v.i.* be startled, take fright
klambu? *n.* mosquito net
klay *n.* millipede
klubi? *n.* Salacea. *Zalacca conferta*
kluc *n.* interior
— *prep.* inside
klbm *v.t.* carry (something) long on one's shoulder
klkol *v.i.* lay one's head down (on something)
klkbnk *n.* butterfly
klkmuk *exp.* big cheeks
klŋnɔŋ *n.* dragon and damsel flies
kljtut *exp.* appearance of something bobbing in and out of view
klidew *exp.* not straight, out of alignment
klid?do? *n.* termite in it's winged life-stage
klp̥et *exp.* small and motionless
klpkap *n.* spider
klptop *n.* cockroach
klpdap *exp.* abundant
klpdop *exp.* abundant
klbeber *n.* * *kulat klbeber* a type of fungus
klmtam *n.* scorpion
kris *n.* kris
krep *adj.* dry
kreta? *n.* car
kranc'es *n.* beetle sp.
krabat *n.* court official
kramat *n.* ascetic
krasāk *exp.* sound of fish taking bait; snatch (something) off (someone)
kru? *v.t.* fish with ichthyocide
krøy *v.t.* sprinkle something down on the ground
krkēr *n.* Brush-tailed porcupine. *Atherurus macrourus* (*Linnæus*)
krcot *n.* sedge. *Lepironia articulata*
krja? *v.t.* work; harry

krdas *n.* edible seed pod of tree sp.
 krntah *n.* lightning bolt
 krnda? *n.* bier, coffin
 krlec *v.t.* cause to emerge
 krba? *n.* carabao
 krbo? *n.* Malayan False Gharial. *Tomistoma schlegelii*
 krmnom *n.* bladder
 krwej *n.* tree, sp. *Dipterocarpus chartaceus*
 krwancej *n.* coral snake
 krwpol *v.i.* 1) come out, emerge. 2) rise, of the sun and moon
 kpe?re? *v.i.* call out when hunting to keep track of partner
 kpou? *n.* tree sp.
 kpjuep *v.i.* blink
 kba? *n.* torso

kbac *v.t.* fish with a pole and line
 kbran *v.t.* fearful lest
 kmaraw *n.* dry season
 kmuc *n.* ghost
 kmus *n.* midge
 kmq *v.i.* finish, end
 — v.supp. exhaustively
 — v.supp. all up; totally
 kmn *n.* niece, nephew
 kmk'om *v.i.* sit
 kmyan *n.* resin. *Styrax benzoin*
 kmpen *n.* wife
 kmbac *n.* fishing rod
 kmmbe *num.* six (counting rhyme)
 kwali? *n.* wok
 kwala *n.* confluence
 kwarga? *n.* family

k^b

k^be? *v.t.* know, understand
 k^bet *v.t.* sting
 k^bat *v.i.* pregnant
 k^bu? *v.i.* 1) vomit. 2) ejaculate
 k^boy *n.* 1) head. 2) upper point
 k^bp *v.t.* lie face down; cover

k^bm *v.t.* achieve, get, manage
 k^bhal *v.t.* recognise, be acquainted with
 k^ben quot. QUOTE, say, sound
 k^bem *n.* odour
 k^bbas *v.i.* 1) die. 2) die down, of wind
 k^bmur *n.* caterpillar; larva

g

gi? *v.t.* give, imperative directed at children and members of the in-group
 gindaj *n.* fabric fashioned into a swag
 geh *v.t.* scratch
 ga= asp. going to
 ga?u? *n.* elder sister
 gajah *n.* elephant. *Elephas maximus* Linnaeus
 gaya? *n.* appearance
 gada? *n.* parent
 gantang *n.* a liquid measure
 gantoj *v.t.* hang, suspend
 gabnrej *top.* place name
 gambar *n.* picture; photo
 gal *v.* 1) flow in small streams. 2) float, following the current
 guc *v.i.* burn with a flame
 gunting *n.* scissors
 — v.t. cut with scissors; cut (short hair)
 gule? *v.i.* turn over
 guruh *n.* thunder
 gurut grit *exp.* appearance of something rolling and bouncing along
 gube? *n.* betel nut pounder

gum *v.t.* winnow
 gon *v.t.* bring; carry
 — v.supp. bring; carry
 gonroj *n.* fishing pole
 go? *v.i.* fell (trees)
 ggch *exp.* sound of approaching rain
 gos *v.t.* peel
 gop *n.* Malay
 gop *v.i.* bite
 gorom *n.* salt
 gom *exp.* started movement, either jumping to a standing position, or falling flat
 ghop *adj.* hot
 ggdak *n.* vine sp.
 gtah *n.* latex
 gtgut *n.* bird sp.
 gdado *exp.* huge
 gdo *adj.* old
 gnen *v.t.* knock
 glis *v.i.* none left
 glet *v.t.* able
 glk *v.i.* laugh
 gloj *v.t.* swallow

glor *v.i.* named
 gres *n.* liver
 grabac *exp.* grab someone from behind, startling them
 grak *v.i.* fall down
 grbok grabak *exp.* sound of something slapping, like rubber thongs which are too big
 grloj *n.* type of cricket
 grgor *exp.* deep rumbling
 grcñ *exp.* a little hair standing up straight

grcñ *exp.* a lot of hair standing up straight
 grphop *v.i. (av.)* eat
 — n (av.) cooked food
 grwnew *exp.* move in and out of view
 gpel *n.* Canarium tree. *Canarium sp.*
 gmarasak *exp.* the sound of walking through forest litter
 gmuk *n.* fat
 — adj. fat
 gwewwet *exp.* move side to side; wag
 gwek *onom.* sound of frog's call

l

gren *v.i.* angry

hj

hjon *exp.* sound of tiger's voice

c

ci *n.* head louse
 ci?gu *n.* teacher
 cika? *n.* gastric discomfort caused by mocking food
 cin *adj.* cooked
 cina? *n.* Chinese
 cim *n.* bird
 ce? *v.i.* disappear
 cec *v.t.* spray
 cep *exp.* sound of rice being winnowed by someone inexperienced
 cece? *v.t.* dislike
 ceyet *v.t.* hold by
 cet *n.* great-grandchild
 ceré? *adv.* always
 ca *v.i./v.t.* cat, consume
 can *n.* pouch
 catak *v.i.* fall apart, detach
 carek *v.t.* tear, rent
 cara? *n.* visual appearance, manner
 caro? *v.t.* eat (vulgar)
 caroj *v.t.* skewer, thread
 capay *v.t.* take hold
 campor *v.t.* mix
 campok *v.t.* throw, toss
 campor *v.t.* add (to fire)
 =ca? *cl.=EM*
 can *v.i.* 1) fit, pass through. 2) (av.) pass away, die
 — v.supp. 1) completely, in total. 2) immediately.
 cal *v.t.* say (something); pronounce; mention
 cu? *n.* grandchild
 cuk *v.t.* 1) pierce through. 2) inject
 cukup *v.i.* enough
 curah *v.t.* pour out, away
 cupok *n.* volume of half a coconut shell
 cop *n.* pit dug in trap construction
 conkil *v.t.* lever (something) up
 cojn *n.* down yonder
 co *n.* dog
 co? *prep.* down at
 cokwen *v.t.* discard
 coc *v.t.* gut (fish)
 coh *v.i.* born
 cokop *v.t.* talk, speak
 cop *n.* Lesser Mouse-deer. *Tragulus javanicus*
 cop *n.* brand, stamp
 c?en *v.i.* happy
 c?ac *v.i.* defecate
 c?um *v.t.* wrap
 c?on *v.t.* roast
 c?co? *v.i.* first; in front
 cher *v.i.* eat
 ckit *v.t.* pinch off
 cnew *v.t.* look down from a height
 cjay *n.* tree sp. *Neobalanocarpus heimii*

cœcūj *n.* elbow
 cœs *v.t.* fart
 cœäp *exp.* sound of something snapping shut
 cœpkinj *n.* hip
 cles *v.t.* slice
 clew *v.t.* (av.) bathe
 claka? *v.i.* ill-omened
 claw *n.* calamitous storm resulting in the inundation of the world. It is caused by the transgression of a taboo
 clok *v.t.* take (something) with one's finger; dip finger into (something)
 cloy *n.* class of monkeys and lesser apes
 cloñ *n.* 1) back. 2) behind
 cloy *v.i.* (av.) die
 critə *n.* story
 — *v.t.* tell a story
 creh *n.* fish
 cret *v.t.* bite off (a piece of food)
 crea *adj.* 1) brittle, of objects. 2) piercing, of sound

cray *v.t.* separate
 cräläp *exp.* sound of someone or something entering undergrowth
 cru? *n.* tray for winnowing rice or yams
 croh *v.t.* meet
 cros *n.* claw; fingernail
 crop *exp.* the sound of something entering water
 crom *adv.* always
 crdek *v.i.* clever
 crner *v.i.* after; behind
 crpek *exp.* pointed oval shape
 crmin *n.* mirror; pane of glass
 crwët *exp.* catch sight of female genitalia
 cpat *adj.* quick
 cman *v.t.* retell a myth
 cmaraw *exp.* sound of something being poured, like water, or rice grains
 cmo? *adv.* more, increasingly
 ctmot *v.i.* negate desire (ritually)

C^b

cök *v.t.* 1) throw (something firm, at a target). 2) plant a cutting or house post in the ground
 cøy *n.* hill

cör *v.t.* burn (something as part of a process)
 cōm *v.i.* swim
 cöör *exp.* be urinated on
 cöjir *exp.* smell of urine, ammonia-like smells

j

ji *pro.* second person pronoun
 ji= *proclit.* second person proclitic
 je=?en *pro.* second person plural
 jake *conn.* in that case
 jakun *n.* Jakun
 jaga? *v.t.* watch over
 jagon *n.* corn
 jaþot *n.* beard
 jaþrom *v.t.* encounter
 jadi? *conn.* so, then
 jala? *n.* net
 jalü *n.* pig, pork. *Sus scrofa*
 jare? *n.* digit
 jarla *n.* thorn
 jam *n.* timepiece; time (duration)
 jampi? *n.* incantation, spell
 jambol *n.* tuft
 jel *v.t.* bark at
 junjug *v.t.* carry (on the head)
 jurukarah *n.* deputy to the mntri?
 jon *v.t.* 1) give. 2) allow

jog *n.* 1) foot. 2) endpoint
 jrjor *v.i.* urinate
 j?ag *n.* bone
 j?ch *v.t.* drink
 j?çy *v.t.* make, do
 j?ji? *adj.* dirty
 j?jo? *adj.* raw, uncooked
 jhat *adj.* bad
 jhor *adj.* darker reddish brown
 jhor *adj.* light russet
 jkös *n.* porcupine. *Hystrix brachyura* (Linnaeus)
 jktek *v.i.* camp out
 jpok *v.t.* look at; observe
 jpknl *n.* span from thumb to forefinger
 jsok *n.* Sambur. *Cervus unicolor* Kerr
 jtek *v.i.* sleep
 jd? *v.i.* 1) good looking. 2) be better, after illness
 jnuh *v.i.* sated
 jlé? *adj.* short

jlawat *n.* fish sp.
 jlök *n.* smoke
 jløj *adj.* long
 jløaw *exp.* appearance of someone thin from illness
 jring *n.* tree sp. with edible seeds. *Pithecellobium lobatum*
 jreh *v.i.* tired, exhausted
 jrag *v.t.* boil water

jrøs *adj.* fast
 jrüh *adj.* deep
 jrkos *quant.* plentiful, laden
 jpon *n.* Japanese
 jme? *v.t.* drink alcohol
 jmo? *n.* vine
 jmoh *ign.* what
 jmrrtek *n.* first basket of the rice harvest
 jwar *v.t.* sew

Jl

=ja *clit.* discourse clitic
 jaça? *adv.* likewise
 jañi? *v.i.* sing
 jañah *n.* blood (childspoke)
 jañpay *n.* small insect

jam *v.t.* sense, smell, feel, taste
 ja? *adv.* very; really
 jaða? *v.i.* suckle
 jað *v.t.* plug, stop

?jn

?nih *v.i.* be ill, be in pain
 ?njac *exp.* sight of something which causes shock or bewilderment

?nat *adj.* ugly, ill-formed
 ?nes *adj.* high
 ?nuk *v.t.* suck; smoke

S

si?sa? *n.* leftovers
 sigaret *n.* tailor-made cigarette
 sijah *v.t.* stop off
 sijan *conn.* as far as
 sijmij *v.i.* go visiting
 sisik *n.* scale (fish, reptiles)
 siran *conn.* while
 siput *n.* snail
 simin *n.* cement
 simpy *n.* woven band
 sey *adj.* thin
 sen₁ *n.* zinc roofing material
 sen₂ *n.* cent
 sener *v.t.* tease by allusion
 sereh *n.* betel vine
 sep *v.i.* suck
 seg *n.* ichthyicide
 sec *v.t.* steal
 — *v.supp.* surreptitiously; do V on the quiet; without permission/knowledge
 së? *adj.* thick
 sahut *v.t.* answer, reply
 sak *v.t.* peel
 sake? *n.* screwpine
 sanjen *v.i.* crave something not had for a long

time
 saja? *adv.* only
 sayar *v.t.* feel for (someone)
 sayor *n.* vegetables
 satu? *num.* one
 satom *n.* rightside
 saden *n.* food, tinned; tinned sardines
 sadara? *n.* relative
 salay *v.t.* smoke
 salop *v.t.* prepare a swidden in a working party
 sar *v.i./v.t.* 1) descend. 2) decrease
 sarek *n.* 1) tomorrow. 2) in the future
 sapu? *v.t.* wipe, anoint, smear
 sabar *v.i.* patient
 sabur *n.* melee
 samel *conn.* while
 same? *v.i.* same; together
 sampay *conn.* until, so
 sawel *n.* leftside
 sac *n.* 1) flesh. 2) meat. 3) material, e.g. body of a winnowing tray. 4) texture, of fabric
 sat *v.i.* have laryngitis
 sar *v.t.* push (something) towards (someone/something)

suk *n.* hair, feathers, fur; fibres around segments of jackfruit
suku? *n.* quarter portion, section
susah *adj.* difficult
susu? *n.* cow's milk
sudu? *n.* spoon
sunot *v.t.* circumcise
sumay *n.* incest
sok *v.t.* go along
sot₁, *pred.* metamorphose into
sot₂ *modal.* must; may
solog *n.* first
— *comm.* at first, in the beginning
som *n.* morning
soc *v.t.* whistle
soy *v.i.* plentiful
soj *n.* endpoint (of something long)
son *clit.* =SC
s'it *v.i.* smell rotten
shpeh *v.i.* disinclined; lazy
sko? *n.* ghost
sknl * *skel k'o* have a headache
sksek *n.* food caught between teeth
skssuk *v.i.* cry until out of breath
skmbaj *n.* servant
snkalan *n.* grinding board
snkyog *n.* bamboo flute attached to arm of windmill
sjepkép *exp.* sound of rice being winnowed by someone experienced
spek *adj.* sweet
sjnot *v.t.* whimper
sjlep *n.* bamboo sp.
syap *v.i.* ready
syol *n.* 1) ground-dwelling bird sp. 2) crest
syok *n.* traces of food
sysuy *v.i.* wear a loincloth
stek *n.* slingshot
sta? *comm.* after
stsot *v.i.* wear a sarong
sdékét *quant.* a little
sdara? *n.* relative
sdac *adj.* cool

y

ye *pro.* first person singular
ye= *procl.* first person
ye? * meaning unknown
yuyuy *n.* gnat
yoh *n.* tortoise

sdér *v.t.* remember; realize
sdom *adv.* only
snih *asp.* in the middle of V-ing, while
snijmig *n.* lover
snaj *v.i.* easy, pleasant
sntep *n.* woven container
slec *exp.* bloodless
slet *v.t.* slide (something) between (two things, so that it is supported)
slew *adj.* be slippery
slalu? *adv.* always
sljew *exp.* the look of a rooster's sparsely feathered neck
slgcac *n.* channel
srut *v.i.* burn
srñ *n.* wild banana sp
srdrér *v.i.* wake up
srñar *v.i.* move, of leaves in breeze
srtep *n.* vine sp.
srwl *n.* pants
s'ro? *v.i.* pine for (someone)
sparoh *adv.* few; some
spew *n.* tree sp.
sphah *n.* four dollars
sprot *n.* same womb
sbaj *adj.* full
sbón *v.t.* keep
sbordj *n.* area
sbroh *quant.* all
smilan *num.* nine
sma? *n.* 1) human being, 2) person
— *ign.* someone; anyone
smajot *n.* spirit of life
smaji *v.t.* request
smut *n.* ant
smul *quant.* plentiful, of fruits with segments
smom *n.* bone marrow
smlay *n.* Semelai
smrykék *exp.* sound of many voices
smbayar *v.i.* pray
swak *v.i.* go; move in customary manner
swara? *n.* voice, call, sound produced by something

yok *v.i.* take; fetch
yp? *comm.* but
ypor *v.i./v.t.* 1) ascend, climb. 2) increase
ypyup *v.i.* cover (oneself) with a blanket

?ye *v.t.* see
?yam *v.i.* cry
?yəg *v.t.* hear
?yot *v.i.* return

?y

— *v.sup.* revert; go back (to V-ing)
?yōm *exp.* sound of wind blowing across a small orifice

t

tihoj *n.* post
tikam *v.t.* stab
tiga? *num.* three
tinal *v.i.* remain
tijo *n.* snake
tijo rbo *n.* rainbow
tida? *neg.* not
tilam *n.* mattress
tiris *v.i.* drip
tipoh *v.i.* deceive; bluff
tipo? *n.* betel service
tibar *v.t.* draw in
timon *n.* cucumber and other similar melons
timol *v.i.* rise to surface of water
te *prep.* to
teh *n.* tea
tet *prep.* to
tah *part.* interrogative particle expressing doubt
tahon *n.* year
tahoy *n.* the day after tomorrow
takar *n.* large water jar
taya? *n.* ladder; stairs
tajoj *v.t.* carry in the mouth
tanjay *n.* stem; handle
tajkol *v.t.* carry (by the stem)
tajam *v.t.* plane
taj *v.t.* weave
taja? *v.t.* ask a question, enquire
tasik *n.* lake
taten *v.t.* remove
tanre? *v.t.* show (something to someone)
tali? *n.* string, cord, rope; strap
tarek *v.t.* pull
tabe? *v.t.* excuse
tabog *n.* container
tabor *v.t.* cast something fine with a sweeping action
tam *v.t.* 1) hit with an open hand. 2) stamp on something with force
tampar *v.t.* slap with an open hand
tampoy *v.t.* hit (with an implement) with force

force
tampoy *n.* wild mangosteen. *Baccaurea reticulata*
tawar *v.t.* offer
tawo *n.* White-handed Gibbon. *Hyalobates lar*
ta? *hort.* let's go!
ta?en *prep.* downward, downstream
tej₁ *n.* 1) ear. 2) (*av.*) betel-vine leaf
tej₂ *v.t.* dish out (food for someone)
ten *n.* instance; time
tu?aki *n.* grandfather
tu?wan *n.* grandmother
tuh *n.* 1) breast. 2) finger peg on a bamboo stringed instrument [kranteg]. 3) right angle fold made when weaving or sewing
tuhpon *n.* 1) Creator of mankind. 2) owner
tukaj *n.* 1) craftsman. 2) bird sp
tukar *v.i.* exchange
tugpl *v.t.* dribble rice
tujku? *n.* hearth stone, tripod
tujuh *num.* seven
tunuk *v.t.* point to, at
tuy *v.t.* place finger on (something)
tut *v.t.* blow; play (a wind instrument)
tunjan *n.* betrothed
tuli? *v.i.* deaf
tulis *v.t.* write
tulah *v.i.* prohibitions against disrespectful behaviour to elders
tulat *n.* third day forward from the present
turjo? *v.t.* give
tupay *n.* squirrel (generic)
tonoy *v.t.* 1) wait. 2) watch over
ts'gbcc *exp.* short and fat (people)
todoj *v.t.* cover
tontot *v.t.* study traditional knowledge
tolog *v.t.* help
ts? *prep.* across at
tob *v.t.* undo; take off
toko? *aff.* indeed
treh *v.t.* tap rubber
tom *n.* base, source

— prep. from
tomok *n.* dagger
than *v.t.* endure, cope, bear
thariban *exp.* huge
thog *n.* hole
tkan *v.t.* hold down
tkel *v.t.* cut
tgah *n.* middle, midst
tgkal *clf.* cut length
tctuc *v.i.* faint
tpi *n.* day(light)
tyap *n.* each
tyaw *v.t.* look down from above
tyek *n.* banana. *Musa sp.*
tytsey *n.* bird sp.
ttap *adv.* for sure, certainly
ttawa? *n.* gong
tduh *v.i.* be shaded
tntu? *v.i.* certain
tlaga? *n.* well
tlaj *v.t.* hold
triyag *top.* Teriang
traj *adj.* clear
tray *v.t.* try
tran *n.* shaft

traw *n.* Long-tailed Macaque. *Macaca fascicularis* (Raffles)
trus *v.i.* straight
— *adv.* immediately without further ado
trund? *n./voc.* uncircumcised male
tron *n.* path, trail, road
tro? *v.i.* have a fever
tros *n.* inner hard wood
trspk *v.t.* listen to
trmpa? *v.t.* catch sight of
tpi? *n.* side, edge
tpui *n.* flour
tben *n.* bank
tbas *v.t.* slash (undergrowth)
tbok *v.t.* carve a space (into something)
tmakaw *n.* tobacco
tmay *n.* previous, earlier
tmay doy *n.* dusk
tmaj *v.* smile, on meeting someone
tmpat *n.* place
tmpoh *num.* seven
twanko? *n.* prince
twanptri? *n.* princess
twar *n.* fish trap
tway *n.* finger blade for reaping rice

thi *n.* hand, paw
thoy *adj.* be big

tham *v.t.* pound; strike downward with a stick
thoh *v.t.* spit

d

dig *n.* bamboo (generic)
de= *procl.* third person plural
deh *pro.* third person plural
dej *adv.* like
denej *n.* wall
dehn *pro.* third person plural S-form
da?, *pred.* exist; there is/are
da?, *neg.* not
dak *n.* 1) water. 2) watery liquid
dak jor *n.* urine
das *n.* panel (mat weaving)
dayaq *n.* darling
dara? *n.* pubescent female
darat *n.* land
dapat *modal.* able, must
dawen! *interj.* Nevermind!
dawon *n.* leaf
dahyah *v.i.* cold, to the point of shivering
də= *cl.* of

duga? *v.t.* attempt, try
duryan *n.* durian
do? *aff.* yes, as for (X)
doy *n.* afternoon
doy *voc.* nick-name for girls
dol *n.* knife handle
dom *aff.* yes
do *prep.* of
dos *v.i.* come; reach; arrive
dol *n.* house
dp? *n.* food vessels
d?en *adv.* 1) two days ago. 2) a long time ago
d?oh *n.* swidden
dkan *n.* Large Bamboo Rat. *Rhyzomis sumatrensis*
dkes *v.i.* near
dydog *n.* log
dyal *n.* above, on top of
ddes *adj.* near

ddə= *cl.* of
dlbh *v.i.* go across
dlbn *n.* 1) tree. 2) wood; stick
dri? *pro.* self
drekdruk *exp.* dirty (body)
dre *n.* rattan

dmch *ign.* what kind of
dmdəm *v.i.* lie down, be lying down
dmpuk *v.t.* follow
dwet *n.* money
dwa? *num.* two

n

nikah *v.t.* marry
ninca *n.* food
nintom *n.* yesterday
nehneh *adv.* just
neg *adv.* before; earlier
nene *n.* sound
na?day *adv.* long ago
naga? *n.* Equatorial Spitting-cobra. *Naja sumatrana*
najam *top.* place name
nas *n.* pineapple
nasip *n.* luck

na?yandeh *imp.* imperative marker
nankon *n.* youth, childhood
namba? *n.* number
nojom *n.* astrologer, seer
no? *dem.* this
npkol *v.t.* play
n?we? *n.* (av.) python
nyaj *interj.* here! summons a dog
nyur *n.* coconut. *Cocos Nucifera L.*
nrnis *n.* leafy plant
nmch *ign.* what class

?n

?no? *dem.* this

?nom *adj.* ripe

hn

hn= *cl.* object proclitic
=hn *cl.* his, hers, its

hne *adv.* shortly
hnən *interj.* agreement

l

lintaq *v.i.* cross; be horizontal to
lima? *num.* five
lek *v.t.* lick
leg *prep.* upwards, upstream
leg *n.* volume
leglong *v.t.* search fruitlessly
lec *v.i.* masturbate (male)
lewah *v.t.* look around or over something
lewnew *v.i.* be unable to finish something
la= *cl.* agent proclitic
— *conn.* because
la? *voc.* nickname for girls
lah *hort.* let us!
lagi? *adv.* again
— *asp.* *adv.* still
laget *v.i.* break off, break in two
lagkah *v.t.* step over
layar *n.* sail

layu? *v.i.* 1) wither. 2) droop, (eyes) from desire to sleep
layur *v.t.* hold over a fire to warm
lat *n.* barricade used to construct a weir when fishing with ichthyocide
lada? *n.* chili
lalag *n.* grass, long sp., generic
largemilar *exp.* writhing
lapan *num.* eight
lawok *n.* cooked meat
lawot *n.* sea
luk *n.* quiver
lan *v.t.* want; desire
luka? *v.t.* wound, injure
— *n.* wound, injury
lutus *n.* segment
— *clf.* rice grains
luluk *n.* waterlogged earth

lubuk *n.* open body of water in lake, separated by swamp forest
lumpot *v.i.* leap
lo? *n.* penis
longke? *n. (av.)* Sambur. *Cervus unicolor* Kerr
lsh *v.i.* go across
luc *asp. adv.* already
lsam *v.* rain
lsong *n.* mortar
lsobs *n.* ant
lyan *n.* space between two objects
lyar *adj.* wild, untamed

lyor *n.* saliva
lnbn *exp.* big
l?lp? *v.i.* be mentally unstable
lher *n.* neck
lrbr *v.i.* oversleep
lpes *n.* 1) tongue. 2) vibrating part of a Jew's harp
lpac *n.* belly; stomach
lbeh *n.* more
lmen *n.* spear
lmoji *n.* tooth
lwär *n.* outside
— *v.t.* direct or project outwards

?I

?le? *v.i.* long (time)
?lele? *adv.* eventually
?lem *adj.* pleasant, used for any good sensation, smelt, tasted, heard or felt
?let *v.i.* extinguish

?lep *adj.* far, distant
?luh *adj.* sharp
?luc *v.i.* pass, free
— *comm.* after

I

risaw *v.i.* fret; worry
ribu? *num.* thousand unit
ribut *n.* wind
reg *v.i.* seek
resen *n.* rations
rencérencéç *exp.* sound of shouting, a commotion
res *n.* 1) root. 2) fibres (in cassava and sweet potatoes)
resjes *n.* twig
rentej *n.* creel
ra?wen *v.t.* leave behind, intentionally
ragi? *n.* pattern
rajin *v.i.* industrious
raja? *n.* sultan
ratus *num.* hundred unit
rantay *n.* chain; weight
rapah *n.* fallen leaves, leaf litter
rabac *v.t.* 1) scratch. 2) dig by hand
rabon *n.* roof ridge
ramay *adj.* many, of people
rap *n.* bed bug, parasite sp.
rah *v.t.* probe for animals
rac *n.* bird sp.
ray *n.* companion
rap *v.t.* tread on (something); thresh by foot
ruji *n.* termite
rus *v.t.* drag

rulus *v.i.* straight
rübün *exp.* unkempt long moustache
ronkop *v.t.* invert in an inverted V shape
roc *v.t.* pull out; uproot
ronog *n.* early afternoon
rompot *n.* weed, any non-cultivar
ro? *n.* basket
roh *n.* branch
rokok *n.* cigarette, tailor-made
ros *n.* female genitalia; vagina
roy *n.* fly
roti? *n.* bread
rodg *n.* group of friends
ron *v.t.* straddle
rol *n.* lorry
rom *prep.* with; and
rökrök *exp.* sound of something large walking through pandanus
röpräp *exp.* sound of something large walking on twigs
ros *v.i.* leap
rss? *n.* fish sp.
rstog *n.* condition associated with yaws
ryal *n.* dollar
ryoh *v.i.* make a noise
rntaj *clf.* classifier for traditional narratives
rpoh *v.t.* hit (with implements like twigs, sticks)

rpok *v.t.* clap
rban *n.* pen, enclosure
rbana? *n.* drum
rbut *v.t.* take (something) off (someone) without permission; snatch
rboj *n.* shoot
rboj *v.t.* bite off

rbol *n.* taro. *Colocasia esculentum*
rmongkòp *exp.* appearance of antlers
rmo? *v.t.* reap rice
rmol *v.t.* male
rmpäl *v.t.* throw (something) on the ground with force
rwaj *n.* inner cavity; stomach

?T

?ris *v.i.* alive

?rey *adj.* many

P

pigag *n.* waist
pit *v.t.* close (eyes)
pinan *n.* arcea palm; betel nut
pintu? *n.* door
pereg *n.* plate
pe? *num.* three
=pa *clf.* =FACT
pa?nah *n.* second-born uncle
pa?itam *n.* fourth-born uncle
pa?ay *v.t.* 1) call. 2) name (an inanimate object)
pa?uda? *n.* fifth-born uncle
pakay *v.t.* use, implement; be acceptable
pakayan *n.* clothes
pagoh *v.t.* move (something) out of reach
pagel *v.t.* call
pasir *n.* sand
pasal *n.* normal, proper
pasuh *n.* water jar
pas?us *v.t.* make a fire
pay *v.t.* set aside; keep aside
— *v.supp.* for now

paya? *n.* stream
payah *v.i.* troublesome
patir *n.* tree sp.
padan *n.* open area
padashal *adv.* in fact
paday *n.* past
panay *v.i.* clever, skilled
pantaj *n.* prohibition
pantaw *v.t.* call (someone)
pala? *adv.* completely, fully, enough
paloh *v.i.* flee
para? *n.* rack for smoking game
parang *n.* gourd
parasot *n.* parachute
papan *n.* board
— *num.* eight (counting rhyme)

phem *v.t.* keep secret; keep quiet about (something)
phor *v.t.* chase out
pkima? *n.* ill-fated
pjayuh *n.* stirrer; paddle
pghulu? *n.* headman
pkke? *v.t.* pick off
pkalan *n.* landing place
ppngrawan *n.* tree sp. *Hopea mengarawan*
pcre? *v.t.* disturb
pjay *v.t.* ignore, neglect
pnakit *n.* illness
pjokel *n.* pick
psupot *v.t.* avoid
pyoh *v.i.* thrash about
pte? *v.t.* place
ptem *v.t.* plant
ptm *n.* night
ptny *n.* tree sp. with edible seed pod. *Parkia biglandulosa*
pt^hac *exp.* sound of a string snapping, breaking
pdaj *n.* sword
pdas *adj.* 1) spicy. 2) smarting or burning sensation
pduli? *v.t.* bother with (someone)
pdo? *v.t.* 1) sprout. 2) cut (a tooth). 3) break out (of a rash)
pdor *v.t.* 1) follow. 2) resemble
pnon *n.* ritual prohibitions
pnjahan *n.* middle
plita? *n.* lamp
ple *n.* fruit
plec *v.i.* emerge, come out
plat *n.* fishes exit point from rushes

plac *exp.* sound or appearance of someone/something emerging
plaw *exp.* sound or appearance of someone emerging
plasit *n.* malevolent shaman
pltaw *exp.* bulging eyes
plto? *v.t.* 1) explode. 2) rupture, of sac
p?la? *n.* animal
 — ign. thing
pri *v.i./t.* free
pri?to? *v.t.* put away
prjka?an *n.* vagina
pre? *adj.* hot
pre? *v.t.* jest
praho? *n.* dugout canoe
prati? *v.t.* watch; look at
pran *n.* storage rack
praw *exp.* sound of something hitting the water
prac *exp.* draw weapon
prac *n.* wing
pru? *num.* six
prokrol *exp.* get up from lying position
propras *exp.* sound of something moving through a tree
prom *n.* lead
prongop *n.* basket trap
prhgoh *n.* tree sp.
prhnti? *v.i.* cease, stop
prnah *adv.* ever
prntoh *exp.* white
pmaleq *n. (av.)* elephant
pwas *v.i.* satisfied, sated

p

p^heker *v.t.* think, cogitate
p^hare *n.* Clouded Monitor. *Varanus bengalensis*
p^hal *v.i./t.* 1) come down, get out of. 2) blow up, of a breeze, wind; arrive (season). 3) descend, into darkness of night, approaching storm. 4) work up (a sweat).

5) play (a card). 6) go fishing
p^har *exp.* sound or appearance of someone leaping down from a height
p^hur *exp.* sound or appearance of someone leaping down from a height
p^hoy *v.t.* fan
p^hor *exp.* white (hair)

b

bih? *v.t.* ignore
bidaj *clf.* classifier for mats
bintay *n.* star
bila? *ign.* when; whenever
bilay *v.t.* count

biras *n.* in-law's spouse
biru? *adj.* blue, deep
beh *neg.* no
beg *v.t.* inspect
ba?i *n.* tuber, rhizome

ba?i₂ *adv.* alone
bah *neg.* no, emphatic
bahaw *top.* Bahau
bageh *n.* Dusky or Spectacled Leaf Monkey. *Presbytis obscura* (Reid)
bagay *adv.* like
banja? *n.* group, kind
baju? *n.* 1) shirt. 2) shed skin, of snakes, cicadas
baseg *n.* Banded Leaf Monkey. *Presbytis metalaphos* (Raffles)
basuh *v.t.* wash
bayag *n.* soul or body spirit
batek *n.* batik
baten *n.* headman
batu? *n.* stone, rock
baneg *n.* Brown Asian Giant Tortoise. *Manouria emys emys*
bantin *n.* fourth day hence
bantay *v.t.* butcher
bantal *n.* pillow
balay *n.* verandah
baraj *n.* thing
bapa? *n.* father
baba *n.* rice plant
bak *v.t.* tie, fasten
bar *v.t.* fasten, secure
buh *v.t.* put; place
buku? *n.* 1) lump, clot. 2) book
buko? *v.t.* open, turn a page
bunsu? *n.* parent's youngest sibling
burgay *n.* aromatic ginger
busu? *n.* parent's youngest sibling
bunuh *v.t.* kill
buntaj *v.i.* off, of food
bul *v.i.* 1) toxic. 2) intoxicating
bulan *n.* 1) moon. 2) month
buruk *adj.* worn out
bubuh *n.* fish trap
boh *n.* elsewhere
bomoj *v.i.* eldest
bos *n.* sugar cane. *Saccharum officinarum L*
boy *neg. imp.* don't!
bom *n.* bomb
bo? *v.t.* carry (a child) on the back in a cloth sling
boy *v.t.* dig
borog *ign.* what (thing)
b?ilak *v.i.* get out of the way, dodge
b?en *neg.* not
bhih *v.i.* sated, full from eating
bhan *v.t.* hew

bhbek *v.i.* clear, of noise
bkira? *v.i.* cook
bkar?n *exp.* ruffled, tousled, hair or feathers
bkaw *n.* flower
bkumpol *v.i.* assemble
bkpol *n.* provisions
blkahi? *v.i.* fight, quarrel
bkrasch *v.i.* whisper
bkro?boc *exp.* sound of a large animal diving into water
bkpol *n.* cigarette smoke
bcpon *top.* location in Tasik Bera
bje? *clf.* general classifier
bjanket *v.i.* contagious
bsi? *n.* 1) metal. 2) blade
bsanar *v.i.* lean against
bsugi? *v.i.* perform ablutions
bsurok *v.i.* shout
bsjet *v.i.* dark, of the night, clouds
byasa? *v.i.* used to, familiar with
btak *adj.* overgrown
btulok *v.i.* set off
btol *v.i.* be true
btoc *v.i.* bored due to loneliness
bt?ng *v.i.* be afraid
bdil *n.* gun
bnih *n.* seed
bnal *n.* tree sp.
bnare? *v.i.* dance
bnam *n.* mountain
bnul *n.* beam
btanay *v.i.* unroll a mat
bnda? *n.* thing
bli? *v.i.* buy
blej *n.* arm
blah *v.t.* split
blahan *n.* blowpipe
blajar *v.t.* study, learn
blas *num.* teens
blayar *v.i.* sail
blu *n.* thigh
blpy *v.t.* look upwards
blhut *exp. (av.)* dark, black
blnlun *exp.* lots of rubbish, scrub; many lice in hair
blldbk *exp.* appearance of someone eating, lips smacking, jowls moving
b?lu? *n.* friend; spouse; companion
bri *n.* forest
brintay *v.i.* stalk, hunt
bregw?n *exp.* a striped pattern
bren *n.* Bren machine gun

bra *n.* Bera River
bras *n.* husked uncooked rice
bradi? *n.* sibling
brapa? *ign.* how much, how many
bramatamey *exp.* all in, in a melee
brus *v.t.* scrub
br?ol *exp.* something big, that looks white or blue, lying still under water
brhluh *v.t.* run
brkas *v.t.* tie, lash something together

m

mingu? *n.* week
mipaj *n.* fine itchy hairs of bamboo
— *v.i.* itch
misay *n.* moustache
miskin *adj.* poor
midur *n.* plant sp. *Goniothalamus lowii*
mirah *adj.* red
meg *n.* cheek
ma= *clt.* irrealis proclitic
mahal *adj.* expensive
mankok *n.* bowl
majar *n.* Black Panther. *Panthera pardus* (*melanistic*) Linnaeus
masa *n.* era
masuk *v.i.* enter
mati?ana? *n.* birth-vampire
mata?are? *n.* sun
marcis *n.* matches
mande *ign.* 1) what sort of. 2) why
mandehmoh *ign.* what
maleg *n.* sky
malaj *n.* side
malan *adj.* intoxicating (betel nuts)
malu? *v.i.* shy; embarrassed
ma=?₁ *clt.* relative clause marker
ma=?₂ *num.* one
manka *conn.* if
mehn *conn.* if
muh *n.* nose
muka? *n.* 1) face. 2) front
muget *n.* great-great-grandparent
musoj *n.* civet
muy *num.* one
mudik *v.i.* go upstream
mudeh *v.i.* easy
muda? *adj.* unripe
mulp? *v.i.* begin

brceŋ *v.i.* having many bones
brsēw *v.i.* wipe off slime
brwaj *n.* Malayan Sun Bear. *Helarctos malayanus* (Raffles)
bpngi? *v.i.* walk
bbalen *n.* windmill with bamboo flutes
bburu? *v.i.* hunt with dog and spear, blowpipe, slingshot
bwol *v.t.* converse, chat

?ma? *n.* mother
?ma?itam *n.* fourth-born aunt
?ma?bapa? *n.* parents

?m

?ma?lag *n.* sixth born aunt
?ma?ambon *n.* first-born aunt
?mot *v.t.* get up on, in; land on

hm

hmoh *ign.* what

W

wen *v.i.* throw away; cast (something)
wec *n.* intestines
wes *v.t.* peel off
wa?ambon *n.* first-born uncle
war *adv.* instead

wunwen *v.i.(av.)* defecate
wo? *adv.* more
wy *n.* knife
whwhoh *v.i.* 1) get up from sleep. 2) be upright
wcwoc *v.i.* coil or curl up (one's body)

References

- Abdullah Hasan. 1974. *The Morphology of Malay*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Adams, Karen Lee. 1986. 'Numeral Classifiers in Austroasiatic', pp. 241–62 of *Noun Classes and Categorization*, edited by C. Craig.
1989. *Systems of Numeral Classification in the Mon-Khmer, Nicobarese and Aslian Sub-families of Austroasiatic*. Pacific Linguistics Series B - 101. Department of Linguistics, Research School of Pacific Studies, Canberra: The Australian National University.
- Adelaar, K. Alexander. 1995. 'Borneo as a cross-roads for comparative Austronesian linguistics', pp. 75–95 of *The Austronesians: Historical and Comparative Perspectives*, edited by P. Bellwood et al.
- Aikhenvald, A. and Dixon, R.M.W. 1998. 'Dependencies between grammatical systems.' *Language* 74.1: 56–80.
- Alien, K. and K. Burridge. 1991. *Euphemism and Dysphemism*. Oxford University Press.
- Alpher, B. 1991. 'Ideophones in Yir-Yoront'. Unpublished ms.
- Alsina, Alex, Joan Bresnan and Peter Sellis (eds). 1997. *Complex Predicates*. Stanford: CSLI Publications.
- Andrews, Avery. 1985. 'The major functions of the noun phrase', pp. 62–154 of *Language Typology And Syntactic Description*, vol. 1: *Clause Structure*, edited by T. Shopen. Cambridge: Cambridge University Press.
- Arbak Othman. 1984. *Tatabahasa Bahasa Malaysia*. Kuala Lumpur: Penerbitan Sarjana Sdn. Bhd.
- Asmah Haji Omar. 1963. 'Bahasa Semang – Dialek Kentakpong.' Unpublished research report. Department of Malay Studies, University of Malaya.
1975. *Essays on Malaysian Linguistics*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
1976. 'The Verb in Kentak Bong', pp. 951–70 of *Austroasiatic Studies*, vol. 1, edited by P. Jenner et al.
1979. 'Languages of Malaysia', pp. 1–76 of *Papers on Southeast Asian languages*, Anthology Series 5, edited by T. A. Llamazon. Singapore: Singapore University Press.
1985. *Susur Galur Bahasa Melayu*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- ed. 1997. *The Encyclopaedia of Malaysia*, vol. 12: *Languages and Literatures*. Kuala Lumpur: Didier Millet.
- Bach, Emmon, Eloise Jelinek, Angelika Kratzer and Barbara H. Partee (eds). 1995. *Quantification in Natural Languages*. Studies in Linguistics and Philosophy, vol. 54. Dordrecht: Kluwer Academic.
- Baharon Azhar bin Raffie'i. 1972. *Some Aspects of the Relationship of the Orang Asli and other Malaysians*. Kuala Lumpur: Department of Orang Asli Affairs.
- Bauer, Christian. 1992a. 'Mon-Aslian contacts.' *Bulletin of the School of Oriental and African Studies* 55: 532–7.
- 1992b. 'Aslian phonetic terminology.' *Orang Asli Studies Newsletter* 10: 2.
- Bellwood, Peter, James J. Fox and Darrell Tryon. 1995. *The Austronesians: Historical and Comparative Perspectives*. Canberra: Department of Anthropology (RSPAS), Australian National University.
- Bellwood, Peter. 1997. *Prehistory of the Indo-Malaysian Archipelago*. Honolulu: University of Hawai'i Press.
- Benjamin, Geoffrey. 1976a. 'Austroasiatic Subgroupings and Prehistory in the Malay Peninsula', pp. 37–128 of *Austroasiatic Studies* vol. 1, edited by Philip N. Jenner et al.
- 1976b. 'An Outline of Temiar Grammar', pp. 129–87 of *Austroasiatic Studies* vol. 1, edited by Philip N. Jenner et al.
1983. 'Map with "Peninsular Malaysia" (map 37) and part of "Southern Mainland Southeast Asia" (map 38)', in *Language Atlas of the Pacific Area*, vol. 2, edited by S. Wurm and S. Hattori.
- 1985a. 'In the long term: three themes in Malay cultural ecology', pp. 219–78 of *Cultural Values and Human Ecology in Southeast Asia*, edited by Karl L. Hutterer, A. Terry Rambo and George Lovelace.
- 1985b. 'On pronouncing and writing Orang Asli languages: a guide for the perplexed (part 1).' *Orang Asli Studies Newsletter* 4: 4–16.
1986. 'On pronouncing and writing Orang Asli languages: a guide for the perplexed (part 2).' *Orang Asli Studies Newsletter* 5: 3–28.
1987. 'Ethnohistorical perspectives on Kelantan's prehistory', pp. 108–53 of *Kelantan zaman awal: Kajian arkeologi dan sejarah di Malaysia*, edited by Nik Hassan Shuhaimi bin Nik Abd. Rahman.
1993. 'Grammar and Polity: the cultural and political background to Standard Malay', pp. 341–92 of *The Role of Theory in Language Description*, edited by W. A. Foley.
- 1997a. 'Affixes, Austronesian and Iconicity in Malay.' Department of Sociology Working Papers. No. 133. National University of Singapore.
- 1997b. 'Issues in the ethnohistory of Pahang', pp. 82–121 of *Pembangunan arkeologi pelancongan Negeri Pahang*, edited by Nik Hassan Shuhaimi bin Nik Abd. Rahman et al.
- (in press). 'The Aslian languages: an assessment', to appear in *Endangered Languages and Literatures of Southeast Asia*, edited by H. Steinhauer and James T. Collins.
- (forthcoming). *Kenaboi: Lost Malayan Language or Forest-Collecting Taboo-Jargon?* Bangi, Selangor: Kajian Bahasa Nusantara, Penerbit Universiti Kebangsaan Malaysia.
- (ms.). 'The Anthropology of Grammar: Self and Other In Temiar.'
- Bishop, Nancy M. 1996a. 'A preliminary description of Kensiw (Maniq) phonology.' *Mon-Khmer Studies* 25: 227–53.

- 1996b. 'Who's who in Kensiw: Terms of reference and address in Kensiw.' *Mon-Khmer Studies* 26: 247–53.
- Bishop, Nancy M. and Mary Peterson. 1994. 'Kensiw Glossary.' *Mon-Khmer Studies* 23: 163–95.
- Blagden, Charles Otto. 1894. 'Early Indo-Chinese influence in the Malay Peninsula as illustrated by some of the Dialects of the Aboriginal tribes.' *Journal of the Straits Branch of the Royal Asiatic Society* 27: 21–56.
1903. 'The Comparative Philology of the Sakai and Semang dialects of the Malay Peninsula.' *Journal of the Straits Branch of the Royal Asiatic Society* 39: 47–63.
1906. 'Language and Comparative Vocabulary of Aboriginal Dialects', pp. 379–472, 481–775 of Skeat and Blagden, *Pagan Races of the Malay Peninsula*, vol. 2.
- Blevins, Juliette. 1995. 'The Syllable in Phonological Theory', pp. 206–44 of *The Handbook of Phonological Theory*, edited by J. Goldsmith.
- Blust, Robert. 1997. 'Nasals and nasalisation in Borneo.' *Oceanic Linguistics* 36.1: 149–79.
- Borer, H. 1988. *Proceedings of the West Coast Conference on Formal Linguistics*, Irvine, CA: University of California.
- Broschart, Jürgen. 1997. 'Why Tongan does it differently.' *Linguistic Typology* 1: 123–65.
- Broselow, E. and J. McCarthy. 1983. 'A Theory of Internal Reduplication.' *The Linguistic Review* 3: 25–8.
- Bulbeck, D. 1996. 'Holocene biological evidence of the Malay Peninsula Aborigines (Orang Asli).' *Perspectives in Human Biology* 2: 37–61.
- Burenhult, Niclas. 2002. *A Grammar of Jahai*. Lund: Department of Linguistics and Phonetics.
- Bybee, J. and S. Fleischman (eds). 1995. *Modality in Grammar and Discourse. Typological Studies in Language* vol. 32. Amsterdam/Philadelphia: John Benjamins Publishing Company.
1995. 'Modality in Grammar and Discourse. An Introductory Essay', pp. 1–13 of *Modality in Grammar and Discourse*, edited by J. Bybee and S. Fleischmann.
- Carey Iskandar. 1975. *Orang Asli: The Aboriginal Tribes of Peninsular Malaysia*. Kuala Lumpur: Oxford University Press.
- Catford, J.C. 1988. *A Practical Introduction to Phonetics*. Oxford: Oxford University Press.
- Chafe, Wallace. 1995. 'The Realis-Irrealis distinction in Caddo, the Northern Iroquoian languages, and English', pp. 349–65 of *Modality in Grammar and Discourse*, edited by J. Bybee and S. Fleischman.
- Chafe, Wallace and Johanna Nichols. 1986. *Evidentiality: the Linguistic Coding of Epistemology*. Norwood: Ablex.
- Chung, Sandra, and Alan Timberlake. 1985. 'Tense, aspect and mood', pp. 202–58 of *Language Typology and Syntactic Description*, vol. 3: *Grammatical Categories and the Lexicon*, edited by T. Shopen.
- Clements G. N. and S. J. Keyser. 1983. *CV Phonology*. Cambridge, Mass.: MIT Press.
- Collings, H. D. 1949a. 'A Temoq word list and notes.' *Bulletin of the Raffles Museum Series B* 4: 69–85.
- 1949b. 'Aboriginal Notes.' *Bulletin of the Raffles Museum Series B* 4: 86–103.
- Collins, James T. 1979. 'Expressives in Kedah Malay', pp. 379–406 of *Southeast Asian Linguistic Studies*, edited by Nguyen Dang Liem. Series C no. 49. Canberra: Pacific Linguistics.
1985. 'Dialek Melayu di Kampung Landai: Menuju penelitian tatabahasa Malaya di Kalangan Orang Asli.' *Dewan Bahasa*, Julai: 476–493.
- Comrie, Bernard. 1985. 'Causative verb formation and other verb-deriving morphology', pp. 309–48 of *Language Typology and Syntactic Description*, vol. 3: *Grammatical Categories and the Lexicon*, edited by T. Shopen.
- Comrie, Bernard and Sandra A. Thompson 1985. 'Lexical Nominalizations', pp. 349–98 of *Language Typology and Syntactic Description*, vol. 3: *Grammatical Categories and the Lexicon*, edited by T. Shopen.
- Coope, A. E. 1978. *MacMillan's Malay-English English-Malay Dictionary*. Student Edition. Kuala Lumpur: Macmillan Malaysia.
- Craig, Colette G. (ed.). 1983. *Noun Classes and Categorization*. Typological Studies in Language vol. 7. Proceedings of a Symposium on Categorization and Noun Classification, Eugene, Oregon. Amsterdam, Philadelphia: John Benjamins Publishing Co.
1992. 'Classifiers in a Functional perspective', pp. 277–301 of *Layered Structure and Reference in a Functional Perspective*, edited by M. Fortescue et al.
- Croft, William. 1990. *Typology and Universals*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.
1991. *Syntactic Categories and Grammatical Relations*. Chicago/London: The University of Chicago Press.
- DeLancey, Scott. 1997. 'Mirativity: The grammatical marking of unexpected information.' *Linguistic Typology* 1–2: 33–52.
- Diffloth, Gerard. 1972. 'Notes on Expressive Meaning.' *Chicago Linguistic Society* 8: 440–7.
1974. 'Body moves in Semai and in French.' *Chicago Linguistic Society* 10: 128–38.
1975. 'Les langues mon-khmer du Malaisie: Classification historique et innovations.' *Asie du sud-est et monde insulindé* 6.4: 1–19.
- 1976a. 'Minor-syllable vocalism in Senoi languages', pp. 229–47 of *Austroasiatic Studies*, vol. 1, edited by Philip N. Jenner et al.
- 1976b. 'Expressives in Semai', pp. 246–64 of *Austroasiatic Studies*, vol. 1, edited by Philip N. Jenner et al.
- 1976c. 'Jah-Hut, an Austroasiatic language of Malaysia', pp. 73–118 of *Southeast Asian Linguistic Studies*, vol. 2, edited by Nguyen Dang Liem.
- 1976d. 'Mon-Khmer numerals in Aslian languages.' *Linguistics* (Special issue: Austroasiatic number systems) 174: 31–8.
1977. 'Towards a history of Mon-Khmer: Proto-Semai vowels.' *Tan'an Ajia Kenkyū (Southeast Asian Studies)* 14: 463–95.
1979. 'Aslian Languages and Southeast Asian Prehistory.' *Federation Museums Journal* 24: 1–16.
1980. 'To taboo everything at all times.' Unpublished ms.

1984. *The Dvaravati Old Mon Language and Nyah Kur*. Monic Language Studies, vol. 1. Bangkok: Chulalongkorn University.
1994. 'i: big, a: small', pp. 107–14 of *Sound Symbolism*, edited by Leanne Hinton, Johanna Nichols and John J. Ohala.
- Difflot, Gerard and Norman Zide. 1992. 'Austroasiatic Languages', pp. 137–42 of *International Encyclopaedia of Linguistics*. Oxford: Oxford University Press.
- Dixon, R. M. W. 1977. 'Where have all the adjectives gone?' *Studies in Language* 1: 19–80.
1994. *Ergativity*. Cambridge Studies in Linguistics 69. Cambridge: Cambridge University Press.
- Dixon, R. M. W. and A. Y. Aikhenvald. (forthcoming). *Adjective Classes: A Cross-Linguistic Typology*. Oxford: Oxford University Press.
- Duangchand, Phaiboon. 1984. 'A Phonological Description of the Kansiv Language. (A Sakai Dialect.)' M.A. Thesis, Mahidol University, Bangkok, Thailand.
- DuBois J. W. 1987. 'The Discourse Basis of Ergativity.' *Language* 63.4: 805–55.
- Dunn, Frederick. 1975. *Rain-Forest Collectors and Traders: A Study of Resource Utilization in Modern and Ancient Malaya*. Monographs of the Malaysian Branch of the Royal Asiatic Society 5. Kuala Lumpur: MBRAS.
- Durie, Mark. 1985. *A Grammar of Acehnese, on the Basis of a Dialect of North Aceh*. Dordrecht: Foris.
1997. 'Grammatical Structures in Verb Serialization', pp. 289–354 of *Complex Predicates*, edited by Alex Alsina, Joan Bresnan and Peter Sells.
- Endicott, Kirk. 1983. 'Slave raiding in the Malay Peninsula', pp. 216–45 of *Slavery, Bondage and Dependency in Southeast Asia*, edited by Anthony Reid.
- Evans, Ivor H. N. 1915. 'Notes on some aboriginal tribes of Pahang.' *Journal of the Federated Malay States Museum* 6: 101–14.
1920. 'Further notes on the aboriginal tribes of Pahang.' *Journal of the Federated Malay States Museum* 9: 16–33.
1937. *The Negritos of Malaya*. London: Frank Cass & Co. by arrangement with Cambridge University Press (1968).
- Evans, N. 1995. 'A-quantifiers and scope in Mayali', pp. 207–70 of *Quantification in Natural Languages*, edited by E. Bach et al.
2000. 'Word classes in the world's languages', pp. 708–32 of *Morphology: a Handbook on Inflection and Word Formation*, edited by Christian Lehmann, John Mugdan and Geert Booij.
2002. 'The true status of grammatical object affixes: evidence from Bininj Gun-wok', pp. 15–50 of *Problems of Polysynthesis*, edited by N. Evans and Hans-Jürgen Sasse. Berlin: Akademie Verlag.
- Evans, N. and Hans-Jürgen (eds). 2002. *Problems of Polysynthesis*. Berlin: Akademie Verlag.
- Fix, Alain G. 1995. 'Malayan Paleosociology.' *American Anthropologist* 97.2: 313–23.
2000. 'Genes, language and ethnic groups: reconstructing Orang Asli prehistory.' *Bulletin of the Indo-pacific Prehistory Association* 19: 11–16.
- Foley, W. A. (ed.). 1993. *The Role of Theory in Language Description*. Trends in Linguistics, 69. Berlin: Mouton de Gruyter.
- Fortescue, M., Peter Harder and Lars Kristoffersen. 1992. *Layered Structure and Reference in a Functional Perspective*. Papers from the Functional Grammar Conference in Copenhagen 1990. Pragmatics and Beyond, New Series 23. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Fox, Barbara (ed.). 1996. *Studies in Anaphora*. Typological Studies in Languages, vol. 33. Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Frawley, William. 1992. *Linguistic Semantics*. New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.
- Friedman, Victor A. 1986. 'Evidentiality in the Balkans: Bulgarian, Macedonian, and Albanian', pp. 108–87 of *Evidentiality: The Linguistic Coding of Epistemology*, edited by W. Chafe and J. Nichols.
- Furtado, J. I., and S. Mori (eds). 1982. *Tasek Bera: The Ecology of a Freshwater Swamp*. The Hague: Dr W. Junk Publishers.
- Gafos, A. 1994. 'The Morphological System of Temiar: A Case Study in "Nonconcatenative" Systems.' Unpublished ms.
- Geniusiene, Emma. 1987. *The Typology of Reflexives*. Empirical Approaches to Language Typology, 2. Berlin: Mouton de Gruyter.
- Gianno, Rosemary. 1990. *Semelai Culture and Resin Technology*. New Haven, CT: Connecticut Academy of Arts and Sciences.
1997. 'Malay, Semelai, Temoq: Semelai concepts of ethnicity in south-central Malaya', pp. 51–83 of *Indigenous Peoples and the State*, edited by Robert L. Winzeler.
- Gil, David. 1994. 'The structure of Riau Indonesian.' *Nordic Journal of Linguistics* 17.2: 179–200.
- Givón, T. 1994. 'Irrealis and the subjunctive.' *Studies in Language* 18.2: 265–337.
- Goldsmith, J. 1990. *Autosegmental and Metrical Phonology*. Oxford and Cambridge: Basil Blackwell.
1995. *The Handbook of Phonological Theory*. Cambridge: Basil Blackwell.
- Gomes, A. 1990. 'Confrontation and Continuity: Simple Commodity Production among the Orang Asli', pp. 12–36 of *Tribal Peoples and Development in Southeast Asia*, edited by Lim Teck Ghee and Alberto G. Gomes.
- Gray, Charles. 1852. 'Journal of a route overland from Malacca to Pahang, across the Malayan Peninsula.' *The Journal of the Indian Archipelago and Eastern Asia* 6: 369–75.
- Haspelmath, Martin. 1997. *Indefinite Pronouns*. Oxford: Clarendon Press.
- Heine, Bernd, and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Hendon, Rufus, J. 1966. *The Phonology and Morphology of Ulu Muar Malay*. New Haven: Department of Anthropology, Yale University.
- Himmelmann, Niklaus. 1996. 'Demonstratives in narrative discourse: A taxonomy of universal uses', pp. 205–54 of *Studies in Anaphora*. Typological Studies in Languages, vol. 33, edited by Barbara Fox. Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Hinton, Leanne, Johanna Nichols and John J. Ohala (eds). 1994. *Sound Symbolism*. Cambridge: Cambridge University Press.

- Hoe Ban Seng. 1964. 'Aboriginal communities at Tasek Bera, southwest Pahang: A case study on the structure of Semelai society.' Report to the Commissioner for Aborigines, Dept. Aboriginal Affairs, Kuala Lumpur. Reprinted 2001 as *Semelai Communities at Tasek Bera: A study of the structure of an Orang Asli society*, edited by A. S. Baer and R. Gianno. Subang Jaya: Center for Orang Asli Concerns.
- Hoffmann, John. 1903. *Mundari Grammar*. Calcutta.
- Hood M. S. 1974. 'An Ethnographical Investigation of the Semelai of Malaysia.' B.Litt. thesis, Oxford University.
1978. 'Semelai Rituals of Curing.' D.Phil. dissertation, Oxford University.
1990. 'Orang Asli of Malaysia: An Overview of Recent Development and its Impact', pp. 141–9 of *Tribal Peoples and Development in Southeast Asia*, Lim Teck Ghee and Alberto G. Gomes.
- Hopper, Paul J. 1986. 'Some Discourse Functions of Classifiers in Malay', pp. 309–26 of *Noun Classes and Categorization*, edited by C. Craig.
- Hopper, Paul and Sandra Thompson. 1980. 'Transitivity in grammar and discourse.' *Language* 56.1: 251–99.
- Horn, Laurence R. 1985. 'Metalinguistic negation and pragmatic ambiguity.' *Language* 61.1: 121–74.
- Huffman, Franklin E. 1970. *Modern Spoken Cambodian*. Yale Linguistic Series. New Haven: Yale University Press.
1986. *Bibliography and Index of Mainland South East Asian Languages and Linguistics*. Newhaven: Yale University Press.
- Hutterer, Karl L., A. Terry Rambo and George Lovelace (eds). *Cultural Values and Human Ecology in Southeast Asia*. Ann Arbor, MI: Center for South and Southeast Asian Studies, University of Michigan.
- Jarkey, Nerida. 1991. 'Serial Verbs in Hmong.' PhD. dissertation, University of Sydney.
- Jenner, Philip N., Laurence C. Thompson and Stanley Starosta (eds). 1976. *Austroasiatic Studies, 2 vols.* (Oceanic linguistics, Special Publication, 13.) Honolulu: University of Hawai'i Press.
- Kähler, Hans. 1946–1949. 'Ethnographische und linguistische Studien von den Orang Laut auf der Insel Rangsang an der Ostküste von Sumatra.' *Anthropos* 41–44: 1–31, 757–85.
- Katamba, Francis 1993. *Morphology*. Macmillan Press Ltd.
- Kemmer, Suzanne. 1993. *The Middle Voice*. Typological Studies in Languages, vol. 23. Amsterdam/Philadelphia: John Benjamins.
- Klavans, J. 1985. 'The Independence of Syntax and Phonology in Cliticization.' *Language* 61: 95–120.
- Koptjevskaya-Tamm, Maria. 1993. *Nominalizations*. Theoretical Linguistic Series. London/New York: Routledge.
- Kruspe, Nicole. (2004). 'Adjectives in Semelai', to appear in *Adjective Classes: A Cross-Linguistic Typology*, edited by R. M. W. Dixon and A. Y. Aikhenvald. Oxford: Oxford University Press.
- (in prep. a). 'A dictionary of Mah Meri, Bukit Bangkong.'
- (in prep. b). 'A preliminary description of the phonology and morphology of Mah Meri.'
- Ladefoged, Peter and Ian Maddieson. 1996 *The Sounds of the World's Languages*. Oxford; Cambridge, Mass.: Blackwell Publishers.
- Lazard, Gilbert. 1997. 'Ergativity.' Review article Dixon 1994. *Linguistic Typology* 1.2: 243–68.
1999. 'Mirativity?' *Linguistic Typology* 3.1: 91–109.
- Lehmann, Christian, John Mugdan and Geert Booij (eds). 2000. *Morphology: a Handbook on Inflection and Word Formation*. Berlin: Mouton de Gruyter.
- Liaw Yock Fang. 1985. *Nahu Malaya Modern*. Singapore: Pustaka Nasional Pte. Ltd.
- Lim R. P., J. I. Furtado and R. J. Morley. 1982. 'General Description of Tasek Bera', pp. 7–12 of *Tasek Bera: The Ecology of a Freshwater Swamp*, edited by J. I. Furtado and S. Mori.
- Lim Teck Ghee and Alberto G. Gomes (eds). 1990. *Tribal Peoples and Development in Southeast Asia*. Special issue of the journal 'Manusia dan Masyarakat'. Kuala Lumpur: Department of Anthropology and Sociology, University of Malaya.
- Llamazon, T. A. (ed.). 1979. *Papers on Southeast Asian Languages*. Anthology Series 5. Singapore University Press.
- Lye Tuck-Po. 2001. (ed.). *Orang Asli of Peninsula Malaysia: A Comprehensive and Annotated Bibliography*. CSEAS Research Report Series No 88. Center for Southeast Asian Studies, Kyoto University.
- Lyons, John. 1968. *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.
1977. *Semantics, vol. 2*. Cambridge: Cambridge University Press.
- Marantz, A. 1982. 'Re reduplication.' *Linguistic Inquiry* 13: 483–545.
- Manley, Timothy M. 1972. *Outline of Sre Structure*. Oceanic Linguistics Special Publication No. 12. Hawai'i: University of Hawai'i Press.
- Matisoff, James A. 1990. 'Bulging Monosyllables: Areal tendencies in South East Asian diachrony.' *Berkeley Linguistics Society* 16: 543–60.
1973. *The Grammar of Lahu*. Berkley: University of California Press.
- McCarthy, John J. 1981. 'A Prosodic Theory of Nonconcatenative Morphology.' *Linguistic Inquiry* 12: 373–418.
1982. 'Prosodic Templates, Morphemic Templates, and Morphemic Tiers', pp. 191–213 of *The Structure of Phonological Representations, part 1*, edited by H. van der Hulst and N. Smith.
- McCarthy, John J., and Alan S. Prince. 1990. 'Foot and Word in Prosodic Morphology: The Arabic broken plural.' *Natural Language and Linguistic Theory* 8: 209–82.
- McGregor, William. 1996. 'Sound Symbolism in Goonyandi, a Language of Western Australia.' *Word* 47.3: 339–64.
- Mintz, Malcolm W. 1994. *A Student's Grammar of Malay and Indonesian*. Singapore: EPB Publishers Pte Ltd
- Mithun, M. 1995. 'On the Relativity of Irreality', pp. 367–88 of *Modality in Grammar and Discourse*, edited by J. Bybee and S. Fleischman.
- Mushin, I. 1995. 'Epistememes in Australian Languages.' *Australian Journal of Linguistics* 15.1: 1–31.

- Nagaraja, K. S. 1999. *Korku Language: Grammar, Texts and Vocabulary*. Institute for the Study of Languages and Cultures of Asia and Africa. Tokyo University of Foreign Studies.
- Nagata, Shuichi. 1992. 'The Semang in South Thailand.' *Orang Asli Studies Newsletter* 10.
- Needham, Rodney. 1974. 'Some Ethnographic notes on the Semelai in Northern Pahang.' *Journal of the Malaysian Branch of the Royal Asiatic Society* 47.2: 123–9.
- Nguyen Dang Liem. 1976. *Southeast Asian Linguistic Studies*, vol. 2. Pacific Linguistics; Canberra.
- Nik Safiah Karim and Ton Binti Ibrahim. 1979. 'Semoq Beri: some preliminary remarks.' *Federated Museums Journal, New Series* 24: 19–38.
- Nik Hassan Shuhaimi bin Nik Abd. Rahman (ed.). 1987. *Kelantan zaman awal: Kajian arkeologi dan sejarah di Malaysia*. Kota Bharu, Kelantan: Perpaduan Muzium Negeri Kelantan.
- Nik Hassan Shuhaimi bin Nik Abd. Rahman et al. 1997. *Pembangunan arkeologi pelancongan Negeri Pahang*. Pekan Malaysia: Muzium Pahang.
- Noonan, Michael. 1985. 'Complementation', pp. 42–140 of *Language Typology and Syntactic Description*, vol. 2: *Complex Constructions*, edited by T. Shopen.
- Noone, H. D. 1939. 'Notes on the Benua Jakun language spoken at Sungai Lenga, Ulu Muar, Johore.' *Journal of the Federated Malay States Museums* 15: 139–62.
- Nowak, B. 1984. 'Can the partnership last: Btsisi' marital partners and development.' *Cultural Survival Quarterly* 8.2: 9–11.
- Osada, T. 1992. *A Reference Grammar of Mundari*. Tokyo University of Foreign Studies: Institute for the Study of the Languages and Cultures of Africa and Asia.
- Palmer, F. R. 1986. *Mood and Modality*. Cambridge: Cambridge University Press.
- Parkin, Robert. 1991. *A Guide to Austroasiatic Speakers and their Languages*. Honolulu: University of Hawai'i Press.
- Payne, John R. 1985. 'Negation', pp. 197–242 of *Language Typology and Syntactic Description*, vol. 1: *Clause Structure*, edited by T. Shopen.
- Peiros, I. 1996. 'Nihali and Austroasiatic.' *Mother Tongue* 2: 75–84.
1998. *Comparative Linguistics in Southeast Asia*. Pacific Linguistics, Series C no. 142. Canberra: The Australian National University, Research School of the Pacific and Asian Studies.
- Pinnow, H.-J. 1963. 'The Position of the Munda Languages within the Austroasiatic Language Family', pp. 140–52 of *Linguistic Comparison in South East Asia and the Pacific*, edited by H. L. Shorto.
- Radhakrishnan, R. 1970. 'A Preliminary Descriptive Analysis of Nancowry.' PhD dissertation, University of Chicago.
- Ratliff, Martha, and Eric Schiller (eds). 1992. *Papers from the First Annual Meeting of the Southeast Asian Linguistics Society, 1991*. Tempe: Arizona State University Program for Southeast Asian Studies.
- Reid, Anthony (ed.). 1983. *Slavery, Bondage and Dependency in Southeast Asia*. St Lucia: University of Queensland Press.
- Reid, L. A. 1994. 'Morphological Evidence for Austric.' *Oceanic Linguistics* 33: 37–72.
- Ruhlen, Merrit. 1986. *A Guide to the Languages of the World*. California: Stanford University.
- Sadock and Zwicky. 1985. 'Speech act distinctions in syntax', pp. 155–96 of *Language Typology and Syntactic Description*, vol. 1: *Clause Structure*, edited by T. Shopen.
- Saha, N., J. W. Mak, J. S. H. Tay, Y. Liu, J. A. M. A. Tan, P. S. Low and M. Singh. 1995. 'Population Genetic Study among the Orang Asli (Semai Senoi) of Malaysia: Malayan Aborigines.' *Human Biology* 67: 37–57.
- Schachter, P. 1985. 'Parts-of-speech systems', pp. 3–61 of *Language Typology and Syntactic Description*, vol. 1: *Clause Structure*, edited by T. Shopen.
- Schebesta, Peter Paul. 1926. 'The jungle tribes of the Malay Peninsula'. *Bulletin of the School of Oriental Studies* 4.2: 269–78.
- Schiller, E. 1990. 'The Typology of Serial Verb Constructions.' *Chicago Linguistic Society* 26.1: 393–406.
1992. 'Autolexical solutions to the problem of "parts of speech" in South East Asian Languages', pp. 397–415 of *Papers from the First Annual Meeting of the Southeast Asian Linguistics Society, 1991*, edited by Martha Ratliff and Eric Schiller.
- Schmidt, Peter Wilhelm. 1901. 'Die Sprachen der Sakei und Semang auf Malacca und ihr Verhältnis zu den Mon-Khmer Sprachen.' *Bijdragen tot de Taal-, Land- en Volkenkunde* 52: 399–583.
1903. 'The Sakai and Semang languages in the Malay Peninsula and their relation to the Mon-Khmer languages.' Translated by W. D. Barnes. *Journal of the Straits Branch of the Royal Asiatic Society* 39: 38–45.
- Shopen, Timothy. (ed.). 1985. *Language Typology and Syntactic Description*. 3 Vols. Cambridge: Cambridge University Press.
- Shorto, H.L., (ed.). 1963. *Linguistic Comparison in South East Asia and the Pacific*. London.
- Sircam, H. S. 1920. 'Original settlement of Pahang. Legend communicated by Haji Sam (of Minangkabau descent).' *Journal of the Federated Malay States Museum* 9.2: 150–1.
- Skeat, Walter William. 1896. 'A Vocabulary of the Besisi dialect.' *Journal of the Straits Branch of the Royal Asiatic Society* 29: 13–31.
- Skeat, Walter William and Charles Otto Blagden. 1906. *Pagan Races of the Malay Peninsula*, 2 Vols. London: MacMillan & Co.
- Sloan, Kelly. 1988. 'Bare Consonant Reduplication: Implications for a Prosodic Theory of Morphology', pp. 319–330 of *Proceedings of the West Coast Conference on Formal Linguistics*, edited by H. Borer.
- Smith, K. 1979. *Sedang Grammar: Phonological and Syntactic Structure*. Canberra: Dept. of Linguistics, Research School of Pacific Studies, A.N.U., Pacific Linguistics Series B no. 59.
- Song, J. J. 1996. *Causatives and Causation: A Universal-Typological Perspective*. London: Longman.
- Spencer, A. 1991. *Morphological Theory*. Oxford: Blackwell.

- Steinhauer, H. and James T. Collins, (eds). (in press). *Endangered Languages and Literatures of Southeast Asia*. Leiden: KITLV Press.
- Svantesson, Jan-Olof. 1983. *Kammu Phonology and Morphology*. Travaux de l'Institut de linguistique de Lund, 18. Lund: CWK Gleerup.
- Ter Mors, C. 1986. 'Affix to X*.' *The Linguistic Review* 3: 275–98.
- Thomas, David and Robert K. Headley Jr. 1970. 'More on Mon-Khmer Subgroupings.' *Linguistics* 25: 398–418.
- Thompson, Sandra A. and Robert E. Longacre. 1985. 'Adverbial Clauses', pp. 169–234 of *Language Typology and Syntactic Description*, vol. 2: *Complex Constructions*, edited by T. Shopen.
- van der Holst, H. and N. Smith. (eds). 1982. *The Structure of Phonological Representations, part I*. Dordrecht: Foris.
- Wazir-Jahan Begum Karim. 1981. *Ma' Betisek Concept of Living Things*. London School of Economics Monographs on Social Anthropology, No. 54. New Jersey: The Athlone Press, Humanities Press Inc.
- Wheatley, Paul. 1961. *The Golden Khersonese: Studies in the historical geography of the Malay Peninsula before A.D. 1500*. Kuala Lumpur: Oxford University Press.
- Wierzbicka, Anna. 1980. *Lingua Mentalis*. Sydney: Academic Press.
- Wilkinson, R. J. 1927. *A Malay-English Dictionary*. Reprinted 1942/43 by Daito Syuppan Kabusiki Kaisya. Japan: Tokyo.
- Williams-Hunt, P. D. R. 1952. *An Introduction to the Malayan Aborigines*. Kuala Lumpur: The Government Printer.
- Winstedt, R. 1957. *Malay Grammar*.
- Winzeler, Robert L. (ed.) 1997. *Indigenous Peoples and the State: Politics, Land and Ethnicity in the Malay Peninsula and Borneo*. New Haven CT: Yale South east Asian studies, No. 46. Yale University Press.
- Wurm, S. A. and Shiro Hattori (eds). 1983. *Language Atlas of the Pacific Area*, vol. 2. Canberra: Australian National University, Australian Academy of Humanities and The Japan Academy.

Index

A

- absolutive 91, 106, 258, 264, 272
 absolutive focus 174
 adjectives 100, 211–14
 dimension 100, 101, 146–8
 colour 101–2
 property 100, 108
 adjuncts 250, 251, 330
 benefactive adjunct 215, 250
 manner adjuncts, *see* adverbs, manner
 temporal adjuncts, *see* adverbs, temporal, time
 verbal adjuncts in adverbial function 315–16
 adnominal, *see* associative phrase
 adverbs 103, 296–316
 aspectual 296–301
 degree, *see* adverbs, quantification
 numerals in adverbial function 314
 frequency, *see* adverbs, quantification
 instrumental orientation 315
 intensity, *see* adverbs, quantification
 lexical adverbs 296–312
 manner 103, 315–6
 quantification 309–12
 replay 301–4
 restrictive 304–6
 similitude 198, 313–4
 temporal 195–6, 250, 307–9
 time 312–13
 affixation 63–87
 forms and processes 64–8
 notation 63
 affixes 62, *also see* reduplication
 -an 'NMZ' 69, 84, 223
 b<>an 'TOG' 84–5, 124–5
 b(r)- 'MID', 'HAVE', 'USE' 82–3, 150–1
 <C> 'IMPERF'/'INDIV'/'NMZ'/'UNIT' 218, 219, 223,

causative 86–7, 155–6
 -i² 'APPL', 'ITER' 84, 135–40
 <mn> 'NMZ' 19, 80–1, 224–5
 mN- 'PERFM' 84, 152–3, 115–17
 +n+ 'NMZ' 76–9, 222
 p- 'CAUS' 83, 130–1
 par- 'CAUS' 70–2, 127–8, 129, 155
 par-/pr- 'xs' 83–4, 225
 +pC+ 'NMZ' 79, 224
 +pn+ 'NMZ' 69–70, 223–4
 partially specified 64, 66
 prespecified syllabic 65, 69–72
 prespecified non-syllabic 76–81
 <r>- 'CAUS' 76, 130–1
 +ra?+ 'COMP' 69
 tar- 'MCAUS' 70–2, 128–9
 t(r)- 'HAPP' 83
 underspecified non-syllabic 62, 65, 72–6
 agent, *see* semantic roles
 Aktionsart 106–7, 116
 anaphora, *see* zero anaphora
 applicative 69, 135–8
 and intransitive verbs 135–6
 directional raising 136
 directed emotion 137–8
 argument coding 258–66
 A-marking 258–61
 cross-referencing 258–9
 proclitic la= 'A' 89–90, 259–61
 in the intransitive clause 267–8, 161–2
 in imperfective clauses 112
 in intransitive and non-verbal clauses 264–6
 in transitive clauses 260–6
 indirect object 263–4
 O-marking 261–3
 ascriptive clause, *see* non-verbal clauses
 ascriptive predictor 103, 269, 273–4, 276–8
 Aslian
 as a linguistic classification 12
 as an ethnic classification 12
 Aslian languages 1, 10–18, 48
 and causatives 87
 and comparative phonology 58–60
 and imperfectives 115–16

Central Aslian 10, 11, 14, 59, 134
 distribution 11
 historical overview 12–14
 internal subgroupings 10–11
 Northern Aslian 10, 30, 58, 134
 population figures 29
 prehistory 16–18
 Southern Aslian 10, 14, 30, 58, 60, 134
 aspect, *also see* adverbs, aspectual
 aspectual modifiers of negated predicates 300–1
 imminent aspect 163–4
 prospective 299
 retrospective 299
 associative phrase 94, 181–2, 202, 203, 213–18
 and compounds 216–8
 functions of the associative phrase 214–5
 locative nominals in the associative phrase 215–6
 attributives and attributive phrases 121, 202, 211–12, 217, 232
 demonstratives as attributives 194–6
 locative prepositional phrases as attributives 211
 relative clauses as attributives 211
 verbs as attributives 212
 Austroasiatic 1, 10, 12, 17, 56, 95, 87, 149, 224, 396
 Austronesian 10, 16, 56, 95, 134
 auto-causative, *see* causatives and middle voice
 auxiliary and supplemental verbs 103
 avoidance speech style 7–10, 80–1, 211, 297–8
 B
 backgrounding 114
 base 63, 80, 81
 Batek Deq 18, 29–30, 58, 148
 Batek Nong 13
 benefactive 156, 215, 250
 Besisi, *see* Mah Meri
 bound 95
 root inflection 87
 bound root 109–10
 bound words 63, *see* clitics
 C
 causal 89, 183, *also see* complex clauses, connective
 causation 124–5, 127, 131
 direct 125, 127
 facilitated 129

imputed 137–8, 159–60, 163
 indirect 126
 mediated/manipulative 130–1
 causatives 124–35
 allomorphs 71, 84
 auto-causative reflexives, *see* middle voice
 cooperative auto-causative 131–4
 derivation of nominals 155–6
 semantics of causative morphemes 127–34
 with the happenstance 146–8
 valency change 125–7
 Ceq Wong 9, 18, 30, 58, 59, 60, 404
 Che' Wong, *see* Ceq Wong
 circumfix 62, 68, 85, 123–4
 classifiers 190, 202, 204, 206–8, 219
 clause types
 basic clauses 269–338
 complex clauses 339–95
 clitic 62, *also see* cliticisation
 absolutive 90–1
 enclitics
 connective enclitic 381–2
 discourse clitics 410–19
 third person possessor 90–1
 proclitics
 absolutive focus 92, 174
 imminent aspect 163–4
 irrealis 161–2
 name 93
 possessive 92
 possessor focus 92
 pronominal 88, 156–61
 verbal clitics 156–64
 cliticisation 64–5, 87–93
 domain of attachment 2, 88
 coda constraints, *see* syllables and syllabification
 coda copy, *also see* imperfective, reduplication
 functions of coda copy 75–6, 110–15
 and nominal to verb derivation 150–2
 collective 123–4
 comitative 246, *also see* instrumental
 Communist emergency 22–3, 57–8, 424, 426
 comparative 69, 146–8
 comparison 147, 244

complement clauses, *also see* complex clauses
 desiderative and achievement predicates 347, 357–9
 paratactic 347, 348–51
 assertion of fact 347, 350
 factive 347, 350
 immediate perception/psyche action 347
 knowledge of identity 350
 nominal equative complements 352–3
 reduced (hypotactic) 347, 351–7
 ability and knowledge 356
 co-referential deletion 351
 internal structure 352–3
 permissive/denial of permission 347, 355–6
 cognitive 347
 predicate types 357–8
 sense 357
 utterance 347, 354
 complementiser 346, 353–4
 complex clauses 383–98
 complement clauses 115, 346–59
 concatenated clauses 373–79
 consecutive 374
 conditional 375
 purposive 374–5
 simultaneous action 116, 375–6
 syntax of the concatenated clause 376–9
 connective clauses 380–95, *see also* adverbial connective clauses 382–91
 causal 383–5
 conditional 385–7
 general features of the clause type 380–2
 opposition 390–1
 temporal 387–90
 serial verb clauses, *see* serial verb constructions
 compound nouns, *see* nouns, compound nouns
 concatenated clauses, *also see* complex clauses
 purposive
 concatenation, *see* complex clauses
 concatenative morphology 66, 82–85
 conditional clauses, *see* complex clauses
 connectives 97, 103, 380–2
 adverbial 382–91
 discourse 391–5

consonants 32–4
 phonetic realisation 34–6
 distribution 48–53
 as onsets 49–52
 as codas 52–3
 glottalised sonorants 36
 prestopped nasals 34–5
 constituent order 4–5, 253–8
 and core arguments 254–6
 and obliques 256
 in intransitive and non-verbal clauses 255–6, 272–3, 277–8
 in perfective/transitive clauses 255
 in universal clauses 254–5
 cooperative auto-causative, *see* causatives

D

decausatives, *see* middle voice
 deictics, *see* deixis; prepositions, locatives *and* directionals
 deixis
 discourse deixis 195–6, 198–201
 spatial deixis 198, 200
 demonstratives 99, 173, 192–201, 202, 203, *also see* pronouns
 deontic modality 293–6
 denial of permission 279, 295
 negative imperative 279, 294–5, 336–8
 prohibition/lack of obligation 279, 295–6
 prohibitive 279, 295–6
 dependent clause, *see* complement clauses
 dependent marking 5
 derivational affixes, *see* affixes
 deverbal nouns, *see* nouns, deverbal nouns
 directionals, *see* prepositions, directionals
 discourse clitics, *see* clitics, discourse
 ditransitive, *see* trivalent verbs
 domain of attachment 65, 66

E

enclitics, *see* clitics
 English loanwords 58
 epenthesis 45–48, 67, 68, 70, 74, 76, 77, 78, 79
 epistemic 279–96
 experiential 279, 281, 285–6
 inferential 279, 281, 282–5
 and affective state predicate 283–4

and intransitive stative predicates 284–5
 mirative 279, 281, 286–94
 and affective state predicates 289
 and discourse functions in traditional narratives 290–1
 and perception and cognition predicates 287–9
 in interrogative clauses 289–90
 hypothetical 279, 282, 292–3
 ergative 90, 106
 excessive agent/performer 83–4, 221, 225
 existential clauses, *see* non-verbal clauses, existential
 existential clauses and possession 275
 existential predictor 103, 269
 expressives 97, 102, 396–402
 phonology and morphology 401–2
 semantics of expressives 397–9
 syntax of expressives 399–401

F

focus
 absolutive, *see* clitics, proclitic
 possessor, *see* clitics, proclitic
 fronting 251, 256–8, 329

G

goal, *see* prepositions, directionals
 grammatical relations 247–53
 core relations 247–9, 251–3
 oblique relations 249–50
 complements 249–50, 254
 adjuncts 250, 254, 332
 grammatical word 62

H

happenstance 83, 140–6
 expressing dreams and thoughts 143
 inability 143
 intensive absorption 144
 with causatives 144–6
 with verbs of speed 146
 head marking 5
 heterosemy 95, 96
 hortative, *see* imperatives
 hypothetical, *see* epistemic modality

I

ideophone, *see* expressive

ignoratives 99, 176–92

- as free-choice pronouns 179, 184–5, 187, 189–90, 192
- as interrogative pronouns 177–8, 179–81, 182–3, 185–6, 188–9, 190–1, 192
- as indefinite and negative pronouns 178–9, 181–2, 183–4, 186–8, 189, 191–2
- ontological categories 177–8, 180, 183, 186, 189, 191–3

imperatives 331–8

- enclitics 411, 417
- formal features of imperative clauses 331–2
- hortative 334–6
- interjections with imperative force 408–9
- negative imperative 296–7, 336–8
- negative imperatives with declarative-like forms 38
- particle 331
- positive imperative 332–6
- positive imperatives with declarative-like forms 334

imperfective 111–17, 221, *also see* Aktionsart

derived with coda copy 111–15

derived with *məN-* 115–17

in complex clauses 116

indefinite pronoun, *see* pronoun, ignorable

indirect object 232–3, 247, 249, 263–4, 332

inferential, *see* epistemic modality

infixes and infixation 67, 86, 111

root infixation 86

instrumental 245–7

intensive 117, 149

intensive absorption, *see* happenstance

interjections 97, 104, 404, 407–10

of negative attitude 407–8

with imperative force 408–9

interrogatives 251–2, 324–33,

alternative interrogative 324, 326–7

information interrogative 324, 328–31, *also see* ignoratives

intonation interrogative 324–6

polar interrogatives 324–8

responses to polar interrogatives 327–8

intonation 325

intransitive verb

irrealis 89, 161–2, 278–96, 334, 336–8

epistemic modality 279, 281–93

deontic modality 279, 293–6, 336–8

formal features of an irrealis clause 279–0

with intransitive verbs 161–2

with transitive verbs 161, 162

iterative 138–40

with a distributive reading 139–40

iterativity 114

J

Jah Hut 10, 11, 12, 93, 53, 59, 60, 67, 115, 134, 148, 206, 224, 337, 396–7, 404

Jahai 11, 13, 53, 59, 67, 76, 134

Jehaic 10

Jakun 12, 18, 29, 55, 396–7

K

Kammu 42, 48, 50, 67, 76, 396

Kensi 13, 30, 53, 58, 59, 134

Kentaq Bong 13, 134

Khasi 10, 15

Khmer 10, 95

kinship terms 23–6, 97–8

Kintaq Bong, *see* Kentaq Bong

Korku 95

L

Lanoh 12, 13, 30

lexemes 61, 94–5

lexicalisation, non-productive patterns 85–7

linguistic affiliation, with Aslian

with Mon-Khmer 14–16

linguistic type 1–10

loan words, *see* English, Malay

locatives, *see* prepositions, locative

locative nominals, *see* nominals, locative nominals

M

Mah Meri 10, 12, 13, 14, 30, 59, 60, 76, 93, 134, 148, 206, 207, 214, 300, 404

main clause 346

Malacca languages 12

Malay 18, 19, 22, 56, 81, 207, 396–7, 398, 402

comparative notes 117, 121, 136, 154, 155, 207, 316, 324, 353

loans 3, 9, 10, 56–8, 67–8, 70, 82, 95, 123, 152, 205–6, 209, 217, 218, 389

Mendriq 13

middle voice 84, 117–23

auto-causative reflexives 118, 122

decausatives 118, 119

intensive experiences 121

reciprocals 122, 155

reflexive reciprocals 119, 122–3
and nominal to verb derivation 153–5
possession 184–5
Mintil 58
mirative, *see* epistemic modality
modal category 164–8
 necessity 165–7
 possibility 167–9
modality, *see* deontic modality; epistemic modality; irrealis
Mon 10, 17, 59, 149
Mon-Khmer 10, 12, 15–16, 18, 67, 87, 134, 152, 207, 224
mood 279
morphology, derivational 110–149
morphophonemic processes 66–8
 glottal deletion 67
 onset deaspiration 68
 /r/ dissimilation 67
 resyllabification 68, 84
 stress reassignment 68, 84
 underspecified vowels, *see* epenthesis
Munda 10, 16, 95, 134, 224
Mundari 396

N

negation 316–24
 aspectual modifiers of negated predicates 300–1
 internal 316–19
 metalinguistic 319–24
 propositional, *see* negation, internal
 negative attitude, expressions of 410–11
negative imperative, *see* imperatives
negators 103
 imperative 316
 metalinguistic 316
 propositional 316
Nicobarese and the Nicobar Islands 10, 15, 66, 76, 87, 224
nominal derivation 218–26
 action 220–2
 agent/attributive nominalisation 224–5
 excessive agent/performer 221, 225
 individuation 93, 218–19
 location nominalisation 222–3
 noun to noun 218–20
 object nominalisation 223–4

process nominalisation 224
result nominalisation 222
state 221, 223
syntax of nominalised clauses 225–6
verb to noun 220–6
nominalisation, *see* nominal derivation
nominals 97–100
 locative nominals 193, 216–17
 nominal to verb derivation 150–7
nonconcatenative morphology 61, 64, 69–80
non-verbal clauses 269–78
 ascriptive clauses 269, 270
 aspect in non-verbal clauses 272–3, 278
 coding strategies 264–6
 constituent order in non-verbal clauses 257–8, 270–1
 equative clauses 269, 270
 existential clauses 269, 273–6
 identificational clauses 269, 270
 locative clauses 269, 270
 metamorphic clause 269, 273, 276–78
 negation in non-verbal clauses 273, 276
 predicate in non-verbal clauses 271–2, 279–80
 subject in non-verbal clauses 271, 276
noun phrases 202–26
 and restrictive modifiers 305, 306
 functions of the noun phrase 203–4
 internal structure 202–3
nouns 97–98
 compound nouns 213–4, 216–18
 deverbal nouns 98, 220–6
 dysphemistic nouns 93, 219
 individuated 218–19
 generic nouns 178–9
 measure nouns 190, 202, 208–10
 reduplication of nouns 221
 unit nouns 207, 219
number, *see* pronouns
numerals 92, 99, 202, 204–6, 207
 and the restrictive modifier 306
 numerals in adverbial function 314–5
 numeral to verb derivation 151–2
Nyah Kur 241

O

objects, 115, *also see* argument coding
 oblique 127
 onomatopoeia 396
 onset, *see* consonants
 onset copy, *see* reduplication
 Orang Asli 29–31
 Orang Kanaq 29
 Orang Kuala 29
 Orang Seletar 29

P

parataxis, *see* complex clauses
 passive 121
 patient, *see* semantic roles
 perfective 112, *see also* *Aktionsart*
 perform, *see* affixes and imperfective
 person, *see* pronouns
 phonemes, *see* consonants, vowels
 phonology 1, 32–60
 phonotactics 48–55
 possession, *also see* middle voice; non-verbal clauses, existential clauses
 ignorative 186
 inalienable 182, 194
 possessive
 construction, *see* associative phrase
 enclitic 90–91, 214
 proclitic 93, 174–5, 247
 preposition, *see* preposition, possessive
 possessor
 focus 93, 175
 marking 252–3, *also see* associative phrase
 predicate types
 ability and knowledge 359
 activity 282
 affective states and bodily sensation 282, 289
 assertion of fact 353
 cognitive disposition 282
 desiderative and achievement 360–2
 factive 353
 perception and cognition 287–9
 permissive/non-permissive 358–9
 sense 360
 utterance 357–8

predicators 103, 231–2, *also see* ascriptive predicator and existential predicator
 prefixes and prefixation 67, 82–4, 111
 prepositions and prepositional phrases 102, 227–46
 as attributives 182, 192, 211, 230
 as peripheral expressions 229, 231–2, 237–8
 as predicatives 229–30
 comitative, *see* instrumental
 core function encoding indirect objects 230–1, 232–3, 238, 263–4
 distributed reading of deictic locatives and directional 236–7, 243
 locatives 212
 deictic locatives 103–4, 227, 229, 233–8, 251–2, 265–6
 neutral locative 231, 230, 233–5
 directionals 227, 238–1
 directionals and deictic locatives 103–4, 241, 352–3
 deictic directionals 227, 229, 239–40
 neutral directional 227, 229, 239
 instrumental 227, 228, 243–5
 possessive 227, 228, 245–6
 source 227, 228, 241–3
 structure of the prepositional phrase 228
 proclitic, *see* clitics
 prohibitive 294–5
 pronominal proclitics 88–89, 158–162, 260–261
 and intransitive verbs 159–61
 first person allomorphs 90
 pronouns 98–99
 and the in-law taboo 172
 augmented 92
 bound pronominal forms 88, 170–1
 collective pronoun 175
 deference 171
 demonstrative 100–1, 173–4, 192–201, *also see* deixis
 locative demonstrative pronouns 193–4
 recognition function 200–1
 exclusive 171
 free pronouns 170–6
 free-choice pronouns 178, 184–5, 187
 ignorative, *see* ignoratives
 indefinite pronouns, *see* ignoratives
 inclusive 171
 minimal 92
 number 93, 171
 person 171
 personal 171–7

reciprocal derivation 155
 reflexive 176
 unidentified agent 120
 prosodic head 65–6, 69
 purposive, *see* complex clauses

Q

quantifier phrase 202, 204–10
 quantifiers, non-numeral 210
 quasi-passive 119
 questions, *see* interrogatives
 quotative marker 403–7

R

recipient 251–2, *also see* semantic roles; indirect object
 reciprocal 117, 120–5, 151, 156–7; *also see* middle voice
 reduplication 67
 coda copy 72–74, 85–6
 light syllable reduplication 66, 81, 150, 221
 of demonstratives 193
 of expressives 402
 of ignoratives 184, 220
 of nominals 220
 of verbs, *see* reduplication, coda copy, light syllable reduplication, onset copy
 onset copy 72, 75
 reflexives 108, 110, 111, 116, 117, 122–3, 132–5; *see also* middle voice
 relative clauses 251, 329–30, 339–46
 as nominal modifiers 212
 as a nominalisation process 344
 formed with the relativiser 340–45
 headed 342–3
 headless 343–6
 juxtaposed 346
 in the formation of avoidance words 343
 relativiser 91, 343
 responses to interrogatives, *see* interrogatives, responses
 resyllabification, *see* syllables and syllabification
 riddles 420–4
 root 62
 bound roots 94, 109
 monosyllabic roots 69, 70, 74, 75, 77, 80, 82, 83
 disyllabic roots 69, 70, 73, 76, 78, 80, 82, 83

S

Sabum 11, 29, 30
 Sakai 12, 13
 Sedang 48, 50
 Semai 12, 13, 53, 60, 76, 87, 93, 95, 134, 148, 396, 398, 402
 Semang 12, 13
 semantic roles 247
 agent 108, 247–8
 experiencer 108, 247, 248, 249
 goal 248
 instrument 248
 location 248
 patient 248–9
 recipient 249–50
 theme 248
 Semoq Beri, *see* Semoq Beri
 Semelai
 at Tasek Bera 21–3
 ethnographic background 23–8
 ethnonym 19
 linguistic situation 28–9
 location 18
 prior documentation 19–21
 Semelaiic 10
 Semnam 11, 29, 30
 Semoq Beri 1, 9, 13, 18, 30, 59
 Senoiic, *also see* Aslian, Central
 serial verb constructions 88–9, 115, 359–73
 general features 361–2
 types 363–72
 associated motion 367–70
 cultural activity 369–70
 manner 363–7
 purposive motion 115, 370–2
 reflexive 115, 368–9
 resultative 372
 synonymic 372–3
 similitude, *see* adverbs
 Sino-Tibetan 10
 source, *see* prepositions
 Sre 67
 stress 40
 in Malay loans 56
 reassignment 69, 84

subcategorisation 107
 subordinate clause
 suffixes and suffixation 62, 68, 84, 111
 syllables and syllabification 1, 40–8, 64–5, 82
 codas and coda constraints in nonfinal syllabites 42–5
 codas, sonorant 42–4
 pre-established 44–5
 epenthesis 45–8, 67, 68
 phonologically underspecified 45–6
 vowels of morphophonemic origin 47–8
 resyllabification 66, 68, 84
 syllable contact constraints 43–5, 66
 syllable structure of loanwords 56–7
 syntactic fusion 347

T

taboo, in-law 172
 tag questions, *see* interrogatives, polar and alternative
 Tai 17, 18
 Tai-Kadai 10
 Tasek Bera 1, 18, 21–3
 Temiar 12, 13, 30, 66, 70, 76, 87, 93, 115, 134, 174, 221, 396–8
 Temoq 13, 14, 18, 22, 29, 30, 174
 templates 64–6, 72–5, 76–7, 78, 79, 81
 temporal
 adverbs, *see* adverbs, temporal
 connectives, *see* connectives clauses
 Temuan 18, 29
 time, *see* adverbs, temporal, time
 transitivity, lexical 100

U

underspecification, *see* affixes and epenthesis

V

valence 131
 monovalent verbs 106
 divalent verbs 101, 107
 trivalent verbs 101, 107, 126–7
 verbs 100–102, 105–69
 active 101, 108, 109, 119, 334
 as attributives 211–12
 body action and processes 108, 125, 127–8, 159–60, 222
 citation form 109
 classification 106–7

emotion 159–60
 in attributive function 212–14
 intransitive verbs 101–2, 105, 108–110
 itable 95
 motion 110, 129, 159–60
 of giving/transfer 125, 250
 postural 108, 112, 222
 process 22
 social interaction 125
 stative 101, 108, 109, 248, 334,
 transitive 100–1, 105, 107, 107–108
 supplemental 103
 verb derivations, *see* affixation and affixes
 verb phrases 89
 and restrictive modifiers 304–5
 verbal clitics, *see* clitics
 verbal word 89, 105–6, 258
 Vietnamese 10
 vocative 172, 203, 331
 vowels 32, 36–9
 distribution 53–6
 final and penultimate syllables 54–5
 prepenultimate syllables 55
 epenthetic 45–8, 55–6
 nasality 37–8,
 non-phonemic 39
 phonetic realisation 38–9
 underspecified, *see* epenthesis

W

word classes 4, 72, 93, 94–104
 membership, prototypical 95–6
 open 95, 97–102
 closed 95–6, 101, 102–4

Z

zero anaphora 260, 266–8, 378
 zero conversion 95, 109