

Ridge, Eleanor Claire (2019) Variation in Vatlongos Verbal Morphosyntax : speaker communities in Southeast Ambrym and Mele Maat. PhD thesis. SOAS University of London. <http://eprints.soas.ac.uk/32205>

Copyright © and Moral Rights for this thesis are retained by the author and/or other copyright owners.

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s.

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

When referring to this thesis, full bibliographic details including the author, title, awarding institution and date of the thesis must be given e.g. AUTHOR (year of submission) "Full thesis title", name of the School or Department, PhD Thesis, pagination.

Variation in Vatlongos
Verbal Morphosyntax:
speaker communities in Southeast
Ambrym and Mele Maat

Eleanor Claire Ridge

Thesis submitted for the degree of PhD

2019

Department of Linguistics
SOAS, University of London

Abstract

This thesis examines morphosyntactic variation in an unusual sociolinguistic context in Vanuatu. Vatlongos (Oceanic, Austronesian) is spoken by communities in the southeast of Ambrym island, and by a peri-urban community near the capital city, who relocated after a volcanic explosion in the early 1950s. Within Southeast Ambrym, the thesis further distinguishes Endu from other Vatlongos-speaking villages on the basis of observed dialectal, sociolinguistic and language-contact differences, especially contact with communities and languages of Northern Ambrym. The sociolinguistic setting of Vatlongos is explored via a survey of speakers from all three communities, looking at implications for language contact, language attitudes and the vitality of Vatlongos.

The verbal morphosyntax of Vatlongos is described through qualitative and quantitative analysis of a >65,000-word corpus of spontaneous and elicited texts recorded during fieldwork, paying careful attention to variation within and between speaker communities. It first outlines the syntax of simple clauses, the tense, aspect and mood categories of the language and the verbal morphology, before moving onto a description of complex verbal constructions: serial verb constructions, complex predicates, subordination and auxiliary verb constructions.

Finally, the thesis examines the frequency of occurrence of these verbal constructions in spontaneous texts across the speaker communities, using chi-square tests and negative-binomial regression modelling to investigate the effects of community and other speaker-level and text-level factors: age, gender, years of education and genre. There are community level differences in the token frequency of auxiliary and serial verb constructions: lower frequency of use of these two constructions is associated with higher level of education in the Anglophone education system.

Both the formal variation and the community level frequency differences are consistent with effects of language shift in Mele Maat, under heavy exposure to Bislama and English.

Acknowledgements

This project has been funded by the Arts and Humanities Research Council (AHRC Doctoral Studentship AH/K5032771). A research visit to Lacito CNRS in January-March 2018 was funded by a mobility grant from Labex EFL.

Attendance at the Vanuatu Languages Conference in July 2018 was part-funded by a SOAS travel grant.

I would like to thank my supervisor Peter Austin for stimulating discussions, detailed feedback and tireless support throughout the course of the PhD. I would also like to thank my external supervisor Cindy Schneider for her insights into Vanuatu language contexts as well as her support and occasionally company during fieldwork, and my second supervisor Lutz Marten for his guidance during the upgrade process and subsequently.

The analysis presented in this thesis has benefited from feedback at many conferences and meetings, most notably the SOAS Grammar Working Group meetings, the South of England Lexical Functional Grammar meetings, the Linguistics Research Student Conference at SOAS, the Austronesian and Papuan Languages and Linguistics conferences, and the Vanuatu Language Conference. It has also greatly benefited from feedback, training and discussions with colleagues at Lacito, CNRS, who I would also like to thank for the warm welcome they gave me in early 2018.

I would like to thank Sophie Yingying Mu, Charlotte Hemmings, Michael Franjeh, Ana Krajinović, Ebany Dohle, Antoinette Schapper and especially Andrew (Tabisini) Gray for feedback and discussions on drafts of this thesis.

I would also like to thank my colleagues at the International Study and Languages Institute at the University of Reading for their enthusiasm and camaraderie over five long summers of intense but enjoyable pre-sessional courses, especially Kate for her kindness and hospitality in Brisbane.

In Vanuatu, I would like to thank the Vanuatu Cultural Centre for their permission and support for this project. I would also like to thank Robert Early and John Lynch at the Pacific Languages Unit at the University of the South Pacific for providing working space in Port Vila during fieldwork, and especially

for their guidance in the initial planning stages of this project and subsequent discussions during fieldwork.

In Port Vila, I would like to thank Sam and Ginny, Kay Johnson and Georgia and their boys, and Claire and James, for their hospitality and support, especially during the aftermath of Cyclone Pam. I would also like to thank colleagues at Save the Children for their tireless hard work and good company during that period, particularly at the Southeast Ambrym base in Moru.

Most importantly, I would like to thank the speakers of Vatlongos who have enthusiastically shared their language, stories and culture with me, not to mention their company, meals and hospitality.

In Mele Maat, I especially want to thank Bell Mansen for his hard work, patience, and humour. I am very fortunate to have worked with such a good and kind man. *Nammesuk pupu, tua navan.*

In Endu, Elder Saksak Ruben and his family welcomed me and my friends into their home, a place full of fun and laughter. I would especially like to thank Saksak for his expertise and enthusiasm for Endu-Vatlongos. *Nammes xamitel toptang. Sipa men!*

In Silimaori (Moru), Elder Simeon and Madleen Ben looked after me like one of their children, Yaxley, Barbara and Tungon. Apart from their relentless hard work, which has been essential to the success of this project, I would like to thank Simeon and Madleen for their kindness, humour and warm hearts. I treasure my time with your family. *Nammes xamitel himetu.*

I would also like to thank the many many participants who contributed their stories and performances to the project. Those who wished to be thanked by name are shown in Table 1. I would also like to thank all those not included by name, and the countless community members who helped the project and myself in other ways, from food, to logistics, to Vatlongos practice. *Nammes xamim vus!*

Finally, I would like to thank my friends and family, especially Grandma, Mum, and Dad, for their love and support during the PhD and always.

Table 1: Named participants who contributed recordings to the project

Code	Name	Village
001	Lik Simelum	Mele Maat (Maat)
002	Chief Albea David	Mele Maat (Maat)
003	Rosie Obed Simelum	Mele Maat (Maat)
004	Ben (Bell) Mansen	Mele Maat
006	Emma Mael	Mele Maat (Nguna)
007	Elsie Kalo	Endu
008	Endu Kindy (performances)	
009	Riana Anna	Endu
011	Joram Mael	Endu
012	Chief Abel Ditamat	Endu
013	Philomen Paul	Endu
014	Windy Saul	Endu (Sameo)
015	Masing Marah	Endu (North Ambrym)
016	Elder Ruben Saksak	Endu
017	Wotan	Endu
018	Rina Joel	Sahuot (Endu)
019	Andrew Kailik	Endu (Sameo)
020	Chief Mael Moses Kondara	Endu
021	Women of Endu (at house building)	
022	Elder Timothy Tania Holiu	Endu
023	Elder Joel Peter	Toak
024	Lilon Paol	Endu (North Ambrym)
025	Maena	Endu (North Ambrym)
026	Elder Simeon Ben	Moru
027	Leiza Houlu	Mele Maat
028	Rael Zekiel	Mele Maat
029	Gina Simelum	Mele Maat
030	Lesbeth David	Mele Maat (Maat)
033	Lydia Mael	Moru
034	Lomael Pong	Penapo (Moru)
035	Annie Azrum	Endu
036	Margaret Jonny	Moru (Endu)
038	Lucy	Baksabue (Bethel)
039	Masel Maki	Moru (Paama)
040	Elenor	Penapo
045	Walter Berry	Mele Maat (Endu)
046	Madleen Ben	Moru (Penapo)
048	Otison Tonney	Moru
049	Alice Bell	Moru
051	Meriam Rosleen	Moru
052	David Takik	Moru
053	Salote Takik	Moru
054	Meriam Taso	Moru (Bethel)
055	Hiram Rit	Taveak
060	Lydia Shem	Sai
061	Obed Jacob	Penapo

Code	Name	Village
062	Elder Philip Mael Walter	Penapo
063	Hinak	Penapo
064	Mary Talivo	Penapo
065	Leah Ruth	Penapo
066	Boys of Moru (playing football)	
067	Bela	Taveak (Toak)
068	Mary Thomas	Taveak (Tanna)
069	Juliet	Taveak
070	Sam Giri	Bethel
074	Leian Maiam Tabong	Moru
075	Elder Eikem Joe	Moru
077	Par Andre	Sahuot
078	Annie Samson	Sahuot
079	Chief Tasis Tom	Utas
080	Elder Sali Hotip	Maat
081	Sapi Ezekiel	Maat
082	Sopa Ezekiel	Maat
083	Lengal	Maat
084	Soule Esal	Maat
085	Cindy Simon	Moru
087	Barbara Simeon	Moru
088	Yaxley Simeon	Moru
090	Mary Tasis	Mele Maat
091	Ali	Mele Maat
093	Billy	Mele Maat
094	Joshua Yen	Mele Maat (Santo)
095	Liz Samuel	Taveak
097	Ernie Moses Simon	Moru
098	Jim Tony	Utas
099	Ruben Thomas	Taveak
100	Peter Resen	Taveak
101	Esra	Taveak
103	Samuel Lohe	Toak
104	Job Samul	Endu
105	Lona Lesi	Endu
106	Mael Tony	Endu
107	Noume	Endu
108	Jenny	Sameo (Ase)
109	Maggie	Sameo (Ranon, North Ambrym)
110	Masing Saul Sesel	Sameo
111	Gabriel	Sameo
112	Moses	Endu
113	Leah Riana Ki	Ase
117	Netty Arsen	Toak
118	Akem George	Ulei
121	Meriam Roy	Moru (Paama)
122	Merry Lulu	Toak

Code	Name	Village
124	Aram Luke	Endu
125	Elder Awel Matin Manses	Utas
128	Ruth Newell	Mele Maat
130	Masel Joshua Petuel	Mele Maat
131	Lamaesapo	Sai (Maat)
134	Tungon John Simeon	Moru
136	Philip Jeremiah	Moru
137	Chief John Daniel Tahul	Endu
138	Jake Tungon	Sameo
139	Andrew Masin	Sameo
140	Lin	Utas (Penapo)
141	Roy Morris	Moru
142	John Fred Masing	Moru
143	Mata James	Utas
144	Jacob Uru	Toak
145	Marie Rotu	Pamel
146	Malap Fred	Pamel
147	Johnson Anhon	Pamel
148	Karolin Malat	Petel
149	Rosna	Endu
150	Joe Tungon	Sameo
151	Chief Isaiah Masing	Ase
152	Jim Moloul Isaiah	Ase
153	Nina	Santo
154	Obed Petuel	Ase
156	Petri	Vameveok
157	Simon Avi	Vameveok
158	Leimari Meteneai	Mele Maat (Mele)
159	Emily Rangimen	Mele Maat
160	Lehi Lucy Willie	Mele Maat
161	Pam Kindy Moru (performances)	
162	Lidi Joseph	Moru
163	Penapo Kindy (performances)	
164	Claudia Rangimen	Mele Maat
165	Kalsa Mael	Penapo
166	Lesbeth Vera Maki	Mele Maat
167	Lesni Simelum	Mele Maat (Penapo)
168	Anya	Mele Maat
169	John Enoch Rangimen	Mele Maat
172	Cyrus	Mele Maat

This thesis is dedicated to the memory of Michael Ridge, a great storyteller, artist and social worker. My Uncle Mick encouraged me to be curious, reflective and adventurous from a young age, and was very supportive of my academic pursuits and work in Vanuatu. His warmth, empathy, and desire to understand everything are an inspiration.

Contents

Declaration	2
Abstract.....	3
Acknowledgements	4
List of Tables	16
List of Figures	19
Orthography, abbreviations and conventions	21
1 Introduction	26
1.1 Vatlongos	28
1.1.1 Vatlongos language	29
1.1.2 Vatlongos speaker-communities	32
1.1.3 Anthropological research in Vatlongos communities	37
1.2 Language change and variation	39
1.2.1 Mechanisms of language change.....	39
1.2.2 Language shift and attrition	44
1.2.3 Language change and variation in the Pacific	46
1.3 Methodology.....	51
1.3.1 Fieldwork.....	51
1.3.2 Data gathering and processing.....	52
1.3.3 Corpus construction and analysis	57
1.3.4 Community outcomes and goals.....	60
1.3.5 Documentation and archiving	62
1.4 Summary.....	63
2 Speaker-community profiles	64
2.1 Social profiles of communities	65
2.2 Language repertoires.....	69

2.3	Language use in different domains	72
2.4	Associations with different languages	78
2.5	Language vitality	83
2.6	Summary.....	86
3	Syntax of simple clauses	87
3.1	Constituent order in simple clauses.....	87
3.1.1	Non-basic word orders and information status	90
3.1.2	Information status of other arguments and non-arguments..	94
3.2	Prepositional arguments and adjuncts	96
3.2.1	Incorporated-preposition verbs	102
3.3	Position of adverbs and adverbials	104
3.3.1	Incorporated-adverb verbs.....	109
3.4	Verbs with additional syntactic constraints	110
3.4.1	Reflexive verb <i>hesi</i>	110
3.4.2	Psycho-collocations	113
3.5	Non-verbal predicates	116
3.5.1	Clauses without verbs.....	116
3.5.2	Copular verb <i>he</i>	118
3.5.3	Locative predicates and existential clauses	121
3.6	Clitics in the clause.....	125
3.6.1	Partitive enclitic <i>=ti</i>	125
3.6.2	Negative clitic <i>ti</i>	128
3.7	Summary.....	142
4	Tense, Aspect and Mood distinctions.....	144
4.1	Applying definitions of tense, aspect and mood to Vatlongos	145
4.2	The extension of Vatlongos TAM categories.....	152
4.2.1	Non-future	157

4.2.2	Prior	161
4.2.3	Immediate future	165
4.2.4	Distant future	172
4.2.5	Imperative and prohibitive moods	178
4.2.6	Apprehensive	179
4.3	Other strategies for marking aspectual and modal distinctions..	180
4.3.1	Completive aspect	181
4.3.2	Durative aspect.....	183
4.3.3	Intensifier	187
4.4	Summary.....	190
5	Verbal morphology.....	193
5.1	Concatenative inflectional morphology.....	194
5.1.1	Subject indexing prefixes.....	197
5.1.2	Negative and apprehensive prefixes.....	209
5.1.3	Object and transitivity suffixes.....	211
5.2	Non-linear morphology.....	217
5.2.1	Verb-initial consonant mutation.....	218
5.2.2	Reduplication	222
5.2.3	Verb stem alternations	227
5.2.4	Irregular verb <i>ammei</i> 'come'	230
5.3	Derivational morphology	230
5.3.1	Nominaliser <i>-en</i>	231
5.3.2	Adjectiviser <i>te-, ta-</i>	232
5.4	Multiple exponence of TAMP	233
5.5	Summary.....	235
6	Serial verb constructions	237
6.1	Defining and describing SVCs	237

6.1.1	Morphosyntactic criteria for identifying SVCs	239
6.1.2	Other factors that correlate with SVCs	243
6.1.3	Typological description of SVCs	245
6.1.4	Conceptual issues	246
6.2	Overview of SVCs in Vatlongos	247
6.3	Formal properties	250
6.3.1	Constituent order	251
6.3.2	Morphological indexing and argument structure	256
6.3.3	Morphological TAMP	259
6.3.4	Comparison with non-contiguous SVCs in Paamese	265
6.4	Functions of SVCs	267
6.4.1	Discourse: focus and comments	267
6.4.2	Adjunct-like modification	269
6.4.3	Prior motion	271
6.4.4	Directional modification	274
6.4.5	Direct quotations	276
6.4.6	Complement clauses	277
6.4.7	Increasing valency	278
6.4.8	Oblique arguments	279
6.4.9	Sub-events within complex events	281
6.4.10	Inchoative aspect	281
6.4.11	Deontic modality	282
6.5	Discussion	282
7	Other strategies for combining verbs	286
7.1	Root-contiguous complex predicates	287
7.1.1	Formal properties	289
7.1.2	Functions of complex predicates	291

7.1.3	Productivity and lexicalisation.....	295
7.2	Subordination.....	300
7.2.1	Complement clauses.....	301
7.2.2	Clausal subjects.....	308
8	Auxiliary verb constructions.....	310
8.1	Formal properties.....	312
8.1.1	Constituent order.....	313
8.1.2	Morphological inflection.....	317
8.1.3	AVCs and prior motion SVCs.....	322
8.2	Functions of auxiliary verbs.....	325
8.2.1	Continuous aspect with 'stay' verbs.....	327
8.2.2	Direction of prior motion with basic motion verbs.....	334
8.2.3	Obligation with Bislama loan word mas 'must'.....	335
8.3	Diachronic developments.....	335
8.3.1	Diachronic origin.....	335
8.3.2	Possible incorporation into the verbal affix template.....	341
8.4	Summary.....	342
9	Variation in frequency of verbal constructions.....	343
9.1	Frequency of constructions as count data.....	347
9.2	Composition of the corpus.....	352
9.3	Effect of community on construction frequency.....	360
9.4	Effect of age, gender, education and genre on construction frequency	366
9.5	Modelling frequency of each construction in a text.....	374
9.6	Summary.....	383
10	Conclusion.....	387
10.1	Notable features of Vatlongos verbal grammar.....	387

10.2	Variation in Vatlongos verbal morphosyntax	392
10.3	Significance of the project.....	396
10.4	Directions for future research.....	401
10.5	Summary.....	404
	References.....	406
	Consent form.....	426

List of Tables

Table 1: Named participants who contributed recordings to the project.....	6
Table 2: Vatlongos orthography used in this project	21
Table 3: Glossing abbreviations.....	23
Table 4: Other abbreviations	25
Table 5: Glossary of regionally-specific terminology	25
Table 6: Lexical variants in Endu dialect.....	35
Table 7: Survey respondents by age, gender and community	65
Table 8: Preferred language in different domains by community.....	74
Table 9: Domains most associated with different languages by community	77
Table 10: Average importance rating of each language (/4) by community	79
Table 11: Positive and negative associations with Vatlongos.....	79
Table 12: Positive and negative associations with Bislama.....	81
Table 13: Positive and negative associations with English.....	82
Table 14: Positive and negative associations with French.....	82
Table 15: Semantic scope of adverbs in different positions in the clause.....	105
Table 16: Semantic classes of adverbs and their distribution	106
Table 17: Phonologically reduced adverbs.....	109
Table 18: Distribution of constituents in different clauses types in relation to negative clitic <i>ti</i>	131
Table 19: Summary of contexts for each morphological TAM category	153
Table 20: Non-morphological marking of aspectual and modal categories	180
Table 21: Independent pronoun paradigm in South-Vatlongos.....	198
Table 22: Complete subject-indexing prefixes for major TAM categories	201
Table 23: Paradigm for subject-indexing prefixes in imperative mood.....	208
Table 24: Paradigm for subject-indexing prefixes in prohibitive mood	209
Table 25: Subject-indexing and negative prefixes in the third person.....	210
Table 26: Prefixes marking verbs as apprehensive	211
Table 27: Object-indexing suffixes.....	214
Table 28: Some semantic roles associated with the object of dynamic transitivised verbs.....	216
Table 29: Environments for verb-initial consonant mutation.....	219

Table 30: Some patterns of verb initial consonant mutation	220
Table 31: Verbs with CV reduplication	223
Table 32: Verbs with CVC reduplication.....	224
Table 33: Verbs with CVCV reduplication.....	224
Table 34: Lexicalised forms derived from rightwards reduplication	225
Table 35: Verb stems undergoing vowel elision with prefixes.....	227
Table 36: Verb stems undergoing vowel elision or losing initial CV syllable ...	229
Table 37: Morphological TAMP dependencies in SVCs.....	260
Table 38: Mood-polarity dependencies in Paamese non-contiguous SVCs	267
Table 39: RRG Interclausal Relations Hierarchy (Van Valin & LaPolla 1997: 481)	283
Table 40: Cline of semantic relations expressed by complex predicates in Oceanic languages (Bril 2007: 302)	284
Table 41: Auxiliary verbs in Vatlongos	312
Table 42: Rate of occurrence and normalised relative frequency of tagged verbal constructions in the corpus	349
Table 43: Contingency table for AVCs by community	361
Table 44: Construction frequency (per 10,000 words) by community and results of chi-square tests	364
Table 45: Construction frequency (per 10,000 words) by age group and results of chi-square tests	367
Table 46: Construction frequency (per 10,000 words) by gender and results of chi-square tests	368
Table 47: Construction frequency (per 10,000 words) by years of education and results of chi-square tests	369
Table 48: Narrow and broad definitions of genre, with word counts.....	371
Table 49: Construction frequency (per 10,000 words) by narrowly-defined genre and results of chi-square tests	372
Table 50: Construction frequency (per 10,000 words) by broadly-defined genre and results of chi-square tests.....	374
Table 51: Significant factors for each construction according to chi-squared tests.....	377
Table 52: Summary of model of frequency of SVCs per text.....	378

Table 53: Summary of model of frequency of AVCs per text	380
Table 54: Summary of model of frequency of zero-marked complement clauses per text.....	381
Table 55: Summary of model of frequency of complex predicates per text	382
Table 56: Summary of model of frequency of prior motion SVCs per text.....	382
Table 57: Summary of significant and near-significant trends in models of construction frequency per text	383

List of Figures

Figure 1: Map of Ambrym and Paama	30
Figure 2: Map of Central Vanuatu, showing Ambrym and Efate islands.....	32
Figure 3: Map of Efate, showing Mele Maat, Mele and Port Vila	33
Figure 4: Map of villages in Southeast Ambrym	34
Figure 5: Workflow for documentation of Vatlongos.....	53
Figure 6: Constituent order in the Vatlongos verbal clause.....	87
Figure 7: Adverb positions in a clause.....	105
Figure 8: Morphological TAM categories in Vatlongos.....	144
Figure 9: Spectrogram of example (365)	186
Figure 10: Vatlongos verbal affix template.....	195
Figure 11: Constituent order possibilities in SVCs with two component verbs	251
Figure 12: Diagram of prior motion direction expressed by auxiliary verbs <i>ha</i> 'go' and <i>mei</i> 'come', relative to deictic centre.	315
Figure 13: Relative position of speaker, addressee and object in (632).....	316
Figure 14: Coherent and incoherent paths expressed by relative ordering of the prior motion auxiliaries	317
Figure 15: Root forms of lexical <i>ti</i> 'stay' in different relative tense and polarity environments	318
Figure 16: Forms of auxiliary <i>ti</i> 'continuous' in different relative tense environments	318
Figure 17: Root forms of lexical <i>ha</i> 'go' in different relative tense and polarity environments.....	319
Figure 18: Forms of auxiliary <i>ha</i> 'GO' in different relative tense environments.	320
Figure 19: Comrie's classification of aspectual oppositions (Comrie 1976: 25)	331
Figure 20: Proposed structure of aspectual categories and extension of 'continuous' label.....	332

Figure 21: Screenshot of a segment of text in the SIL FLEx database showing morpheme-level and word-level glossing, with sentence-level translations and tags [20141027a_n01m001_56].	344
Figure 22: Histogram of normalised relative frequency of SVCs per 10,000 words in each text	349
Figure 23: Histogram of normalised relative frequency of SVCs per 10,000 words in each text	350
Figure 24: Boxplots of normalised relative frequency per 10,000 words in each text for each construction	351
Figure 25: Composition of texts in the corpus by age group, community, gender, genre, and years of education.	353
Figure 26: Composition of words in the corpus by age group, community, gender, genre, and years of education.	354
Figure 27: Mosaic plot of words in the corpus by community and age group.	355
Figure 28: Mosaic plot of words in the corpus by community and gender.	356
Figure 29: Mosaic plot of words in the corpus by community and years of education	357
Figure 30: Mosaic plot of words in the corpus by community and genre	358
Figure 31: Mosaic plot of words in the corpus by gender and genre	359

Orthography, abbreviations and conventions

In this project, Vatlongos is transcribed using a practical orthography that takes elements from the orthography of Gary J. Parker’s (1970a) *Southeast Ambrym Dictionary* and the Vatlongos Bible translation (Wycliffe Bible Translators 2015), shown in Table 2. Allophones and phonetic variants of the phonemes are described in (Parker 1968a). The orthography is designed to be as phonemic as possible, while maintaining similarity to the spelling systems of the other languages of literacy in the community, Bislama and English (see Chapter 2). Bislama orthography is used to represent sounds that only occur in loans (marked with an asterisk (*) in Table 2). Following community preferences, this orthography does not use any diacritics. This results in some loss of information in the representation of vowels: the grapheme ‘e’ represents the mid and low front vowels /e/ and /æ/, but the /æ/ is somewhat marginal. The only digraph is ‘ng’ for /ŋ/, as this digraph is used in both Bislama and English.

Table 2: Vatlongos orthography used in this project

Grapheme	Phoneme (IPA Symbol) ¹	Description
a, A ²	a	low central unrounded vowel
b, B	^m b	prenasalised voiced bilabial stop
d, D	ⁿ d	prenasalised voiced alveolar stop
e, E	e	mid front unrounded vowel
	æ	low front unrounded vowel
f, F	*f	voiceless labiodental fricative
g, G	^ŋ g	prenasalised voiced velar stop
h, H	h	voiceless glottal fricative
i, I	i	high front unrounded vowel
j, J	*dʒ	voiced post-alveolar affricate
k, K	k	voiceless velar stop
l, L	l	alveolar lateral
m, M	m	bilabial nasal
n, N	n	alveolar nasal
ng, Ng	ŋ	velar nasal
o, O	o	mid back rounded vowel
p, P	p	voiceless bilabial stop

¹ (Decker 1999)

² The grapheme ‘a’ also represents the /æ/ in forms based on the very common basic motion verb *ha* ‘go’. When it undergoes verb-initial consonant mutation to occur as *ba*, *ma* or *va* the vowel is realised as /æ/. Rather than create another homograph for the homophones *he* ‘copular’ and *he* ‘go to’, the ‘a’ grapheme is used for all the inflected forms of this root.

Grapheme	Phoneme (IPA Symbol)¹	Description
r, R	r	alveolar tap or trill
s, S	s	voiceless alveolar fricative
t, T	t	voiceless alveolar stop
u, U	u	high back rounded vowel
v, V	β	labial fricative
w, W	*w	labiovelar approximant
x, X	χ	velar fricative
z, Z	*z	voiced alveolar fricative

Punctuation and capitalisation follow English conventions. In addition, disfluencies in speech are marked with a comma, and the trailing ellipsis (...) is reserved for a distinctive prosodic pattern associated with durative markers (§4.3.2).

Glossed examples follow the conventions of the Leipzig glossing rules (Bickel, Comrie & Haspelmath 2015). Each morpheme in the transcription line has a corresponding gloss using an English translation for approximate lexical meaning, and the abbreviations for grammatical meanings in Table 3. Each example is followed by a free translation in English. Examples judged ungrammatical are preceded by an asterisk (*).

The transcription and gloss lines are aligned by grammatical word. Morpheme boundaries within words are marked with a hyphen (-), and in cases of cumulative exponence, different semantic components are separated with a period (.). When a lexical meaning is glossed by more than one word in English, these are separated with an underscore (_). Clitics are analysed as independent grammatical words (§3.6), so are aligned separately, but preceded or followed by an equals sign (=) to indicate whether the clitic is prosodically dependent on the preceding or following word. Complex predicates (§7.1) are analysed as a single grammatical word, so they are treated as a single word for alignment, but component lexemes are separated by a space according to the orthographic preferences in the community. The boundary between the component lexemes is marked with a plus sign (+) in the gloss line.

Person-number features of an argument marked on a head are followed by the grammatical function of the argument they index, e.g. an object (2SG.OBJ) or possessor (3SG.POSS). When no grammatical function is specified, the person-

number features index a subject. When the absence of a formative marks a certain cell of a paradigm, the feature values of that cell are prepended to the root, separated by a period (e.g. 3SG.NFUT.see). This is to avoid committing to a zero-morpheme analysis. In the English translations, the third-person singular is translated with singular *they/them/their* unless the gender of the referent is clear from the context.

The translation is followed by a code in [square brackets] identifying the session, text and segment number, in the format YYYYMMDDa_b01c002_3, e.g. 20141105e_n01e012_04. This begins with the date the recording was made, followed by a letter ('a') to distinguish multiple sessions in a day. 'b' is a letter indicating text type: c=conversation; h=oral history; i=interview; k=commentary; m=metadata survey; n=narrative; p=procedural; s=song; , t=public speech; v=poem; x=elicited examples. 01 is a different number for each recording made in a single session. c is a letter indicating the community membership of the speaker(s): e= Endu; m= Mele Maat; s= Ase-Taveak. 002 is the speaker code listed in Table 1 (or the main speaker in multi-participant texts, or this is skipped for texts with a large number of equal contributors). Finally, '3' is the segment number within the recording.³

Literacy materials produced in VESP workshops to support the new national curriculum (§1.1.1), are coded Y1X2_3. 1 is the school year, X is a letter code for medium (R=reader, P=poster), 2 is the number in the series of materials (which is the same for materials in all vernacular languages included in the VESP project), and 3 is the sentence number in the text.

Glossing abbreviations are in Table 3; other abbreviations used in the thesis are given in Table 4; and Table 5 is a glossary of regionally-specific terms that occur in glosses and elsewhere in the thesis.

Table 3: Glossing abbreviations

Gloss	Meaning
1	first person
2	second person

³ Segmentation is discussed in §1.3.2.

Gloss	Meaning
3	third person
ADJZR	adjectiviser
APR	apprehensive mood
BIS	Bislama loan verb
CL	possessive classifier stem
COME	prior motion towards deictic centre, grammaticalized 'come'
CONT	continuous aspect (progressive and habitual)
CONTR	contrastive
COP	copula
DFUT	distant future
DR	drinkable and mats/bedding (possessive classifier)
DU	dual number
DOM	domestic: plants, animals, land (possessive classifier)
ED	edible (possessive classifier)
EXCL	exclusive person
FUT	future (ambiguous between IFUT and DFUT)
GEN	general (possessive classifier)
GO	prior motion away from deictic centre, grammaticalized 'go'
HES	hesitation
IFUT	immediate future
IMP	imperative
INCH	inchoative aspect
INCL	inclusive person
INSTR	instrument
LOC	locative
MED	medial
MIN	minimiser
NEG	negative polarity
NFUT	non-future tense
NMLZR	nominalizer
OBJ	object
OBL	oblique
PART	partitive
PC	paucal number
PL	plural number
POSS	possessor
PRI	prior
PROX	proximal (deixis)
REAL	realis (in Vatlongos, non-future and prior)
RED	reduplication
REL	relative clause marker
SG	singular number
SUB	subordinator
SUBJ	subject
TR	transitiviser
UNKN	unknown
V2	subsequent verb in a complex predicate

Table 4: Other abbreviations

Abbreviation	Meaning
0Comp	Zero-marked complement clause
Aux	Auxiliary verb
AVC	Auxiliary verb construction
Comp	Complement clause (either zero-marked or marked with a complementiser)
corpus	Chapters 3–8 and 10: the corpus of interlinearised texts Chapter 9: the corpus of spontaneous interlinearised texts tagged for quantitative analysis
CPred	Complex predicate
E	Event time
NP	Noun phrase
PNG	Papua New Guinea
PP	Prepositional phrase
R	Reference time
RRG	Role and Reference Grammar
S	Speech time
subcorpus	Chapters 3–8 and 10: the corpus of spontaneous interlinearised texts tagged for quantitative analysis Chapter 9: subsections of the corpus of spontaneous interlinearised texts tagged for quantitative analysis
SVC	Serial verb construction
SVO	Subject-verb and Subject-verb-object constituent order
TAM	Tense-Aspect-Mood
TAMP	Tense-Aspect-Mood-Polarity
TSit	Situation time
TT	Topic time
TU	Utterance time
VESP	Vanuatu Education Sector Program
VNSO	Vanuatu National Statistics Office
VP	Verb phrase

Table 5: Glossary of regionally-specific terminology

Word	Meaning
bushnut	Sea almond
kava	A plant (<i>Piper methysticum</i>) and the narcotic drink prepared from its roots.
laplap	The national dish of Vanuatu, prepared by grating root vegetables or banana, soaking in coconut milk and baking in banana leaves.
nakamal	Men's meeting house where kava is prepared and drunk. Single young men often sleep here.
ni-Vanuatu, ni-Van	A person from Vanuatu, their nationality
unu	A fool (this term is found in Ambrym languages)

1 Introduction

This thesis focuses on an unusual sociolinguistic situation to explore questions concerning language change and variation in morphosyntactic structure. The Vatlongos language is spoken in the villages of Southeast Ambrym where it originates historically, but also in Mele Maat, a peri-urban community on the outskirts of Port Vila, the capital of Vanuatu, on Efate island, around 150 kilometres south of Ambrym. The people of Maat village on Southeast Ambrym migrated to Mele Maat in 1951, following a volcanic explosion and its aftermath. For several generations, Vatlongos speakers in Mele Maat have therefore been separated from the rest of the language community and subject to different sociolinguistic pressures, more exposed to urbanisation, globalisation and national languages (especially Bislama and English) than their peers on Ambrym.

This well-documented split in the speaker community makes Vatlongos a feasible quasi-experimental setting for investigating questions about the existence and pathways of internal language divergence. While the effects of urbanisation and globalisation on language change and shift have been described in many different settings, it is rare to have such a clear comparison case: the villages of Ambrym can, to some extent, constitute a ‘control group’, being shielded from the full impact of the language-contact and sociolinguistic changes affecting the Mele Maat community.

Morphosyntactic features have been relatively neglected in work on language change, compared to phonetic, phonological, and lexical developments, or strictly morphological change. Sound changes in the lexicon and the organisation of morphological paradigms underlie the comparative method used for reconstruction in historical linguistics (Nichols 1996). Whether morphosyntactic change is more or less prone to language change and the effects of language contact is a question of some controversy in the historical literature (e.g. Dunn et al. 2007; Donohue & Musgrave 2007). The study of sound changes in progress was foundational in variationist approaches to language change, (Labov 1963 and following), and the standard tools developed to address this kind of variation have predisposed variationist

research to focus on binary variants that occur in clearly delineated contexts (Kiesling 2011). Phonetic variants of a phoneme, or variants of a morpheme that occur in a particular cell of a paradigm, are straightforwardly addressed through this framework, but many areas of variation in morphosyntax are more difficult to conceptualise in this way, especially the distributions of multi-functional constructions.

The main contribution of this thesis addressing this problem is to be found in modelling the frequency of verbal constructions by word count (Chapter 9). It is argued that this is a valid first step for identifying subtle morphosyntactic differences in language varieties, and identifying some of the sociolinguistic factors involved, especially for understudied languages where an outsider linguist is unlikely to have strong intuitions about precise contexts for variation. It is hoped that the results of this analysis can be used to isolate the functions and contexts where variation is concentrated, and be a useful starting point for research in other communities.

The project aims to address these questions through documenting and describing an endangered and understudied language. It aims to combine documentary and variationist techniques in a way that is mutually reinforcing, as advocated by Meyerhoff (2017). Language documentation aims to create a lasting, multi-purpose record of a language, that is as diverse and large as possible, while being maximally accessible and useful to the speech community, other linguists and researchers in other disciplines (Woodbury 2003; Gippert, Himmelmann & Mosel 2006). This usually involves creating a corpus of annotated audio and video recordings, and perhaps written texts, along with an apparatus of detailed metadata and description (Lüpke 2005; Himmelmann 2006). This project has therefore involved creation of a diverse corpus of transcribed and translated recordings. Observation of variation in corpus has informed the description of verbal grammar that forms the bulk of this thesis. A subcorpus of spontaneous texts has been annotated for tokens of morphosyntactic constructions and cross-referenced with sociolinguistic metadata, constituting the data for quantitative analysis of variation in Chapter 9.

This chapter introduces the Vatlongos language, its speaker communities, and previous research on the language and its speakers. Then it discusses relevant aspects of the literature on language variation and change. Finally, it describes the project's methodology, as well as community and documentation outcomes.

1.1 Vatlongos

Vatlongos is one of the (estimated) 138 vernacular languages of Vanuatu, all of which are part of the Oceanic branch of Austronesian (François et al. 2015: 2–9). With a population of around 275,000 (VNSO 2017), Vanuatu is the most linguistically diverse country in the world, measured by languages per person (Crowley 2000: 50; Malau 2011: 306; Meyerhoff 2017: 532).

Due to Vanuatu's colonial history, there are two languages of education: English and French. Before achieving independence in 1980, Vanuatu was jointly administered by British and French colonial governments as the New Hebrides, a system known as the Condominium. Despite recent attempts to unify the curriculums (Vanuatu Ministry of Education 2012), the two language-based education systems are largely separate; the regional distribution of the two languages of education reflects the missionary history of the islands.

The official language of Vanuatu is Bislama, a branch of Melanesian Pidgin English that emerged in the context of blackbirding, the primarily forced movement of labour from Melanesian islands to plantations across the Pacific (Crowley 1990). Bislama is now used in law, politics, media, and increasingly education. It is the lingua franca of Vanuatu and is spoken by the vast majority of the population. Rather than the colonial languages of English and French, competition with Bislama poses the greatest threat to Vanuatu's vernacular languages, as it is increasingly used in the key domain of family life (François 2012: 103–106; François et al. 2015: 12). Nonetheless, Vanuatu remains a highly multilingual society. Most Ni-Vanuatu speak three or more languages; 94% of the population speak at least one vernacular language in addition to Bislama, and most report higher use of vernacular than official languages (VNSO 2013).

Vatlongos is situated within this wider language ecology, and speaker communities are differently affected by national trends and contexts. This section describes existing linguistic research on Vatlongos, the speaker-communities differentiated in this project, and anthropological research in Vatlongos-speaking communities.

1.1.1 Vatlongos language

In previous literature Vatlongos is known by the toponym Southeast Ambrym. The name ‘Vatlongos’ was chosen during language development workshops for a Bible translation project (Wycliffe Bible Translators 2015). Most speakers refer to it as ‘language of the Southeast’ or simply by the Bislama term *lanwis* (‘language’, meaning vernacular languages as opposed to Bislama, English and French). However, the name ‘Vatlongos’ is becoming increasingly popular.

Most linguistic research on Vatlongos is based on data gathered by Gary J. Parker in the 1960s: his analysis in three articles on phonology, verb inflection and the morphophonemics of inalienable nouns, and a dictionary of around 1800 lexemes (Parker 1968a; 1968b; Parker 1970a; 1970b). Crowley’s (2002a) sketch grammar supplements Parker’s analysis with original material from a single speaker. Parker’s Vatlongos data has been the basis of linguistic research covering topics of regional and theoretical significance: verb-initial consonant mutation (Crowley 1991); reduplication (Inkelas & Zoll 2005: 54–57); and conditions on sound change (Lynch 2008; Blevins & Lynch 2009).

Southeast Ambrym is linguistically and culturally distinct from the rest of Ambrym island. This was remarked by Paton (1971: 119; 1979: viii) and Parker (1968a: 81; 1968b: 27), comparing Vatlongos to Lonwolwol in Northern Ambrym. Research on other Ambrym languages confirms these impressions (von Prince 2012: 2–3; Franjeh 2012: 22–23). Vatlongos is most closely related to Paamese (Parker 1968a: 81–82; Crowley 1982: 8; 1991: 183; Clark 2009: 3); speakers sometimes describe Vatlongos and Paamese as the same language, although they are mutually unintelligible (Crowley 2000: 58–59; 2004a: 7–8). Blevins & Lynch (2009) reconstruct subject-indexing prefixes and lexical items for Proto-Paamese-Vatlongos (Proto-Paamese-Southeast Ambrym).

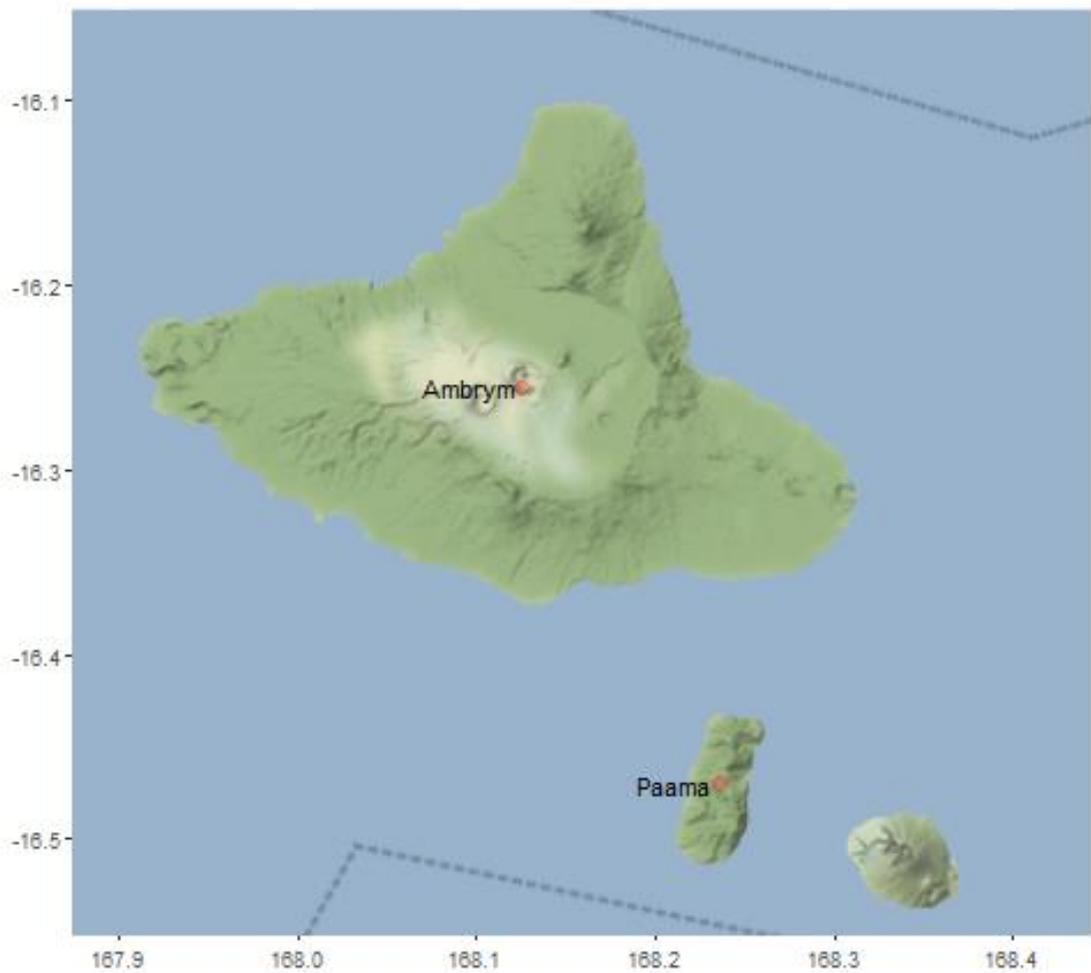


Figure 1: Map of Ambrym and Paama⁴

Clark (2009: 3–4) classifies Vatlongos genetically within Central Vanuatu, descending from Proto North and Central Vanuatu. However, the genealogical grounding of this grouping is not uncontroversial: Lynch (2001) instead suggests that Central Vanuatu languages group to the South, with South-Efate and the languages of Southern Vanuatu and New Caledonia; Ross et al. (2008: 8–10) accept Lynch’s genealogical hierarchy, but also include North-Central Vanuatu as a contact-induced linkage; while François et al. (2015: 11) reject a tree-model to allow for both these groupings. Lynch et al. (2002: 112–114) situate Vatlongos within Central Vanuatu, the Nuclear Southern Oceanic linkage,

⁴ Maps were produced using terrain map tiles from Stamen (2018), and the following R packages: ggmap (Kahle & Wickham 2013); ggplot2 (Wickham 2016); ggrepel (Slowikowski 2018).

the Southern Oceanic linkage, and Central/Eastern Oceanic, a primary subgroup of Oceanic. Vatlongos has the major typological characteristics of canonical Oceanic languages as described by Ross(2004): it has nominative-accusative alignment, SVO constituent order, and is head-marking, with subject-indexing by verbal prefixes, and object-indexing by verbal suffixes.

Parker's research has had little impact on the language community: very few speakers are aware of the dictionary and it has not led to standardisation. Church-based linguistic interventions have a greater impact. The main base for Presbyterian missionaries in the early twentieth century was Paama; Paamese was used as the language of education in mission schools and Bible-study groups for many decades (Frater 1922). The dominance of Paamese in religious life meant that many Vatlongos speakers learnt Paamese during this period, a pattern that Crowley (2000: 122) observed was increasing at the end of the twentieth century. Taveak, the southern-most village in Southeast Ambrym (Figure 4), was the site for the first mission in the region; there could have been some dialect-levelling towards Vatlongos as spoken in Taveak. However, overall, Vatlongos does not seem to have been drastically impacted by the work of missionaries.

A Bible translation was completed in 2015 (Wycliffe Bible Translators 2015), and as yet does not seem to have had any linguistic effects. For example, most speakers have a relaxed attitude to orthography and find a range of systems acceptable. It remains to be seen whether growing use of the Vatlongos Bible, and recent developments in vernacular education, will lead to greater standardisation. As one of the most widely spoken local languages in Vanuatu, Vatlongos is included in the Vanuatu Education Sector Programme (VESP), developing vernacular materials to support the national curriculum (Vanuatu Ministry of Education 2012). Vatlongos materials for the first year of primary school became available in 2017, and during fieldwork I took part in workshops translating materials for Years 2 and 3 with Ben Mansen and Simeon Ben. These texts were included in the corpus for qualitative analysis. However, it may be many years before primary schools in Southeast Ambrym can fully implement the new materials. Observations in 2017 showed patchy uptake of the Year 1

materials. Teachers and head-teachers identified a need for training in the transition from teaching English-based literacy; this is an important focus for future work supporting the community's language aims.

1.1.2 Vatlongos speaker-communities

Three communities are distinguished in this project. Firstly, the villages of Southeast Ambrym are separated from Mele Maat, a peri-urban community whose original members relocated from Ambrym to Efate island, outside the capital city Port Vila, in the 1950s.

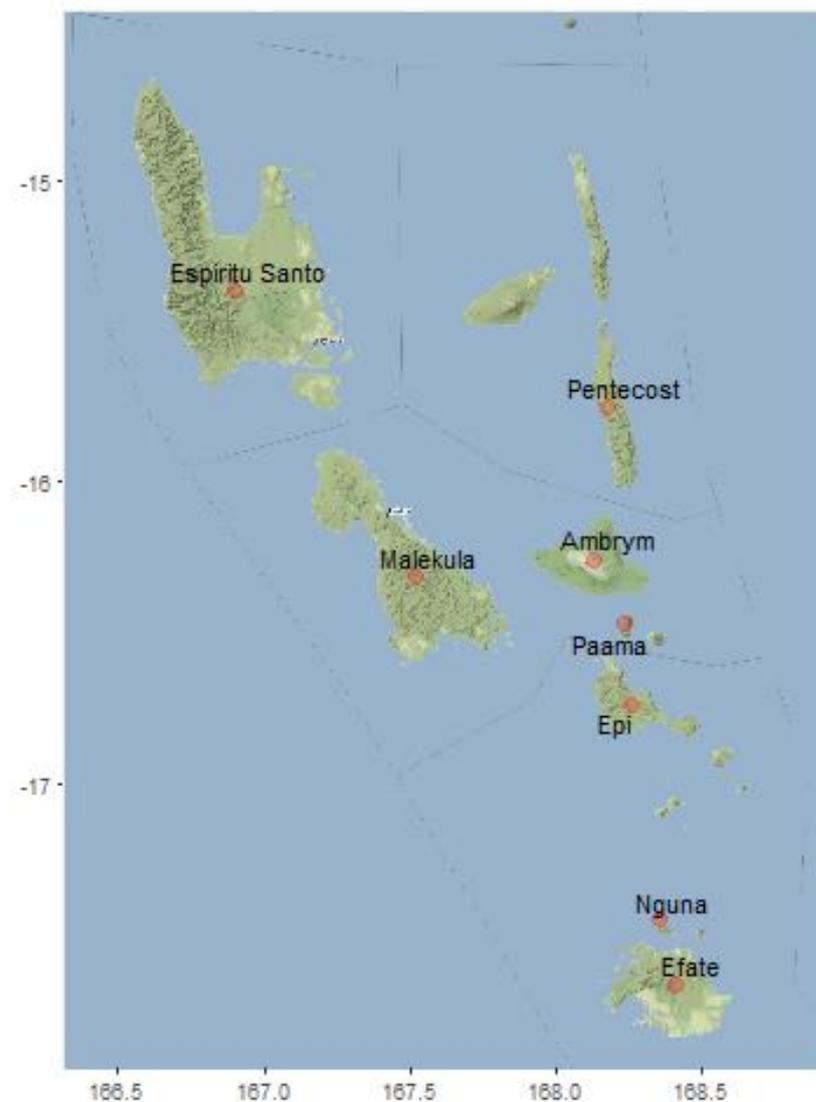


Figure 2: Map of Central Vanuatu, showing Ambrym and Efate islands

This community relocated from Maat village, Ambrym, in 1951 following a volcanic explosion. Due to months of heavy ashfall, the population of Southeast Ambrym was relocated to Epi island. At Christmas in 1950 they were struck by

a cyclone, triggering a tsunami and landfalls, killing many people. Elder Solomon from Maat village was in Port Vila and arranged for members of his village to move to Efate to work on a trader's plantation. They founded the village of Mele-Maat, combining the name of their original village on Ambrym with their closest neighbouring village on Efate, Mele (Figure 3). Mele Maat is now a large and well-established settlement, which elects its own provincial councillor for Efate (Simon 2013). Although the rights of the Mele Maat community are, in theory, well-established by various payments and custom ceremonies, there are still conflicts over land rights both within Mele Maat and with adjacent Mele village (Johansen 2012). The historical separation of Mele Maat from Southeast Ambrym, and the different sociolinguistic pressures on Mele Maat as a satellite community in the urban cash-economy, are all reasons for predicting differential language use.

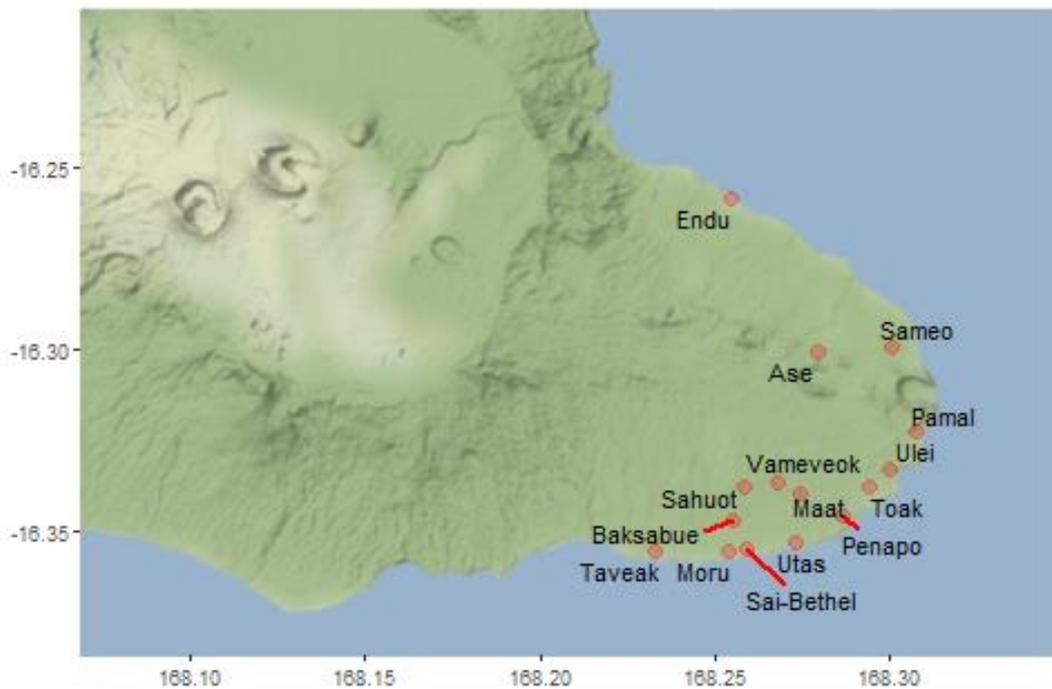


Figure 3: Map of Efate, showing Mele Maat, Mele and Port Vila

In addition, the Southeast Ambrym community has been further divided to distinguish Endu, the northernmost village, from other villages in Southeast Ambrym (Figure 4). Compared to the other villages, Endu is more closely integrated with the rest of Ambrym, with close ties to Northern Ambrym

communities through a long history of intermarriage and bilingualism in North Ambrym language. It also faces different sociolinguistic pressures, due to increasing numbers of international tourists staying in the village to visit the adjacent volcano. Tourists rarely stay elsewhere in Southeast Ambrym, except to pass through the airport in Ulei.

Figure 4: Map of villages in Southeast Ambrym



The most important reason for separating Endu from the other villages in Southeast Ambrym is reported linguistic distinctiveness. Speakers in Endu and the rest of Southeast Ambrym report that Endu has its own dialect (cf. Tonkinson 1968: 16), here called Endu-Vatlongos as opposed to South-Vatlongos. These dialects are mutually intelligible to a large extent. Only a few elderly speakers in Endu have any comprehension problems with South-Vatlongos, but South-Vatlongos speakers do report some difficulty in understanding Endu-Vatlongos. Language consultants from other communities found recordings of Endu speakers more difficult and time-consuming to transcribe, but could transcribe most Endu-Vatlongos forms faithfully and translate them into Bislama. Comprehension difficulties in one direction across dialect boundaries have been described for a marginal dialect of Apma (Central

Pentecost) (Gooskens & Schneider 2016; Schneider 2017; Schneider & Gooskens 2017).

The most salient differential features of Endu-Vatlongos are lexical. Frequent lexical variants are summarised in Table 6. There are also some phonological tendencies that are recognised as indicative of Endu-Vatlongos, but none are unique to Endu speakers, who do not consistently use these forms to the exclusion of South-Vatlongos alternatives. For many lexemes in Vatlongos, there is a choice between using forms with /o/ or forms with /a/; Endu speakers use the /o/ form more than others. In this respect, as well as many others, Endu-Vatlongos seems to be a conservative dialect (it is usually the /o/ forms that are listed in Parker’s (1970a) dictionary). Differences in the frequency of some grammatical variants are also characteristic of Endu-Vatlongos; this thesis discusses differences in the use of the negative clitic (§3.6.2.10), variants of subject-indexing prefixes (§5.1.1.2), and patterns of verb-initial consonant mutation (§5.2.1), as well as differences in the frequencies of verbal constructions (Chapter 9).

Table 6: Lexical variants in Endu dialect

Gloss	South-Vatlongos form	Endu-Vatlongos form
3SG.NFUT.finish ⁵	<i>bus</i>	<i>buis</i>
2SG	<i>xouk</i>	<i>exo</i>
again	<i>mun</i>	<i>men</i>
1PC.INCL	<i>ratel</i>	<i>ratil</i>
one, such	<i>tava</i>	<i>top</i>
very	<i>tavatang</i>	<i>toptang</i>
on and on	<i>maa</i>	<i>ve</i>
PROX	<i>ak</i>	<i>xal</i>
MED	<i>xai</i>	<i>xei</i>
1PC.EXCL	<i>xametel</i>	<i>xamtel</i>
REL, SUB	<i>xa</i>	<i>xe</i>

The other villages of Southeast Ambrym are treated as a single community, which I refer to as Ase-Taveak, named after the first and last villages travelling from North to South (Figure 4). However, this masks some dialectal differentiation within South-Vatlongos. Tonkinson (1968: 16) observes village-

⁵ This form frequently occurs as a marker of completive aspect (§4.3.1).

level dialectal differences recognised by speakers without hindering communication; as well as Endu, he mentions Toak and Sameo as having the most noticeable dialectal differences from the others. Speakers in Sameo and Ase seem to have some of the features associated with Endu-Vatlongos, including patterns of verb-initial consonant mutation (§5.2.1). Some elderly speakers in Toak also have a distinctive pattern of verb-initial consonant mutation, using /r/ in non-future affirmative environments, which is otherwise the most stable context for this morphological process. Toak also seems to have distinctive prosodic features, which speakers describe as a different ‘tune’. A favourite example is *tavatang* ‘very’, where stress unexpectedly appears on the second syllable, perhaps due to a lexically-determined long vowel. Prosodic micro-variation would be a fruitful area for future research.

Finally, Vatlongos is spoken in other parts of Port Vila and the other major urban centre in Vanuatu, Luganville on Santo island. None of the speakers included in the corpus primarily identified with these urban communities, but some had lived in these areas. It would also be interesting to investigate any differences in the Vatlongos spoken in these communities, similarities with Mele Maat, and whether speakers returning from long-term stays in these areas have any distinctive speech patterns. Meyerhoff & Walker (2007) find frequency differences in the use of grammatical variants by speakers they term ‘urban sojourners’.

Crowley (2000: 70) estimated the speaker population of Vatlongos at 3700. This was based on extrapolation from the regional population of areas known to speak the language in the 1989 census, increased to reflect overall population growth. Although a detailed survey of all speakers of the language is beyond the scope of this project, I am inclined to estimate a lower figure for several reasons. Firstly, population growth is variable across Vanuatu. While Crowley (2002a: 660) put the population of villages in Southeast Ambrym at around 2000, the 2016 mini-census found a figure of 1,549 (VNSO 2017), showing either a decline over 17 years, or that Crowley’s estimate was too high. Crowley (2002a: 660) also states that 40% of speakers live in Mele Maat, on the basis of Tonkinson’s (1979: 112) estimate in the late 1970s. Based on a total speaker

population of 3700, this would put the population of Mele Maat at 1480, although recent anthropological work estimates 1000 (Johansen 2012: 4). In any case, not all residents of Mele Maat speak Vatlongos: those who have married in often do not learn it, and increasingly neither do children. It is even harder to estimate the number of speakers living in other parts of Vila and Vanuatu. Taking all these data into account, I propose that 3,700 is probably the upper bound for possible speaker numbers, which could be as low as 2,500.

The current sociolinguistic contexts of each community are discussed in Chapter 2.

1.1.3 Anthropological research in Vatlongos communities

A wealth of details are available about the original relocation and the progress of the Mele Maat community since the late 1960s, as well as social conditions in Southeast Ambrym, thanks to the anthropologist Tonkinson (1968; 1979; 1981; 1982; 1985; 2011). Much of his research has been influential in the anthropology of the region, especially in describing the role of *kastom* (Bislama 'custom' or 'culture') in Melanesia (Lindstrom 2008; Acciaioli 2013).

An important point for this project is that the separation of Mele Maat from the island communities is not absolute. Tonkinson (1968: 271–273) explains how links between Mele Maat and Southeast Ambrym were maintained by letters and visits in both directions in the early years of resettlement. The Mele Maat settlers retained their land rights in Maat, Ambrym, and the Southeast Ambrym community used Mele Maat as a base in Port Vila to visit for work, healthcare, or to escape threats of sorcery on Ambrym (Tonkinson 1981: 83). This was still the case in 1978, when Tonkinson (1979: 110–116) noted that 30% of residents in Mele Maat were non-Maat Ambrymese. However, he also observed that Mele Maat hosted fewer guests than in 1967, as casual visitors to Efate now had a wider range of options to stay with other Southeast Ambrym community members who settled in Port Vila. A higher proportion of young men were marrying outside the Vatlongos-speaking community, a more significant sign of a break with the Southeast Ambrym community. Today Mele Maat is still a base for visitors to Port Vila from Southeast Ambrym.

Tonkinson (1968: 173; 1979; 1985) stresses continuity in Mele Maat's 'Ambrymese' identity. The maintenance of Vatlongos, when communication external to the village is conducted in Bislama or English, is evidence of this continued link. Cultural identification with the 'homeland' seems to have been particularly strong in the first few years after relocation, when the arrangement was still thought a temporary one, and may have waned in the following years. In 1967, Tonkinson (1968: 272–273) observed a 'weakening of community spirit'. By 1978, the community seemed to have embraced an urban lifestyle: 85% of men and 64% of women worked in Port Vila, rather than in subsistence gardening or plantation labour, the kinds of employment available on Ambrym (Tonkinson 1979: 111). Consequently, there has been increasing interaction with outsiders.

While Tonkinson's early research was concerned with the success or failure of the relocation venture, more recent anthropological work would see the mobility of Vatlongos speakers within the framework of globalisation (Lindstrom 2008: 162). Mele Maat faces Port Vila's urban sprawl and a huge rise in tourism in the capital. The Vanuatu National Statistics Office (VNSO) reports an increase in the total number of tourists per year from 82,019 in 1995 to 328,861 in 2014 (VNSO 2018). Moreover, the 'Mele Cascades', a waterfall near Mele-Maat, has become a tourist attraction (Tourism Vanuatu 2017). Mele Maat is in a very different situation from 1967, when few tourists visited Vila, let alone Mele Maat (Tonkinson 1968: 131). In the mixed economy, Mele Maat families must maintain 'gardens' for traditional foods, and earn money to purchase goods and services of the urban economy. Mele Maat residents discuss these challenges in recent anthropological research (Johansen 2012: 31–34).

The wealth of anthropological research focusing on the Vatlongos-speaking community, and especially Mele Maat, is very useful to addressing the central questions of this project. It is rare to be able to date a division in a linguistic community so precisely, or to have access to such detailed information about the circumstances during and following a relocation.

1.2 Language change and variation

This section situates the project within wider frameworks for understanding language change and variation. First, it discusses purported mechanisms in relation to structural and syntactic change. In an endangered-language context the impact of language contact and attrition is also relevant, especially for Mele Maat. Finally, §1.2.3 summarises findings of previous research on language change and variation in a Pacific context, and motivates the morphosyntactic topics that are the focus of this thesis.

1.2.1 Mechanisms of language change

Language change is central to many linguistic disciplines, especially historical linguistics, studies of language contact, and variationist sociolinguistics. Mechanisms of language change have been described in different ways that reflect the aims, objects of study, and assumptions of these subdisciplines. Comparative historical linguistics takes the broadest linguistic and temporal perspective: it traditionally pertains to entire languages as complex and separable systems, and is interested in genetic relationships between individual languages (Harrison 2003: 214). The ‘tree model’ assumes that languages diverge when speakers are socially and physically separated from each other (François 2014: 162–6). Innovations on one side of the split are not shared across the divide, and eventually the accumulation of separate innovations produces mutually incomprehensible varieties: separate languages which share a parent. The recent and well-documented migration of the Mele Maat community away from the rest of the Vatlongos-speaking community is therefore an opportunity to examine early stages of possible language divergence. However, as François (2014: 165) points out, total splits in language communities occur very rarely. Even in fairly extreme situations of migration after a natural disaster, like Mele Maat’s, communication with other speaker communities is usually maintained to some degree (§1.1.3, §2.1).

Specifically excluded from the comparative method of historical linguistics is language change resulting from contact between languages. Following Thomason & Kaufman (1988), much research in language contact has focused on the linguistic effects of bilingualism (Muysken 2013), distinguishing between

borrowing – the influences of a speaker’s second language on their first language – and interference through shift – the effects of a speaker’s first language on their second. Language change in contact situations is thus figured as resulting from the interaction of grammars as psychological linguistic systems, which face cognitive pressures towards convergence when individual speakers speak multiple languages (Silva-Corvalán 1994; Ross 2001: 149; Enfield 2003: 16–17; Heine & Kuteva 2005: 264). Aside from convergence, these cognitive pressures can also lead to simplification in language contact situations.

While early work on this topic assumes a model of sequential language learning, with a starting point of native monolingualism, there is increasing recognition of more complex multilingual situations where different languages are acquired simultaneously and/or specialised to particular domains (Croft 2000: 90–94). Research from this perspective is more relevant to the multilingual environment of Vatlongos speakers, and other communities in Vanuatu and the Pacific.

The closest perspective on language change in progress is provided by variationist sociolinguistics. This framework focuses on the distribution of linguistic variables, loosely defined as different ways of ‘saying the same thing’, or linguistic choices available to a speaker (Labov 1963; Kiesling 2011: 13–25). Early variationist research followed Labov in trying to identify the social categories that correlate with linguistic variables, which usually meant those relevant in an industrialised Western society, including social class, age and gender. Subsequent research has taken a more ethnographic approach by looking at the linguistic behaviour of individuals within closely-observed social networks (e.g. Milroy & Milroy 1985; Eckert 1989). A variant can spread quickly when it is associated with overt or covert prestige, or becomes emblematic of identity with a particular social group: this latter function has been suggested to be an especially important factor in the diversification of Melanesian languages (Laycock 1982; Crowley 1992a: 198–200; François 2011: 50–52; François 2012: 92). When an innovative variant is propagated throughout the speech community, it can become dominant and eventually replace an older variant.

This is assumed to be the basic mechanism underlying historical change (François 2014: 168), for example in models of language change based on evolution (Croft 2000) or epidemiology (Enfield 2003). Considering evolutionary frameworks from the perspective of typology, Evans (2016) discusses various obstacles to testing this ‘macro-from-micro hypothesis’, including the difficulty in bridging the gap between the methods, models, and areas of interest of typologists and variationists.

An important insight of these models is that change progresses at the level of utterances: linguistic forms used in context. Changes to grammar, understood as an individual’s cognitive model of language, arise from their exposure to linguistic utterances. The divergent individual languages that are the object of historical linguistics are a further abstraction over the grammars of numerous individual speakers distributed in space and time. For this reason, the basic methodology of this project is corpus analysis. The best place to look for very early signs of language divergence resulting from the separation of the Mele Maat and Ase-Taveak communities is in the distribution of variables in collections of linguistic utterances, which can be both a cause and an effect of changes in the cognitive grammars of individual speakers.

Structural features, especially morphosyntactic ones, have a controversial role in theories of language change (Muysken 2013: 720), and particularly in historical linguistics. Some work in historical linguistics has suggested that structural features change more slowly than the lexicon, and are only transferred between languages in situations of extreme contact, and therefore could be used to establish relationships between languages at a greater time depth than the comparative method. For example, Nichols (1995) identifies a number of structural features as diachronically stable. On this basis, Dunn et al. (2005; 2007; 2008) use structural features of Oceanic and Papuan languages in Island Melanesia to reconstruct their relationships to a greater time depth than has been achieved with lexicon-based methods. Silva-Corvalán (2008: 214–216) argues that abstract syntactic features are never transferred between languages, although many of the features she excludes might be considered syntactic, or at least structural, in other frameworks.

Others have argued that structural features are prone to convergence in contact situations (Thomason & Kaufman 1988: 54–55; Ross 1996; François 2011; 2012), so are either not useful for deducing language history, or can only reflect areal rather than genetic relationships (Donohue & Musgrave 2007). These conflicting views on the rate and circumstances of structural change perhaps reflect differences in the rate of change of different structural features: Greenhill et al. (2017) aim to identify the rate at which different structural features change within Island Melanesia, and find that while some structural features are very stable over time, others change rapidly.

These debates about structural and morphosyntactic change are usually from the viewpoint of historical linguistics, aimed at finding out how useful structural features are for determining historical relationships between languages. Most of the literature on structural change therefore takes individual languages as a unit of comparison, and requires typological features to be assigned to a language as a whole. Less work has been done on morphosyntactic change at the level of intra-language variation. This is partly because morphosyntactic features can be difficult to construe as sociolinguistic variables (Kiesling 2011: 15; Meyerhoff 2014a: 99). Variant forms of a morpheme can fairly straightforwardly be modelled as binary choices relating to a cell in a paradigm, as in §5.1.1, although this is also complicated by differing patterns of syncretism. However, it can be very difficult to precisely delineate a context for a morphosyntactic construction, like the multifunctional serial verb constructions (SVCs) described in Chapter 6. Much variationist work on morphosyntactic variables has therefore been restricted to those morphosyntactic phenomena that are most neatly constrained to an easily formulated semantic context, ideally with a binary distinction between two syntactic variants. Studies on English syntactic variation have focused on binary variables that are either lexically restricted, like the dative alternation with ten frequent three-place predicates (Gries 2003a) and particle placement with transitive phrasal verbs (Gries 2003b), or semantically restricted, like the simple past and present perfect as alternative means to express past situations (Hundt & Smith 2009).

While this model is adequate for some morphosyntactic phenomena, speakers also face more complex linguistic choices, and the resulting variation should also be studied. Kiesling (2011: 23) points out that the VARBRUL-based statistical tools that quickly became dominant in variationist sociolinguistics have sometimes led researchers to reduce more complex patterns to binary variables.

A more serious problem is the use of semantic or pragmatic meanings to delineate a context for a morphosyntactic variable, when the choice to express a certain semantic or pragmatic function can itself be a locus of variation (Enfield 2003: 18; Meyerhoff 2014a: 99). For example, Crowley (2002b: 67) observes that there is a strong discourse preference in Paamese to express direction of motion by serialising a basic motion verb whenever motion from one place to another is expressed. Restricting a study to contexts where direction of motion is expressed, for example by comparing the use of verb serialisation and prepositions, might reveal interesting patterns of variation, but would not allow for a comparison of the discourse preference for expressing a category in different language varieties. For example, we may want to build on Crowley's observation by finding out whether Paamese expresses direction of motion more often than Vatlongos, if speakers of Vatlongos and Paamese express this category more frequently than speakers of other languages, or if speakers who are more exposed to English are less likely to express this category.

In order to avoid losing information about variation of this kind, Chapter 9 explores the overall frequency of tokens of verbal constructions in the corpus by word count, in relation to community membership and other sociolinguistic factors. This can be thought of as summing over a great number of individual linguistic choices that result in the overall frequency of a construction. The investigation in Chapter 9 could form the basis for more fine-grained future research, isolating particular functional contexts for variation, but it does not make *a priori* predictions about what those contexts might be. This approach is especially appropriate for under-described languages, as the process of tagging the corpus for quantitative analysis has proceeded in tandem with the qualitative descriptive analysis in Chapters 3 to 8. These descriptions have

greatly benefited from the ability to run concordances on constructions, and their interaction with word-level and morpheme-level phenomena. Quantitative analysis of the overall frequency of constructions thus capitalises on data gathered as part of the descriptive analytic process.

One consequence of the dominance of historical linguistics as a framework for investigating structural change in language, is a tendency to try to deduce sociolinguistic contexts in the past from their linguistic effects in the present (Ross 2001: 160–161; Ross 2003; Toulmin 2009; Muysken 2013: 723–725). The aim of this project is to do the opposite. We have identified a contemporary sociolinguistic context that exhibits many features related to language change and divergence in the literature: separation of a community, intense multilingualism, and exposure to urbanisation and globalisation. The ability to compare data from two communities that are more and less exposed to these factors is a unique advantage of the situation of Vatlongos speakers. The project aims to find out what the effects of these factors are on Vatlongos as spoken in Mele Maat, long before the emergence of wholesale language divergence and restructuring that is usually when language change becomes of interest to historical linguists.

1.2.2 Language shift and attrition

Apart from normal processes of language change in response to separation and language contact, Vatlongos in Mele Maat may also be vulnerable to processes associated with language shift (Fishman 1991; Pauwels 2016) and attrition (Seliger & Vago 1991; Schmid 2011) in situations of language-endangerment. European colonialism has been a major cause of accelerated global language loss (Mühlhäusler 1996: 17–21; Crystal 2000; Mufwene 2002; Simons & Lewis 2013: 13–16), and certainly Vanuatu’s colonial history has determined the contact languages that are most threatening to the future of Vatlongos. However, Vatlongos in Mele Maat is now impacted by urbanisation, which has been identified as a key and growing threat to endangered languages in global surveys (Simons & Lewis 2013: 16; Lewis & Simons 2016: sec. 2.1). Aside from the existential threat posed by economic incentives for language shift, linguistic

effects of language contact also seem to be accelerated in urban contexts (Heine & Kuteva 2005: 28).

Language attrition often manifests as linguistic simplification, one of the most identifiable forms of which is phonological and morphological erosion (Schmidt 1985: 231; Heine & Kuteva 2005: 91). At the level of morphology, attrition has been associated with tendencies for less allomorphy, synthesis, fusion and grammatical concordance, as well as a structural shift from morphological to syntactic expression of categories (Maher 1991: 68; 1996: 286–289). For Silva-Corvalán (1991: 152), simplification implies a form being used in a larger number of contexts, as documented in contexts of language attrition in Australia (Donaldson 1985: 137; Schmidt 1985: 46–76). This shift can be observed in the frequencies of competing forms until one replaces the other, a process that appears to be in progress in Vatlongos subject-indexing prefixes (§5.1.1.2). Syntactic simplification can be evident in a tendency to avoid syntactically complex constructions (Schmidt 1985: 121; 1991: 119; Maher 1991: 68; Mühlhäusler 1996: 291). This is one reason why this thesis focuses on syntactic constructions involving more than one verb (Chapters 6 to 8).

A very different factor affecting language change in endangered language contexts is ideological linguistic purism (Hill & Hill 1980). Dorian (1994) argues that conservative attitudes in speaker communities can be an obstacle to language maintenance and revitalisation of endangered languages. Florey (2004) describes how the puristic urge to maintain a conservative variety of a threatened language can accelerate language loss in a ‘language shift cycle’. Much of the literature on this topic focuses on documenting attitudes in speaker communities and assessing them as an obstacle (or not) for language maintenance (Wertheim 2003; Cru 2016; Albury & Carter 2018). Comments on the linguistic effects of these attitudes have observed that lexical, rather than grammatical, features are most salient for language purists (Kroskrity 2010: 200). The results of the sociolinguistic survey described in Chapter 2 show that conservative attitudes are most common in Endu; some possible linguistic consequences are observed in Chapters 5 and 9.

1.2.3 Language change and variation in the Pacific

Research on language change and variation in the Pacific reveals that some aspects of the models of language change discussed above do not fit neatly with patterns of language use in this regional context. Models of language change in contact situations and language attrition often assume that bilingualism arises through sequential language learning: that speakers learn a single language as their L1 and subsequently learn an L2. However, the traditional language ecology in many Pacific contexts involves widespread multilingualism, involving a mix of languages associated with very small social groupings and regional lingua francas, with more than one language learned from childhood (Mühlhäusler 1996: 33–47; Ross 1996: 181; François 2012). Thurston (1992) distinguishes between esoteric languages – those that function as markers of identity for small groups at the village level – and exoteric languages – those that function as lingua francas for a larger region. He argues that esoteric languages tend towards greater differentiation, while exoteric languages face pressures towards simplification. Thus, rather than the cognitive pressures of bilingualism in adults, it is the role played by an individual language in a multilingual context that determines the direction and intensity of language change in language contact in the Pacific.

One tendency that arises from this pattern of multilingualism is the often-observed opposing trajectories of formal and structural change. Because esoteric languages play an important role in identifying members of small groups, there is a propensity for small-scale differentiation that parallels other cultural practices in this region (Laycock 1982; François 2011: 49; 2012: 92). However, lexical and phonological features of language that are more salient to speakers are more valued as a symbol of identity (Schmidt 1991: 122), and thus more likely diverge. Structural features, on the other hand, are likely to converge in these patterns of language contact: Ross (1996; 2001; 2003; 2007) developed the concept of ‘metatypy’ to describe the large-scale structural syntactic and semantic borrowing between languages that often occurs in these intensely multilingual environments. This results in groups of neighbouring languages with very different lexicons but similar grammars (Crowley 2004a: 6; François 2009; 2012; 2014: 174–182). Given the new contact situation of the

Mele Maat community, structural developments are a good place to look for early signs of language change, but it is important to examine structural features that show variation in the region given the broad grammatical similarities between Oceanic languages.

Another idiosyncrasy of many Melanesian contexts today, relevant to Vatlongos and processes of language change, shift and loss, is the role of Melanesian Pidgin English. Unlike many other colonial contexts, the most important contact language for vernaculars in modern Vanuatu is not the colonial language itself, but a contact variety whose grammatical structure is heavily influenced by Oceanic languages (Walsh 1984; Crowley 1990). In some ways, this might shield vernacular languages in Melanesia from convergence with 'Standard Average European' (Mühlhäusler 1996: 283). Aside from the historical origins of the contact variety's structure, there is evidence that structural features of an individual vernacular language can be transferred into contemporary Bislama (Meyerhoff 2009). Regional variation in Bislama also reflects features of different vernacular languages (Meyerhoff 2000: 39; Crowley 2004b: 8), so convergence is likely to involve speakers recreating patterns of their vernacular language in Bislama, rather than, or as well as, vice versa.

On the other hand, the ease with which speakers of Oceanic languages acquire Melanesian Pidgin English makes it more disruptive of pre-colonial linguistic practices, replacing local lingua francas and removing some of the communicative incentives to learn the vernaculars of other communities (Mühlhäusler 1996: 75–138). Similarities between the grammatical systems of Oceanic languages and Melanesian Pidgin English do not mean there is no room for convergence in language contact, and in fact might make structural convergence more likely. Jenkins (2005) describes convergence of vernacular languages towards Tok Pisin, the Papua New Guinea variety of Melanesian Pidgin English. Heine & Kuteva (2005: 263) find examples supporting the observation that structural borrowing is most prominent between structurally similar languages.

Observations of variation in Vanuatu languages are mostly found in the introductions to descriptive grammars and dictionaries, and are aimed at delineating regional dialects.⁶ Crowley (2004a) reflects on the issue of dialect boundaries in Oceanic languages, though mostly from a metalinguistic perspective, rather than examining intralectal evidence, while the Suru Kavian dialect of Apma (Central Pentecost) has been the subject of more focused work on intra-dialectal intelligibility (Gooskens & Schneider 2016; Schneider 2017). Introductions to descriptive works often also include researcher impressions of pressure from Bislama.⁷ Quantitative approaches have been used to address the frequency and distribution of loan words from Bislama in some vernacular languages (Crowley 2004c; Early 2004; Budd 2011; Meyerhoff 2014b). Meyerhoff's (2000; 2002) work on Bislama in a variationist framework has found that discourse factors are more important than speaker-level social factors for the distribution of certain morphosyntactic variables, although gender is important in the use of some conversational strategies, including apologies (Meyerhoff 1999; 2003). However, Bislama's status as a contact variety means it is not directly comparable with vernacular languages, so Meyerhoff's work on variation in Tamambo (Meyerhoff 2009) and Nkep (East Santo) (Meyerhoff 2015; 2017) is a better parallel for this project's methods in addressing variation in Vatlongos.

Given the major structural similarities between Oceanic languages (Ross 2004), this project focuses on areas of linguistic structure where language-internal variation is more likely to be found. One way of identifying such areas is to focus on domains of grammar where extensive inter-language variation has been attested in descriptive work on Vanuatu languages. Another is to look at grammatical topics that prior research on language-internal variation in Vanuatu has addressed. Finally, from a historical linguistic perspective, Greenhill et al. (2017) have tried to identify structural features that change at different rates in the linguistic history of Island Melanesia. Many of the

⁶ e.g. in Paamese (Crowley 1982: 8–10; 1992b: x–xvi), Apma (Central Pentecost) (Schneider 2010: 7–11), and Daakaka (West Ambrym) (von Prince 2015: 9–10).

⁷ e.g. in Tamambo (Malo) (Jauncey 2011: 7–9) and Daakaka (West Ambrym) (von Prince 2015: 10–12).

parameters they find to change at a ‘fast’ or ‘medium’ pace are relevant to topics in this project. On the basis of these kinds of evidence, the focus of this project is restricted to the grammar of verbs in Vatlongos, as verbal grammar seems to show greater degrees of morphosyntactic variation in Oceanic languages.

Chapter 3 describes the syntax of simple clauses in Vatlongos, and is an important foundation for understanding the morphosyntactic features described in later chapters. §3.6 looks closely at the behaviour of the partitive and negative clitics, as Vanuatu languages vary in their expression of negation, using simple, double and sometimes triple marking to express different kinds of negation (Early 1994: 395–422; Budd 2010; Vossen & van der Auwera 2014; Barbour 2015). Negation has long been acknowledged as a linguistic domain that is prone to characteristic patterns of change described as ‘Jespersen cycles’ (Jespersen 1917). Clause-initial negation changes at a fast rate in Greenhill et al.’s model, and morphological marking of negation at a medium rate.

Chapter 4 describes the extension of Tense-Aspect-Mood (TAM) categories that are marked morphologically in Vatlongos into different contexts, as well as some periphrastic strategies for marking additional categories. TAM categories are described differently in grammars of Vanuatu languages. Even probable cognate forms in related languages are described as marking different categories. For example, forms based on /te/ are described as non-recent past tense in North Ambrym (Franjeh 2012: 118–120), perfective aspect in Apma (Schneider 2010: 174–178), and ‘distal mood’ in Daakaka (von Prince 2015: 250–255). Although some descriptive differences arise from linguists’ analytical choices, these very different analyses show the scope for reanalysis of TAM markers by speakers of languages as well. The MelaTAMP project (Krifka et al. 2018) aims to explore the extension of TAM and polarity (TAMP) categories in several languages of Vanuatu through targeted storyboard stimuli.

Chapter 5 focuses on verbal morphology in Vatlongos. Subject-indexing, TAMP, and transitivity or object-indexing markers in Vanuatu languages are variously described as affixes, clitics and particles (Lynch, Ross & Crowley 2002: 45; Ross 2004: 495; Ridge 2017c). Pawley (2003) argues that the Oceanic verb complex is subject to centripetal and centrifugal forces of grammatical

change, with grammatical elements either becoming more tightly grammatically and phonologically bound to the verb root (resulting in affixation), or becoming more independent (resulting in particles). This is clearly an area of intra-language and inter-language variation and change in Vanuatu languages, and some changes appear to be in progress in the Vatlongos subject-indexing prefixes (§5.1.1.2). Many features associated with verbal morphology change at a medium or fast rate in Greenhill et al.'s model, including the marking of the subjects of transitive verbs, the existence of transitivity morphology, and variation in marking of arguments in different TAM environments.

Verb-initial consonant mutation is another morphological feature that varies between Vanuatu languages (Crowley 1991). §5.2.1 shows that this appears to be undergoing change in Vatlongos, in a way that may eventually lead to the loss of some variants, especially with less frequent verb roots. Synchronically, this involves a proliferation of variants in single cells of the paradigm. As this process is restricted to certain verb lexemes, it is an area of the project that relied on elicitation more than corpus analysis.

Chapters 6 to 8 focus on complex verbal constructions, especially serial verb constructions (SVCs) and related forms. SVCs show a great deal of variation between Oceanic languages (Crowley 2002b; Brill & Ozanne-Rivierre 2004), and seem to be prone to diachronic change through language contact, grammaticalization, fossilization, and loss (Crowley 2002b: 169–214; Aikhenvald 2006). The existence of SVCs changes at a medium rate in Greenhill et al.'s model. The only more fine-grained feature was the use of SVCs for causatives, a fast-moving parameter. The constructions I am describing as 'complex predicates' in §7.1 are not fully productive, and diachronically probably derive from a more closely bound verb-serialisation strategy than the one described in Chapter 6. They could be analysed as a form of verb compounding, and Greenhill et al. find that a regular process for verb compounding is a fast-changing feature. In Chapter 8, I argue that Vatlongos auxiliary verb constructions (AVCs) have emerged from SVCs, and are still extending into the semantic domain of prior motion. Greenhill et al. found that aspectual and modal auxiliaries – which are found in Vatlongos – are fast

changing features, whereas tense auxiliaries change slowly. Variation in the frequency of these constructions in the corpus is discussed in Chapter 9.

1.3 Methodology

In order to address the research questions while furthering the goals of the Vatlongos speaking community, the most important component of this project's methodology has been qualitative and quantitative corpus analysis (Lüpke 2005). This section describes the two periods of fieldwork and the workflow for building the corpus. Corpus analysis is briefly discussed here, but the rationale for the choice of models and tools for quantitative analysis is described in more detail in Chapter 9. I then summarise the outcomes of the project in terms of the community's language goals, and give an overview of the results of the project as a contribution to language documentation.

1.3.1 Fieldwork

Fieldwork for this project consisted of two trips of six months, from October to April 2014-2015 and 2016-2017. A shorter trip in July 2018 to attend the Vanuatu Languages Conference in Port Vila allowed for additional follow-up and dissemination of the results of the project. Permission to conduct the project was obtained at a national level from the Vanuatu Cultural Centre. Chief Albea gave permission for the project to proceed in Mele Maat. As far as possible, permission was also obtained from chiefs of individual villages in Southeast Ambrym, but logistically this was not always possible. Overall, the project was very well received and enthusiastically supported by the speaker community, especially when distribution of community outcomes commenced during the second field trip.

Researcher time was split between Port Vila and Ambrym. More time in total was spent in Port Vila, both to allow easier data processing and analysis with reliable access to power, and because recording proceeded more slowly in Mele Maat, due to difficulty in scheduling recording sessions around urban working hours. On Ambrym, I stayed with Elder Saksak Ruben and his family in Endu, and with Elder Simeon and Madleen Ben and family in Silimaori, a small hamlet attached to Moru village (Figure 4). These three also acted as language assistants on Ambrym, arranging recording sessions and video screenings,

taking part in elicitation sessions, conducting metadata questionnaires, and distributing files. In addition, Simeon and Madleen did transcription for the project, a task they were well prepared for by their previous experience supporting the Bible translation project. In Mele Maat, Chief Albea nominated Ben (Bell) Mansen for this work, and he performed all these tasks there.

There were various obstacles and logistical issues that delayed the progress of the project, especially during the first field trip. Illness meant that the first trip to Ambrym (October 2014) was restricted to Endu, and interrupted all work for most of January 2015. In March 2015 Vanuatu was struck by Cyclone Pam, affecting all Vatlongos-speaking communities. Researcher time was therefore diverted to recovery efforts for the last six weeks of the field trip, although we also recorded accounts of this historically important event, documenting the damage caused by the Category 5 cyclone. The second trip was considerably more efficient, but due to administrative delays I was not able to visit Ambrym in 2016.

Nevertheless, we visited every Vatlongos-speaking village, both to make recordings, and to distribute results of the research through video screenings, sharing videos on SD-cards and USB sticks, and sharing eBooks and an eBook-reader app. The project involved a total of 368 recording events with 172 participants (including corporate participants, e.g. football teams, Table 1). 215 recordings have time-aligned transcriptions and translations into Bislama. Of these, 122 sessions have been interlinearised, cross-referenced to a lexical database, and translated into English. The success of the project was largely thanks to the enthusiastic reception of the speaker communities, and the dedication, expertise and patience of Bell, Saksak, Simeon and Madleen.

1.3.2 Data gathering and processing

The large proportion of recording events that were transcribed and processed was also due to an efficient workflow, using a set of freely available software tools, many of which are produced and supported by the Summer Institute of Linguistics (SIL International). This process is described in detail here, divided into stages of data gathering, file processing and linguistic processing. The workflow was designed to support the construction of a corpus in a format that

supported the analytical goals of this project (§1.3.3), to allow for easy, cheap and rapid distribution of outcomes for the community (§1.3.4), and with a view to data longevity for the purposes of documentation and archiving (§1.3.5). The different stages of the workflow and how they inform each other are schematised in Figure 5.

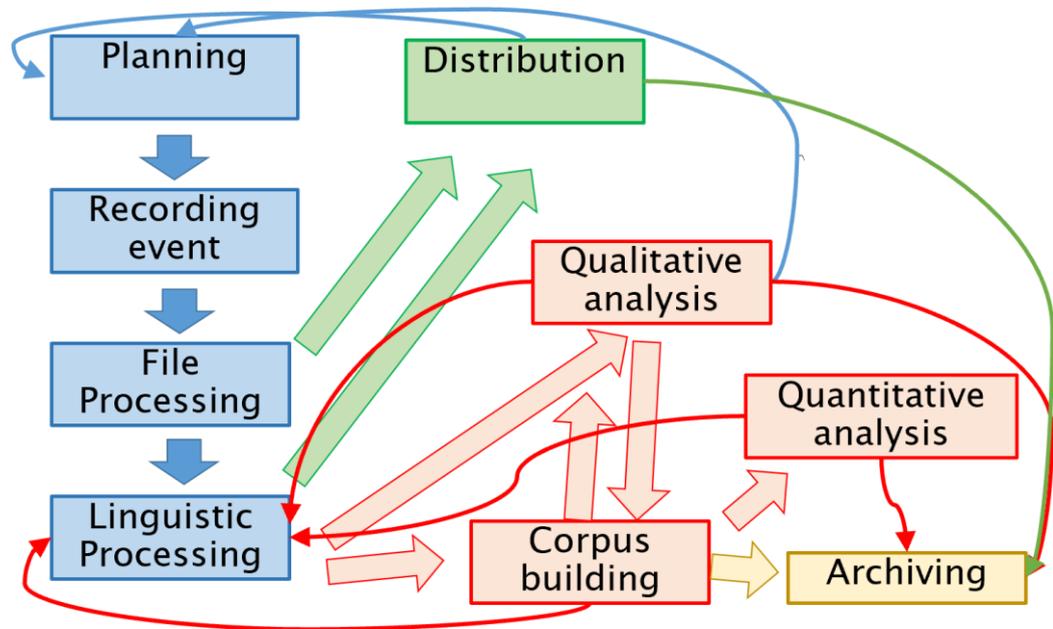


Figure 5: Workflow for documentation of Vatlongos

The first stage in data gathering was planning recording sessions. This involved recruiting individual volunteers through word of mouth, announcements at Church services, and liaising with chiefs and women’s groups in different villages. Increasingly, this also involved responding to requests from community members to record individual’s stories and community events, especially in the second period of fieldwork when distribution of videos from the first trip began in earnest. More carefully controlled language-focused recordings, using elicitation tools and stimulus materials, were handled differently, targeting different communities, age groups and genders. In line with advice from the Vanuatu Cultural Centre and other linguists working in Vanuatu, speakers and language assistants were paid for these types of activities, as well as transcription and conducting metadata interviews. Community-led recording events, where speakers themselves chose the content of the recordings, were not paid, but audio and video files were shared with the

speakers wherever possible and desired. This distinction was understood and respected by speakers, who responded with enthusiasm to the chance to record stories and events. Planning also involved discussions with individual speakers about the aims of the project and how their data could be used, and responding to participants' questions.

During the recording event, audio (in wav format) and video were recorded separately. A range of microphones were used for different recording settings: stereo wherever possible, but either shotgun or lapel microphones, depending on speaker preference, in more noisy environments (especially to avoid too much traffic noise in Mele Maat recordings). Lapel microphones were also preferred for interviews and other two-person recording events, and for situations where the speaker needed to move around, in which case the recorder was placed in a basket hanging around the speaker's neck. Each recording session began with a brief audio recording of consent, which also functioned as a sound test. The speaker gave their name and village, and consented to take part in the project and be recorded and/or videoed, all in Bislama for greater transparency (a handful of elderly speakers preferred to do this in Vatlongos). The recording then began after a signal clap for syncing audio with video.

After the recording, a metadata interview was conducted, including sections on detailed consent for distribution of the data,⁸ personal history and language repertoire, use and attitudes (Chapter 2). Again, this was conducted in Bislama for all but a few elderly speakers, for greater transparency and for ease of coding. Bislama is also a language associated with official processes and governmental and non-governmental surveys in Vanuatu. This was done after the main recording event to avoid biasing the speaker (e.g. by drawing attention to the use of Bislama), and so that the detailed consent could be given in relation to the content of a specific recording. Sometimes the interview immediately followed the recording event, and sometimes the language assistants returned later at a convenient time for the participant, in which case

⁸ See the appendix.

it was recorded on a mobile phone. The metadata interview was optional, but we tried to record at least the detailed consent section as far as possible. When the speaker was a child, both their consent and a parent's consent was recorded. Overall, speakers were keen to share their stories widely. Every speaker was happy to have their stories shared with researchers and the Vanuatu Cultural Centre, and used in community materials. Around two thirds of speakers were happy to have their story shared without restrictions on the internet, and this proportion grew as speakers became more familiar with the internet and Facebook, which were becoming more available in Vanuatu and Southeast Ambrym over the course of this project.

File processing began by uploading the audio and video files into SayMore (Moeller 2014) session folders, filling out metadata for the time, place and participants, and a summary of the topic of the recording. SayMore's file naming conventions also ensured consistency. HandBrake (2017) was used to convert the video files to mp4 format. Next the audio files were autosegmented in SayMore, generating an eaf file. In ELAN (MPI Nijmegen 2018), each segment was labelled with a session code and segment number. From SayMore, these labels were exported to Audacity (2017), where individual mp3 files were generated for each segment. These files were then transferred to language assistants' mobile phones for transcription. Because audio and video were recorded separately, the high-quality audio also had to be linked to the video. The high quality audio was trimmed to match the output from the video recording in Audacity (2017), and then linked to the video in Avidemux (Mean 2017), with the signal clap trimmed from the beginning of the video. The output from Avidemux was then processed in HandBrake to produce a high quality mp4 for archiving, a medium quality mp4 for playback on video projectors, and a highly compressed mp4 for distribution to community members with touchscreen phones. Finally, the SayMore project was regularly backed up to external hard drives and, whenever possible, online.

The first stage of linguistic processing was transcription and translation into Bislama. Language assistants listened to the mp3 file for each segment on a loop on their mobile phones, copied the session code and segment number from

the file name, and then transcribed each segment and translated it into Bislama in notebooks. Initially, the researcher and each language assistant went through this process together and discussed issues like orthography and the importance of including repetitions and hesitations in the transcription; this latter concept went against the grain of Madleen and Simeon's previous experience as Bible translators but was quickly grasped. Preparing the files in this way meant that time-alignment was not lost, and allowed for independent transcription, which was much more efficient than requiring the presence of the researcher (Jung & Himmelmann 2011: 206), especially given that researcher time was split between two locations. However, it was also important to have regular discussions and feedback sessions. Language assistants were encouraged to make notes of any comments or concerns they had on the facing page of the transcription, and to include any corrections to be included in distribution formats, especially translations of Bislama code switches.

The next stage was for the researcher to type up all transcriptions and translations into the ELAN file via SayMore. As I became more proficient in Vatlongos, some frequent transcription spelling errors could be corrected at this stage and noted for discussion in feedback sessions, and increasingly the texts could be punctuated as well, following conventions of English, as this is the language of education for nearly all speakers (§2.2). This file was then exported from SayMore as a flextext file, and imported into Fieldworks Language Explorer (FLEX Development Team 2018). In FLEX, the texts were interlinearised and linked to a growing lexical database. This process facilitated consistent spelling without obfuscating real variation in speech: it was straightforward to replay the segments in SayMore to confirm where there were real phonological variants or just spelling errors. The researcher also added English transcriptions in FLEX, paying attention to the Bislama translations, transcriber notes, and increasingly to the Vatlongos original as my knowledge, and the lexical database, grew. A number of syntactic constructions were tagged in the note field (Chapter 9). Throughout these processes, questions arising were typed up in the note field of utterances or lexemes. These notes were then discussed in regular sessions with the language assistants, or used to plan elicitation materials between field trips. In total, 169

texts consisting of 66,351 words and 5,894 unique word forms were added to the FLEx database.

Finally, the researcher listened to the metadata interviews and coded them in a spreadsheet. Most questions were closed, and were coded as binaries or multi-level factors, but having the audio recordings allowed any extra observations from participants to be transcribed in a note field. Answers to the four open questions pertaining to language attitudes were transcribed and then tagged (§2.4). Recording the metadata interviews also allowed for follow up training with language assistants, especially on the importance of avoiding biasing the responses. The metadata questionnaire firstly provided a detailed picture of the sociolinguistic context in each community (Chapter 2), and secondly allowed the relationship between sociolinguistic factors and linguistic features in the corpus to be explored (Chapter 9).

1.3.3 Corpus construction and analysis

A wide variety of text types were recorded in accordance with community and speaker priorities. Narrative genres included custom stories, fable-like children's stories with animal protagonists, accounts of disasters, oral histories of villages and missions, autobiographies (especially recounting employment history), and some funny anecdotes. Procedural texts included explanations of how to obtain and prepare foods and construct traditional products and buildings. There were also some custom performances of sand drawings and rope drawings, sometimes with descriptions or related narratives. A descriptive technique that proved fruitful was the use of running commentaries (Margetts 2011). For these, the same speaker videoed an event and provided running commentary; events included football matches and a custom wedding. Another genre was formal speeches at custom ceremonies. A variety of songs were recorded, some in Vatlongos, some in archaic or nonsense varieties associated with songs in Melanesia (Lindstrom 1990: 121–122), and some in Bislama and English as well. These included children's songs, hymns, and songs recounting historical events. Some nursery-rhyme-like poems or chants were also recorded. A few interviews and casual conversations were recorded, but the corpus is mostly monologic. This was partly due to community preference for

recording narratives, and partly due to language assistant preference, as transcribing a conversation is considerably more challenging. It also suited the aims of this project which was especially interested in complex predicates and clauses, which are more likely in monologues without interruptions from other speakers.

Researcher-involved genres mostly consisted of elicitation via translation from Bislama, usually with sets of elicitation questions prepared in advance. These also included grammaticality judgements on Vatlongos sentences, sometimes in short narrative contexts, and other questions about speakers' metalinguistic knowledge. During the first period of fieldwork, some sections of Johnston's (1980) elicitation tool for grammar and basic vocabulary in Oceanic languages was also used, using Peter Budd and Kay Johnson's Bislama translation. Non-verbal stimuli materials included a Frog Story (Mayer 1969), which has been used to study the expression of motion events (Huang & Tanangkingsing 2005), a semantic domain often expressed with SVCs. We also videoed versions of the Max Planck Institute Cut and Break clips (Bohnemeyer, Bowerman & Brown 2001), using the same semantic dimensions but tools and objects more relevant to a Vanuatu context. This was in the hope of targeting contiguous SVCs, on the basis of an earlier project that found these constructions were common in this domain in Apma (Ridge 2013). Vatlongos complex predicates (§7.1) proved to be less productive and less diverse than their Apma equivalents.

Finally, the corpus contains some written materials, including literacy materials developed during the VESP workshops, translated from Bislama (§1.1.1). Some recordings were made of speakers reading Vatlongos literacy materials aloud. Example sentences from Parker's (1970a) dictionary were processed in FLEx prior to the first field trip, and transcriber corrections were also included in a separate text.

Storing the corpus in FLEx was extremely helpful in conducting the qualitative analysis of Vatlongos verbal grammar described in Chapters 3 to 8. FLEx structures interlinear texts in a four-level Emeld model (EMELD Project 2000; Hughes, Bird & Bow 2003; Bow, Hughes & Bird 2003), which allows the

texts to be accessed by morpheme, word and phrase, and includes notes at the phrase level. The ability to conduct simple and complex concordances at all of these levels, and lexical concordances by entry, sense or form, allowed analytical hypotheses to be quickly tested against the growing corpus, and meant analytical changes could be updated across all texts. FLEx also allows for fairly complex morphological analyses to inform an automatic parser. The ability to construct affix templates, to define phonological environments for allomorphs, and link forms and rules to TAM categories all proved helpful to this project, and meant that interlinearisation became increasingly efficient thanks to the semi-automatic parsing function.

In the early stages it was useful to interlinearise some elicited and prompted texts, and some non-spontaneous genres such as children's songs were also included at community request. However, interlinearisation increasingly targeted texts that were suitable for quantitative analysis. In Chapters 3–8, 'corpus' is used to refer to all the texts in the corpus that have been interlinearised; the subsection of the corpus that has also been tagged for quantitative analysis is referred to as the 'subcorpus'. The first criteria for inclusion in the subcorpus was complete metadata. Speakers who did not want to complete the metadata interview, or where the information was not complete due to technical or human error, were not included. Secondly, non-spontaneous genres were excluded: songs, elicitation sessions, responses to short stimuli, and recitations of written texts. An exception was made for the Frog Story, which was a less controlled guided narrative, and seemed fairly similar to the animal narrative genre (see discussion in §9.4). Finally, the subcorpus was inevitably limited by time constraints, and an attempt was made to prioritise texts for transcription and processing with the aim of including texts from across the three communities, with a good mix of genres, and speakers of different genders, ages and education levels. The remaining imbalances in the corpus are discussed in §9.2. In total, this left a subcorpus for quantitative analysis of 104 texts from 70 speakers, consisting of 47,854 words.

Quantitative analysis of the corpus was handled in R (2018) and RStudio (2018). The corpus was exported from FLEx as generic XML, and then imported

into R as linked dataframes, using the *interlineaR* package (Loiseau 2018). R provided more powerful tools for frequency counts, for example, searches for all logical combinations of morpheme forms, lexemes, word forms and syntactic tags. It also made it possible to cross reference the corpus with the results of the metadata questionnaire, and investigate the relationship between linguistic variation and community membership, as well as other sociolinguistic factors (Chapter 9, §5.1.1.2). The specific tests and models used in the quantitative analysis are discussed in detail in Chapter 9.

1.3.4 Community outcomes and goals

While enthusiasm for the project in the community was generally high, there were very few concrete goals suggested by community members in the early stages. However, as the researcher took the initiative with a few community materials, more suggestions and clearer goals arose, some of which will hopefully be addressed in the future.

The most popular community outcome was video, and this was also the easiest to produce as it did not require any linguistic processing. During the second field trip, video screenings were organised in all villages of Southeast Ambrym. These were screened using a white sheet hung from a string, and a cheap LED projector running off precharged battery packs, playing files from a hard drive. This allowed the screenings to take place without access to power, although only after dark to ensure visibility. A medium sized portable speaker was also used, but as speakers are very popular items, there was normally a better option available in each village, much appreciated when there was a large attendance. In total 24 screenings took place, with an estimated total attendance of 925. A projector and hard drive with video files were left with the language assistants in Mele Maat, Moru and Endu so that further video screenings could be held, and they were still working in 2018, although for shorter periods of time. In Mele Maat, speakers requested Bislama subtitles to help some younger speakers and outsiders to also access the villages, so for transcribed texts these were generated and included on the hard drive.

Small video files of the recordings were also distributed directly to community members. Due to the time and power required to link the high-

quality audio and process the videos, these files could not be distributed immediately in the field, but instead were sorted into folders by village and distributed on SD cards bought by the researcher in Port Vila and sold at cost. This was a popular initiative, and in total 80 SD cards were distributed in this way. Speakers could then delete videos that were not of interest to them, and had the benefit of a reliable piece of hardware that was difficult to purchase on Ambrym. Videos were also transferred directly to 10 speakers' phones, USB sticks or SD cards. All the speakers I contacted in 2018 still had the video files on their SD cards and valued them highly. Videos of custom stories, community histories, very funny stories, and videos of elderly speakers were most prized, especially in the sad cases where a speaker had passed away.

Many participants were also keen to have their stories shared more widely beyond the community. To this end, subtitles in Vatlongos and either Bislama or English were generated in Elan for texts that had been processed in FLEx, and some of these were uploaded to YouTube (with permission) (Ridge 2017a). These links were then shared on Facebook, the main way that Vatlongos speakers access the internet, especially young people in Mele Maat. Lots of speakers commented with their appreciation of these videos, and they have been viewed in Vanuatu, Australia, New Zealand, USA, UK and France. It is hoped that in future, online videos can be accessed via an archive instead.

EBooks were also produced for 32 texts, using Calibre (Goyal 2018). The texts were chosen on the basis of their interest, and at request. They were lightly edited to remove some hesitations and repetitions, grouped into paragraphs for an easier reading experience, and transcriber corrections were incorporated where a speaker made a mistake or used a long stretch of Bislama. These were distributed on SD cards along with an installation file for AI Reader (Neverland 2018), a free, small eBook reader app. These were well received, especially by teachers at the primary schools.

Some teachers requested a dictionary to help them in implementing the new vernacular education policy (Vanuatu Ministry of Education 2012). Although it was not ready during the main fieldwork trips, a dictionary app was produced from the FLEx database using the Dictionary App Maker (SIL

International 2018), and distributed to some teachers and other community members during August 2018. For now, this mostly consists of simple English glosses, although definitions in Vatlongos and translations in Bislama and English were produced for some of the most frequent words during 2016 fieldwork, when it was not possible to reach Ambrym. This will hopefully be developed in future; the app contains my contact details for feedback and suggestions.

A major change that was made to the scope of the project in response to community goals was the inclusion of Endu-Vatlongos. Originally the project was conceived as involving a comparison of Vatlongos in Mele Maat and South-Vatlongos, focusing only on the linguistic effects of the relocation. However, speakers in Endu were keen to have their dialect included in the project, especially due to fears around dialect loss. They have also asked for a future project to focus on documentation of Endu-Vatlongos, prioritising recording the eight remaining speakers of the most conservative variety of the dialect, as well as a dedicated dictionary app, and a phrase-book with English and French to help tourists communicate with locals.

1.3.5 Documentation and archiving

At the time of writing, the complete corpus has not yet been archived. However, decisions about software and data formats have been taken with a view to ease of archiving, and with the intention that this project constitutes language documentation as well as description and analysis. All the linguistic software tools use transparent XML formats. Storing the files in SayMore allows the entire project to be exported as an IMDI file for archiving, including metadata. More generally, choosing to use existing linguistic software and standard data formats as far as possible, rather than using bespoke data structures that only suit the aims of this project, makes the data more useful to other linguists, and more accessible for a range of other purposes.

So far, videos, community outcomes and the FLEx database have been submitted to the Vanuatu Cultural Centre, and texts in the subcorpus are in the process of being uploaded to the Pangloss Collection (Ridge 2018a). The Pangloss Collection allows for easy, rapid access to individual examples in

context, as data is presented in interlinear format linked to audio, in addition to subtitled videos. I also plan to archive with Paradisec, which holds deposits on many languages of Vanuatu.

1.4 Summary

This project aims to find out if there are morphosyntactic differences in the Vatlongos spoken in Mele Maat, a relocated peri-urban community, and villages in Southeast Ambrym, focusing on the verbal grammar of the language as the most likely area for variation to be found. It also investigates morphosyntactic differences in Endu-Vatlongos. The project's findings could have implications for the early stages of language divergence, the effects of relocation and urbanisation on endangered or minority languages, and the role of morphosyntactic variation in language change. These questions are firstly addressed through corpus-informed description of the verbal morphosyntax of Vatlongos, paying careful attention to attested variation, which is also explored through targeted elicitation. Evidence of differences in construction frequency are explored through quantitative analysis of a tagged corpus of spontaneous texts.

Chapter 2 describes the results of a sociolinguistic survey of speakers, in order to give a clearer picture of social context and language ecology of each speaker-community. Chapters 3 to 5 describe the syntax of simple clauses, the Vatlongos TAM categories and the verbal morphology respectively. Chapters 6 to 8 describe complex verbal constructions: SVCs, complex predicates, subordination and AVCs. Finally, Chapter 9 investigates variation in the token frequency of these verbal constructions in the subcorpus.

2 Speaker-community profiles

This chapter reports the results of a sociolinguistic survey across the three speaker communities. This survey has two main aims. One is to investigate the sociolinguistic context of each community. This is an important part of the documentation of the language, and informs hypotheses about language divergence between the communities, for example comparing the level of exposure to the national languages and other vernaculars in each community, finding out how much contact is maintained between communities, and identifying differences in language attitudes. The other aim is to gather information about each speaker which can be cross-referenced with the subcorpus, to find out how linguistic behaviours correlate with the speaker's sociolinguistic metadata. Only a few factors had enough variation to be useful for this latter aim, and the results are discussed in Chapter 9. Only speakers who contributed a recording completed the metadata survey, which skews the characterisation of the communities that emerges: for example, members of the Mele Maat community who do not speak Vatlongos are not included in this survey.

Of the 169 participants involved in the project, 114 completed a full metadata questionnaire. Some were excluded for deliberate reasons, such as corporate entities (for example football teams and kindergarten classes), and those who contributed non-linguistic or tightly constrained performances (such as sand-drawings or songs). There were also some participants who did not complete the questionnaire for logistical reasons or technical failures.

More speakers were interviewed in Ase-Taveak than in Endu or Mele Maat (Table 7). In Endu this is a reasonable reflection of population size, but in Mele Maat it indicates a difference in lifestyle. Speakers in Mele Maat were often engaged in employment with fixed hours, so it was more difficult to organise recording sessions, even though more time was spent in Mele Maat than on Ambrym in an attempt counterbalance this. Overall there is a good balance of genders, but more women than men were interviewed in Mele Maat. Mele Maat had the highest average age, followed by Endu, and this reflects fewer younger speakers contributing to the corpus in these communities, partly due to lack of

confidence in their language abilities, and also the time spent at school and in work by young people in Mele Maat, where only older people were available to make recordings during working hours. Altogether the questionnaire was made up of 68 questions and took around ten to twenty minutes to complete.

Table 7: Survey respondents by age, gender and community

	All	Men	Women	Average Age
Ase-Taveak	67	35	32	44
Endu	21	10	11	50
Mele Maat	26	9	17	53
Overall	114	54	60	47

The questionnaire consisted of five sections: detailed consent, demographic information, language background, language choices and language attitudes. This chapter is structured around the content of last four sections, focusing on the dynamics that are expected to be most relevant to language divergence. Taking a direct approach to gauging language attitudes can be problematic, as self-reported attitudes do not straightforwardly align with linguistic behaviour, and can be influenced by a number of factors in project design and context (Garrett 2000: 28–66; Garrett, Coupland & Williams 2003: 24–50). Bislama was chosen as the least marked choice for conducting a questionnaire in the Vanuatu language context, and to avoid compounding possible biases in favour of Vatlongos (as the focus of the project) and English (as the language of the researcher). The survey was audio-recorded to avoid bias in favour of the languages used in written contexts, as well as to save time and ensure consistency in coding. Where observed language-use differs from attitudes expressed in the survey, this is commented on below.

The results of this survey have been described in more detail elsewhere (Ridge 2018b), but this chapter also includes results from two additional speakers in Mele Maat, interviewed in July 2018, and some clarifications and corrections from other respondents.

2.1 Social profiles of communities

Many questions in the survey targeted what Haugen (1972: 336) describes as the ‘linguistic demography’ of each community: social factors that could have an effect on language use. The survey therefore addressed employment, education,

religion, community activities, and, importantly, mobility between the Mele Maat and island communities.

If employment is broadly defined as ways to obtain money and sustenance, most speakers in all three communities have several forms of employment; nearly all speakers maintain a 'garden', growing staple foods and cash crops like kava, in addition to more formal jobs or commercial enterprises. This survey only asked speakers to identify their main form of employment; while some richness of data is lost in this approach, it does reveal interesting differences between the communities. More than half of respondents in Endu, and 42% of respondents in Ase-Taveak identify maintaining a garden as their main form of employment, compared to just 15% of respondents in Mele Maat. Instead, nearly 40% of Mele Maat respondents name 'housework' as their main form of employment: for Mele Maat speakers, this usually means housekeeping in a hotel or for expatriates in Port Vila, although unfortunately it was not always clear if they were referring to maintaining their own household, so these two very different forms of employment are unavoidably amalgamated here. On the island, 'housework' was restricted to elderly speakers who could no longer maintain a garden.

More unusual forms of employment are also found in the survey, especially in Mele Maat. More than a quarter of Mele Maat respondents named a unique or rare role, restricted to only one or two respondents. This reflects the more varied job opportunities in the capital and includes skilled jobs and direct sales roles, such as a nurse, a mechanic, a market vendor and the manager of a kava bar. In contrast, the rarer jobs in Ase-Taveak include a butcher, a sailor, and selling kava to trading ships to be sold on in Vila. In Endu, these jobs include tourist-facing roles, such as tour guide and guesthouse owner.

Respondents in Mele Maat tended to have received more education, with an average of 8.3 years of education, compared to 6.7 and 6.4 years in Ase-Taveak and Endu respectively. Ase-Taveak and Endu are close to the national average of 6.8 years (UNDP 2016), suggesting few speakers make it past compulsory primary education (Crowley 2000: 79; 2005). The higher average in Mele Maat shows that more speakers progress to secondary school, and is also skewed by

some individuals who have reached tertiary level education. This reflects higher motivation to prioritise education in order to access employment opportunities in Port Vila, and higher purchasing power to pay for school fees, which are usually the greatest expense for most families in Vanuatu. Self-reported literacy rates are very high in all communities, but especially in Mele Maat where everyone interviewed said they could read (compared to 99% and 95% in Ase-Taveak and Endu). Bislama is the main language of literacy, despite only recently being officially taught in schools (Vanuatu Ministry of Education 2012). Only a few elderly speakers on the island described themselves as non-literate.

The Vatlongos speaker population is overwhelmingly Presbyterian Christian: 78% of respondents describe themselves as Presbyterian, the denomination of the first missionaries to reach this part of Ambrym (Frater 1922). This religious background accounts for the prevalence of English as the language of education, with only the very small Catholic population in the village of Pamal using French in education. The main competition to the dominance of the Presbyterian church comes from the Seventh Day Adventist (SDA) church (13%). It is especially popular in Mele Maat where more than a quarter of respondents identified as SDA. The community in Mele Maat is more exposed to international missionaries in Vila and is a driver of religious innovation on the island as well, for example by fundraising for a church building in Bethel village. Apart from the Presbyterian and SDA churches, most other religious affiliations are restricted to a single village or family group. For example, Christian Fellowship was only found in Endu, Catholicism only in Pamal, and the two adherents to Bahá'í were related to each other.

As well as engaging in multiple forms of employment, most members of the community hold one or more voluntary roles, usually through the churches or in the hierarchy of chiefs. There are roughly similar proportions of respondents involved in the church (61%) or acting as chiefs (17%) in all the locations. The data supports an impression that the church and other communal activities are less important in Mele Maat. More than a quarter of speakers in Mele Maat described themselves as having no community role (compared to only 6% in Ase-Taveak), and this included younger people, unlike on the island where only

the very elderly made such a claim. No-one in Mele Maat said they performed manual labour for the wider community, whereas on the island contributing goods for fundraisers or labour for communal projects was considered a defined role (10%). In casual conversation, Mele Maat's lack of community spirit was often decried and blamed on urban lifestyles. The Endu survey, although a smaller sample, also suggested less involvement in community activities, with 14% declaring no community role, and only 50% of respondents involved in church activities. Church activities include dedicated women and youth groups, and the Presbyterian Women's Missionary Union (PWMU) is a major force in the community, regularly organising fundraisers, projects and religious events. 56% of women on the island gave positions in women's groups as their main role in the community, while a still substantial 44% of women in Mele Maat are also involved in these groups.

It is important to consider mobility between the communities, because if extensive contact is maintained this is likely to curb language divergence. 83% of speakers in Mele Maat have visited Ambrym at some point. Many had been sent to stay with relatives as children, both to relieve childcare pressures in Mele Maat and to encourage them to learn Vatlongos and cultural practices. This pattern of exposure to vernacular languages for children in urban families is observed by Crowley (1990: 387). Similarly, more than half of respondents in Ase-Taveak have at some point visited Mele Maat. Less than a quarter of respondents in Endu have visited Mele Maat, perhaps because many speakers in Endu tend to reside in the peri-urban settlement of Bladinia when visiting Port Vila, staying with family from North Ambrym. Reciprocal mobility shows potential for speech patterns to be shared between Mele Maat and the island communities, and should temper any expectations of stark linguistic differentiation between the communities.

Another reason to examine patterns of mobility is that urban lifestyles and patterns of language contact in towns have been hypothesised as a factor encouraging language change and possibly language shift for the Mele Maat community. Speakers on the island were therefore also asked if they had visited any part of Port Vila, and an impressive 92% of respondents on the island have

visited Port Vila at some point in their lives. While the overall figure does not give us much information about the extent of respondents' exposure to urban living, it does suggest that the rural environment of Southeast Ambrym is not insulated from effects of urbanisation and globalisation.

2.2 Language repertoires

For these questions, speakers were asked to self-assess their competence in Vatlongos, Bislama, English and French on a four-point scale (none, some, good, fluent). They were also asked about any other languages they knew.

As everyone taking the survey had already agreed to contribute at least one recording in Vatlongos to the project, these figures are not an accurate reflection of overall levels of fluency in Vatlongos in the wider community, especially in Mele Maat. However, the decision to report competence as 'fluent' rather than merely 'good' could be an indicator of confidence in command of the language. Interestingly the lowest proportion of reported fluency in Vatlongos is in Endu (81% compared to 88% in Mele Maat and 94% in Ase-Taveak), which also has the lowest reported fluency in Bislama (76% compared to 92% in Mele Maat and 81% in Ase-Taveak). This indicates the dominance of North Ambrym language in many families in Endu, where women have married in from the North. This lack of confidence is perhaps also related to a preoccupation with language purity in Endu (Table 11), as many speakers believe that Endu-Vatlongos is the true original language of Southeast Ambrym, and that knowledge of this true language is being lost due to mixing with South-Vatlongos. Schneider (2017: 8) cites similar beliefs about Suru Kavian, a threatened dialect of Apma language (Central Pentecost, Vanuatu). Only in Mele Maat do a higher proportion of speakers consider themselves fluent in Bislama (92%) than in Vatlongos (88%). However, the very high overall rate of fluency in Bislama (83%) indicates high levels of bilingualism in Vatlongos and Bislama across the three communities. This means that there is potential for cognitive and communicative pressure towards convergence with Bislama structures (Muysken 2013; Fernández, Souza & Carando 2017), and these influences are likely to affect speakers in all communities, not just in Mele Maat.

The dominance of Bislama has not curtailed knowledge of other indigenous languages. Despite the ease of communicating in Bislama, there is still prestige attached to knowledge of other Vanuatu languages, and learning them is seen as an important mark of family ties and friendship. Equally, the sample includes nine respondents who speak another Vanuatu language as a native language, and learnt Vatlongos as an adult. All of these speakers are women who have married into Vatlongos-speaking communities, where they have lived for between 12 and 55 years. Four considered themselves 'fluent' speakers of Vatlongos, four 'good', and one described themselves as only speaking 'some'. All said that they spoke Vatlongos with their children, five in addition to another Vanuatu language, and one in addition to Bislama as well as another Vanuatu language.

As expected given the high incidence of intermarriage with speakers of North Ambrym, speakers in Endu have by far the highest rate of knowledge of other Vanuatu languages at 81%. Two thirds of respondents in Endu speak North Ambrym, while others speak Fanbyak, another language of Northern Ambrym. Others are also familiar with West Ambrym, Paamese and South Pentecost languages.

Although much less than in Endu, a sizeable minority (37%) of speakers in Ase-Taveak are familiar with another Vanuatu language. 20 of these speakers know Paamese, the most closely related language to Vatlongos, spoken a short boat ride away from the South coast of Ambrym. While it is particularly likely to be known by older speakers, who experienced the use of Paamese as a mission language, Crowley (2000: 122) comments on the continuing growth of the non-reciprocal spread of Paamese as an additional language in Southeast Ambrym, as well as Epi. Other languages known in Ase-Taveak are from the surrounding region: North Ambrym, West Ambrym and Epi.

Mele Maat has the lowest proportion of speakers who are familiar with another Vanuatu language, but still nearly a third at 31%. While Paamese is the most popular second Vanuatu language, others know languages located nearer to Mele Maat, such as Ifira-Mele, the language of Mele village, and Nguna, the language of an island off North Efate. These languages, especially Ifira-Mele

which is a Polynesian outlier, are less closely related to Vatlongos, suggesting high willingness and effort on the part of Mele Maat residents who have learned them, and close integration with Efate communities when these are parental or family languages for respondents. In fact, this was a conscious language policy planned by Elder Solomon, the founder of Mele Maat, in the early days of the community. John Enoch Rangimen explains how Mele Maat children were sent to different villages on Efate to study in the hope that they would learn the local languages and help the community as a whole to integrate into the life of their new island:

'the thinking behind sending them to different villages to study was that they'd stay in these places and at the same time they'd learn the language. It was like that so when they got together with different people, they knew the language of this place, they knew the language of that place, they knew the language of that place. An old man they called Elder Solomon, it was him who had this idea. It made it so lots of people from Mele Maat village, they knew the language of different villages. The language of Nguna, people here know it. The language of Eratap, people here know it. The language of Ifira, people here know it. The language of Mele, people here know it, it's like that. And it was a good idea, because they did it because they wanted us to be Efate people and we are truly Efate people because people know the language of different places.' [20170413a_h01m169_35-45]

Due to Vanuatu's colonial history, ni-Vanuatu today can choose to study in the English or French education system (Crowley 2000). Due to the prevalence of Presbyterianism in their community, Vatlongos speakers are overwhelmingly Anglophone in their language of education. 89% of respondents identified English as their language of education: 88% in Ase-Taveak (which includes the only Catholic/Francophone village), 95% in Endu and 88% in Mele Maat. The Mele Maat rate reflects the greater choice of schooling in Port Vila. There is an increasing trend in Mele Maat to put children into the French education system rather than the Anglophone primary school in Mele. Reasons cited for this decision include higher perceived educational standards in Francophone schools, and an impression that students in the Francophone system end up speaking better English as well as French. Surprisingly few speakers mentioned a desire for their children to speak French as a motivation for this decision.

However, high rates of enrolment in English language education are not reflected in expressed confidence in speaking English. Only around a third of speakers described themselves as ‘good’ or ‘fluent’ in English, while nearly two thirds only claim to speak ‘some’ English. In Mele Maat, nearly one in five respondents described themselves as ‘fluent’ in English, compared to only 7% of respondents on the island. This reflects the greater opportunities for education, travel and exposure to international tourists and expatriates in Port Vila. This higher exposure seems to lead to higher standards for self-assessment in the Mele Maat community: the 60% of respondents in Mele Maat describing themselves as only speaking ‘some’ English includes many people with a strong grasp of conversational English. Only two of the 114 respondents describe themselves as ‘good’ or ‘fluent’ speakers of French, while three quarters know no French at all. It will be interesting to see if this changes in Mele Maat as more students pass through the Francophone system.

2.3 Language use in different domains

Lewis and Simons (2010: 105; 2016: chap. 4), following Fishman (1991), discuss the importance of domains or functions as indicators of language vitality: if a language is used in more domains for more functions it is more valuable to speakers, who are therefore more likely to maintain and transmit the language. They advocate a goal of stable multilingualism involving ‘clearly defined functional assignments of the languages in an ecology’ (Lewis & Simons 2016: sec. 4.6). To some extent, this is a good description of the typical language ecology in Vanuatu: local languages (like Vatlongos), Bislama, English and French are all specialised to different domains, in a complicated example of polyglossia (Hudson 2002). However, not all languages are equally protected by distribution across different domains. Mühlhäusler (1996: 274) warns that traditional languages in the Pacific have lost their power when used in fewer and less prestigious domains, especially when education is conducted in a metropolitan language. That said, the results of this survey show that, at least in Ase-Taveak, Vatlongos is moving into domains beyond oral communication with family and friends, a promising sign for vitality.

Speakers were asked which language(s) they use in different situations, such as speaking to their parents, sending a text message, or reading the Bible. Speakers could name more than one language for any given situation. Table 8 shows the most popular language choices in each community for each situation. The percentage of speakers in the community who gave that answer is shown below, including speakers who named that language among others. The different contexts are arranged from informal, oral contexts on the left, to more formal written contexts on the right. Although school is a partially oral context, it is closely associated with literacy and formal education so appears to the right along with school books.

The overall trend is to associate Vatlongos with intimate and oral settings, Bislama with official services and emerging technologies, and English with formal and written language use. This pattern is replicated in each community, but with different cut-off points for each language. Ase-Taveak communities use Vatlongos in more formal domains like church, and only associate English with school textbooks. In Mele Maat, Bislama was the language most highly associated with friendship, while English is the top choice for school, reading the Bible and other books as well as textbooks. English is also the top choice for music, suggesting higher engagement with international media and culture in Mele Maat than elsewhere. Endu falls somewhere between Ase-Taveak and Mele Maat. Vatlongos and Bislama are in close competition in non-family oral domains. A lower proportion of speakers associated Vatlongos with friendship than in Ase-Taveak, and Bislama was a close second at 52%. In religious settings, Bislama (62%) narrowly beats Vatlongos (57%). Although some speakers in Endu said it would be better to use Vatlongos in church, they justified the use of Bislama given the number of women and children who are not fluent in Vatlongos. On the other hand, in commercial settings Vatlongos (67%) surpasses Bislama (57%) (unlike other villages in Southeast Ambrym, Endu has numerous well-stocked stores, which may explain the difference from Ase-Taveak in this domain). At the other end of the scale, Endu respondents associate English with books beyond textbooks, and both Bislama and English with the semi-oral school setting.

Table 8: Preferred language in different domains by community

	Father	Mother	Siblings	Spouse	Children	Friends	Church	Store	Hospital	Radio	Music	Text message	Email	Letter	Newspaper	Bible	Other book	School	Schoolbook	
AT	V	V	V	V	V	V	V	B	B	B	B	B	B	B	B	B	B	B	B	E
%	94	94	93	76	94	81	69	76	96	97	93	79	78	88	96	82	81	49	78	
E	V	V	V	V	V	V	B	V	B	B	B	B	B	B	B	B	E	BE	E	
%	86	81	86	90	90	57	62	67	76	100	95	81	43	81	81	71	57	43	76	
MM	V	V	V	V	V	B	B	B	B	B	E	B	BE	B	B	E	E	E	E	
%	92	92	92	73	69	92	85	100	100	96	88	96	65	77	96	81	83	69	96	

V = Vatlongos, B = Bislama, E = English

In all communities, Vatlongos is the preferred language choice for communicating with family members, but a lower proportion of speakers in Mele Maat use Vatlongos to communicate with their spouse and children, suggesting interrupted transmission. These figures are much higher than the percentage of urban households using a local language as their main communication vehicle according to the 2009 census: the national figure is only 21% (VNSO 2009: 167). It is likely that a much lower percentage of speakers in Mele Maat use Vatlongos as the primary family language than suggested in responses to this questionnaire. In informal discussions speakers in Mele Maat complained that children did not speak or understand Vatlongos, necessitating the use of Bislama. In my own observations while some children knew very little Vatlongos, others were shy about speaking the language themselves but were able to respond to questions and prompts.

In Ase-Taveak, Vatlongos also seems to be moving into more formal domains. It was the top choice for church services, which is a recent development: the Presbyterian church in Southeast Ambrym took the decision to conduct services in Vatlongos where possible, partly in response to the work on the Vatlongos language Bible which was published in 2015 (Wycliffe Bible Translators 2015). In addition, the recent policy shift in education to allow local languages or Bislama to be used in the first three years of primary school (Vanuatu Ministry of Education 2012) is starting to be implemented in Vatlongos in Ase-Taveak primary schools. Whether this policy is ultimately successful will now depend on the individual primary schools delivering the courses. Mele Maat primary school is using Bislama, as it also serves children in Mele who speak Ifira-Mele as their local language. It is not clear which language will be used in Endu: some speakers would like separate editions of the textbooks in Endu-Vatlongos, while others believe Bislama will be more practical, as some children speak North Ambrym rather than Vatlongos. The kindergarten in Endu uses Endu-Vatlongos as much as possible, in contrast to the kindergarten in Mele Maat, which has used Bislama and English to the exclusion of Vatlongos since it was established more than thirty years ago.

Bislama is highly associated with both written and spoken media, and is also the preferred language for written communication via letters, text messages and email. Crowley (2000: 90) discusses the dominance of Bislama in radio and its historical origins. Since the questionnaire was composed, Facebook has become a popular social media platform, especially for young people in Mele Maat. Although it was not included in the questionnaire, observed language use on the platform shows a preference for using English and Bislama; English is preferred for status updates, while comments are often in Bislama. Use of Vatlongos is usually restricted to emblematic phrases such as 'thank you' and 'good morning', often side by side with the same greeting in English or Bislama.

English is dominant in written and educational contexts. For all communities, English is the most frequent association with school textbooks, but the percentage is much higher in Mele Maat at 96%, while many speakers on the island associate schoolbooks with Bislama, and some with Vatlongos. One reason for the very high proportion in Mele Maat might be that Mele Maat primary school and kindergarten have received second-hand English language textbooks donated by tourists and non-governmental organisations. While most speakers in Mele Maat also associate English with general language use in schools, this is a closely contested domain on the island. Bislama edged out the other languages in Ase-Taveak with 49%, compared to around a third for both English and Vatlongos, while in Endu both English and Bislama were associated with school by 43% of speakers. These figures show that the recent language policy changes (Vanuatu Ministry of Education 2012) reflect the perceived reality in the classroom, where most teachers have always supported English language teaching with Bislama and local languages, even when this was not officially condoned (Crowley 2000: 79; 2005; Willans 2011).

While Table 8 gives a good idea of the languages most used in different contexts, it does not give a full picture of the domains most associated with each language, especially the roles of French and other Vanuatu languages. Table 9 shows the five domains most often linked with each language in each community. Father and mother are grouped together as 'parents' where they

are equally ranked, but in Endu, where exogamous marriages with North Ambrym speakers are common, more respondents speak Vatlongos with their father than with their mother, and vice versa for other Vanuatu languages.

As already seen, Vatlongos is most closely connected with family members and friends. Bislama is closely related to media and written communication. English is allied to formal written domains, but also music. Music is also the main domain where speakers are exposed to French and other Vanuatu languages: Crowley (2000: 91) discusses the variety of languages in music broadcast in Vanuatu. Respondents in Mele Maat seem to listen to more international music, with half of speakers listening to music in French and 88% to music in English. Apart from music, French patterns with English in being associated with media and education. Other Vanuatu languages resemble Vatlongos in being connected to family domains. In Endu, Vatlongos faces competition from North Ambrym for these core domains.

Table 9: Domains most associated with different languages by community

Vatlongos	Bislama	English	French	Other
Ase-Taveak				
Parents ⁹ (94%)	Radio (97%) Hospital (96%)	Schoolbooks (78%) Other books (45%)	Music (18%) Radio (7%) Newspaper (6%)	Music (24%) Parents (6%) Siblings (6%) Spouse (4%) Friends (4%)
Children (94%)	Newspaper (96%)	Music (45%) Bible (36%)	School (4%) Books ¹⁰ (4%)	
Siblings (93%)	Music (93%)	Newspaper (34%)		
Friends (81%)	Letter (88%)			
Spouse (76%)				

⁹ Where father and mother are associated with a language by the same proportion of respondents, they are collapsed as 'parents'.

¹⁰ Where school books and other books are associated with a language by the same proportion of respondents, they are collapsed into a 'books' category.

Vatlongos	Bislama	English	French	Other
Endu				
Spouse (90%)	Radio (100%) Music (95%)	Schoolbooks (76%)	Music (38%) Radio (10%)	Music (38%) Spouse (24%)
Children (90%)	Text message (81%)	Other books (57%)	Other books (5%) ¹¹	Mother (19%)
Siblings (86%)	Letter (81%)	Music (57%)		Father (14%)
Father (86%)	Newspaper (81%)	Bible (48%)		Siblings, Friends (14%)
Mother (81%)		Newspaper (48%)		
Mele Maat				
Parents (92%)	Hospital (100%)	Schoolbooks (96%)	Music (50%) Radio (46%)	Music (31%) Parents (8%)
Siblings (92%)	Store (100%) Text message (96%)	Music (88%) Other books (81%)	Newspaper (15%) School books (12%)	Siblings (8%) Children (8%) Letter (8%)
Spouse (73%)	Radio (96%)	Bible (81%)		
Children (69%)	Newspaper (96%)	School (69%)	Email (12%)	
Friends (58%)				

2.4 Associations with different languages

For these questions, respondents were asked to rate the importance of Vatlongos and the three official languages of Vanuatu on a four-point scale, and then to explain their rating. This question was deliberately open-ended to avoid leading respondents in a particular direction. The chance to ‘give reasons’ meant that different factors involved in ‘importance’ judgements could be captured, without limiting or pre-empting responses. The answers from this survey could be used as the basis for statements in a more detailed survey taking a Likert-scale approach (Garrett, Coupland & Williams 2003: 40–42). A four-point scale was used rather than the more usual five, both because respondents were not always familiar with the convention of rating scales, and because four levels were easy to distinguish orally.¹²

¹¹ These were the only domains where French was mentioned by respondents in Endu.

¹² *impoten tumas* ‘very important’; *impoten* ‘important’; *impoten smol* ‘a bit important’; *no impoten* ‘not important’

The average scores for each language in each community are shown in Table 10. Vatlongos has the highest overall score showing very positive attitudes to the language. Although research in other communities in Vanuatu has found that Bislama is viewed less positively than English (Schneider 2017: 8), Bislama and English are rated very similarly by Vatlongos speakers, with Bislama slightly higher in Mele Maat than on the island. This perhaps shows that negative attitudes towards Bislama (Thomas 1990: 246; Mühlhäusler 1996: 86–92) are eroding. Bislama is more positively assessed than other branches of Melanesian Pidgin English, which have not received the same level of official recognition (Crowley 1990: 15–17). French is also rated slightly more highly in Mele Maat, where there is a growing move towards education in the Francophone system.

Table 10: Average importance rating of each language (/4) by community

	Vatlongos	Bislama	English	French
Ambrym	3.3	2.7	2.8	2.0
Endu	3.1	2.7	2.9	2.0
Mele Maat	3.7	3.0	3.0	2.2
Overall	3.4	2.8	2.9	2.1

Respondent's reasons for importance ratings were transcribed, and grouped into themes under different tags. The most frequent positive and negative themes associated with each language are shown in Table 11 to Table 14, along with an example answer, and the percentage of speakers in each community who appealed to that topic.

Table 11: Positive and negative associations with Vatlongos

Positive associations					
Tag	All%	AT%	E%	MM%	e.g.
Identity	46%	54%	38%	31%	'it's my language', 'it's our language'
Culture	11%	12%	10%	8%	'it's our culture'
Secrecy	11%	3%	0	42%	'if I'm telling a secret someone else won't understand'
Maintenance	8%	9%	0	12%	'to keep the language strong'
Understanding	8%	7%	14%	4%	'I can understand everything'

Negative associations					
Tag	All%	AT%	E%	MM%	e.g.
Corrupted	4%	1%	14%	4%	'not our real language anymore'
Not Written	1%	1%	0	0	'you can't write it'
Not Widely Known	1%	0	5%	0	'some people don't know it'
Not Good	1%	1%	0	0	'it's not good'

The most frequent positive association with Vatlongos is identity, which is closely related to the second most frequent answer of 'culture' (usually expressed with the Bislama term *kastom*). However, the association with identity is more widespread in Ase-Taveak than in Endu or Mele Maat. In Endu this might be due to competition with North Ambrym language as a token of identity; North Ambrym is more closely related with *kastom* practices, such as carvings, taboos and dances. In Mele Maat the most popular association is very different: the ability to keep secrets from outsiders. Compared to speakers on the island, respondents in Mele Maat are more likely to interact with non-speakers of Vatlongos and therefore benefit from this function of the language. These are common associations for local Vanuatu languages (*lanwis* in Bislama). Meyerhoff (2000: 33) finds similar trends in her research in Malo and Santo: she mentions language maintenance, emphasising family roots, and maintaining privacy when speaking around outgroup members as instrumental and affective motivations for choosing to speak in *lanwis*. Schneider (2017: 8) finds that speakers of Suru Kavian dialect of Apma strongly valued their dialect as a marker of identity, community and home.

The most frequent negative association with Vatlongos is the theme of language corruption: the idea that the language has already changed and degraded too much to be worth speaking. This narrative is especially prevalent in Endu. The situation in Endu closely parallels Schneider's (2017: 10) findings for Suru Kavian: she discusses her informants' fears around the widely held language ideology that the dialect 'is changing and will be lost'.

Table 12: Positive and negative associations with Bislama

Positive associations					
Tag	All%	AT%	E%	MM%	e.g.
Communication	31%	19%	48%	46%	'you can talk to other people'
National Status	22%	28%	5%	19%	'it's Vanuatu's language'
Understanding	5%	6%	10%	0	'I understand straightaway'
Widely Known	4%	3%	0	12%	'lots of people know it'
Friendship	4%	3%	5%	4%	'I use it with my friends'
Negative associations					
Tag	All%	AT%	E%	MM%	e.g.
Not Secrecy	4%	4%	0	4%	'it reveals hidden things'
Not Identity	4%	6%	0	0	'it's not my language'
Not Clarity	4%	0	14%	4%	'things aren't clear'
Not Culture	1%	1%	0	0	'makes us forget our culture'

The most frequent positive association for Bislama is its value for communication. This association is more salient for speakers in Endu and Mele Maat, who are more likely to interact with non-speakers of Vatlongos. The national status of Bislama as a language that unites Vanuatu is another important association. Bislama played an important role in the movement for independence and in unifying Anglophone and Francophone regions (Thomas 1990). Second to this is its value for communication. Meyerhoff (2000: 33) also identified these two themes in her research in Santo and Malo.

Interestingly, Bislama was considered easy to understand by some speakers, but unclear by others, especially in Endu. This may reflect a difference between the experience of a listener and a speaker in a contact language like Bislama: while listeners are unlikely to be derailed by unfamiliar lexical items given the limited vocabulary of extended pidgins, a speaker has to explain concepts that might be expressed by a single lexical item in other languages in their repertoire. The other main theme emerging in the negative associations is the idea that Bislama does not express identity or culture. This idea that Bislama, and pidgins and creoles more generally, are inherently foreign and colonial (Walsh 1984; Mühlhäusler 1996: 98–101), is belied by the role of Bislama 'as a focus for anti-colonial discontent' in the run up to independence (Thomas 1990: 238), and these views are much less common than the idea that Bislama represents Vanuatu.

Table 13: Positive and negative associations with English

Positive associations					
Tag	All%	AT%	E%	MM%	e.g.
Tourism	19%	19%	33%	8%	'to talk to tourists'
Education	16%	19%	5%	15%	'for children at school'
Travel	11%	6%	0	31%	'to travel overseas'
Communication	11%	12%	14%	4%	'to talk to other people'
Employment	4%	3%	0	8%	'have to know it for work'
Negative associations					
Tag	All%	AT%	E%	MM%	e.g.
Not Known	13%	15%	19%	4%	'I don't know it'
Difficulty	6%	6%	14%	0	'it's too hard'
Not Understanding	3%	3%	0	4%	'I don't understand'

The main associations with English are tourism and travel, with an interesting distinction between respondents on the island and in Mele Maat. Speakers on the island associated English with tourists; a third of speakers in Endu made this connection, showing the impact of Endu's role as a base for tourists visiting the volcano. On the other hand, a third of speakers in Mele Maat saw English as useful for travelling abroad, especially to Australia, reflecting a more internationally mobile community than on the island. English was also associated closely with education, though surprisingly few speakers connected it with employment opportunities; this connection was mentioned by a slightly higher proportion of respondents in Mele Maat.

The main negative associations are to do with lack of knowledge of English, and perceived difficulty in learning it. Lots of speakers in Endu, especially, believe English is inherently difficult, perhaps because they are more likely to experience the challenge of interacting with English-speaking tourists.

Table 14: Positive and negative associations with French

Positive associations					
Tag	All%	AT%	E%	MM%	e.g.
Tourism	14%	12%	24%	12%	'sometimes French tourists come'
Education	10%	12%	0	12%	'children have to learn it at school'
Travel	3%	3%	0	4%	'if you go to New Caledonia'
Employment	2%	1%	0	4%	'can help you find work'

Negative associations					
Tag	All%	AT%	E%	MM%	e.g.
Not Known	55%	55%	67%	46%	'I don't know it'
Difficulty	5%	7%	5%	0	'too hard'
Not Employment	1%	1%	0	0	'doesn't help you find work'

As with English, a major association with French is tourism (especially in Endu), though it is less closely linked with travel and education. Although some speakers associate it with employment opportunities, one speaker argues that it is not useful for finding work, at least in comparison to English. For most respondents French is not important because they do not know it themselves.

2.5 Language vitality

The absolute speaker numbers (somewhere between 2500 and 3700, §1.1.2) may appear to indicate that Vatlongos is vulnerable from an international perspective; as the UNESCO report on language vitality and endangerment points out, 'a small speech community is always at risk' (Brenzinger et al. 2003: 8). However, from a Vanuatu perspective, Vatlongos is one of the most robust languages in the country, based on speaker numbers alone. François et al. (2015) put Vatlongos in the top 20% of languages in Vanuatu by speaker numbers based on the figure of 3700. Even if the real speaker population is nearer 2500, it is still comfortably in the top third. Most languages in Vanuatu are limited to one or two villages with a few hundred speakers, but are still robustly transmitted to children (François 2011; François et al. 2015).

Using the Expanded Graded Intergenerational Disruption Scale (EGIDS) developed by Lewis & Simons (2010; 2016: chap. 5), *Ethnologue* (2018) rates Vatlongos as a level 5 'developing language' and therefore 'safe'. This reflects that it is used orally by the whole community and in written form by part of it: the publication of the Vatlongos Bible in 2015 (Wycliffe Bible Translators 2015) raised it into this category. However, it remains to be seen how much the Vatlongos Bible will be used; around a third of speakers mention Vatlongos as one of the languages they access the Bible in, but Bislama and English were mentioned by more speakers in all three communities (Table 8).

Nevertheless, this seems to be an accurate assessment of the vitality of Vatlongos in Ase-Taveak communities, and in fact the new developments in education policy (Vanuatu Ministry of Education 2012) could push it up to a level 4 'educational' language, where literacy is taught through an institutionally supported education system. However, the new policy only affects the first three years of primary education, and whether this policy change has a significant effect on the vitality of the language will again depend on uptake by the community. Lewis and Simons (2016: sec. 5.2.5) stress the importance of community-based institutional support to achieving this level of vitality.

However, applying the EGIDS criteria to the vitality of the language in the other speaker-communities gives less encouraging results. Unlike in Ase-Taveak, Vatlongos is not used in church in Endu, and there are no concrete plans to implement the new curriculum in Vatlongos either. In Endu, Vatlongos was rated as less important than in Ase-Taveak or Mele Maat, and fewer respondents connected Vatlongos with their identity than elsewhere in Southeast Ambrym. Three respondents in Endu felt that Vatlongos has become corrupted and can no longer function as a marker of identity, and this ideological stance was also raised by other Endu-Vatlongos speakers outside of the formal survey. In Endu, Vatlongos is in competition with North Ambrym language for the key domain of the family, and although children and young people all appeared to understand Vatlongos, fewer were confident enough to contribute recordings to this project (Figure 27). Overall, Vatlongos in Endu is probably at level 6a 'vigorous' on the EGIDS (Lewis & Simons 2016: sec. 5.2.7), but could quickly become threatened depending on the decisions made by young people now.

In Mele Maat, the vitality of the language is again much less robust than in Ase-Taveak. Here there are no plans to use Vatlongos in education, although a third of speakers mention the Vatlongos Bible. In Mele Maat, the language is not being transmitted orally to all children, making it a level 6b 'threatened' language, the equivalent to a 'vulnerable' language on the UNESCO scale (Brenzinger et al. 2003). This supports the observation made by Simon & Lewis (2013), summarising the results of applying EGIDS to the languages in

Ethnologue, that urbanisation is now the main driver of language endangerment.

An interesting feature of the Mele Maat situation is that transmission seems to be delayed, rather than completely disrupted. One of the questions on the survey asked about age of acquisition, and it was designed to be asked only to non-native speakers, such as those marrying in to the community. However, the language assistants also asked the question of native speakers, and the results were revealing: while most speakers in Ambrym and Endu said they learnt Vatlongos at age 1 or 2, speakers in Mele Maat often said they learnt the language as late as 8 or 10. This may reflect higher standards for 'language learning', due to lower confidence in the urban community, but I did observe that only older children and teenagers were comfortable speaking Vatlongos in Mele Maat, though younger children may have already acquired a passive competence in the language.

Discussions around the survey, and other conversations with community members, suggested that while many children do not speak or understand Vatlongos, many go on to learn and use the language as teenagers. Young adults who had acquired Vatlongos in this way spoke about their motivations: they often felt a need to have a language connected to their island identity, or a private language around strangers. Sometimes being around speakers of other Vanuatu languages at secondary school led them to see the value of Vatlongos. This might be an example of how languages can be maintained by the wider linguistic environment (Mühlhäusler 1996): by seeing how a parallel local language is a valuable resource to other groups, Mele Maat teenagers see the potential of Vatlongos. A similar pattern of acquisition has been reported for the Kwarandzhey in Tabelbala (Algeria; Songhay) (Souag 2010: 30). Thurston (1992: 135) describes teenage speakers of Amara (New Britain, Papua New Guinea; Austronesian) unexpectedly embracing the highly endangered language and using it to 'show off' to peers in other language groups. The mechanisms and linguistic effects of this pattern of acquisition would be an interesting question for future research.

2.6 Summary

Much of the information collected via this survey can inform hypotheses about differential language use in the three communities. The reported reciprocal mobility between communities is likely to lessen the extent of linguistic divergence between Mele Maat and the island communities, but other features of the social characteristics of each community point to stronger motivation for language shift to Bislama and English in Mele Maat than in Southeast Ambrym, especially greater opportunities for education and formal employment, including internationally. However, nearly all respondents on the island have experienced some exposure to urban life, so are not entirely shielded from these kinds of incentives.

The overall profiles of language repertoires and language choices in different domains reveal a greater exposure to Bislama and English in Mele Maat and, to a lesser extent, Endu, than in Ase-Taveak. Speakers in Mele Maat in particular engage with international media, and associate English with international travel. These findings support the idea that Vatlongos could be more influenced by Bislama and English in Mele Maat than in Ase-Taveak, leading to differential trends of convergence with the national languages in the different communities. Parallels and differences with structures in Bislama and English are highlighted in the description of Vatlongos verbal morphosyntax in Chapters 3–8.

Despite the dominance of Bislama in a number of domains, knowledge of local languages remains prestigious, and a sizeable minority of speakers are familiar with a second Vanuatu language. In Endu, Vatlongos is in direct competition with North Ambrym in the family domains, and as a marker of identity and culture, which poses a different challenge to the language vitality of Vatlongos in Endu than in the other communities.

The lower vitality of Vatlongos in Mele Maat and Endu, is an important factor to consider in the discussion of linguistic features and variation in speech patterns in the rest of this thesis. It is possible that Vatlongos in these two communities could undergo processes of language attrition, which could result in simplification of grammatical systems and avoidance of complex structures.

3 Syntax of simple clauses

A verbal clause in Vatlongos minimally consists of an inflected verb, plus several optional elements. Figure 6 shows the order of these elements, including those found in constructions discussed in later chapters. Discourse prominent arguments and adjuncts are discussed in §3.1.1 and §3.1.2. Auxiliary verbs are discussed in Chapter 8. The negative clitic can appear before or after an object noun phrase (NP), and never occurs with the partitive clitic, which is why these are marked in *italics* in the diagram, to show that only one of these clitics can appear in a single clause (§3.6). Other arguments and modifiers include prepositional arguments and adjuncts (§3.2), complement clauses (§7.2.1) and the subsequent verb in an SVC (Chapter 6). These last elements can appear in any order with respect to each other, so it is assumed that they are adjoined to the verb phrase (VP), even when they are selected by the verb.

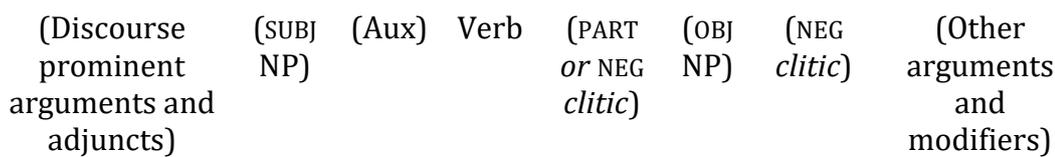


Figure 6: Constituent order in the Vatlongos verbal clause

This chapter also describes minor subclasses of verbs that display different syntactic behaviour, including verbs which appear to have developed from incorporation of prepositions (§3.2.1) and adverbs (§3.3.1), a reflexive verb (§3.4.1), and psycho-collocations (§3.4.2). Non-verbal predicates are also possible, either expressed with verbless clauses, or introduced with the copular verb; these are discussed in §3.5 along with locative predicates and existential clauses. Finally, §3.6 discusses clitics, especially the behaviour of the negative clitic in different clause types.

3.1 Constituent order in simple clauses

Vatlongos has a basic constituent order of SVO, which is canonical for Oceanic languages (Ross 2004), and the most common constituent order in languages of Vanuatu (Lynch, Ross & Crowley 2002: 49). In Vatlongos clauses where both subject and object are expressed with NPs, SVO is the least morphologically

marked constituent order and the most neutral pragmatically, two criteria for identifying basic word order (Mithun 1992).

Vatlongos is a head-marking language (Nichols 1986), with some relationships between a verb and its arguments marked on the verb. Subject person and number agreement features are coded by subject cross-indexing prefixes, and the person of a singular pronominal object is indicated by the object suffix, a pro-index (Haspelmath 2013). As is common in head-marking languages, pronominal arguments need not appear as separate constituents in the clause when their agreement features are marked on the clausal head (Nichols 1986: 107): the subject prefixes optionally co-occur with a coreferential NP, while the object suffix is in complementary distribution with a post-verbal object NP. A minimal clause can therefore consist of a single verb. Although SVO is the basic constituent order when both arguments are expressed with an NP, this is relatively infrequent compared to in languages that do not allow zero anaphora of pronominal arguments. The low frequency of transitive verbal clauses with both arguments coded as NPs is a common obstacle to identifying basic constituent order in Oceanic languages in a way that is valid for typological purposes, so not necessarily a very revealing parameter (Ross 2004: 494; Lynch, Ross & Crowley 2002: 49).

The roles of argument constituents in head-marking languages have been questioned. In languages with extensive head-marking on the verb and polysynthetic languages, it has been argued that NPs expressing arguments are more similar to coreferential adjuncts or dislocated constituents than true arguments in strictly configurational languages (Jelinek 1984; Nichols 1986: 107; Mithun 1992: 58–59; Austin & Bresnan 1996; Baker 1996; Baker 2001). Ross (2004: 515–534) discusses the implications of typical Oceanic verb marking strategies for the analysis of argument constituents in the clause, and most of his observations are relevant to Vatlongos. Although Vatlongos does have a dedicated clause-initial position for discourse prominent arguments (§3.1.1, §3.1.2), the choice to include an overt argument constituent, especially an independent pronoun, is more marked than in languages where argument constituents are obligatory.

An NP can consist of a single noun, a noun with modification (adjectival, possessive, prepositional or clausal), or an independent pronoun. Simple, complex or pronominal NPs have identical distributions.

Examples (1), (2) and (3) each show a subject NP preceding an intransitive verb: in (1) the subject is a complex NP, here consisting of a noun and a possessive classifier, in (2) the NP is a single noun and in (3) it is a single pronoun. These different NP types are exemplified in the same order for each of the constituent orders.

(1) *horamue na-n pat*
 boy CL.GEN-3SG.POSS 3SG.NFUT.sleep
 'his son slept' [20141105e_n01e012_04]

(2) *tatal met*
 snake 3SG.NFUT.die
 'the snake died' [20141106f_n01e018_31]

(3) *xalu lu-ba¹³*
 3DU 3DU.NFUT-NFUT.go
 'they went' [20150219b_n01s001_71]

Example (4) shows an intransitive verb as an independent clause, without any separate constituent referring to the subject.

(4) *lu-ba*
 3DU.NFUT-NFUT.go
 'they went.' [20141212n_n01s046_03]

Examples (5), (6), and (7) show a subject-verb-object constituent order with transitive verbs.

(5) *meletin xil la-gal laplap na-n*
 person PL 3PL.NFUT-NFUT.buy laplap CL.GEN-3SG.POSS
 'People buy his laplap.' [20141030a_p01m004_38]

(6) *vatiang bas moletin*
 wind 3SG.NFUT.hit person
 'The cyclone struck people.' [20141117a_n01m003_26]

¹³ Some verbs with initial /h/, /v/, /t/, /x/, /k/ or a vowel can undergo verb-initial consonant mutation in certain TAMP and subject person-number environments, most frequently taking a prenasalised initial consonant (/b/, /d/, /g/) in affirmative non-future environments. This process is discussed in section 5.2.1.

(7) *inou na-laxat xametel*
 1SG 1SG.NFUT-watch 1PC.EXCL
 'I supervised us.' [20170224a_n01s141_94]

Examples (8), (9) and (10) show a subject NP preceding a transitive verb, while the object is coded with a pronominal object suffix.

(8) *moletin xil la-ga-ni*
 person PL 3PL.NFUT-NFUT.eat-3OBJ
 'people eat it.' [20141028b_p01m003_20]

(9) *koh mi-gur-i*
 heron 3SG.NFUT-NFUT.take-3OBJ
 'The heron got it.' [20141220d_n01s082_08]

(10) *Inou na-pus-i*
 1SG 1SG.NFUT-see-3OBJ
 'I saw it.' [20150414b_h01e016_70]

Examples (11), (12) and (13) show a transitive verb preceding its object NP, without a subject NP.

(11) *Lu-pus tatal ak*
 3DU.NFUT-see snake PROX
 'They saw this snake.' [20141220g_n01s080_31]

(12) *i-kas tuvava*
 3SG.DFUT-wash baby
 'She'll wash the baby.' [20150303g_p01e009_21]

(13) *lu-dup xil*
 3DU.NFUT-NFUT.shoot 3PL
 'they shot them' [20141220g_n01s080_15]

Example (14) shows a transitive verb as an independent clause.

(14) *La-gul-e.*
 3PL.NFUT-NFUT.cover-3OBJ
 'They cover it.' [20141028b_p01m003_17]

Different transitivity classes of verbs are discussed in §5.1.3, in relation to object and transitivity suffixes.

3.1.1 Non-basic word orders and information status

Different word orders are possible under certain information status conditions, when arguments of the verb can be left- or right-dislocated (Lambrecht 1996: 181–184). In Vatlongos, left dislocation of a subject or object to establish a new

or contrastive topic in the discourse is very common. A topic can be defined as a given referent that a proposition is about (Lambrecht 1996: 117–119; Erteschik-Shir 2007: 13). The use of NP topics in clause-initial position in Vatlongos matches Lambrecht’s (1996: 176) observations that left dislocation of a constituent is used to express a referent that is accessible in the discourse context as an active topic, that is, it marks a switch topic rather than a continued topic (Erteschik-Shir 2007: 4). Right dislocation is much rarer in Vatlongos, and only used for discourse repair, ensuring that the listener has understood the intended referent in what is sometimes referred to as an ‘afterthought’ construction (Mithun 1992: 49). Dislocation is also marked by prosodic cues such as a pause, although these have not been investigated closely.

Example (15) shows the subject NP appearing clause-initially as a topic, separated from the rest of the clause by a prosodic break, marked with a comma in the transcription. Impressionistically, the topic constituent also tends to have a higher pitch and faster rate of speech than the rest of the clause.

- (15) *Xamem, ma-pus vetei ...*
 1PL.EXCL 1PL.EXCL.NFUT-see breadfruit..
 ‘Us, we see breadfruit (*is a good food*) [20150303d_p01e016_22]

While the first and second persons are always accessible and available as topics (Erteschik-Shir 2007: 11–12), example (16) shows how a third-person referent is first introduced into the discourse, then appears in initial position separated from the verb by a prosodic break to show that it is a new topic. This position is typical of the second mention of a referent.

- (16) *masta ne ok ise-n Smet.*
 master of PROX name-3SG.POSS Smith
Smet, bemei mi-sekhan mi xamel
 Smith 3SG.NFUT.come 3SG.NFUT-shake_hands to 1DU.EXCL
 ‘This one’s boss was called Smith. Smith, he came and shook hands with us.’ [20170224a_n01s141_18-9]

When the object appears in a different position due to its information status, it is also cross-referenced with a resumptive pronominal object suffix on the verb, which is strong evidence that these are not basic word orders. As the object suffix cannot co-occur with a NP in the post-verbal object position, this shows that there is an anaphoric relationship with a left-dislocated constituent, rather

than the grammatical agreement that is found between a subject NP in subject position and the subject agreement markers on the verb. This is very similar to the distinction between subject and object agreement morphology in Chichewa (Bantu) described by Bresnan & Mchombo (1987).

In examples (17) and (18), an object NP appears at the beginning of the clause and the verb is marked with an object suffix. In (19) three object NPs are fronted in contrastive focus.

- (17) *Kava igak mi-gur-i*
 corrugated_iron here 3SG.NFUT-NFUT.take-3OBJ
 ‘The corrugated iron here, it [*the cyclone*] took it out.’
 [20150419c_h01s004_25]
- (18) *his sa-n, tas gur-i ba*
 banana CL.DOM-3SG.POSS sea 3SG.NFUT.take-3OBJ 3SG.NFUT.go
 ‘His banana plant, the sea took it away.’ [20170119g_n01s087_07]
- (19) *langmas, ta rat-pus-i;*
 insane_person just 1PC.INCL-see-3OBJ
titamol, ta rat-pus-i;
 dwarf just 1PC.INCL-see-3OBJ
holesok xil tavusien xe
 thing PL all REL
mi-sa, rat-pus-i.
 3SG.NFUT-bad 1PC.INCL-see-3OBJ
 ‘Madmen we’ve seen; dwarves we’ve seen; all bad things we’ve seen.’
 [20141107b_n01e020_10]

An independent personal pronoun that is coreferential with the object can also be fronted in contrastive contexts, as in (20). The independent pronoun strictly refers to singular referents, whereas both the pronominal object suffix and subject agreement prefixes can refer to plural inanimate referents (§5.1.3), as in the first clause of the example.

- (20) *ueili xil vusien xa dik,*
 pig PL all REL 3SG.NFUT.be_here
u-has-i, e xi, o-naa-vas-i =ti
 2SG.DFUT-kill-3OBJ but 3SG 2SG.IFUT-NEG-NEG.hit-3OBJ =NEG
 ‘All the pigs here you’ll kill, but him, don’t kill him.’
 [20170222d_n01s152_86]

When the object is not in a person-number combination that can be indexed with an object suffix (§5.1.3), the independent pronoun object must appear after the verb in addition to the fronted independent object pronoun:

3.1.2 Information status of other arguments and non-arguments

Objects of prepositions, possessors, and locations can also appear in information status privileged positions in the clause.

Objects of prepositions can be fronted, in which case a preposition marked with a resumptive object suffix (§3.2) appears in-situ.

- (26) *muli-lau* *muli-he* *rin* *ak* *e*
 2DU.DFUT-hunt 2DU.DFUT-go_to side PROX but
rin *ak* *mul-naa-va* *ti* *e-n*
 side PROX 2DU.DFUT-NEG-go NEG LOC-3SG.OBJ
 ‘You two must hunt over this side but this side you mustn’t go in’
 [20141220g_n01s080_11]

Prepositional objects can also be right-dislocated; the in-situ preposition is again marked with a pronominal object suffix.

- (27) *na-sak* *mi-ni,* *puiteh* *ok*
 1SG.NFUT-try to-3OBJ door PROX
 ‘I struggled with it, this door.’ [20150415a_h02s125_28]

The possessor of the subject often appears at the beginning of a clause as a topic:

- (28) *xalu,* *nep* *na-lu* *tavuol*
 3DU knife CL.GEN-3DU.POSS 3SG.NFUT.absent
 ‘Them, they didn’t have a knife.’ [20141027a_n01m001_63]
- (29) *atou* *xalu* *xa* *na-leh* *xalu* *iaxai,*
 woman DU REL 1SG.NFUT-take 3DU MED
mama *na-lu* *be* *tatal*
 mother CL.GEN-3DU.POSS 3SG.NFUT.COP snake
 ‘These two women that I took, their mother is a snake.’
 [20141107b_n01e020_20]

Example (30) shows the possessor of the NP complement of the copula at the beginning of the clause as a topic, while the referent of the subject of the copular verb is clarified by right-dislocation.

- (30) *xamem,* *taa-ve* *vonine-mem* *=ti,* *igak*
 1PL.EXCL 3SG.NFUT.NEG-NEG.COP place-1PL.EXCL.POSS =NEG here
 ‘Us, it isn’t our place, here.’ [20141117a_n01m003_02]

Adjunct NPs referring to a location often appear at the beginning of a clause as an overt stage topic, a physical or temporal location that the truth of the

proposition is evaluated against (Erteschik-Shir 2007: 10–11). These stage topics are followed by any subject NP, as in (31). This is confirmation that the clause-initial position reserved for information-status privileged constituents is separate from the usual subject position in the SVO constituent order, that is, it is not the case that every subject NP is a topic or focus.

- (31) *Metimel tahal Moru, amel xil te-he hat*
 village of.LOC Moru nakamal PL 3SG.PRI-COP three
 ‘In the village of Moru, there are three nakamals.’
 [20141218g_h01s075_01-02]

Example (32) has two fronted constituents and a subject NP. The first NP, ‘a home or household’, is a stage topic, setting the scene for the following comment. The second NP, ‘woman’, is the possessor of the subject. It is being contrasted with the ‘man’ in the comparative SVC (§6.4.1). This could suggest that there are separate positions for a topic and a contrastive focus (Erteschik-Shir 2007: 48–51) constituent in the Vatlongos clause, but there are not enough examples with more than one fronted constituent to confirm this. Alternatively, the stage topic could be analysed as a non-verbal predicate by itself, and it might function as a topic at the level of discourse rather than within the clause.

- (32) *Tim tei mu haoshol tei,*
 home one or household one
atou, pol-ien na-n
 woman work-NMLZR CL.GEN-3SG.POSS
mahulong mahulong mi-lii toromue
 3SG.NFUT.big 3SG.NFUT.big 3SG.NFUT-beat man
 ‘In a home or household, the woman, her work is much greater than the man’s.’ [20150303g_p01e009_03]

A fronted constituent often has a whole-part relationship with an argument in the clause. In (33) the zero-coordinated subject NP ‘the men and the little boys’ are a subset of the referents of the fronted pronoun *xamem* ‘1PL.EXCL’. In (34), the object of the complex predicate which is the second verb in an SVC, ‘one side of the corrugated iron’, is a part of the fronted constituent ‘my house’.

- (33) *xamem, tiramue xil tutut horamue xil la-pat*
 1PL.EXCL man PL little boy PL 3PL.NFUT-sleep
 ‘Us, the men and the little boys slept.’ [20150226a_n01s098_36]

meaning of the predicate rather than the preposition, there is a strong case for analysing the PP as an argument rather than an adjunct. For example Bugenhagen (2010: 457) analyses PPs in Tuam (Oceanic, PNG) as oblique arguments when their semantic role is tightly linked to the predicate in question. It is not controversial to posit the existence of PP arguments in Oceanic languages: Pawley (1973: 117) assumes the existence of oblique arguments introduced with prepositions in Proto-Oceanic.

PP adjuncts, on the other hand, provide information which is not central to the predicate's meaning and might describe any kind of situation equally well. For example, they can provide information about the time or place of the situation as a whole, or the speaker's attitude to the situation.

Arguments expressed as a PP are often referred to as obliques, but I am avoiding the term here, because different definitions have been suggested which would make the term less specific than the PP constituents under discussion in this section. For example, Andrews (2007: 157–160) responds to the difficulty of determining the argument status of PPs by suggesting that all PPs should be described as obliques, only some of which have argument status. On the other hand, Foley (2007: 369–370), in the same volume, treats 'oblique' as a feature assigned to NP arguments that require specific additional 'predicators', including adpositions, but also case affixes, clitics or serial verbs. In Vatlongos, SVCs can also play a role in introducing oblique arguments in this sense (§6.4.8).

The main prepositions in the Vatlongos clause are closely associated with certain semantic roles, but can introduce both arguments and adjuncts. Similar semantic extensions are described for the cognates of these prepositions in Paamese (Crowley 1982: 195–210), except that Paamese does not have a dedicated instrumental preposition, and instruments are instead introduced by the most semantically general *eni* preposition, equivalent to Vatlongos *e*.

The preposition *mi* 'to, for' is used for a cluster of semantic roles that are usually borne by animate participants: recipient (36), (37), beneficiary (38), (41), interlocuter (39), (40), and caused experiencer (41). It is particularly

common with verbs of giving or speaking, suggesting that these verbs select for a recipient or interlocuter oblique argument. *Mi* takes the same pronominal object suffixes as a CV transitive verb (Table 27).

- (36) *ma o-sa-ni mi-ni*
 then 2SG.NFUT-give-3OBJ to-3OBJ
 ‘Then you give it to him.’ [20141106d_p11e016_21]
- (37) *na-sa-ni mi xil*
 1SG.NFUT-give-3OBJ to 3PL
 ‘I gave it to them.’ [20141121a_n01m041_88]
- (38) *mate-pol mi-ni*
 1PL.EXCL.PRI-work for-3OBJ
 ‘We worked for him.’ [20141117a_n01m003_28]
- (39) *u-hiteni mi xil*
 2SG.DFUT-tell to 3PL
 ‘You’ll tell them’ [20150419c_h01m004_58]
- (40) *xi mi-tumul-ni mi inou*
 3SG 3SG.NFUT-narrate-TR to 1SG
 ‘He told me about it.’ [20141211a_n01s046_3]
- (41) *Ma pisen-i mi-nou*
 then 3SG.NFUT.show-3OBJ to-1SG.OBJ
 ‘Then she showed it to me.’ [20141212g_p01s046_12]

Mi can also be used for the target of greetings at different times of day as in (42), which could be an extension of the benefactive or interlocuter functions of the preposition.

- (42) *bos mahis mi xir tavusien*
 3SG.NFUT.good afternoon to 1PL.INCL every
 ‘Good afternoon to all of us.’ [20150118b_n01m090_01]

The preposition *e*, *ena* is mainly used for locations, and other uses seem to be extensions of this basic meaning: it is also used for goals and sources, as well as certain types of instruments, purposes and stimuli. As a marker of location, it mostly introduces adjuncts of place giving information about where the situation predicated by the verb takes place:

- (43) *mei i-kkah e tim*
 COME 3SG.DFUT-make_laplap LOC home
 ‘She’ll come and make laplap at home.’ [20141116b_c01m_05]

- (50) *Mi-stal* *e* *hospitel.*
 3SG.NFUT-come_out LOC hospital
 'She came out of hospital.' [20170224a_n01s141_155]

This preposition can be used for certain types of instruments which can be figured as places that encompass or are brought against the object:

- (51) *atuli ak mi-dalxat rop e he-n tei*
 girl PROX 3SG.NFUT-NFUT.hold rope loc hand-3SG.POSS one
 'The girl holds the rope in one hand.' [20141127a_x02m004_20]

- (52) *guppas-i e simen*
 3SG.NFUT.smash-3OBJ LOC cement
 'She smashes it on the cement.' [20141127a_x02m004_27]

Finally, *e* is used for some participants that are not included in the usual inventory of semantic roles (Van Valin & LaPolla 1997: 85–86), but whose meanings are closely related to the specific predicates involved and are therefore analysed as arguments rather than adjuncts. Thus, *e* is used with the loan adjective *sori* 'sorry' when it is used as a predicate with the copular verb (§3.5.2), to introduce the object of pity or sorrow that takes *for* in English. This is equivalent to the target of an emotion verb in Van Valin and LaPolla's inventory (1997: 115).

- (53) *tuneli xil di la-be sori e xil*
 brother PL CONT.REAL 3PL.NFUT-NFUT.COP sorry LOC 3PL
 'The brothers are sorry for them.' [20141208b_n01s044_49]

E is also consistently used to introduce the third argument of *tutou* 'help': the task or area of aid introduced by English *with*. This could be described as a purpose.

- (54) *uhia mi-dutou-ni xamem holu e skul fi*
 wild_yam 3SG.NFUT-NFUT.help-TR 1PL.EXCL lots LOC school fee
 'Wild yam helps us a lot with school fees.' [20141106d_p06e016_18-9]

- (55) *Hii dutou-ni nou e neta xil holu*
 God 3SG.NFUT.help-TR 1SG LOC thing PL lots
 'God helps me with lots of things.' [20141218f_h01s074_19]

When *e* has a pronominal third-person singular object, it takes a suffix *-n*. This inflected form is also incorporated into verbs (§3.2.1).

- (56) *ut xa rali-ngngel e-n*
 place REL 1DU.DFUT-rest LOC-3OBJ
 'the place that we'll rest in.' [20141105f_p01e013_24]

The third main preposition used at the clausal level is *ni* which introduces instrumental participants. It could be derived from proto-Oceanic instrumental preposition **kini* (Pawley 1973: 142).

- (57) *lu-soxat-i ni uvuei sap xil mun*
 3DU.NFUT-cover-3OBJ INSTR leaf different PL again
 'They covered it back up with different leaves.'
 [20141027a_n01m001_90]

When it occurs without a following NP it indexes a third-person singular object that can be deduced from the context:

- (58) *ha mal-tihi neta tei a-mel ni*
 GO 1DU.EXCL.FUT-slice thing one CL.ED-1DU.EXCL.POSS 3OBJ.INSTR
 'We'll go slice a food of ours with it.' [20141027a_n01m001_65]

The preposition *ve* is used for nominal purposes (59), including beneficiaries (60) and goals (61). It is also used as a subordinating connective meaning 'in order to' or 'because'.

- (59) *o-sep men-ni ve neta xa mi-sa*
 2SG.NFUT.speak+laugh-TR for thing REL 3SG.NFUT-bad
 'You make fun of them for something bad' [20170217e_n01s142_06]

- (60) *na-be traehat ve famli na-van*
 1SG.NFUT-NFUT.COP try_hard for family CL.GEN-1SG.POSS
 'I work hard for my family.' [20150118b_n01m090_10]

- (61) *ma-kaakau ve hu*
 1PL.NFUT-walk for hill
 'we walked towards the hill' [20141121a_n01m041_32]

Like locative *e*, *ve* takes the *-n* suffix for a third-person singular pronominal object.

- (62) *Na-pre ve-n*
 1SG.NFUT-pray for-3SG.OBJ
 'I prayed for it.' [20170413a_h01m169]

A less frequent preposition is *re, ra* 'from', probably derived from proto-Oceanic **tani* (Pawley 1973: 142). This is perhaps being replaced with the more general locative *e*, which is used for sources as well as goals.

(63) *Song gur rat futbol re Jacky*
 Song 3SG.NFUT.take out football from Jacky
 ‘Song takes the ball from Jacky.’ [20141215f_k01s026_26]

Like *mi*, *re* takes the same pronominal object suffixes as a CV transitive verb (Table 27):

(64) *Jacky gur rat-i re-ni*
 Jacky 3SG.NFUT.take+out-3OBJ from-3OBJ
 ‘Jacky takes it from him’ [20141215f_k01s026_25]

The prepositions *usil* ‘about’ and *xat* ‘on’ seem to have emerged from verbs used in complex predicates (§7.1.3).

3.2.1 Incorporated-preposition verbs

Some very frequent verbs are derived from combinations of *ha* ‘go’ and *ta* ‘stay’ with the preposition *e*, and obligatorily take arguments that are more like objects of a preposition than of a transitive verb.

The verb *he* ‘go to’ is always followed a locative NP or a preposition that usually occurs after the locative preposition *e*, such as *nesau* in (66). It occurs once every 72 words in the subcorpus.

(65) *ma-be nahou*
 1PL.EXCL.NFUT-NFUT.go_to garden
 ‘we went to the garden’ [20150410a_h01s121_09]

(66) *lat-be nesau*
 3PC.NFUT-NFUT.go_to up
 ‘they went up’ [20141107b_n01e020_14]

To refer to a third-person singular pronominal location, speakers use the verb root *haen*, diachronically derived from the combination of *ha* ‘go’ and the inflected preposition *e-n* ‘LOC-3SG.OBJ’, which is pronounced as a single syllable.

(67) *ma-baen*
 1DU.EXCL.NFUT-NFUT.go_there
 ‘We went there.’ [20141220g_n01s080_37]

Similarly, *te* and *taen* are derived from a combination of the locative preposition *e* and the verb *ta* ‘stay’. The form *te* could equally be derived from *ti* ‘stay’ (§3.5.3), but the form of *taen* confirms that these are based on *ta*. *Te* occurs once every 192 words in the subcorpus.

- (68) *nou* *nat-te* *Penapo*
 1SG 1SG.PRI-be_at Penapo
 ‘I lived in Penapo’ [20141212g_p01s046_07]
- (69) *venu* *xa* *de* *nesau*
 volcano REL 3SG.NFUT.be_at up
 ‘the volcano that’s up there.’ [20141105f_p01e013_110]
- (70) *rute* *xa* *skul* *daen*
 place REL school 3SG.NFUT.be_there
 ‘the place where the school is’ [20170413a_h01m169]

There is contradictory evidence about the status of these roots as simple verb lexemes or analysable contractions of a verb and a preposition. On the one hand, speakers from Mele Maat allow the negative clitic to follow the locative argument of *te* and *he*, treating it as a nominal object rather than an oblique (§3.6.2).

- (71) *taa-re* *hat* *=ti*, *be* *pesin*
 3SG.NFUT.NEG-NEG.be_there stone =NEG 3SG.NFUT.go_to dish
 ‘It [a laplap] wasn’t on the stones, it went in a basin’
 [20150129c_i01m095_23]

There are a few examples where the negative clitic *ti* follows the pronominal object form of an incorporated-preposition verb, suggesting it is a simple verb root in these examples. This could be evidence of a change in progress in this area of Vatlongos syntax.

- (72) *inou na-bit* *xamil mul-naa-vaen* *=ti*
 1SG 1SG.NFUT-NFUT.say 2DU 2DU.FUT-NEG-NEG.go_there =NEG
 ‘I told you not to go there.’ [20141220g_n01s080_61]

But this example from the same speaker in the same narrative splits the verb *haen* into its component parts *ha* and *en*:

- (73) *rin ak mul-naa-va* *=ti e-n*
 side PROX 2DU.FUT-NEG-NEG.go =NEG LOC-3SG.OBJ
 ‘This side you won’t go to.’ [20141220g_n01s080_11]

Most speakers use *ha* ‘go’ and a locative NP or PP, instead of *he*, when a negative clitic is required, as in (74). Similarly, when an adverb is in post-verbal position, the verb *ha* is used followed by a PP argument headed with *e* in examples (47) and (48) above.

(74) *tati bit ral-naa-va =ti igak*
 dad 3SG.NFUT.say 1DU.INCL-NEG-NEG.go =NEG here
 ‘Dad said for us not to go here.’ [20141220g_n01s080_22]

Speakers did not accept any of these incorporated-preposition verbs with the nominalising suffix *-en* (§5.3.1), one diagnostic for verbal status.

3.3 Position of adverbs and adverbials

Adverbs can be defined as a syntactic class of lexemes that are restricted to adverbial functions of modifying predicates and propositions, especially situations and events. This means that they primarily modify clauses or verbs, and perhaps also adjectives and other adverbs although this rarely happens in Vatlongos. They can be distinguished from adverbials: words from other lexical classes and larger syntactic constituents used for the same purposes (Ernst 2002: 7–8). As adverbs are often derived from other lexical classes, it can be difficult to distinguish, for example, a noun being used as an adverbial, from an adverb that is derived from an also extant noun.

Adverbs and other adverbials can appear in various positions in a clause, but the main syntactic restriction is that no adverb can intervene between a verb or a preposition and its object. Only the partitive and negative clitics can appear between the verb and the object. Figure 7 is excerpted from Figure 6 above, and shows where adverbs can appear relative to the main constituents of a clause. The negative clitic *ti* can also appear before the object, but this is not shown in Figure 7 as it does not affect adverb position: no adverb can appear between a verb and its object. As mentioned above, ‘other arguments and modifiers’ includes PPs, complement clauses and the subsequent verb phrase in an SVC, which are all syntactically adjoined to the VP resulting in a complex VP. In fact, the post-verbal and clause-final positions can be thought of as occurring in this same position, rather than preceding or following it. As shown in Table 16, they are not distinguished by any semantic classes of adverbs.

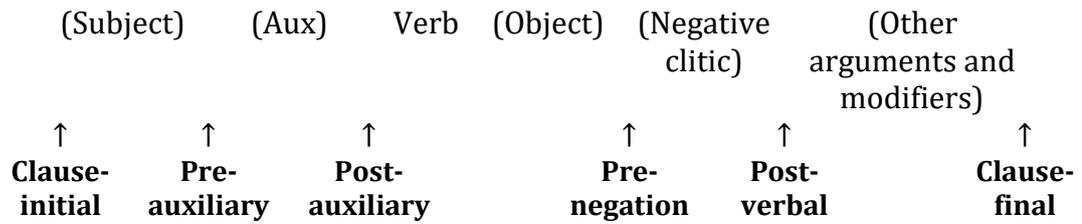


Figure 7: Adverb positions in a clause

These positions were targeted via elicitation of grammaticality judgements in eleven frame sentences for 21 adverbs, including affirmative and negative statements with intransitive and transitive verbs, PP arguments, SVCs and AVCs. It was found that adverbs appear in these different positions depending on their intended semantic scope, as shown in Table 15. However, as all the constituents except the verb are optional, the position and scope of an adverb in any specific example from the corpus is often ambiguous. When there is no auxiliary in a clause, and an adverb appears between the subject and the verb, I refer to this position as ‘pre-verbal’.

Table 15: Semantic scope of adverbs in different positions in the clause

Adverb position	Scope of the adverb
Clause-initial	Whole clause (especially its relationship to preceding clauses)
Pre-auxiliary	Predicate (excluding subject and clause-level adjuncts)
Post-auxiliary	Predicate except aspectual information from the auxiliary
Pre-negation	Predicate except negation
Post-verbal	Predicate except information in any subsequent arguments/adjuncts
Clause-final	Whole clause

The scope of the adverb in the ‘pre-negation’ position, where it appears before the negative clitic *ti*, is perhaps better thought of as being determined by the scope of the negative clitic rather than the adverb. The meaning of an adverb in this position is negated, while if it appears after the negative clitic the adverb modifies the negated situation.

(75) *Mael taa-xa viak tavatang =ti*
 Mael 3SG.NFUT.NEG-NEG.eat taro very_much =NEG
 ‘Mael doesn’t eat taro very much.’ [20170326g_x01s026]

(76) *Mael taa-xa viak =ti tavatang*
 Mael 3SG.NFUT.NEG-NEG.eat taro =NEG very_much
 ‘Mael really doesn’t eat taro’ (i.e. he refuses to eat taro all the time, too much) [20170326g_x01s026]

As the scope of adverbs corresponds to their position in the clause, the distribution of an adverb can be predicted by its semantic class, as shown in Table 16. A relationship between the syntactic distribution of adverbs and their semantic features is well established cross-linguistically (Jackendoff 1972: 47–107; Ernst 2000; 2002).

Table 16: Semantic classes of adverbs and their distribution

Semantic type	Examples	Possible positions
Discourse markers - showing how the clause relates to situations in other clauses.	<i>Maxani</i> ‘then suddenly’ <i>Tama</i> ‘then’	Clause-initial * Pre-auxiliary * Post-auxiliary * Pre-negation * Post-verbal * Clause-final
Modal – evaluating the likelihood of a situation	<i>Xos</i> ‘maybe’	Clause-initial Pre-auxiliary * Post-auxiliary * Pre-negation Post-verbal Clause-final
Absolute time and place – positioning the situation in time, relative to utterance time; and place, relative to deictic centre	<i>Igak</i> ‘here’ <i>Nogonoh</i> ‘yesterday’ <i>Taxeak</i> ‘now’	Clause-initial * Pre-auxiliary * Post-auxiliary * Pre-negation Post-verbal Clause-final
Relative time – positioning the situation in time, relative to other situations in the discourse	<i>Mu</i> ‘first’ <i>Mun</i> ‘again’ <i>Ngamu</i> ‘already’ ¹⁴ <i>Sung</i> ‘first’ <i>Tamu</i> ‘still’ <i>Tuei</i> ‘already’ <i>Vari</i> ‘at once’	* Clause-initial Pre-auxiliary Post-auxiliary * Pre-negation Post-verbal Clause-final

¹⁴ *Ngamu* ‘already’ is exceptional as it cannot appear in the pre-auxiliary position, although I have not identified any semantic differences with *tuei* ‘already’ which can appear in the pre-auxiliary position. Perhaps this is a kind of syntactic specialisation of two competing synonyms, although in the corpus both are overwhelmingly more likely to appear in post-verbal or clause-final position.

Semantic type	Examples	Possible positions
Relative time proximals – positioning the situation in time relative to other situation in the discourse, and pointing out this situation	<i>Tuok</i> ‘then, at that point’ <i>Sumok</i> ‘just, just then’	* Clause-initial * Pre-auxiliary * Post-auxiliary * Pre-negation Post-verbal Clause-final
Degree – showing the extent of a gradable situation	<i>Tavatang</i> ‘very much’ <i>Vuteili</i> ‘a bit’ <i>Pupu</i> ‘a lot’	* Clause-initial * Pre-auxiliary * Post-auxiliary Pre-negation Post-verbal Clause-final
Restriction – distinguishing this situation from other possible situations	<i>Kestang</i> ‘just, only’ <i>Tang</i> ‘just, only’	* Clause-initial Pre-auxiliary * Post-auxiliary * Pre-negation Post-verbal Clause-final
Manner – specifying the manner of a situation	<i>Kamanon</i> ‘fast’	* Clause-initial * Pre-auxiliary * Post-auxiliary * Pre-negation Post-verbal Clause-final

It is possible for other lexical classes to function as discourse markers as well, indicating semantic and temporal relationships with situations in other clauses, but cross-linguistically discourse markers are usually adverbs (Urgelles-Coll 2010: 1).

Relative time proximals are listed as a separate class from relative time adverbs due to the coincidence of their more restricted distribution and their morphological form: both seem to incorporate the proximal demonstrative *ak*, *ok*. The demonstrative always follows the NP it modifies, so perhaps the relative time proximals are also required to follow the constituent they modify, in this case a verba constituent.

There are additional restrictions on some adverbs based on their semantics. For example, some adverbs cannot cooccur with stative predicates or negative predicates (e.g. *mu* ‘first’, *tamu* ‘still’). The adverb *mu* ‘first’ cannot occur in post-auxiliary position with the continuous auxiliary *ti*, but can follow the prior motion auxiliaries (Chapter 6). Degree adverbs can only scope over parts of the

clause that have a gradeable semantic component, so when a PP or serialised VP is not gradeable, the clause-final position is not available to this type of adverb. In (77) the subsequent verb in the SVC is only contributing path information, so it is infelicitous for *tavatang* ‘very much’ to appear at the end of the clause.

- (77a) *Mael mi-kaakau tavatang be bien*
 Mael 3SG.NFUT-walk too_much 3SG.NFUT.go_to sea
 ‘Mael walks to the sea too much.’
- (b) ? *Mael mi-kaakau be bien tavatang*
 Mael 3SG.NFUT-walk 3SG.NFUT.go_to sea too_much

The modal adverb *xos* ‘maybe’ is unusual in being able to appear in multiple positions in a single clause, to emphasise the speaker’s uncertainty or hesitancy (241). In example (78), *xos* appears between a discourse marker and the subject, then between the subject and the verb, and then again after the object. In (79), *xos* occurs at the beginning of the clause and after the object before the PP. Both these examples were taken from recordings using the *Frog, where are you?* storyboard, and gives each speaker’s hypothesis about why the boy is holding his nose after looking into a rat hole; there is no picture of the rat scratching his nose, or the boy and the rat bumping noses, so these are the speakers’ conjectures.

- (78) *e xos asu xos mi-kkal vulkusi-n xos*
 and maybe rat maybe 3SG.NFUT-scratch nose-3SG.POSS maybe
 ‘And maybe the rat scratched his nose’ [20141212h_n01s046_35]
- (79) *Xos mi-sen vulkusi-n xos e vulkusi*
 maybe 3SG.NFUT-bang nose-3SG.POSS maybe LOC nose
na tut haramue ak
 CL.GEN little boy PROX
 ‘Maybe it banged its nose against the nose of the little boy’
 [20150118a_n01m090_32]

There are some common polysemies with adverbs and other lexical classes that can confuse the interpretation of their distribution. The adverbs of restriction *kestang* ‘just, only’ and *tang* ‘just, only’, as well as *mun* ‘again’, are homonymous with nominal quantifiers and therefore sometimes appear to occur in the pre-verbal position when they are in fact modifying the subject (where *mun* has the meaning of ‘too’ rather than ‘again’). Other adverbs (e.g. *vari* ‘at once’, *kestang* ‘just, only’) can also occur as the second part of a complex predicate, and thus

intervene between a verb and its object, although they always take the transitivising suffix *-ni* (§7.1).

As discussed above, other word classes and syntactic constituents can be used as adverbials, but there seems to be a preference for them to appear in clause-initial or post-verbal positions, especially if they are syntactically complex. The most restricted position is the post-auxiliary position, where only adverbs of relative time can appear. There are no examples of syntactically complex constituents appearing here, so it may be that there is a syntactic as well as a semantic restriction on this position.

Many adverbs in Vatlongos are phonologically light with a CV syllable structure, or have CV variants, as shown in Table 17. These are cognate with ‘bound modifier’ clitics in Paamese (Crowley 1982: 229–337), but do not affect word stress in adjacent words in Vatlongos so are not analysed as clitics.

Table 17: Phonologically reduced adverbs

<i>Li</i>	‘alas, strong emotion’
<i>Mu</i>	‘first’
<i>Sung, su</i>	‘after, then’
<i>Tang, ta</i>	‘just, only’
<i>Tuei, tu</i>	‘already’

Adverbs can also appear with non-verbal predicates. In examples in §3.5.1, they follow, (110), (111), (113), or precede (117) the predicate, including appearing between the topic and the predicate (114).

3.3.1 Incorporated-adverb verbs

There are a few verbs that appear to be derived from incorporation of *tang* ‘just only’ or its phonologically reduced form *ta*: *hat* ‘just go’, *tat* ‘just stay, stay put’ and *tit* ‘just stay, stay put’.

(80) *mei* *ni-hat*
 COME 1SG.DFUT-just_go
 ‘I’ll come and just go’ [20150305d_n01s109_14]

(81) *ni-tat* *tass-ok*
 1SG.DFUT-just_stay self-1SG.POSS
 ‘I’ll just stay by myself.’ [20150303a_x01e016_29]

(82) *mei ral-tit i-mak*
 COME 1DU.INCL.DFUT-just_stay 3SG.DFUT-like_this
 'We'll come and stay like this.' [20150223a_n01m096_59]

Unlike some of the incorporated-preposition verbs in §3.2.1, these are fully unverbated in all TAMP environments, and are followed by the negative clitic *ti*:

(83) *ni-naa-vat =ti*
 1SG.DFUT-NEG-NEG.just_go =NEG
 'I won't just go.' [20150305d_n01s109_16]

They can cooccur with the full adverb *tang* if another word intervenes:

(84) *li-tit vote tang*
 3PL.DFUT-just_stay quiet just
 'They'll just stay quiet.' [Y3P3_17]

Further evidence that these are well established lexicalised compounds is that *tit* and *tat* are the lexical sources for the minimising continuous auxiliary verbs (Chapter 8). However, only *hat* can be nominalised with the nominalising suffix *-en* (§5.3.1).

3.4 Verbs with additional syntactic constraints

Some verbs place additional syntactic constraints on the clauses they appear in: the reflexive verb *hesi*, and psycho-collocations: a class of verbs pertaining to emotion and mental state that require a possessed body part as their syntactic subject.

3.4.1 Reflexive verb *hesi*

A reflexive verb *hesi* 'be happy' differs from other transitive verbs in that it must take an independent pronoun object that is coreferential with the subject. It cannot appear without an object and never takes the *-ni* transitiviser, so appears to be basically transitive with additional restrictions. It cannot take an object suffix, or be followed by any other NP, even one that is co-referential with the subject.

Examples (85) to (88) show how this works with some different subject persons and numbers.

However, as free pronouns are often used for animate objects this does not rule out a non-reflexive reading, as in (93), which could be ambiguous without context.

- (93) *di* *la-gal* *xil*
 CONT.REAL 3PL.NFUT-NFUT.buy 3PL
 ‘They [*people from Paama*] were buying them [*pigs*].’
 [20141215b_h01s062_06]

There is also a reflexive verb that is restricted to appearing as the subsequent verb in complex predicates (§7.1). *Bin* follows verbs of consumption or perception, and means that the initial verb was performed to completion: to eat until full in (94), to drink until dead, i.e. drown, in (95) and feel exhausted in (96).

- (94) *a-ga* *bin-uk* *a?*
 2SG.NFUT-NFUT.eat complete-2SG.OBJ yes
 ‘You’re full eh?’ [20141116b_c01m028_33]

- (95) *meletin* *ak* *ba* *mun* *bin* *xi* *tang*
 person PROX GO 3SG.NFUT.drink_lots complete 3SG just
 ‘this person just went and drowned himself’ [20170217a_n01s038_28]

- (96) *la-long* *bin* *xil*
 3PL.NFUT-feel complete 3PL
 ‘they were worn out’ [0170413a_h01m169_129]

Transitive verbs participating in these reflexive or reciprocal constructions need not be considered a separate lexical class, as suggested by Parker (1970a: viii). Rather, *hesi* ‘happy’ is a syntactically transitive verb with an irregular lexical constraint that the object must be coreferential with the subject, and expressed as an independent pronoun rather than with an object suffix.

Crowley (1982: 70) describes a class of verbs with the same constraint in closely related Paamese, including *sii* ‘happy’ and *vinii* ‘do to excess/death’. Lamén, the language of a small island off North Epi, has a class of reflexive verbs, which are especially to do with motion (Early 2002: 676–677). Moysé-Faurie (2008: 109, 114–115) gives a summary of other Oceanic languages with similar classes of verbs, including languages in Vanuatu, New Caledonia, Fiji and Papua New Guinea. While they are usually described as ‘reflexive verbs’, she points out that they are restricted to describing change of orientation or

emotion and excessive behaviours, suggesting they may be better analysed as expressing a subset of middle situations.

3.4.2 Psycho-collocations

A small class of verbs that mainly express emotions or mental states take a default subject that is marked for possession by the experiencer. The possessed subjects are mostly body parts, although the meaning of some of them is unknown to contemporary speakers, or are not transparent when used in these constructions. They are all inalienably possessed bound nouns. Similar verbs in Daakaka (West Ambrym) have been described as ‘psycho-collocations’ by von Prince (2017a), who draws parallels between these structures in many Oceanic languages and similar structures in South East Asian languages. A similar strategy is used in North Ambrym (Franjeh 2012: 221–222). Psycho-collocations (described as ‘body part metaphors’) involving the liver and the insides have been reconstructed for Proto-Oceanic (Osmond 2016), and put forward as a typological feature distinctive of Austronesian languages in Central/Eastern Indonesia (Klamer 2002: 369).

Hei ‘like, want’ (97), *bong, vong* ‘forget’ (98) and *kat, xat* ‘angry’ (102) all take a default possessed subject *e* which can mean ‘inside’ or ‘heart’, but speakers do not feel it has a meaning beyond its grammatical function in these constructions. Cognates of these verbs taking the same possessed subject are also found in Paamese (Crowley 1995a: 409), and a lexeme with a similar derivation is also a common possessed subject in Daakaka (von Prince 2017a: 7). The verb *kakan* ‘afraid’, (99), (100), (101), takes a possessed subject *uli* which speakers do not associate with a specific meaning, although it has a cognate in Paamese which Crowley (1995a: 407) translates as ‘seat of emotions’. According to Parker (1970a: 37) it could also be used with the verb *sa* ‘bad’ to mean ‘feel shame’, a combination which has fallen out of use.

- | | | | | |
|------|--|---------------|---------------|------------|
| (97) | <i>e-m</i> | <i>i-hei</i> | <i>u-ling</i> | |
| | inside-2SG.POSS | 3SG.DFUT-want | 2SG.DFUT-put | |
| | <i>na</i> | <i>ohia</i> | <i>hu</i> | <i>top</i> |
| | hes | wild_yam | new | one |
| | ‘You’ll want to put a new wild yam down.’ [20141106d_p05e016_05-6] | | | |

(98) *e-ok* *bovong-ni* *is-en* *mun*
 inside-1SG.POSS 3SG.NFUT.forget-TR name-3SG.POSS again
 ‘I forget his name again.’ [20150219b_n01s001_95]

The verb is always marked for third-person singular subject agreement, regardless of the number and person of the experiencer, as shown in the plural examples.

(99) *uli* *xalu* *mi-kkan* *turei* *xi-e*
 UNKN 3DU 3SG.NFUT-afraid already 3SG-CONTR
 ‘They were scared already now.’ [20141220g_n01s080_30]

(100) *uli-mem* *mi-kkan-ni* *simen* *xil-e*
 UNKN-1PL.EXCL.POSS 3SG.NFUT-afraid-TR cement PL-CONTR
 ‘We were scared of all this concrete.’ [20150226a_n01s098_61]

The experiencer can also be included in the clause as an NP but is in the clause-initial topic or focus position (§3.1.2), rather than being the grammatical subject; it does not trigger subject agreement.

(101) *xi*, *uli-n* *mi-kkan*
 3SG UNKN-3SG.POSS 3SG.NFUT-afraid
 ‘He was scared.’ [20150226a_n01s098_51]

(102) *inou* *e-ok* *gat-i*
 1SG inside-1SG.POSS 3SG.NFUT.angry-3OBJ
xa *e-ok* *gat-i*
 SUB inside-1SG.POSS 3SG.NFUT.angry-3OBJ
 ‘I was very very angry with him.’ [20150419e_h01m128_36]

The nominalised forms of these verbs are compounds of the uninflected subject and the uninflected verb, followed by the nominaliser *-en*, as in examples (103), (104) and (105). However, these are fairly infrequent: all the examples are elicited or translated from Bislama.

(103) *xi* *di* *mi-kaakau* *ba*
 3SG CONT.REAL 3SG.NFUT-walk 3SG.NFUT.go
del *xat* *ekkat-en* *na-n*
 3SG.NFUT.with affect angry-NMLZR CL.GEN-3SG.POSS
 ‘He was walking along with his anger.’ [Y2R4_19]

(104) *O-ling* *ut* *etan* *ak*
 2SG.NFUT-put place down PROX
danga *del* *ehei-en*
 3SG.NFUT.stay_behind 3SG.NFUT.with like-NMLZR
 ‘You leave this world behind with love.’ [Y3R8_150-1]

- (110) *Na-ratel* *tang?*
 CL.GEN-1PC.INCL.POSS just
 “Ee,
na-mtel *na-mtel!*”
 no CL.GEN-2PC.POSS CL.GEN-2PC.POSS
 “Is this just ours?”
 “No, it’s yours, it’s yours!” [20141106f_n01e018_17]

Numerals can also be used as the predicate of a clause, as in (111). The previous clause describes one of two harnesses breaking.

- (111) *ma* *tei* *takes*
 then one only
 ‘then there was only one.’ [20170224a_n01s141_124]

PPs can also be used predicatively:

- (112) *Inou* *tahal* *Abrim*
 1SG of.LOC Ambrym
 ‘I am from Ambrym’ [20150129c_i01m095_08]

NPs can also be used predicatively. In examples (113) and (114), the first NP is the subject and the second is the predicate. The NPs are delineated with [square brackets]. In the first example there is a prosodic break after the long first NP suggesting it is in clause-initial topic position to further emphasise the revelation at the end of a custom story about the origin of yam.

- (113) [*huram xa xir di ra-ga-ni*]
 yam REL 1PL.INCL CONT.REAL 1PL.INCL.NFUT-NFUT.eat-3OBJ
 [*xosali ak*], [*tatal*] *tang*
 today PROX snake just
 ‘The yam that we eat these days, it’s just snake.’
 [20141106f_n01e018_40]

- (114) [*ngan-ak*] *xos* [*vueili hokkorong tang*]
 one-PROX maybe pig small just
 ‘This one is maybe just a small pig’ [20150305f_p01s110_37]

NPs can also be used existentially, especially as existential presentatives (Givón 1983: 25–35) introducing the characters in a narrative:

- (115) *Haromue* *tei* *e* *kuli* *e* *frog.*
 boy one and dog and frog
 ‘There’s a boy and a dog and a frog.’ [20141212h_n01s046_03]

- (116) *Va* *tei,* *maletin* *tei*
 time one person one
 ‘Once upon a time, there was a person’ [20150304f_n01e024_01]

Pronouns can also be used existentially in this way, and the modality of the existence can be modified by adverbs, like *xos* in example (117).

- (117) *xos* *tava* *mun* *mu* *xos* *i-he*
 maybe one more or maybe 3SG.DFUT-COP
tovuol *mun*
 absent again
 ‘Maybe there’ll be another one [*a cyclone*] or maybe there won’t be anymore’ [20150410a_h01s121_42]

3.5.2 Copular verb *he*

A copular verb *he* can be followed by an adjective, possessive classifier or numeral phrase. It is also used with NPs, including pronouns, both for nominal predicates like those expressed in verbless clauses, and for equational meanings, where the NP is referential and the subject and predicate can be switched without a change in meaning (Dryer 2007: 233). *He* is described as an equational verb by Parker (1970a: vii). The copular verb is also used as a strategy for incorporating loan verbs.

He is analysed as a verb because it takes verbal prefixes (§5.1) and undergoes verb-initial consonant mutation (§5.2.1), occurring as *be*, *ve* and *me* in different TAMP environments. However, as it cannot function as a predicate independently, it is the complement of the copular that is considered the predicate in these clauses. This analysis is confirmed by the behaviour of the negative clitic *ti* which always occurs after the complement of the copular, whereas it can occur before or after an object NP (§3.6.2.3).

Example (118) shows the copular verb *he* used with an adjective phrase.

- (118) *be* *kokkorong* *tang*
 3SG.NFUT.COP small just
 ‘It’s only small.’ [20141106d_p04e016_05]

In example (119), the copular verb is followed by a possessive classifier.

- (119) *be* *a* *famli* *xil* *ak*
 3SG.NFUT.COP CL.ED family PL PROX
 ‘It’s these relatives’ food.’ (‘It’s for these relatives.’)
 [20150219a_p01m001_50]

In example (120), the copular verb is twice used with numerals. The first negated numeral phrase is modified with restrictive *takres* ‘only’, and shows

that the negative clitic *ti* follows the complement of the copular verb, an argument for its predicative status.

- (120) *venu* *taa-ve* *tei* *takres* =*ti*,
 volcano 3SG.NFUT.NEG-NEG.COP one only =NEG
be *lu*.
 3SG.NFUT.COP two
 ‘The volcano isn’t just one, it’s two.’ [20141105f_p01e013_121]

Example (121) shows the copular verb followed by a noun in a simple NP.

- (121) *Ngan-ak* *be* *hoxalite*.
 one-PROX 3SG.NFUT.COP end
 ‘That’s the end.’ [20141212l_n01s046_7]

Example (122) shows the copular verb followed by a pronoun in a simple NP. This is an example of an equational clause.

- (122) *tati* *na-e* *be* *inou*
 dad CL.GEN-3PL.POSS 3SG.NFUT.COP 1SG
 ‘Their Daddy is me.’ [20141107d_p01s022_23]

The copular verb *he* is also used as a strategy for incorporating loan verbs from Bislama, as in (123). In this section, I have glossed the verbs borrowed from Bislama BIS for clarity.

- (123) *skul* *te-he* *klos*
 school 3SG.PRI-COP close.BIS
 ‘The school closed.’ [20150219b_n01m001_007]

This copular strategy is also found in Paamese (Crowley 1982: 171–172; 1991: 192), Bierebo (Budd 2011) and other languages of Epi (Early 2004). Vatlongos is the northernmost language to exhibit this areal feature, which is not found in languages of North or West Ambrym (Franjeh, p.c.; Krifka, p.c.; Von Prince, p.c.). Following Wohlgemuth’s (2009: 102–117) typology of verbal borrowings, this is a light verb strategy, the second most common borrowing strategy in his cross-linguistic sample. The most common strategy is direct insertion (Wohlgemuth 2009: 87–92), where a loan verb is treated identically to any other verb in the language, and this is also a possibility in Vatlongos. Example (124) shows the same speaker in the same narrative using the direct insertion strategy a few clauses later.

- (124) *xamem mat-naa-skul =ti*
 1PL 1PL.EXCL.PRI-NEG-study.BIS =NEG
 'We didn't go to school'. [20150219b_n01s001_009]

The direct insertion strategy is common with well-established loan words like *skul* 'study, go to church' whereas the copular strategy is preferred with more recent borrowings, and nonce borrowings or one-word code switches (Wohlgemuth 2009: 53). For example, in (125), the verb *rekotem* 'record' is very specific to the unusual situation of the recording event:

- (125) *o-biteni e xi be rekotem-ni*
 2SG.NFUT-NFUT.speak and 3SG 3SG.NFUT.COP record.BIS-TR
 'You spoke and she recorded it.' [20141121a_n01m041_83]

A similar pattern is observed for Paamese and Epi languages, but Vatlongos seems to allow more variability in the choice of strategy, whereas direct insertion is restricted to older loans in languages to the South. A higher proportion of verbal borrowings use the copular strategy in texts by speakers from Ase-Taveak and Endu than Mele Maat (Ridge 2017b).

The behaviour of the Vatlongos transitivity suffix *-ni* (§5.1.3) and the Bislama transitive suffix *-Vm* (Crowley 2004b: 77–81) complicate the characterisation of the copular strategy as an example of a light verb strategy. All Bislama verbs seem to be borrowed as intransitive stems, requiring the use of the transitivity suffix *-ni* to take an object NP, or to be understood as transitive. In this regard they behave like Vatlongos intransitive verbs, and the strategy partially resembles direct insertion, in that they take a suffix like this subset of native verbs.

However, most speakers also regularly inflect the verbs with the Bislama transitive suffix *-Vm*. In example (125) the Bislama verb *rekotem* could be further analysed as a root *rekot* and the suffix *-em*, marking it as transitive. Example (126) shows the same Bislama root *tij* 'teach' borrowed without the Bislama transitive suffix with an intransitive meaning, and marked with both the Bislama and Vatlongos transitive suffixes when followed by an object NP. This shows that the meaning contribution of the Bislama suffix is maintained when it is borrowed.

- (126) *na-be* *tij* *ena* *Sade skul.*
 1SG.NFUT-NFUT.COP teach LOC Sunday_school
Na-be *tijim-ni* *lesen* *tei*
 1SG.NFUT-NFUT.COP teach.BIS-TR lesson one
 'I teach at Sunday school. I teach a lesson' [20150118b_n01m090_25-26]

This is an example of the much rarer verbal borrowing strategy of paradigm insertion, where a verb is borrowed with verbal inflectional morphology from the donor language. Wohlgemuth (2009: 118) finds that this strategy 'only occurs in intensive contact situations, involving widespread bilingualism', and this is certainly true of the language contact situation for Vatlongos (§2.2).

3.5.3 Locative predicates and existential clauses

Vatlongos uses a set of three locative verbs, *ti*, *te*, and *ta*, roughly glossed as 'stay' or 'be at', to express locative predicates: clauses which predicate the location of the subject. Two of these verbs are also used in existential clauses, asserting the existence of the subject or introducing a referent to the discourse. This is a common functional overlap cross-linguistically (Dryer 2007: 240–241). The locative verb *te* differs from *ti* and *ta* in requiring a locative NP immediately following the verb. This is because it appears to be derived from a combination of *ta* and the locative preposition *e* (§3.2.1). *Ti* and *ta* are distinguished by more subtle differences in their distribution. Both are lexical sources for continuous auxiliary verbs (§8.2.1), so their lexical and discourse-functional distribution is worth discussing here.

Ti is the most frequent locative verb occurring once every 100 words in the subcorpus. It usually functions as a locative predicate, followed by an NP or PP expressing a location:

- (127) *mei* *lu-di* *tim*
 COME 3DU.NFUT-NFUT.stay home
 'They came and stayed at home.' [20141027a_n01m001_12]
- (128) *atuli* *xalu* *lu-di* *bien*
 girl DU 3DU.NFUT-NFUT.stay sea
 'The two girls were at the sea.' [20141107b_n01e020_02]

- (129) *inou nat-ti Santo*
 1SG 1SG.PRI-stay Santo
 'I was in Santo.' [20150219b_n01m001_57]

The kinds of locations that follow *ti* are usually fairly large, imprecise areas that are readily identifiable in context. When the location is recoverable from the context or discourse, *ti* can be used independently, as in (130). For this reason, *ti* is analysed as an intransitive verb, and any following locative NP does not function as an object.

- (130) *lu-mmei tim; mei lu-di,*
 3DU.NFUT-come home COME 3DU.NFUT-NFUT.stay
 'They came home; they came and stayed [*there*]'
 [20141027a_n01m001_19]

However, even when the location is recoverable from context, a pro-form of the locative preposition can be used:

- (131) *Xamem ma-di e-n*
 1PL.EXCL 1PL.EXCL.NFUT-NFUT.stay LOC-3OBJ
 'We stayed in it [*a hole*]' [20150219b_n01m001_16]

Ti can also be used without a location phrase, especially at the beginning of narratives (132) or procedural texts (133), to establish the background for events or procedures. It can be difficult to translate this function into English, but, depending on the context, it can be roughly glossed as 'live', 'be there' or 'hang around'.

- (132) *Mata-di vongien,*
 1PC.EXCL.NFUT-NFUT.stay night
 'We were there at night,' [20150303h_n01e106_01]

- (133) *taem xa ma-di, ma-di*
 time REL 1PL.EXCL.NFUT-NFUT.stay 1PL.EXCL.NFUT-NFUT.stay
ma tas mes,
 then sea 3SG.NFUT.dry
 'When we're there, we're there and the tide goes out,'
 [20141028b_p02m003_1]

It is also frequently used immediately after the introduction of characters in a narrative, before the events of the narrative begin:

- (134) *tumen* *lu,* *tei* *ise-n* *mai,*
bird two one name-3SG.POSS pigeon
tei *ise-n,* *na,* *koh.*
one name-3SG.POSS HES heron
Lu-di *maa...* *lu-be* *bien*
3DU.NFUT-NFUT.stay on_and_on 3DU.NFUT-NFUT.go_to sea
‘There were two birds, one was called pigeon and one was called
heron. They were living... they went to the sea.’
[20141220d_n01s082_01-03]
- (135) *mai* *xal* *as,* *lu-di* *maa...*
pigeon with ant 3DU.NFUT-NFUT.stay on_and_on...
‘There was a pigeon and an ant, they were living...’
[20170119h_n01s088_03]

In both these examples *ti* is followed by the durative discourse marker *maa* ‘on and on’ (§4.3.2). During a narrative, *ti* is often marked for durative with repetition as well to mark the passing of time, without any locative constituent.

- (136) *La-di* *la-di*
3PL.NFUT-NFUT.stay 3PL.NFUT-NFUT.stay
la-di *maa...*
3PL.NFUT-NFUT.stay on_and_on
‘They carried on and on...’ [20150226a_n01s098_05]

Ta is less frequent than *ti* occurring every 217 words in the subcorpus. Parker’s (1970a: 28, 33) dictionary does not have separate entries for the lexical and auxiliary sense of *ti* and *ta*, and his proposed distinctions between *ti* and *ta* are discussed in more detail in relation to the continuous auxiliaries in §8.2.1. His observation that *ta* marks situations as relatively more ‘definite as to time, place or other circumstances’ is perhaps true of lexical *ta*. Compared to *ti*, *ta* is more likely to occur without a locative NP or PP, and instead refers to a place that is definite in the sense that it is accessible from the discourse or wider context of speech. In example (137), the understood location is the location of the speech act, and in (138) it is the current setting of the narrative. In (139), the location, school, is accessible from the previous clause, while in (140) a locative NP is in clause-initial position, and refers back to Tongoa in the previous clause.

- (137) *xouk* *a-da*
2SG 3SG.NFUT-NFUT.stay
‘you are here’ [20150419c_h01m004_54]

- (138) *sup taa-ra ti*
 chief 3SG.NFUT.NEG-NEG.stay NEG
 ‘The chief wasn’t there.’ [20170222d_n01s152_17]
- (139) *na-mmei mun skul; ma na-da.*
 1SG.NFUT-come again school then 1SG.NFUT-NFUT.stay
 ‘I came back to school; then I stayed there.’ [20170119f_n01s136_27]
- (140) *Rute nen ak, nine-n tei da*
 place of_it PROX mother-3SG.POSS one 3SG.NFUT.stay
 ‘In this same place, an aunt of his lived.’ [20141208a_n01m045_15]

Example (510) below shows *ta* modified with a subsequent verb in an SVC that expresses specific temporal bounds. As a subsequent verb in SVCs (§6.4.4), *ta* can often be translated as ‘away’ or ‘to one side’, as in (247) below.

Similarly, *ta* is used to question an exact location: every speaker who decided to translate the title of the frog story ‘Frog, where are you?’ used *ta*:

- (141) *frog o-da xavi?*
 frog 2SG.NFUT-NFUT.stay where
 ‘Frog, where are you?’ [20150118a_n01m090_15]

Ta also often occurs with the adverbs *tang* ‘just’ and *tamu* ‘still’:

- (142) *mama na-da tang*
 mum 1SG.NFUT-NFUT.stay just
 ‘Mum I’m just here.’ [20170220i_n01e149_09]
- (143) *Nat-ta tamu tim sa-van*
 1SG.PRI-stay still home CL.DOM-1SG.POSS
 ‘I was still living at home.’ [20141211a_n01s046_02]

Like *ti*, *ta* can be marked for durative and used to show time passing, in example (144) in instructions for growing yam.

- (144) *ma i-ta i-ta ve...*
 then 3SG.DFUT-stay 3SG.DFUT-stay on_and_on
 ‘They it will carry on and on...’ [20141106d_p10e016_10]

Ta can also be used in existential clauses, to assert the existence of the subject or introduce it into the discourse.

- (145) *Wajman tei na-metel da*
 watchman one CL.GEN-1PC.EXCL.POSS 3SG.NFUT.stay
 ‘We had a watchman.’ [20150303h_n01e106_17]

The high frequency and semantically bleached discourse functions of these two 'stay' verbs are probably factors in their reanalysis as auxiliary verbs.

3.6 Clitics in the clause

A clitic can be defined as an independent syntactic word that is phonologically or prosodically dependent on another word (Zwicky & Pullum 1983; Toivonen 2001: 37; Bickel & Nichols 2007: 175). Their phonological form resembles an affix, while their distribution resembles an independent function word (Spencer & Luís 2012: 140).

There are two distinct clitics in the Vatlongos clause that have the same phonological form: /ti/. This is cross-linguistically a canonical form for a clitic, as it is a monomoraic CV syllable, and forms a single prosodic word with the element it attaches to (Spencer & Luís 2012: 132). However, in some positions in the clause the negative clitic can be a prosodically independent word, unlike the partitive enclitic which is always prosodically dependent on the preceding word.

However, the distribution of the two clitics is very different. The negative clitic is more similar to an independent function word in that it can appear in many positions in the clause, attaches to members of many different lexical classes and attaches to the boundary of a syntactic constituent rather than its head word (Bickel & Nichols 2007: 175–176; Spencer & Luís 2012: 134). The partitive clitic nearly always attaches to a verb, making it more similar to an affix; the cognate form in Paamese is analysed as a verbal suffix (Crowley 1982: 144). However, the Vatlongos partitive can also follow NPs, affirming its status as a clitic.

3.6.1 Partitive enclitic =*ti*

The partitive enclitic =*ti* usually modifies a predicate and appears immediately after the verb.

- (146) *inou na-ma na-barang =ti*¹⁵
 1SG 1SG.IFUT-IFUT.go 1SG.IFUT-walk_around =PART
na-kaakau =ti e
 1SG.IFUT-walk =PART LOC
aelan xil ak
 island PL PROX
 ‘I’m going to travel a bit, I’m going to visit these islands a bit.’
 [20141208a_n01m045_07]

With transitive verbs the predicate scope can be interpreted as restricting the number of an object, as in (147). This wide scope is one of the criteria for cliticness (Spencer & Luís 2012: 134). On the other hand, a narrow predicate scope is obligatory when the object is a singular pronoun, as in examples (153) and (155), or an NP that is explicitly marked singular, as in (635) below.

- (147) *Mal-tup =ti tumen xil a-ratel*
 1DU.EXCL.FUT-shoot =PART bird PL CL.ED-1PC.INCL.POSS
 ‘We’ll shoot some birds for us all.’ [20141220g_n01s080_07]

One reason for analysing this form as a clitic rather than an adverb is that it intervenes between a verb and its object, a position that is not available to other adverbs (Figure 7).

It is also phonologically dependent on the word it attaches to and part of the same prosodic word, a criterion for cliticness (Spencer & Luís 2012: 132). This is demonstrated by the shift in word stress that sometimes occurs when the partitive clitic *=ti* is used. Main word stress in Vatlongos falls on the final syllable if that syllable has a consonant coda or a long vowel (analysed as bimoraic syllable structures), and the penultimate syllable if the final syllable has a short vowel (analysed as monomoraic syllables), as in (148).

- (148) 'la.væ
la-va
 3PL.IFUT-IFUT.go
 ‘they’ll go’ [20170413e_n01m030_25]

¹⁵ To reflect their status as independent grammatical words, clitics are aligned separately in examples, but the prosodic link with the element they attach to is marked with the clitic boundary marker =.

When the partitive clitic =*ti* attaches to a word ending in a short vowel, the main stress therefore shifts onto that syllable:

- (149) la.'væ.ti la.'va.ti mal
la-va =*ti* *la-va* =*ti* *mal*
 3PL.IFUT-IFUT.go =PART 3PL.IFUT-IFUT.eat =PART dragon_plum
 'They'll go and eat some dragon plums.' [20170217a_n01s038_05]

When it attaches to a word ending in a consonant coda or long vowel, there is no shift in word stress. However, the clitic itself never takes a word stress of its own, showing that it is phonologically dependent on the word it attaches to.

As it is phonologically attached to the verb in most of these examples, partitive *ti* could be analysed as a suffix, but it can also attach to nouns. This fulfils the requirement for a clitic to attach to words of a range of categories (Bickel & Nichols 2007: 175). There are only two examples where =*ti* attaches to a noun in the corpus, both from the same speaker in Mele Maat. Example (150) shows that partitive =*ti* attaches immediately to the right of the noun, before the plural quantifier *xil*. Example (151) also shows the expected stress shift, confirming that *ti* is part of the same phonological word.

- (150) *bit* *atou* =*ti* *xil* *la-ve* *joen*
 3SG.NFUT.want woman =PART PL 3PL.IFUT-IFUT.COP join
 'It [*a work scheme*] wanted some of the women to join'
 [20170413e_n01m030_8]

- (151) ru.'te.ti
rute =*ti*
 place =PART
 'some places' [20170413e_n01m030_57]

Note that neither of these modified nouns function as an object: if they did, the partitive would attach to the verb as in the other examples. In (150) *atou* is the subject of the subordinate clause; in (151) *rute* is a locative oblique argument.

Apart from a partitive meaning, =*ti* is also used as a general marker for polite suggestions as in (152), especially for polite use of the imperative (153). Examples like (154) and (155) demonstrate a bridging context for reanalysis of a partitive to a politeness marker, where the pragmatic intention is to minimise the labour involved in the suggested task, or the inconvenience of the request.

- (152) *Ha ralo-xul =ti e tas*
 GO 1DU.INCL.IFUT-swim =PART LOC sea
 ‘Let’s go have a swim in the sea.’ [20141027a_n01m001_095]
- (153) *ha pus =ti xouk*
 GO IMP.see =PART 2SG
 ‘Go and have a look at yourself.’ [20141212f_n01s054_14]
- (154) *na-bit ha ral-sei =ti hoei*
 1SG.NFUT-NFUT.want GO 1DU.INCL.FUT-collect =PART bushnut
 ‘I want us to go collect some bushnuts.’ [20141212n_n01s046_2]
- (155) *bit va-pus =ti aexo*
 3SG.NFUT.want 3SG.IFUT-see =PART you
 ‘He wants to see you a bit.’ [20141107b_n01e020_15]

In Parker’s dictionary, the clitic I am calling ‘partitive’ is described as ‘indicating that action is contingent on favourable circumstances, such as approval of addressee’ (Parker 1970a: 33). Although most uses of the clitic can be accounted for with a partitive interpretation, and its extension into politeness marking, it also frequently cooccurs with the matrix verb *hit* or *hitene* ‘want’, occasionally in situations where politeness is not a relevant factor, as in example (635) below. Parker’s ‘favourable circumstances’ analysis could be a better fit here, or the distribution could reflect a lexical collocation between the partitive and the matrix verb.

The partitive clitic has a fairly low frequency, occurring once every 665 words in the subcorpus. It is nearly entirely restricted to verbs, with the quantifier *rute* ‘some, half’ used more frequently in the NP for a similar meaning. However, its existence in Vatlongos has some typological ramifications for the analysis of the related negative clitic *ti*. Vossen and Van der Auwera (2014: 73–74) list Vatlongos (Southeast Ambrym) as an example of an Austronesian language where a cognate of *ti* only has a negative meaning. Removing Vatlongos from the list leaves only one language in this category, which represents the final stage in a Jespersen Cycle.

3.6.2 Negative clitic *ti*

Like many Oceanic languages in Melanesia, Vatlongos marks negation discontinuously (Lynch, Ross & Crowley 2002: 51; Barbour 2015: 435). Double marking of negation is especially common in Vanuatu (Vossen & van der

Auwera 2014: 64–74), normally with an older prefix and a post-verbal element with more varied distribution, often derived from a partitive. As well as the negative prefixes (§5.1.2), negation in Vatlongos is marked with a clitic *ti*, and most negative clauses would be ungrammatical without this clitic. However, the variable positioning of the negative clitic in relation to object NPs, and its optionality in some constructions, suggests that the prefix and the clitic do not constitute a single discontinuous morpheme (or circumfix); a common analysis for double-marking of negation. Additional evidence for their separate morphemic status is that *ti* is also used in prohibitive utterances, where it co-occurs with the prohibitive prefixes (Table 24). The distribution of *ti* in the prohibitive does not differ from *ti* in negative polarity, so the prohibitive is not described separately here.

Like the partitive clitic, negative *ti* is part of the same prosodic word as the syntactic word it attaches to, shifting the main word stress onto a final short syllable. However, the negative clitic can attach to a wider range of word types than the partitive clitic. It can attach to verbs (156), nouns (157), adjectives (158), quantifiers (159) and adverbs when they appear in the final position of a complex predicate (160) (§7.1). It also shows the expected stress shift with loan nouns (161). Unlike the partitive clitic, negative *ti* can follow the object, and attaches to an NP rather than the head noun itself: it follows the quantifier *xalu* in (159), rather than preceding the quantifier like partitive =*ti* in (150) above. Negative *ti* is therefore a more canonical clitic than the partitive, as its position is defined by syntactic constituency rather than attaching to a head word like an affix (Spencer & Luís 2012: 134).

(156) li,naa.va.'si.ti
li-naa-vas-i =*ti*
 3PL.DFUT-NEG-kill.NEG-3OBJ =NEG
 'They mustn't kill it.' [20170222d_n01s152_53]

(157) ,a.te.'li.ti
ateli =*ti*
 basket =NEG
 '{We won't carry} baskets' [20150307a_x01s046_22]

- (158) ,te.me.'le.ti
temele =*ti*
 easy =NEG
 ‘{It’s not} easy’ [20150225a_x02s046_01]
- (159) xa.'lu.ti
xalu =*ti*
 DU =NEG
 ‘{We didn’t wash the} two {dresses}.’ [20150228a_x01s026_14]
- (160) ku.'hi.ti
kuhi =*ti*
 properly =NEG
 ‘{It wasn’t good} enough’ [20170413e_n01m030_20]
- (161) ti.'tʃa.ti
tija =*ti*
 teacher =NEG
 ‘{I’m not} a teacher.’ [20170406a_n01m164_57]

In fact, *ti* seems to form a single phonological word with the word it attaches to, as phonological segments can also be elided when the clitic attaches to a transitive verb: the pronominal object suffix *-i* is optionally deleted before negative *ti*, as in example (181) below. However, occasionally negative *ti* does occur as an independent phonological word taking main word stress. This is only on the rare occasions when *ti* is fronted to mark focus over existential subject negation, as in (213) below.

As in many other Vanuatu languages, the negative clitic usually occurs post-verbally. This is why similar clitics or particles are often glossed ‘NEG2’ in descriptive grammars, as there is usually strict relative ordering of the negative particles involved (Budd 2010: 513). However, Vatlongos *ti* can appear in different positions in the clause, depending on clause type and semantic scope, which is why both the prefix and the clitic are only glossed ‘NEG’ here.

Examples (162) and (163) show that *ti* appears immediately after an intransitive verb.

- (162) *li-naa-pat* =*ti*
 3PL.DFUT-DFUT.NEG-sleep NEG
 ‘They won’t sleep.’ [20141028a_x01m002_15]

Clause type	Constituents before <i>ti</i>	Constituents after <i>ti</i>	Section
SVCs	<i>Object NP of first verb</i>	<i>Object NP of first verb</i> Subsequent VP	6.3.3

3.6.2.1 Transitive verb phrases

In transitive VPs, the distribution of the negative clitic *ti* is not predictable. It can appear before (165) or after (166) an object NP, and speakers accept either order as grammatical in any context. Of the 80 negative clauses in the subcorpus that include an overt object NP, *ti* appears before the object in 50 examples, and after the object in 30.

(165) *puskat taa-toxol =ti asu*
 cat 3SG.NFUT.NEG-catch =NEG rat
 ‘The cat didn’t catch a/the rat.’ [20170112b]

(166) *puskat taa-toxol asu =ti*
 cat 3SG.NFUT.NEG-catch rat =NEG
 ‘The cat didn’t catch a/the rat.’ [20170112b]

However, there are clear trends in the corpus relating the position of the clitic to the semantic role of the object determined by the verb’s argument structure. These preferred orders with particular verbs were confirmed in elicitation: the preferred orders were always given first, although the alternative order is still accepted as grammatical and identical in meaning. *Ti* follows patient objects which are highly affected by the situation expressed by the predicate. For example, in the corpus *ti* always follows the object of a ‘eat’:

(167) *netak naa-xa xatel =ti*
 this_one 3SG.FUT.NEG-eat 3PC =NEG
 ‘This thing won’t eat them.’ [20141220g_n01s080_51]

(168) *ra-taa-a huram =ti*
 1PL.INCL-NEG.NFUT-eat yam =NEG
 ‘We don’t eat yam’ [20141106f_n01e018_41]

Ti also follows objects that are created as a result of the action expressed by the verb: such as a rolled dumpling (169) or a woven fence (170).

(169) *la-taa-vei vloh xil =ti*
 3PL.NFUT-NFUT.NEG-weave fence PL =NEG
 ‘They didn’t weave fences.’ [20150219b_n01m001_34]

- (170) *a-naa-pis siboro =ti naa-maxai*
 2SG.IFUT-NEG-roll dumpling =NEG 3SG.IFUT-like_that
 ‘You won’t roll a dumpling like that.’ [20141212g_p01s046_11]

On the other hand, stimulus objects of perception verbs always follow the negative clitic in the corpus:

- (171) *lu-taa-pus =ti tumen*
 3DU.NFUT-NFUT.NEG-see =NEG bird
 ‘They didn’t see any birds.’ [20141220g_n01s080_20]
- (172) *mu-taa-long =ti nou*
 2DU.NFUT-NEG.NFUT-hear =NEG 1SG
 ‘You didn’t hear me’ [20141220g_n01s080_59]

Similarly, *ti* always precedes the object of the verbs of cognition *kil* ‘know’ and *nenem* ‘think, think about’, both of which can also take complement clauses.

- (173) *nat-naa-kil =ti sepini-en na Ahii*
 1SG.PRI-NEG-know =NEG speak-NMLZR CL.GEN God
 ‘I didn’t know the word of God.’ [20150118b_n01m090_38]
- (174) *la-taa-nnem =ti xil*
 3PL-NFUT.NEG-think =NEG 3PL
 ‘They didn’t think about them.’ [20141117a_n01m003_38]

Theme objects, which undergo a change of location or possession, usually precede the negative clitic, but *ti* can precede a generic theme object with a negated partitive interpretation.

- (175a) *ni-naa-sa nep na-van =ti mi xamil*
 1SG.DFUT-NEG-give knife CL.GEN-1SG.POSS =NEG to 2DU
 ‘I won’t give you my knife.’ [20141027a_n01m001_72]
- (b) *la-taa-sa =ti mani mi xil*
 3PL.NFUT-NFUT.NEG-give =NEG money to 3PL
 ‘They didn’t give them any money.’ [20170413a_h01m169_130]
- (176a) *li-naa-xur horel =ti e-n*
 3PL.DFUT-NEG-take crayfish =NEG LOC-3SG.OBJ
 ‘they wouldn’t get crayfish there.’ [20150304f_n01e024_06]
- (b) *la-taa-xur =ti atou*
 3PL-NFUT.NEG-take =NEG woman
 ‘They didn’t take any wives.’ [20150223a_n01m096_58]

The goal object of *kamuet* ‘find’ has the same distribution in relation to the negative clitic:

- (177a) *lu-taa-kamuet* *frog* =*ti*
 3DU.NFUT-NFUT.NEG-find frog =NEG
 ‘They didn’t find the frog.’ [20141212h_n01s046_43]
- (b) *na-taa-kamuet* =*ti* *pol-ien*
 1SG.NFUT-NFUT.NEG-find =NEG work-NMLZR
 ‘I didn’t find any work.’ [20150223a_n01m096_27]

The referential status of the object has an effect on negative marking in other Vanuatu languages. Closely related Paamese also has differential negative marking for generic objects, marking examples like the (b) versions of (175)–(177) with the negative prefix but no negative/partitive suffix *-tei* (Crowley 1982: 140–141), while other transitive VPs are marked with both. In the Lolovoli dialect of North-East Ambae, the negative particle *tea* follows a specific object NP, but precedes a non-specific object NP (Hyslop 2001: 262). Both examples of ‘non-specific’ objects in the description are also generic.

However, as well as conveying that the object is generic, this ordering in Vatlongos seems to be associated with a more precise interpretation: it conveys that no part of the generic set of referents is affected by the verb, a similar meaning to English *any* seen in the translations. This meaning perhaps arises from the clitic’s origin as a partitive: partitive =*ti* appears in this immediately post-verbal position when it scopes over an object (§3.6). Example (178) shows a generic theme object appearing before the negative clitic in a situation where negative partitive meaning is irrelevant: people in the past were in the habit of not wearing trousers, the question of how many trousers or how much trouser they wore is irrelevant.

- (178) *la-taa-kave-ni* *torosis* =*ti*, *la-pis* *tamu*
 3PL.NFUT-NFUT.NEG-wear-TR trousers =NEG 3PL.NFUT-roll still
 ‘They didn’t wear trousers, they still wore penis sheaths.’
 [20141107b_n01e020_11]

Other evidence that it is the generic negative partitive meaning, rather than a simple generic, associated with this word order comes from examples where the scope of the generic is restricted by other elements in the clause: an adverb incorporated into a complex predicate in (179), and a zero-marked relative clause in (180). These kinds of restrictions have been referred to as ‘triggers’ or

‘sub-triggering’ in the literature on cross-linguistic *any* (Krifka 1995; Dayal 1998; Dayal 2004).

- (179) *la-taa-leh kuh* =*ti* *sel*
 3PL.NFUT-NFUT.NEG-take+properly =NEG shell
 ‘They didn’t get a good amount of shells’ [20170220g_n01s148_07]
- (180) *ma-taa-pol-ni* =*ti* *neta va-he holu*
 1PL.EXCL.NFUT-NFUT.NEG-do-TR =NEG thing 3SG.IFUT-COP lots
 ‘We didn’t do a lot of things.’ [20170412b_n01m167_13]

When an object suffix is used, the suffix *-i*, and its allomorphs *-a* and *-e*, (Table 27) are usually dropped when the negative clitic *ti* appears (181), though sometimes it does still appear (182), (183). The *-ni* allomorph is not lost before *ti* (184). This is similar to Bierebo (Epi), where the object enclitic does not cooccur with the negative marker, which is also hypothesised to derive from a partitive (Budd 2009: 381–382).

- (181) *inou na-taa-pus* =*ti*
 1SG 1SG-NFUT.NEG-see =NEG
 ‘I didn’t see it.’ [20141117a_n01m003_12]
- (182) *lu-taa-kamuet-i* =*ti*
 3DU.NFUT-NFUT.NEG-find-3OBJ =NEG
 ‘They didn’t find it.’ [20141212h_n01s046_49]
- (183) *o-naa-tuv-a* =*ti*
 2SG.IFUT-NEG-shoot-3OBJ =NEG
 ‘You won’t shoot him.’ [20150305h_h01o111_18]
- (184) *taa-xa-ni* =*ti*
 3SG.NFUT.NEG-NEG.eat-3OBJ =NEG
 ‘He didn’t eat it.’ [20170330h_x01s026]

At first glance, the presence or absence of the object suffix seems to align well with the correlation of semantic roles with the relative order of the negative clitic and an object NP. However, there are not enough tokens in the corpus to confirm this correlation. Example (185), where the stimulus object of *pus* ‘see’ is coded with the object suffix, shows that other factors might be influencing the presence or absence of the suffix: in this case the object suffix is a resumptive pronominal referring to the NP in the matrix clause from within a relative clause.

- (190) *taa-ve* *eilep pupu =ti*
 3SG.NFUT.NEG-NEG.COP big very =NEG
 'It isn't very big.' [20141119c_c01s039_07]
- (191) * *na-taa-ve =ti eilep*
 1SG.NFUT.NEG-NEG.COP =NEG big
 * ('I am not big') [20161111b_x01s026]
- (192) *taa-ve vatiang =ti*
 3SG.NFUT.NEG-NEG.COP wind =NEG
 'It wasn't a cyclone.' [20141117a_n01m003_19]
- (193) * *xi taa-ve =ti Elta*
 3SG 3SG.NFUT.NEG-NEG.COP =NEG Elder
 * ('He isn't an Elder.') [20161111b_x01s026]

The negative clitic follows even syntactically complex NPs, as in (194) which is also interrupted by the hesitation marker *na*. However, example (195) shows the negative clitic interrupting the object NP, following the head noun but preceding a modifying PP.

- (194) *taa-ve na prans te vatiei tei =ti*
 3SG.NFUT.NEG-NEG.COP HES branch of tree one =NEG
 'It wasn't um the branch of a tree.' [20150118a_n01m090_47-48]
- (195) *inou na-taa-ve sup =ti tahal Edu*
 1SG 1SG.NFUT.NFUT.NEG-NEG.COP chief =NEG from Endu
 'I am not the chief of Endu.' [20141223b_h01s046_10]

The negative clitic also interrupts an NP in the same way in example (196), with the verb *xoni* 'be like'. The negative clitic follows the head noun and quantifier, but precedes a relative clause. This could be an argument for analysing the verb as a kind of copular, as it also is obligatorily followed by a complement. However, unlike copular *he*, *xoni* can also take clausal complements.

- (196) *taa-xoni tati xil =ti xa di la-bang*
 3SG.NFUT-like daddy PL =NEG REL CONT.REAL 3PL.NFUT-hang_around
 'It's not like the men who hang around.' [20150303g_p01e009_18-9]

This suggests that the complement of the copular verb is more closely bound to the verb than an oblique or even an object: the distribution of the negative clitic respects the role of the complement as a predicate.

3.6.2.4 *Loan verbs introduced with the copular verb he*

The negative clitic *ti* also follows the complement of the copular verb *he* when it is used to incorporate loan verbs from Bislama:

(197) *na-taa-ve* *jenis* =*ti*
1SG-NEG.NFUT-NEG.COP change.BIS =NEG
'I didn't change.' [20150223a_n01m096_35]

(198) *la-taa-ve* *fiksimap-ni* =*ti*
3PL.NFUT-NEG.NFUT-NEG.COP fix.BIS-TR =NEG
'They didn't fix it' [20150419e_h01m128_54]

More surprisingly, the negative clitic can also follow any object of the incorporated loan verb (199), even when the object NP is made up of Bislama loanwords, so that the entire copular construction has the same distribution as a Vatlongos transitive or transitivised verb.

(199) *o-naa-ve* *litim* *nou* =*ti*
2SG.IFUT-NEG-NEG.COP lead.BIS.TR 1SG =NEG
'You won't lead me.' [20170218b_n01s144_08]

(200) *la-taa-ve* *fiksimap-ni* *evri* *samting* =*ti*
3PL.NFUT-NEG.NFUT-NEG.COP fix.BIS-TR every.BIS thing.BIS =NEG
'They didn't fix everything.' [20150419e_h01m128_57]

3.6.2.5 *Reflexive verbs*

The negative clitic always follows the object of the reflexive verb *hesi* 'be happy', which is always coreferential with the subject (§3.4.1) (see also example (89)).

(201) *Mama na-mem* *xil la-taa-vesi* *xil* =*ti*
mother CL.GEN-1PL.EXCL.POSS PL 3PL.NFUT-NFUT.NEG-NEG.happy PL =NEG
'Our mothers aren't happy.' [Y3R8_73]

3.6.2.6 *Complement clauses*

The negative clitic always precedes any complement clauses (§7.2.1).

(202) *a-taa-kil* =*ti* [*mu-be* *vahis* *vari?*]
2SG.NFUT-NFUT.NEG-know =NEG 2PL.NFUT-NFUT.COP how_many at_once
'Don't you know how many of you there were?'
[20141121a_n01m041_65]

As this pre-verbal position of the clitic is also associated with existential negation of the subject, perhaps these sentences have a literal interpretation similar to ‘I don’t have any insides that like this.’

3.6.2.8 Existential negation of subject

The clitic can also occur between the subject NP and the verb to convey existential negation of the subject referent, as shown by the different interpretation of (210) and (211).

(210) *moletin* *taa-kkesae* =*ti*
 person 3SG.NFUT.NEG-lost =NEG
 ‘The person wasn’t lost.’ [20161111b_x01s026]

(211) *moletin* =*ti* *taa-kkesae*
 person =NEG 3SG.NFUT.NEG-lost
 ‘No person was lost.’ [20161111b_x01s026]

The difference in meaning also affects subject agreement. The plural pronoun in (212) takes singular agreement because a negated plural is singular for the purposes of agreement.

(212) *xil* =*ti* *taa-viteni*
 3PL =NEG 3SG.NFUT.NEG-NEG.speak
 ‘None of them spoke.’ [20141219c_n01s033_13]

This position where *ti* modifies the subject probably derives from its marginal use as a partitive marker following nouns, but it cannot cooccur with a post-verbal negative *ti*.

Harder to explain are a few examples where it precedes the subject, which seems to indicate focus on the negative existential component of the proposition. In this position the clitic is an independent phonological word taking full word stress.

(213) *ti* *xamem* *taa-met* *vari*
 NEG 1PL.EXCL 3SG.NFUT.NEG-die at_once
 ‘None of us died at all.’ [20150414b_h01e016_95]

The negative clitic scoping over a subject NP is also found in a clause subordinated to lexical negative verb, *sakras* ‘be unable to’.

Vatlongos are verbs derived from incorporation of adverbs, and the variable positioning and optionality of the negative clitic to mark different clause types and scopes of negation. The syntactic behaviour of simple clauses outlined in this chapter is important to the analysis of different verb combining strategies discussed in later chapters.

4 Tense, Aspect and Mood distinctions

There are four main morphological TAM categories in Vatlongos, as well as morphologically marked imperative and prohibitive moods, and a less commonly used apprehensive mood. Each category is used to encode situations that could be distinguished on the basis of tense, aspect or mood, so deciding which of these distinctions is primary for each category on the basis of its ‘extension’ to a set of contexts (Dahl 1985: 3) is not straightforward. Tense, aspect and mood are closely interconnected and because of the potential for overlap in the extension of TAM categories, defined by different TAM strands (for example use of tense categories to mark modal distinctions), deciding which of the three TAM strands is prominent in Vatlongos is an even greater challenge (Bhat 1999: 93).

I analyse the four main morphological TAM categories in Vatlongos as relative tenses, as they are best distinguished from each other by their temporal extensions. The imperative, prohibitive and apprehensive are described as moods as they unite a coherent modal concept with a pattern of inflectional expression (Bybee, Perkins & Pagliuca 1994: 181). However, all the TAM categories can be further categorised into two broad semantic domains: those which assert the actuality of the situations they describe, and those describing possible situations. This is similar to the distinction between realis and irrealis that is marked in many Oceanic languages, and has been reconstructed for Proto-Oceanic (Lynch, Ross & Crowley 2002: 84; Ross 2004: 500). I have therefore labelled these higher-level groupings ‘realis’ and ‘irrealis’ (Figure 8).

Realis		Irrealis				
Prior	Non-future	Immediate future	Distant future	Imperative	Prohibitive	Apprehensive

Figure 8: Morphological TAM categories in Vatlongos

While prior and non-future only describe situations that actually occur in the real world, the two future relative tenses and the three morphological moods can all be characterised as irrealis, as they all express non-actual situations (Table 19). This level of organisation is probably a consequence of a more

mood-prominent system in the history of Vatlongos; Crowley (1982: 131; 1991: 184) analyses Paamese as a modal language. It is unusual to use these terms to refer to sets of morphologically-marked TAM categories, rather than individual moods. However, Crowley (1982: 131–140) distinguishes realis mood from an array of categories he labels with the umbrella term ‘irrealis’ in Paamese, as discussed by Budd (2009: 398–399). This use of the realis-irrealis distinction is similar to Givón’s (1994: 167–168) communicative definition of realis and irrealis propositional modalities as strong assertions of factual truth versus weak assertions of possibility, likelihood, uncertainty or (un)desirability.

In this chapter, I use ‘situation’ as a general term for the different states of affairs that can be expressed by clauses containing a predicate, including events, states, activities, and processes.

I begin by discussing definitions of tense, aspect and mood, and apply them to the extensions of the Vatlongos TAM categories, before giving detailed examples of how each category is used.

4.1 Applying definitions of tense, aspect and mood to Vatlongos

Tense and aspect are both concerned with time: tense is usually defined as indicating the location of a situation in time (Comrie 1985: 1, 9; Bybee 1985: 28), while aspect is concerned with the internal temporal structure of situations (Comrie 1976: 3; Bybee 1985: 28; Dahl 1985: 24). Given the close relationship between tense and aspect, Timberlake (2007: 315) suggests a general category of ‘temporality’. This is especially useful because there are major differences in how the definitions of tense and aspect are applied by different linguists, and the temporal distinctions made in Vatlongos fall into a contested area that could be characterised as relative tense (Comrie 1985: 56–82) or viewpoint aspect (Bhat 1999: 58–59).

Reichenbach’s (1947) model has been extremely influential in modelling tense. He characterises different tenses as relationships between speech time (S, the here-and-now of speech), event time (E, the time at which the predicated situation takes place) and a reference time (R, some other salient point in time). For some scholars, any grammaticalized category that marks a sequential

relationship between these three points falls into the category of tense. Thus Comrie (1985: 130) models tense categories as the logical possibilities of the Reichenbachian model, and for Bybee (1985: 28) tense places a situation relative to 'either the moment of speech or another point in time' (i.e. 'R'). On this view, there are two kinds of tenses: absolute tense (Comrie 1985: 36–55) relates event time E to the moment of speech S, while relative tense (Comrie 1985: 56–82) relates event time E to another temporal reference point R (also described as deictic and non-deictic tense, Dahl 1985: 25). However, scholars who support the existence of a category of relative tense tend to assume that relative tense is a subsidiary system: either it coexists with absolute tenses (Comrie 1985: 60), or R needs to be explicitly stated in the same sentence, perhaps in a matrix clause (Bhat 1999: 20–21).

A rival definition would restrict tense to categories that relate situations to speech time, denying the possibility of relative tenses (Dahl 1985: 25). Klein's (1994: 3–9) articulation of this definition has become influential in formulations of the tense-aspect distinction. He usefully discusses the characteristics of the Reichenbachian points in time, suggesting more precise terminology: utterance time (TU) for S; situation time (TSit) for E; and topic time (TT) for R. These terms are both more general – including non-speech-based utterances and situations other than events – and more specific, showing how discourse and information status determine the meaning of TT or R. Topic time is therefore the time that the utterance is 'about' and the truth of the proposition evaluated against, similar to a topical referent (Lambrecht 1996: 117–119; Erteschik-Shir 2007: 13). Where Klein departs from the first definition of tense is in claiming that only the relationship between TU and TT is relevant to tense; the relationship between TT and TSit comes under aspect. This means that Comrie's relative tenses should all be analysed as aspect rather than tense. Klein defines perfect, perfective, imperfective and prospective aspects straightforwardly in these terms.

However applying this model of aspect to other distinctions traditionally analysed as aspectual operators, such as progressive and iterative (Timberlake 2007: 304), requires a particular model of lexical aspect determined by

individual predicates (Klein 1994: 8–9, 37–48). Bohnemeyer (2014) argues that while categories in some languages conflate relative tense with viewpoint aspect, like the English present-perfect that is central to Klein's (1992) analysis, some languages do have 'purely' aspectual and/or tense-based categories in this domain. He puts forward a model that accounts for these distinctions by including relations with a 'perspective time' in addition to Klein's three neo-Reichenbachian times, and greater attention to the composition of situations, especially the runtime of pre-states and post-states in a causal chain.

The result of these competing definitions is conflicting applications of the same terminology. For Bybee (1985: 159–161) and Bybee et al. (1994: 54), 'anterior' and 'perfect' are equivalent terms and should both be included under tense, whereas Bhat (1999: 21, 31, 49) distinguishes the non-deictic tense 'prior' from the combination of deictic and non-deictic tense 'perfect', which he also analyses as an aspectual category in aspect-prominent languages.

These rival definitions have also affected the analysis of temporality in Bislama, the national language of Vanuatu. Holm's (1988: 150–154) discussion of the TAM systems of pidgins and creoles describes the anterior as a 'tense' which shows an event precedes the reference time denoted by the unmarked verb. Although he does not describe the unmarked verb's time reference as 'tense' it could also be analysed as a relative tense: TSit is simultaneous with TT. While this fits with Comrie's definitions of relative tense, it would be considered unusual because there is no system of absolute tense it interacts with. Meyerhoff (2000: 175), however, follows Klein's definition, taking the irrelevance of TU to the interpretation of the anterior to mean that the anterior in pidgins and creoles is an aspect, not a tense.

The temporal distinctions made in Vatlongos are in exactly this area that seems to fall between tense and aspect. Rather than relating situation time to utterance time, they relate it to topic time, as shown in examples like (221) and (222) in the immediate future and (223) in the prior.

Example (221) describes events that were about to happen two weeks before the recording was made, when an earthquake struck. The immediate

Although it seems clear that the relative order of TSit and TT is more relevant to temporality in Vatlongos than that between TSit and TU, there is still some doubt as to how this should be described. Some linguists would analyse it as relative tense, whereas for others it should be analysed as aspect. However, it positions the situation as a whole in time and is not concerned with the internal temporal structure of the situation, for example, the boundaries of the situation, whether it is punctual or ongoing, or whether it recurs over time. Instead, some of these traditional aspectual categories are marked with AVCs or other periphrastic strategies summarised in §4.3. Following Bhat (1999: 102), the use of these ‘indirect means’ is evidence that aspect is not the prominent TAM strand in Vatlongos. These less grammaticalized means for expressing aspectual distinctions are also compatible with all the morphological marked TAM categories: examples (224)–(227) show that all four relative tenses can be marked with the continuous auxiliary, which expresses progressive and habitual aspects (§8.2.1). That these aspectual categories are orthogonal to morphological TAM suggests that the morphological paradigms mark something other than aspect.

(224) *di nate-pol e Santo*
 CONT.REAL 1SG.PRI-work LOC Santo
 ‘I worked in Santo.’ [20150219b_n01m001_057]

(225) *mama na-n di mi-til e-n*
 mother CL.GEN-3SG.POSS CONT.REAL 3SG.NFUT-sew inside-3SG.POSS
 ‘His mother [used to] sew with it.’ [20141211a_n01s046_7]

(226) *Ti na-me kontiniu ejukesen level na-van*
 CONT 1SG.IFUT-IFUT.COP continue.BIS education level CL.GEN-1SG.POSS
 ‘I was going to be continuing my education level’
 [20170406a_n01m164_07]

(227) *xir ti ri-las*
 1PL.INCL CONT 1PL.INCL.DFUT-clear_land
 ‘We’ll clear the land.’ [20141219c_n01s033_06]

Another key feature of many tense systems cross-linguistically is the expression of temporal distance or remoteness (Comrie 1985: 83–102; Dahl 1985: 120–128), which Timberlake (2007: 315) calls ‘metrical’ tense. This seems to be the major distinction between the use of the immediate future and distant future in

Vatlongos, although the distinction is subjective rather than objective (Dahl 1985: 123–125), and therefore not very rigid (Comrie 1985: 90).

Definitions of mood or modality in the linguistics literature vary more than tense or aspect, and often resort to listing different kinds of modality rather than defining the concept itself. Palmer (1986: 9–14) gives an overview of earlier approaches to the topic, which often focus on identifying types of modality such as epistemic, deontic, or evaluative. Timberlake’s (2007: 315) definition is especially useful for encompassing and distinguishing the various different kinds of modality discussed by other authors: modality considers ‘alternative realities mediated by an authority’. Thus, the primary modal distinction is between realis and irrealis situations: those that happen(ed) in the real world as opposed to those that happen(ed) in a non-actual world (a branching timeline), which are merely possibilities in the actual world.

Usually, ‘fictive’ worlds, such as those of stories, are treated as their own real world, so fictional situations can be expressed with realis categories. This is certainly the case in Vatlongos, as in example (228), where the non-future is used to describe what the rat and the crab get up to in the fictive universe of a children’s story.

(228) *asu xal oum lu-pol-ni nahou*
 rat with crab 3DU.NFUT-do-TR garden
 ‘The rat and the crab made a garden.’ [20141119b_n01s036]

Authority is central to the formulation of modality because identifying an event as real, or an irrealis event as more or less likely, involves a judgement on the part of the speaker. Irrealis modality has a close relationship with future tense, because no future event can be guaranteed to happen in the real world. Talking about future situations always involves intention, prediction or contingency, so future tenses are generally acknowledged to have some inherent modal properties, and in tense-prominent languages often have a range of irrealis modal functions (Dahl 1985: 103; Palmer 1986: 216–218; Bhat 1999: 65; Timberlake 2007: 306).

Although different authors divide modality into a number of different types, Timberlake’s (2007: 316–329) three-way distinction is the most useful for

applying to situations in Vatlongos. Epistemic modality is to do with knowledge of events: certainty, probability and the authority for the knowledge. Directive modality assigns responsibility for the state of the world. For example, the imperative mood confers responsibility for ensuring a particular reality comes to pass onto the addressee(s). Directive modality includes obligations, also known as ‘deontic’ modality, and self-directed obligations such as purposives. Finally, Timberlake identifies ‘contingency’ as a third kind of modality, where responsibility for one situation is assigned to another situation. The typical expression of contingency is through conditionals, and he further distinguishes three kinds of conditionals. Firstly, ‘general conditionals’ discuss frequent, expected contingencies,¹⁷ like wanting to make laplap in (229).

(229) *Laplap,* *xos* *xa* ***la-bit***
 laplap if SUB 3PL.NFUT-NFUT.want
 [*la-pol-ni,*] ***la-leh*** *anien* *nen,*
 3PL.IFUT-do-TR 3PL.NFUT-take food for_it
 ‘Laplap, if they want to make it, they get the ingredients for it.’
 [20141028b_p01m003_02-03]

‘Counterfactual conditionals’ talk about alternative histories that cannot be realised in the real world:

(230) ***na-tanga*** *etan,* *buluk* ***va-has-nou***
 1SG.IFUT-stay_behind down bull 3SG.IFUT-kill-1SG.OBJ
 ‘If I’d stayed down, the bull would have killed me.’
 [20170126c_n01e025_18]

Finally, ‘potential conditionals’ discuss possible contingencies where the speaker does not commit to whether or not the contingency occurs in the real world,¹⁸ as in (231).

(231) *xos* *xa* *taetal* *weiv* ***i-mmei***
 if SUB tidal wave 3SG.DFUT-come
e ***mu-loh*** ***mu-ha***
 then 1PL.EXCL.DFUT-run 1PL.EXCL.DFUT-go
 ‘If a tidal wave comes, we’ll run away.’ [20150310a_h01s114_18]

¹⁷ Also referred to as habitual or generic conditionals (Thompson, Longacre & Hwang 2007: 255).

¹⁸ Also referred to as hypothetical conditionals (Thompson, Longacre & Hwang 2007: 256)

Alternative definitions of modality focus instead on the primacy of speaker attitudes and opinions to the concept of modality (Dahl 1985: 26; Palmer 1986: 2, 16; Bhat 1999: 63). The three different kinds of modality all involve speaker judgements on, for example, the probability of a situation, what should happen, or how situations depend on each other. However, all of these can also be further modified by straightforward speaker attitudes about whether a situation is good or bad, often called 'evaluatives' (Palmer 1986: 119–121). The apprehensive mood in Vatlongos combines an epistemic meaning – that a situation is likely to happen – with an evaluative meaning, that the event is unwanted (Timberlake 2007: 329).

4.2 The extension of Vatlongos TAM categories

Typologically, most Oceanic languages are described as mood-prominent, making a main distinction between realis and irrealis contexts inherited from Proto-Oceanic, and having additional aspectual distinctions with no marking of tense (Lynch, Ross & Crowley 2002: 84; Ross 2004: 500). However, this characterisation seems to be based on the definition of tense as absolute rather than relative. Even in languages which make a binary distinction, the extension of modal categories into different contexts is difficult to tease apart from temporal distinctions, as discussed by Barbour (2011) for Neverver (Malekula), and Budd (2009: 351–358, 406–417, 444–445) for Bierebo (Epi), and other Vanuatu languages.

Modality and the realis-irrealis divide are important to TAM distinctions in Vatlongos. The partial morphological paradigms for imperative, prohibitive, and apprehensive all mark modal categories, with each primarily associated with a single irrealis modal context. The imperative and prohibitive express directive modality, while the apprehensive expresses a combination of epistemic and evaluative modality.

The analysis of the four main TAM categories as relative tenses, and the hierarchical structure shown in Figure 8 where sets of categories are labelled as 'realis' and 'irrealis', are both based on the extension of each morphological paradigm into different contexts. Table 19 summarises the different TAM contexts associated with each morphological TAM category. The contexts are

divided into temporal, epistemic, directive or contingency distinctions, and there is also information about the cooccurrence of each TAM category with other relevant morphosyntactic operators: the aspectual auxiliaries and negative polarity. Where more than one TAM category can be used, but one is more frequent or more closely associated with that context, the others are marked with a smaller tick. An example for each TAM category in that context is given in brackets.

Table 19: Summary of contexts for each morphological TAM category

	Context or distribution	Realis		Irrealis				
		PRI	NFUT	IFUT	DFUT	IMP	PROH	APR
Temporality	Prior to TT (223) (239)	✓	✓					
	Simultaneous with TT (238)		✓					
	Immediately after TT (221) (315)			✓	✓			
	Some time after TT (312)				✓			
Epistemic modality	Real situations (232) (256)	✓	✓					
	Deducing prior situations (240)		✓					
	Possibilities			✓	✓			✓
	Predictions (277) (311)			✓	✓			
	In relative clauses describing non-factive entities (293)			✓	✓			
	Subordinate to non-factive matrix verbs (284) (327)			✓	✓			
	Alternate realities (counterfactuals) (305) (339)			✓	✓			
	Unwanted possibilities (349)							✓
Directive modality	Commands (343) (345)				✓	✓		
	Prohibitions (347)				✓		✓	
	Obligation (308) (330)			✓	✓			
	Suggestions (281)			✓	✓			
	Ability and permission (327)			✓	✓			
	Intentions (297) (328)			✓	✓			
	Purpose (subordinate clauses) ¹⁹ (283) (318)			✓	✓			
	Intentions and instructions in indirect speech (303) (304) (338)			✓	✓			

¹⁹ A combination of self-directed directive modality and contingency (Timberlake 2007: 319).

	Context or distribution	Realis		Irrealis				
		PRI	NFUT	IFUT	DFUT	IMP	PROH	APR
Contingency	General conditionals (protasis and apodosis) (229)		✓					
	Potential conditionals (protasis and apodosis) (341)				✓			
	Counterfactual conditionals (protasis and apodosis) (230) (343)			✓	✓			
Syntax	With CONT auxiliary	✓	✓	✓	✓	✓	✓	✓
	With INCH auxiliary	✓	✓	✓	✓			✓
	NEG polarity	✓	✓	✓	✓			

Table 19 shows a very clear divide between prior and non-future on the one hand, and immediate and distant future on the other. Prior and non-future are used to express realis contexts – situations that occur(ed) in the real world - while immediate and distant future are used to express irrealis contexts, events that have not yet happened or might never happen in the real world. Immediate and distant future are used for a variety of irrealis contexts beyond the temporal future, which, as discussed above, is itself closely related to irrealis modality. The non-future is also associated with two less strictly realis contexts: deducing prior events from current evidence and general conditionals. Both apply to situations that are thought to occur in the real world, although without the speaker authoritatively guaranteeing that the situation has taken place at a specific time. It seems that in the Vatlongos modal system, both these contexts are included under the umbrella of realis modality.

Bybee et al. (1994: 236–239) argue against the validity of a binary distinction between realis and irrealis modalities cross-linguistically, pointing out that each is usually expressed by multiple markers in a single language and that different languages divide the conceptual space in different ways that sometimes violate the basic definitions of ‘real’ and ‘unreal’ situations. However, the regional prominence of this division has become widely accepted in the years since their survey, and this divide can account for certain features of the Vatlongos TAM system that are not predicted by the grammaticalization paths they discuss, especially the restriction of the prior to non-future

situations and the ‘late-developing’ uses and formal integration of the immediate future.

Although the diachronic consequences of the Proto-Oceanic realis-irrealis divide are evident in Vatlongos TAM categories, trying to characterise all the distinctions between the main four TAM categories through a modal lens runs into difficulties. Looking at Table 19, it is easier to distinguish prior from non-future, and immediate future from distant future, on temporal grounds than by any modal extensions: prior is only used for situations before topic time, and immediate future is only used for situations that immediately follow topic time. Although the two future tenses are somewhat specialised for different irrealis functions – notably, the distant future is always used for potential conditionals – there is usually some variation in which category a speaker chooses to encode an irrealis situation.

Crowley (1982: 131) analyses closely related Paamese as a modal language, but Paamese does not have an equivalent to the Vatlongos prior paradigm (though there is a ‘completive’ enclitic). Moreover, his distinction between immediate and distant ‘anticipatory’ moods (equivalent to Vatlongos futures) is couched in temporal descriptions: ‘connected to time of utterance’ or ‘not connected to time of utterance’. In later descriptions, Crowley (1991: 184) describes these moods as immediate and distant irrealis, but the distinction between them is still temporal: ‘an unreal event is about to become real’ or ‘will, or may, become real at some time in the future’.

The closest parallel to the Vatlongos prior paradigm is the West Ambrym languages. Both Daakie (Krifka 2011; Krifka 2016: 580–583) and Daakaka (von Prince 2015: 250–255; von Prince 2017b) have a TAM category described as ‘distal’, marked by a particle preceding the verb that has a similar phonological form to the Vatlongos prefixes. The West Ambrym distal markers can be used for the distant past, have a default temporal meaning, and like the Vatlongos forms are particularly common in subordinate clauses (von Prince 2015: 250). However, the West Ambrym distal has a role in irrealis contexts such as conditionals and counterfactuals (which in Vatlongos are usually expressed with the distant future), and unlike the Vatlongos prior it cannot be negated. It

therefore seems that the Vatlongos prior has retained the temporal component of the West Ambrym category, without its modal character.

Instead, the Vatlongos prior is strictly associated with realis situations. In some languages a prior or anterior category can combine with the future tense to report situations preceding a future situation (Bhat 1999: 22), or to express conditionals (Holm 1988: 164). However, in Vatlongos, the prior can only be used to report situations that have or have not occurred in the actual world. As the future is inherently irrealis, the prior is confined to situations preceding a TT that is simultaneous with or prior to TU. To talk about a situation prior to an irrealis event speakers are obliged to use the periphrastic completive strategy with the verb *vus* 'finish' (§4.3.1). Although I have avoided using formal evidence to deduce relationships between the TAM categories, there is some morphosyntactic evidence to support the idea that prior is grouped with non-future in this realis category. For most speakers, the prior and non-future both take the non-future form of the continuous auxiliaries *ti*, *ta*, *tit*, *tat* and the directional auxiliary *ha* 'go' (§8.1.2): compare the *di* form in examples (224) and (225) with the *ti* form in (226) and (227). The *di* form of the auxiliary is therefore glossed REAL, including prior and non-future.

Parker's (1968b; 1970a) analysis of TAM in Vatlongos differs both in the choice of terms and the structure of the system outlined here, so it is worth clarifying how these proposed categories map onto the previous description. Non-future is equivalent to Parker's aorist or indicative; prior to his past; and distant future to future. Parker describes the immediate future as a subjunctive mood, and he accounts for patterns of syncretism with the non-future for verbs with certain initial consonant mutation patterns by saying that some stem classes do not inflect for subjunctive in all person-number combinations. Finally, the forms described as apprehensive mood here are labelled 'potential', and are described as having a similar meaning to the subjunctive, but 'more definite'.

The verbal morphology of Vatlongos is complex and exhibits overabundance and syncretism. There are different patterns of syncretism in different person-number combinations, which is one reason I have not

discussed formal evidence for the hierarchical relationship proposed in Figure 8. While in some parts of the paradigm realis and irrealis contexts are opposed to each other, in other areas the non-future and immediate future are grouped together, or the prior is grouped with the two futures. As described in Chapter 5, the same cell in a paradigm can be represented by multiple forms, and a single form can represent different cells. This means that in certain person-number combinations the exact TAM category marked by a prefix is only disambiguated by cooccurrence with an irregular verb-stem alternation or by the verb’s participation in an SVC or AVC (Chapters 6 and 8). As a result, lots of examples used in this section are syntactically complex, and glosses may appear inconsistent.

The rest of this section exemplifies how each TAM category is used.

4.2.1 Non-future

The non-future is used for realis situations: situations currently happening and situations in the past. It can be thought of as the ‘default’ TAM category (Dahl 1985: 19), used to describe situations with unmarked TAM.

In these two examples the non-future is used for situations happening at utterance time: (232) is during Chief’s Day celebrations, and (233) is from commentary on a football match.

(232) *xosali, xametel mat-bituei*
 today, 1PC.EXCL 1PC.EXCL.NFUT-NFUT.gather_together
 ‘Today, we gather together’ [20150305f_p01s110_24]

(233) *David mi-loh ba-ni futbol*
 David 3SG.NFUT-run 3SG.NFUT.go-TR football
 ‘David runs to the football’ [20141215f_k01s026_10]

When describing current situations, the non-future frequently cooccurs with the continuous auxiliary *ti, di* (§8.2.1).

(234) *“e, xouk di o-pol-ni mi-dep xi-ak?”*
 hey 2SG CONT.REAL 2SG.NFUT-do-TR 3SG.NFUT-NFUT.how 3SG-PROX
 ‘{The crab said,} “Hey, what are you doing now?”’
 [20141027a_n01m001]

(235) *rumali ok di lu-dihi xi-e*
 3DU PROX CONT.REAL 3DU.NFUT-NFUT.cut 3SG-CONTR
 ‘These two are cutting here.’ [20150305f_p01s110_08]

the old woman knew it was there. Example (241) shows that the non-future can also be used for less confident hypotheses about past situations: the parents of a worried child think a devil might have taken their child. While the repeated use of the adverb *xos* ‘maybe’ minimises the authoritativeness of the statement (§3.3), the verb is still marked with the realis non-future.

- (240) *“tovolih ak xi-ak xa mei*
 old_woman PROX 3SG-PROX REL COME
mi-leh rat vueili a-ralu xi-ak.”
 3SG.NFUT-take out pig CL.ED-1DU.INCL.POSS 3SG-PROX
 “It was this old woman who came and took away our pig.”
 [20141027a_n01m001_111]
- (241) *lu-bit “xos netak*
 3DU.NFUT-NFUT.say maybe this_one
vebangen xos po netak”
 devil maybe 3SG.NFUT.carry this_one
 “They said “maybe a devil maybe took him”
 [20141105e_n01e012_13]

Apart from situations in a narrative, the non-future is also used for habitual and typical situations.

- (242) *tawi Leimaku pol-ni mak*
 in-law Leimaku 3SG.NFUT.do-TR 3SG.NFUT.like_this
 ‘Sister-in-law Leimaku makes it [*laplap simboro*] like this.’
 [20141212g_p01s046_20]
- (243) *Toak mi-sepin mi-sap vuteili,*
 Toak 3SG.NFUT-speak 3SG.NFUT-different a_little
Edu mi-sepin mi-sap vuteili
 Endu 3SG.NFUT-speak 3SG.NFUT-different a_little
 ‘In Toak they speak a bit differently, in Endu they speak a bit differently.’ [20141107d_p01s022_10]

These habitual or recurrent activities also frequently co-occur with the continuous auxiliary *ti, di* (§8.2.1), including habits in the past that no longer occur at utterance time, as in example (225) above.

- (244) *vehakut xamem mol Abrim xil*
 always 1PL people Ambrym PL
di ma-bas vueili
 CONT.REAL 1PL.NFUT-NFUT.kill pig
e melengien te sup
 LOC day of chief
 ‘We people of Ambrym always kill pigs on Chief’s Day.’
 [20150305f_p01s110_34]

The non-future can also be used to describe typical procedures with a non-specific subject.

- (245) *Moletin* *bit* *va-a* *oum,*
 person 3SG.NFUT.want 3SG.IFUT-eat crab
be *bien,* *be* *metteh,*
 3SG.NFUT.go_to sea 3SG.NFUT.go_to creek
ba *mi-gil-i*
 GO.REAL 3SG.NFUT-NFUT.dig-3OBJ
 ‘[When] someone wants to eat crab, they go the sea, they go to the creek, they go dig them up.’ [20141216h_p01s046_02]

This use for typical procedures or situations is also employed for alternatives when only one would be true in a specific instance.

- (246) *samtaem* *ma-be* *openem* *del* *pre*
 sometimes 1PL.EXCL.NFUT-NFUT.COP open.BIS 3SG.NFUT.with pray
mu va *tei* *ma-ba* *ma-las* *tang*
 or time one 1PL.EXCL.NFUT-NFUT.go 1PL.EXCL.NFUT-clear_land just
 ‘Sometimes we start with a prayer or sometimes we just clear the land.’ [20141030a_p01m004_08]

- (247) *mei* *mi-so-ni* *da* *mu* *be* *usum-ni*
 COME 3SG.NFUT-put-3OBJ 3SG.NFUT.stay or 3SG.NFUT.COP use-TR
 ‘they come and put it [money] away or they use it.’ [20141030a_p01m004_42]

This use for typical situations also extends to very probable possibilities in subordinate clauses. More usually, verbs subordinated to *kil* ‘can, know how to’ take one of the future tenses, but in (248) the non-future is used to show that this is a frequent occurrence in the real world, not just a possibility. This use with *kil* could be related to the use of an auxiliary *save* ‘know’ in Bislama for habitual meanings by some speakers (Crowley 2004b: 99).

- (248) *venu* *mi-kil-a* *bas-i*
 volcano 3SG.NFUT-can-3OBJ 3SG.NFUT.hit-3OBJ
 ‘the volcano can damage it [the yam]’ [20141106d_p07e016_05]

The non-future is used with stative predicates for descriptions and defining attributes.

- (249) *inou* *na-be* *misin* *woka* *tei*
 1SG 1SG.NFUT-NFUT.COP mission worker one
 ‘I am a mission worker.’ [20141218f_h01s074_15]

The prior is used for notably past situations (256), especially when the exact time is explicitly stated as an adjunct (257), (258), (259).

- (256) *lat-he* *presin* *e* *Kwinslan*
 3PL.PRI-go_to prison LOC Queensland
 ‘They went to prison in Queensland.’ [20150305h_h01o111_32]
- (257) *nat-hur-i* *e* *tu* *aklok*
 1SG.PRI-take-3OBJ LOC two o’clock
 ‘I gave birth to her at two o’clock.’ [20150413a_h01s122_03]
- (258) *moletin* *hau* *tavatang* *tahal* *Edu* *late-met* *pat* *1951*
 person lots very_much from Endu 3PL.PRI-die sleep 1951
 ‘Lots and lots of people from Endu passed away in 1951.’
 [20150414b_h01e016_91]
- (259) *lat-hur-i* *te-he* *hospitel* *nangis*
 3PL.PRI-take-3OBJ 3SG.PRI-go_to hospital before_yesterday
 ‘They took him to hospital the day before yesterday.’
 [20170412b_n01m167_10]

The prior is also used for notably past habitual or iterative situations, either alone (260), (261), or with the continuous auxiliary *ti*, *di* (224), (262).

- (260) *late-pol* *ena* *enjin* *tena* *tavak*
 3PL.PRI-work LOC engine of tobacco
 ‘They [men abducted by blackbirders] worked on the tobacco engine.’
 [20150305h_h01o111_40]
- (261) *xamem* *mat-naa-skul* *=ti*
 1PL.EXCL 1PL.EXCL.PRI-NEG-school =NEG
 ‘{During World War Two} we didn’t go to school.’
 [20150219b_n01m001_009]
- (262) *di* *lat-he* *usum* *te-xoni* *saak*
 CONT.REAL 3PL.PRI-COP use.BIS 3SG.PRI-be_like shark
 ‘They used it [a traditional war canoe] like a shark.’
 [20150228g_n01s100]

The prior is also used to express facts about the past which are no longer true. Example (263) explains that Moru used to have three *nakamals* (meeting houses and family units), whereas now it only has one.

- (263) *Metimel* *tahal* *Moru*, *amel* *xil* *te-he* *tol*
 village of Moru nakamal PL 3SG.PRI-COP three
 ‘In Moru village, there were three nakamals.’
 [20141218g_h01s075_01-02]

Narratives often begin by setting the scene in the prior, establishing a past reference time, before moving into the non-future.

(264) *tetiamu xa tetiamu kokot, mometu xil lat-ti*
 before SUB before long_time ancestor PL 3PL.PRI-stay
tamu Vesoso igoxor. La-di vuma...
 still Vesoso over_there 3PL.NFUT-NFUT.stay on_and_on
 ‘A long long time ago, our ancestors still lived over there in Vesoso. They lived there until...’ [20150305h_h01o111_02]

During narratives, a specific prior time is often established in a subordinate ‘when’ clause, with the main clause in the non-future.

(265) *taem xa wol wo te-ta,*
 time REL world war 3SG.PRI-stay
moletin xil ba la-gil vul
 person PL GO.REAL 3PL.NFUT-NFUT.dig hole
 ‘When the World War was happening, people went and dug holes [*to hide from planes*]’ [20150219b_n01m001_10-11]

Even without an explicit subordinator, the use of the prior seems to have this function, as in:

(266) *nou nat-te Penapo,*
 1SG 1SG.PRI-stay_at Penapo
ma di na-pis-ni mak
 then CONT.REAL 1SG.NFUT-roll-TR 3SG.NFUT.like_this
 ‘[When] I lived in Penapo, then I would roll them [*laplap simboro*] like this.’ [20141212g_p01s046_07-08]

One of the most common uses of the prior is to report situations that happened before the current topic time in a narrative, which are relevant to the topic time. In (267), the story has now reached the morning after Cyclone Pam, but Elder Saksak needs to say how he came to have a chainsaw with him, to explain how he cleared the road back to the house.

(267) *nat-po jenso e ba*
 1SG.PRI-carry chainsaw and GO.REAL
nate-pat nat-tanga klasrum
 1SG.PRI-sleep 1SG.PRI-stay_behind classroom
 ‘I’d brought a chainsaw and gone and slept back in the classroom.’
 [20150414b_h01e016_49]

This function of the prior is especially useful for narrative repair, when a detail missed earlier in the storyline suddenly becomes very important to the

narrative. In (268) the bushnuts the boys had gathered earlier are vital to stopping the snake that is now circling their house. In (269) the story of how the rat came to have a tail is nearing its climax when the speaker remembers to tell us that the rat did not have a tail to start with.

(268) *Ena horamue xalu ok lute-leh huit tavu xil*
 and boy DU PROX 3DU.PRI-take fruit bush_nut PL
 ‘And the boys had picked bushnuts.’ [20141220g_n01s080_48]

(269) *asu xale-n te-he tovuol*
 rat tail-3SG.POSS 3SG.PRI-COP absent
 ‘The rat didn’t have a tail.’ [20141219c_n01s033_21]

The prior is especially frequent in relative clauses, to give background information or identify a participant in terms of earlier events in the narrative.

(270) *frog [xa lute-hur-i tetiamu] ma danga tang*
 frog REL 3DU.PRI-take-3OBJ before then 3SG.NFUT.stay_behind just
 ‘The frog they owned earlier just stayed behind.’
 [20150118a_n01m090_76]

(271) *ma duxoh melengien [xa te-ling xat-i]*
 then 3SG.NFUT.reach day REL 3SG.PRI-put+on-3OBJ
 ‘Then it got to the day they’d set for it [*the ceremony*].’
 [20150219b_n01m001_33]

(272) *mi-leh rat kava xil [xa*
 3SG.NFUT-take+out corrugated_iron PL REL
mate-tii te-kkoxol wido xil]
 1PL.EXCL.PRI-nail 3SG.PRI-block window PL
 ‘It [*the cyclone*] took off the corrugated iron that we’d nailed to the windows.’ [20150419d_h01s046_23]

(273) *masta Utie, ngan xa [mate-pol mi-ni],*
 master Utie the_one REL 1PL.PRI-work for-3OBJ
mi-sa sip na-n ba
 3SG.NFUT-send ship CL.GEN-3SG.POSS 3SG.NFUT.go
 ‘Master Utie, the one we worked for, sent his ship over.’
 [20141117a_n01m003_28]

As with the subordinate ‘when’ clauses mentioned above, the prior can be used by itself in a zero-marked relative clause to give this background information.

- (274) *la-be* *visal* *metalo* *tei*
 3PL.NFUT-NFUT.go_to near white_person one
[te-te *Vemavili]*
 3SG.PRI-stay_at Vemavili
 ‘They went to white man [*who*] lived at Vemavili.’
 [20141215b_h01s062_08]

Finally, the prior is sometimes used with stative verbs to describe participants, as in the relative clause in (275). In the third-person singular this function is ambiguous with the adjectiviser *te-*, *ta-* (§5.3.2), as adjectives can be used predicatively without the copular verb *he* in a relative clause (276).

- (275) *moletin* *xil* *xa* ***late-lep***
 person PL REL 3PL.PRI-big
 ‘the big men [*i.e. bosses*]’ [20141119c_c01s039_47]
- (276) *mi-kop-ni* *hat* *xil* *xa* ***te-lep***
 3SG.NFUT-throw-TR stone PL REL 3SG.PRI/ADJZR-big
 ‘It threw big stones’ [20150226a_n01s098_46]

4.2.3 Immediate future

The immediate future tense is used for situations in the near future or that happen immediately after TT, and many different irrealis contexts. Where TT is earlier than TU (i.e. the future in the past), then these contexts are often irrealis, in that they talk about branching timelines that are not, or might not be, part of actual history. In most of these contexts speakers can alternatively choose to use the distant future; the distance component of the meaning is intensional rather than having rigid boundaries (Comrie 1985: 90–91).

The immediate future is used to show that a situation is about to happen, often used with the adjunct *sangas* ‘near, nearly’.

- (277) “*Mama!* *To-he* *klasrum* *vena*
 mother 2PC.IMP-go_to classroom because
vatiang ***va-mmei*** ***va-xehik*** *xiak!*”
 wind 3SG.IFUT-come 3SG.IFUT-strong now
 “Mother! Go to the classroom because the cyclone is going to get strong now!” [20150414b_h01e016_11]
- (278) *sangas* ***va-he*** *fiftin* *yias*
 near 3SG.IFUT-COP fifteen years
xa *na-pol-ni* *pol-ien* *ak*
 REL 1SG.NFUT-do-TR do-NMLZR PROX
 ‘Soon it will be fifteen years that I’ve been doing this work.’
 [20141116a_p01s026_03]

- (283) *moletin xil di la-pol ves suse tangan*
 person PL CONT.REAL 3PL.NFUT-do find road in_order_to
xa [la-kamuet selen va-xo e-n]
 SUB 3PL.IFUT-find money 3SG.IFUT-from inside-3SG.POSS
 ‘People find ways to make money out of it [*yam*].’
 [20141106d_p01e016_09-10]

The immediate future frequently occurs in subordinate clauses of non-factive matrix verbs that do not make a commitment to the reality of their complement clauses.

- (284) *e-m bei [o-va-ni]*
 inside-2SG.POSS 3SG.NFUT.like 2SG.IFUT-IFUT.eat-3OBJ
 ‘You want to eat it.’ [20141106d_p11e016_11]

Similarly the immediate future occurs in subordinate clauses of verbs that denote impossibility, either by their lexical meaning, like the verb *sakras*, *saxaras*, *saxa* ‘cannot, unable to’ (285), (286), or by the meaning of the verb in context, as in (287). The loan verb from Bislama *fas* ‘stuck, get stuck’ here means that the road is too littered with debris to reach the house.

- (285) *mi-saxa [rato-pat]*
 3SG.NFUT-cannot 1PC.INCL.IFUT-sleep
 ‘We couldn’t sleep.’ [20150419e_h01m128_20]

- (286) *mi-ngngel, mi-sakras [va-ngngel ve-he ovataem]*
 3SG.NFUT-rest 3SG.NFUT-cannot 3SG.IFUT-rest 3SG.IFUT-COP overtime.BIS
 ‘She sleeps but she can’t sleep too long.’ [20150303g_p01e009_17]

- (287) *susei tena terak te-he fas va-mmei tim*
 road of truck 3SG.PRI-COP stuck 3SG.IFUT-come home
 ‘The road was blocked from reaching the house.’
 [20150413b_h01s123_09]

Any verbs subordinated to the immediate future marked verb after a verb of impossibility are also marked for immediate future.

- (288) *taxiak o-sakras [o-pus meu*
 now 2SG.NFUT-cannot 2SG.IFUT-see megapode
[va-te sangas tim xal tuto]]
 3SG.IFUT-be_at close home with chicken
 ‘Now you can’t see the megapode close to the village with the chicken.’
 [20141212f_n01s054_19]

- (294) *ngan xa [va-viles sepini-en*
 one REL 3SG.IFUT-turn say-NMLZR
na Hii va-mmei ve-he
 CL.GEN God 3SG.IFUT-come 3SG.IFUT-COP
sepini-en na-r]
 say-NMLZR CL.GEN-1PL.INCL.POSS
 ‘Someone who translates the word of God into our language {will find it hard}.’ [20141116a_p02s026_09]
- (295) *bit [va-kamuet mesal tova [va-he insaet]]*
 3SG.NFUT.want 3SG.IFUT-find space some 3SG.IFUT-go_to inside
 ‘He wanted to find a way in.’ [20141220g_n01s080_46]

The immediate future is especially associated with clauses subordinate to *hit*, *hitene* ‘want, say’. When the ‘want’ reading of the verb is relevant, the immediate future can express wished for or intended situations that immediately follow utterance time (296), or the current reference time which could be in a fictive narrative (297), or in the past (635). It is used regardless of whether the stated desire comes to pass (296), (635), or not (297). These are examples of self-directed directive modal contexts, where the self is the subject of the verb of desire (Timberlake 2007: 319).

- (296) *inou na-bit [na-miteni va-mak]*
 1SG 1SG.NFUT-NFUT.want 1SG.IFUT-IFUT.speak 3SG.IFUT-like_this
 ‘I want to say this...’ [20150419c_h01m004_54]
- (297) *tatal ak bit [va-a xalu]*
 snake PROX 3SG.NFUT.want 3SG.IFUT-eat 3DU
 ‘The snake wanted to eat them.’ [20141220g_n01s080_31]

The use of *hit*, *hitene* ‘want’ can be extended to corporate (298) or inanimate subjects (299), (300), where it is more difficult to give a desiderative reading to the verb. In these cases, the matrix verb seems to be bleached of lexical semantics and instead seems to have a grammatical function emphasising the immediacy of the action.

- (298) *Prespiterian joj bit [va-hur]*
 Presbyterian church 3SG.NFUT-want 3SG.IFUT-take
idepedens na-n]
 independence CL.GEN-3SG.POSS
 ‘The Presbyterian Church wanted to gain its independence.’
 [20141218g_h01s075_32]

- (299) *taem xa vatiang bit [va-mmei],*
time REL wind 3SG.NFUT.want 3SG.IFUT-come
bemei e vongien
3SG.NFUT.come LOC night
‘When the cyclone was going to come, it came at night.’
[20150419c_h01m004_04]
- (300) *e tu aklok xos, ut bit [ti va-lan]*
LOC two o’clock maybe place 3SG.NFUT.want CONT 3SG.IFUT-dawn
‘At around two o’clock, it was about to get light.’
[20150413a_h01s122_03]

The extension of the use of a verb meaning ‘want’ to these contexts could suggest that the verb *hit* is on an established grammaticalization path from desire to future meaning (Bybee, Perkins & Pagliuca 1994: 256). There are also a couple of examples in the corpus where the expected subject agreement prefix on *hit* does not appear, suggesting that the verb is undergoing phonological reduction in line with its bleached semantics. This process is similar to that proposed for the development of AVCs in Vatlongos (§8.3.1), although with a subordination structure as its starting point (§7.2.1), rather than SVCs (Chapter 6), resulting in different TAM marking patterns on the grammatical and lexical verbs.

- (301) *bit [lato-pat]*
NFUT.want 3PC.IFUT-sleep
‘They were about to sleep.’ [20170222f_n01s153_48]

The immediate future is also frequently used in reported speech subordinated to *hit*, *hitene* with the sense of ‘say’, when reporting utterances with directive modality, either for reported intentions (302), (303), or reported instructions (304). The use of the immediate future in these context reflects a cross-linguistic tendency for the same grammatical strategy to be used for both purpose clauses and the complement of verbs of wanting and ordering (Bybee, Perkins & Pagliuca 1994: 230), although the distant future can also be used in these contexts.

- (302) *taem xa turis bitene [va-mot*
time REL tourist 3SG.NFUT.say 3SG.IFUT-fall
ena tavon plen]
LOC place plane
‘When a tourist says they’re going to land at the airport’
[20141105f_p01e013_004]
- (303) *tei mun tahal Limen bitene*
one again from Limen 3SG.NFUT.say
[va-mmei va-pus nou]
3SG.IFUT-come 3SG.IFUT-see 1SG
‘Another one from Limen said he’d come and see me.’
[20150223a_n01m096_54]
- (304) *na-biteni mi-ni na-bitene [ha va-ha*
1SG.NFUT-NFUT.say to-3OBJ 1SG.NFUT-NFUT.say GO 3SG.IFUT-go
va-leh na neta, tapolen e NDMO]
3SG.IFUT-take HES what tarpaulin LOC NDMO
‘I said to him to go and get some um, tarpaulin from the NDMO
[*National Disaster Management Office*].’ [20150419e_h01m128_58]

The immediate future is also used to discuss counterfactual situations, for example to talk about events that appeared to be imminent but then did not happen:

- (305) *sip mei va-so xil e ut,*
ship COME 3SG.IFUT-put 3PL LOC shore
mei taa-so xil =ti
COME 3SG.NFUT.NEG-put 3PL =NEG
‘The ship came to put them ashore, it didn’t come put them there.’
[20150305h_h01o111_35]

The immediate future can also be used for counterfactual conditionals, to discuss alternate timelines if an event in the past had happened differently.

- (306) *vahit ena naa-ve tavu xa xiak,*
if but 3SG.FUT.NEG-COP bushnut REL PROX
ratel vus rato-met
1PC.INCL all 3PC.INCL.IFUT-die
‘If it wasn’t for this bushnut, we’d all be dead.’
[20141220g_n01s080_57-58]

The immediate future can also be used as a counterfactual main clause to express regret or pity that the situation did not happen, similar to English ‘if only’. In example (307), the foolish (*unu*) brothers would not have drowned if only they had looked up at the fruit tree, rather than down at its reflection.

- (307) *Ti lat-ketteh ve-he nesau, e*
 CONT 3PC.IFUT-look 3SG.IFUT-go_to up but
xatel lat-taa-ketteh ve-he nesau.
 3PC 3PC.NFUT-NFUT.NEG-look 3SG.IFUT-go_to up
 ‘If only they’d been looking up, but they didn’t look up.’
 [20170217a_n01s038_30]

Finally, the immediate future is sometimes used to express obligations in conjunction with the Bislama loan-verb *mas* ‘must’, although the distant future is more common for this function.

- (308) *ngan-ak mas la-vas vueili tei*
 one-PROX must 3PL.IFUT-IFUT.kill pig one
 ‘These ones must kill a pig.’ [20150305f_p01s110_24]

The many irrealis contexts that the immediate future expresses encourage an analysis as a modal category, but there is great overlap with the distant future for all these functions. As these two categories are almost only distinguished by their different temporal extensions, they are described as primarily relative tenses.

4.2.4 Distant future

The distant future is used to report situations in the distant future some time after the topic time, as well as many irrealis functions, where it overlaps with the immediate future.

The distant future is typically used for situations in the future that are not about to happen immediately: it can describe situations that will occur later in the day, tomorrow, or in the following weeks, months or years:

- (309) *xosali mahis, mati-anien i-tel xatel*
 today afternoon 1PC.EXCL.DFUT-eat 3SG.DFUT-with 3PC
 ‘This afternoon we’re going to eat with them.’
 [20150305f_p01s110_30]
- (310) *mei mu-he intaviu-ni xi*
 COME 1PL.EXCL.DFUT-COP interview-TR 3SG
 ‘We’ll come interview her {tomorrow at eight o’clock}’
 [20150223a_n01m096_12]
- (311) *meul-ien na holesok xil i-mmei mun*
 live-NMLZR CL.GEN thing PL 3SG.DFUT-come again
 ‘The life of everything will come back {in the next three or four months}.’ [20150414b_h01e016_82]

modal functions, either expressing predictions (317) or purposes (318), (319), (320).

(317) *la-bitene* [*vatiang* *ok* *i-has*
 3PL.NFUT-NFUT.say wind PROX 3SG.DFUT-hit
xir *ena* *maakue* *ok*]
 1PL.INCL LOC morning PROX
 ‘They said this cyclone would strike us in the morning.’
 [20150410a_h01s121_17-18]

(318) *inou* *na-be* *reti*
 1SG 1SG.NFUT-NFUT.COP ready
tangan [*ni-he* *maket* *Vila*]
 in_order_to 1SG.DFUT-go_to market Vila
 ‘I was ready to go to market in Vila.’ [20141119c_c01s039_35]

(319) *ma-ba* *ma-lang-ni* *anien* *xa*
 1PL.EXCL.NFUT-NFUT.go 1PL.EXCL.NFUT-look-TR food REL
mu-hur-i *i-mmei* *i-tutou-ni* *xamem*
 1PL.EXCL.DFUT-carry-3OBJ 3SG.DFUT-come 3SG.DFUT-help-TR 1PL.EXCL
 ‘We went to look for food that we’d bring back and help us.’
 [20150410a_h01s121_28]

(320) *i-sak-ni* *meul-ien* *i-hos.*
 3SG.DFUT-make-TR live-NMLZR 3SG.DFUT-good
 ‘It [*keeping pigs in fences*] would make life better. {*But people didn’t want to*}.’ [20150219b_n01m001_28]

One of the most common uses of the distant future in narratives set in the past is for rhetorical questions denoting an impossibility or insurmountable difficulty.

(321) *ma-kuk-ni* *be* *ut* *lelep* *e*
 1PL.EXCL.NFUT-cook-TR but place muddy and
mu-soxat-i *i-tep?*
 1PL.EXCL.DFUT-cover-3OBJ 3SG.DFUT-how
 ‘We cooked it but the place was muddy and how would we cover it up?’
 [20141121a_n01m041_44]

(322) *E* *visi* *ha* *i-pou-e?*
 but who GO 3SG.DFUT-carry-3OBJ
 ‘But who would go get it? [*food in the kitchen during a cyclone*]’
 [20150419d_h01s046_44]

Unsurprisingly, the distant future tense is very often used for making plans (323), discussing the consequences of those plans (324) and planning how to respond to anticipated future situations (325), (326). According to Dahl (1985: 105), this is the most typical use of future tenses cross-linguistically.

- (323) *lu-ga* “*ei, rali-pol-ni* *i-tep?*”
 3DU.NFUT-NFUT.say hey 1DU.INCL.DFUT-do-TR 3SG.DFUT-how
 ‘They said “Hey, what shall we do?”’ [20141027a_n01m001_121]
- (324) *ral-kikei* *ena* *tahei* *i-mmei*
 1DU.INCL.FUT-sing and wave 3SG.DFUT-come
ma *mei* *i-hur rat* *tovoli*
 then COME 3SG.DFUT-take+out old_woman
 ‘We’ll sing and a wave will come and carry away the old woman.’
 [20141027a_n01m001_122]
- (325) *taem xa* *ngan* *i-mmei* *mak*
 time that one 3SG.DFUT-come like_this
e *u-tal xat-i*
 then 2SG.DFUT-catch+on-3OBJ
 ‘When one comes this way then you grab it.’
 [20141027a_n01m001_42]
- (326) *Vatiang* *i-mmei* *ma inou*
 wind 3SG.DFUT-come then 1SG
ni-po *anien* *sung ni-loh* *ni-hat*
 1SG.DFUT-carry food first 1SG.DFUT-run 1SG.DFUT-just_go
 ‘The cyclone will come and I’ll get the food and just run over.’
 [20150414b_h01e016_21]

The distant future is also used in the subordinate clauses of non-factive matrix verbs including *hit* ‘want’, although less commonly than the immediate future.

- (327) *Ahii mi-kil-a* [*i-he* *helpem mi xir*]
 God 3SG.NFUT-can-3OBJ 3SG.DFUT-COP help.BIS to 1PL.INCL
 ‘God can help us.’ [20170406a_n01m164_65]
- (328) *na-bit* [*ni-hiteni* *i-mak*]
 1SG.NFUT-NFUT.want 1SG.DFUT-say 1SG.DFUT-like_this
 ‘I want to say this’ [20141106d_p01e016_12]
- (329) *visuvong* *ma-bit* [*mu-he* *mitim-ni*
 tomorrow 1PL.EXCL.NFUT-NFUT.want 1PL.EXCL.DFUT-COP meet-TR
xi *ena* *eapot* *ena* *eit* *aklok*]
 3SG LOC AIRPORT LOC eight o’clock
 ‘Tomorrow we want to meet her at the airport at eight o’clock.’
 [20150223a_n01m096_11]

On the other hand, the distant future is much more common than the immediate future for expressing obligation with the Bislama loanword *mas* ‘must’.

- (330) *taem xa melengien na sup xil*
time REL day CL.GEN chief PL
i-mas mati-has vueili tei
3SG.DFUT-must 1PC.EXCL.DFUT-kill pig one
‘When it’s Chief’s Day we have to kill a pig.’ [20150305f_p01s110_36]
- (331) *asu i-mas ha i-a-ni*
rat 3SG.DFUT-must GO 3SG.DFUT-eat-3OBJ
‘When you leave food somewhere, the rat has to go eat it.’
[20141219c_n01s033_14]
- (332) *u-mas u-mmei u-sa xouk mi Ahii*
2SG.DFUT-must 2SG.DFUT-come 2SG.DFUT-give 2SG to God
‘You must come give yourself to God.’ [20150118b_n01m090_24]

In conditionals, the distant future can express obligation independently without the use of *mas* ‘must’ after a non-future marked protasis. If the apodosis was also marked non-future it would only express the modality of a contingency. The distant future is therefore marking the deontic modality of an obligation.

- (333) *Xos xa tame atuli xil duxoh ten,*
if SUB father girl PL 3SG.NFUT-reach ten
li-lang ves vuei i-he ten
3PL.DFUT-look+for pig 3SG.DFUT-COP ten
‘If the bride has ten uncles, they have to find ten pigs.’
[20150219a_p01m001_29]
- (334) *Xos xa o-leh rat-i ma u-hiteni*
if SUB 2SG.NFUT-take+out-3OBJ then 2SG.DFUT-say
‘If you took it out then you must say so!’ [20141027a_n01m001_114]

The distant future is also used for future possibilities with the adverb *xos* ‘maybe’.

- (335) *xos i-he tovuol mun, ra-taa-kil =ti*
maybe 3SG.DFUT-COP absent again 1PL.INCL.NFUT-NFUT.NEG-know =NEG
‘maybe there’ll be no more, we don’t know.’ [20150410a_h01s121_42]

Like the immediate future, the distant future can be used for reported intentions (336) and instructions (337), (338).

- (336) *avu bitene [mei i-kkah e tim]*
grandparent 3SG.NFUT.say COME 3SG.DFUT-make_laplap LOC home
‘Grandma said she’ll come make laplap at the house.’
[20141116b_c01m_05]

- (337) *lat-bitene* **[ni-ha** **ni-pol-ni**
 3PC-NFUT.say 1SG.DFUT-go 1SG.DFUT-do-TR
kontrak *tei* **i-tel** *xatel]*
 contract one 3SG.DFUT-with 3PC
 ‘They said for me to go do a contract with them.’
 [20170406a_n01m164_44-45]
- (338) *a-long-e* **[na-bit**
 2SG-hear-3OBJ 1SG.NFUT-NFUT.say
[u-naa-va =*ti* *vatit* *vak]]*
 2SG.DFUT-NEG-NEG.go =NEG tree banyan
 ‘You heard me tell you not to go to the banyan tree.’
 [20141106f_n01e018_23]

The distant future is also used, along with the immediate future, for situations that did not happen in the actual (or fictional) timeline, especially for mistaken thoughts and predictions.

- (339) *haromue* *ak* *mi-nnem-i*
 boy PROX 3SG.NFUT-think-3OBJ
bit **[kuli** *sa-n* **i-met]**
 3SG.NFUT.think dog CL.DOM-3SG.POSS 3SG.DFUT-die
 ‘The boy thought his dog would die.’ [20141212h_n01s046_21]
- (340) *ma-bitene* **[i-he** *tistroem* Vanuatu *vari*,
 1PL.EXCL.NFUT-think 3SG.DFUT-COP destroy Vanuatu at_once
Vanuatu **i-he** *tavorel* mun]
 Vanuatu 3SG.DFUT-COP absent again
 ‘We thought it [*the cyclone*] would destroy Vanuatu completely, there would be no more Vanuatu.’ [20150419c_h01m004]

The distant future is often used in conditionals. In potential conditionals situations in both the protasis and the apodosis are marked for distant future, without making a commitment to the situations actually occurring:

- (341) *xos* *xa* **u-tuv-a** *ma*
 if SUB 2SG.DFUT-shoot-3OBJ then
let *ha* **i-posou**
 lead GO 3SG.DFUT-bang
ena *aeankot*, *let* **i-tilomun.**
 LOC iron_coat lead 3SG.DFUT-return
 ‘If you shoot him then the bullet will go bounce off his iron coat and the bullet will come back.’ [20150305h_h01o111_18]

(342) *xos xa mu-hur-i i-he nesau*
 if SUB 2PL.DFUT-take-3OBJ 3SG.DFUT-go_to on_top
ma tas i-has sip i-ta tak
 then sea 3SG.DFUT-hit ship 3SG.DFUT-stay just
 'If you take him upstairs then the sea will just keep the ship here.'
 [20150305h_h01o111_30]

Both the protasis and the apodosis are also marked for distant future for counterfactual conditionals: contingencies that did not occur in the past and their outcomes.

(343) *ni-sepin e xi mun i-he*
 1SG.DFUT-speak but 3SG again 3SG.DFUT-COP
alaut ma na-taa-sepin =ti
 aloud.BIS so 1SG.NFUT-NFUT.NEG-speak =NEG
 'If I'd spoken, he'd have got loud too so I didn't speak.'
 [20150419e_h01m128_23]

The distant future is also occasionally used to give orders, where it seems to be politer than the morphological imperative. This could be an influence of Bislama, where the imperative includes a subject index, unlike the zero morpheme for singular imperative in Vatlongos. On the other hand, as the distant future is already associated with obligations, this could have always been an option in Vatlongos. It is common for the same grammatical strategy to be used for future and imperative cross-linguistically (Bybee, Perkins & Pagliuca 1994: 273–274).

(344) *u-ha u-po tutut retio*
 2SG.DFUT-go 3SG.DFUT-carry little radio
 'Go and get the little radio.' [20150419e_h01m128_46]

4.2.5 Imperative and prohibitive moods

The imperative and prohibitive moods are used to give orders to (not) do something. They are therefore examples of directive modality where the responsibility for a situation (not) happening in the real world is conferred to the addressee(s).

(345) *Ale, kikei-ni!*
 okay 2SG.IMP.sing-TR
 'Okay, sing it!' [20141220d_n01s082_27]

additional example in a subsequent paper (Crowley 1983: 271). Crowley (1991: 184) comments that both the Paamese and Vatlongos forms ‘translate as ‘might’ or ‘lest’ and also occur in warning constructions (‘watch out or VP’).’

4.3 Other strategies for marking aspectual and modal distinctions

Vatlongos has periphrastic strategies for expressing many aspectual distinctions, and the deontic modality of obligation. The functions of the ‘minimiser’ form of the continuous auxiliary, and repetition of the predicate for intensification could be considered evaluative modalities. Some of these are discussed in later chapters, but others are discussed here. The aspectual and modal categories and the associated strategies are summarised in Table 20.

These distinctions are less central to the Vatlongos TAM system than the morphological categories discussed above. Cross-linguistically, less central TAM categories are often marked with of periphrasis (Bhat 1999: 102) and ‘alternative ways of expression’ (Dahl 1985: 23). These periphrastically-marked categories can therefore be further analysed as more or less grammaticalized, and more or less central to the Vatlongos TAM system, according to the availability of alternative means of expression. Thus, the continuous, which is overwhelmingly expressed using the auxiliary verb *ti*, is more central than the durative, which can be expressed by any combination of three strategies.

Table 20: Non-morphological marking of aspectual and modal categories

TAM distinction	Strategy	Discussion
Continuous (progressive and habitual) aspect	Auxiliary <i>ti, ta</i> ‘stay’	§8.2.1
Continuous and minimiser	Auxiliary <i>tit, tat</i> ‘just stay’	§8.2.1
Inchoative aspect	Auxiliary <i>mei</i> ‘come’	§8.2.2
	<i>ammei</i> ‘come’ as V1 in non-contiguous SVC	§6.4.10
Completive aspect	<i>vus</i> ‘finish’ in independent clause with 3SG subject, often with other discourse markers e.g. <i>ma</i> ‘then’, <i>di</i> ‘then’	§4.3.1

TAM distinction	Strategy	Discussion
Durative aspect	Repeated VP or predicate (often combined with one of the other durative strategies)	§4.3.2
	Discourse marker <i>ma, mo, ve, vuma</i> ‘until’ with distinctive prosodic contour	
	<i>ha</i> ‘go’ in an independent clause with 3SG subject	
Intensifier	Repeated VP or predicate	§4.3.3
	Repeated VP or predicate connected with the general subordinator <i>xa</i>	
Obligation (a deontic modality)	Loan verb <i>mas</i> as an auxiliary verb	§8.2.3
	Loan verb <i>mas</i> as V1 in non-contiguous SVC	§6.4.11

4.3.1 Completive aspect

Completive aspect is marked periphrastically with the verb *vus* ‘finish’, taking third-person singular subject-indexing in an independent clause. The subject is understood to be the situation in the previous clause, and *vus* is marked for the same TAM category as the situation it modifies. It marks a situation as being completed and finished. A lexical verb meaning ‘finish’ is a frequent source for grammatical morphemes marking the completive aspect cross-linguistically (Bybee, Perkins & Pagliuca 1994: 56), and *finis* ‘finish’ is also used as a post-verbal modifier marking completive in Bislama (Crowley 2004b: 103–104).

(351) *hat mi-gan bus, la-si.*²⁰
stone 3SG.NFUT-NFUT.burn 3SG.NFUT.finish 3PL.NFUT-stick
‘The stones finish heating up, they pick them up with a stick.’
[20141028b_p01m003_16]

(352) *na-sikel-ni bus, ti*
1SG.NFUT-weigh-TR 3SG.NFUT.finish then
na-so-ni di oei
1SG.NFUT-put-3OBJ 3SG.NFUT.stay water
‘I finish weighing it, then I put it in water.’ [20141105b_p01s009_04]

²⁰ In §4.3 it is the strategy under discussion, rather than the verb being modified, that is marked in **bold**.

(353) *inou* *ni-pol-ni* *pases* *na-van*
 1SG 1SG.DFUT-do-TR fare CL.GEN-1SG.POSS
i-vus *ti* *ni-kil-a*
 3SG.DFUT-finish then 1SG.DFUT-can-3OBJ
[ni-husil *mama* *xil]*
 1SG.DFUT-follow mother PL
 'I'll finish making enough money for my fare then I'll be able to go
 with the women.' [20141119c_c01s039_30]

In these examples, this strategy resembles an SVC (Chapter 6). Serialisation of a verb meaning 'finish' to mark completive aspect is a common strategy in Oceanic languages (Cleary-Kemp 2015: 137), and the Paamese cognate of *vus* is serialised for this function (Crowley 2002b: 82). However, in Vatlongos *vus* 'finish' is often separated from the preceding clause it modifies by a discourse marker such as *ma* 'then'. It is also frequently followed by the discourse marker *di, ti* 'then'.

(354) *lu-leherat* *tan* *ba* *ma* *bus*
 3DU.NFUT-take_out earth 3SG.NFUT.go then 3SG.NFUT.finish
di *lu-leherat* *uvuei*
 then 3DU.NFUT-take_out leaf
 'They finished taking the earth out then they took off the leaves.'
 [20141027a_n01m001_102]

The completive *vus* often follows a verb marked for durative with one of the strategies discussed below, but never modifies a verb marked with the continuous auxiliary.

In narratives and procedural texts, which are often concerned with the sequence of events and actions, the completive is frequently used in the 'head' of a tail-head linkage structure, a common discourse organisation strategy in the region. The last clause (tail) in a section of discourse is repeated as the first clause (head) of the next section (Early 1994: 454; Hyslop 2001: 426–427; Thieberger 2004: 324–325; de Vries 2005; Schneider 2010: 240; Jauncey 2011: 376). This strategy is closely linked to information structure: when the situation is expressed in a 'tail' it is usually new information; then it is repeated as old information in the 'head', the topic of a new section of discourse. When the situation is first reported in the tail, it is presented as unfinished; then, when it appears in the head, it is often marked with the completive to show it is over

The subject NP (359), (360), or any auxiliary verb (361), (362), are not repeated. Unlike the completive, the durative can co-occur with the continuous auxiliary *ti, di* (361), and cooccurs with the directional auxiliaries (362).

(359) *Us mus mus mus mus*
 rain 3SG.NFUT.rain 3SG.NFUT.rain 3SG.NFUT.rain 3SG.NFUT.rain
 ‘It kept raining.’ [20141117a_n01m003_20]

(360) *Rute xil la-loh la-loh la-loh la-loh*
 some 3PL 3PL.NFUT-run 3PL.NFUT-run 3PL.NFUT-run 3PL.NFUT-run
 ‘Some of them kept running.’ [20150226a_n01s098_18]

(361) *Di lu-kas-i lu-kas-i maa...*
 CONT.REAL 3DU.NFUT-wash-3OBJ 3DU.NFUT-wash-3OBJ on_and_on
 ‘They kept on washing it.’ [20141106f_n01e018_37]

(362) *ba lu-lau lu-lau lu-lau maa...*
 GO.REAL 3SG.NFUT-hunt 3SG.NFUT-hunt 3SG.NFUT-hunt on_and_on
 ‘They went on hunting on and on.’ [20141220g_n01s080_18]

A possible origin for this interpretation would be zero-coordination. Repetition with coordination is one way of showing that a situation continued over a long time, as repeating an activity or process will take longer.

A common discourse strategy is to report the situation once without any durative marking, then repeat the verbal constituent after a prosodic break.

(363) *lul mi-lul.*
 earthquake 3SG.NFUT-quake
Mi-lul mi-lul mi-lul maa...
 3SG.NFUT-quake 3SG.NFUT-quake 3SG.NFUT-quake on_and_on
 ‘An earthquake quaked. It kept quaking on and on...’
 [20150303h_n01e106_03]

Repetition is used both for durative aspect and for intensification. This does not lead to ambiguity as the meaning is determined by the lexical aspect of the predicate. When processes and activities are repeated, the repetition is interpreted as durative aspect. When stative predicates are repeated, the repetition is interpreted as intensification. This is because states are already ‘expected to continue by inertia’ (Timberlake 2007: 284), so durative marking on a stative predicate would be superfluous; an intensification meaning is inferred instead. Processes, on the other hand, ‘require an input of energy’ to

- (368) *di* *la-storian* *ve...*
 CONT.REAL 3PL.NFUT-chat on_and_on
 ‘They were chatting on and on’ [20141105e_n01e012_05]
- (369) *Lata-ba* *lata-ba* *lata-ba* *ve...*
 3PL.NFUT-NFUT.go 3PL.NFUT-NFUT.go 3PL.NFUT-NFUT.go on_and_on
 ‘They kept going on and on’ [20141107b_n01e020_14]
- (370) *mi-til* *e-n* *vuma...*
 3SG.NFUT-sew LOC-3SG.OBJ on_and_on
 ‘She kept on sewing on it {until I came}.’ [20141211a_n01s046_07]
- (371) *xamem* *ma-di* *ma-di*
 1PL.EXCL 1PL.EXCL.NFUT-NFUT.stay 1PL.EXCL.NFUT-NFUT.stay
ma-di *vuma...*
 1PL.EXCL.NFUT-NFUT.stay on_and_on
 ‘We kept on living there {until we came back}.’
 [20141215b_h01s062_34]

The use of durative aspect with *ti, di* ‘stay’ is especially common when setting the scene for a narrative as in (365). The characters are introduced, and then ‘stay on and on’ until the events of the narrative begin.

These discourse markers are usually associated with distinctive prosodic features: an iconically motivated long duration with high pitch and volume ending with a falling contour. These features are seen in Figure 9, a spectrogram of example (365), and are marked with an ellipsis (...) in transcription.

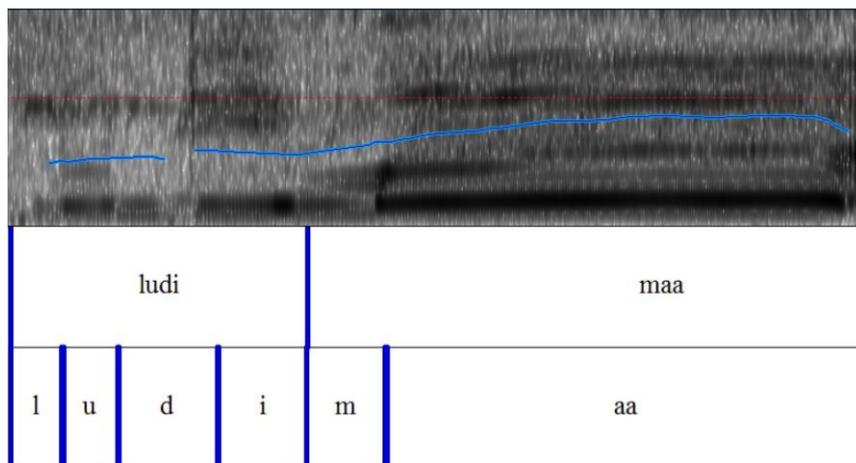


Figure 9: Spectrogram of example (365)

A third strategy for marking durative aspect is the use of *ha* ‘go’ with a third-person singular subject in an independent clause; here it has a meaning of ‘it went on’ or ‘the situation continued’. This is usually combined with the other strategies:

- (377) *besi xi, besi xi, besi xi*
 3SG.NFUT.happy 3SG 3SG.NFUT.happy 3SG 3SG.NFUT.happy 3SG
 ‘They’re very very happy.’ [20141105f_p01e013_98]

When an adjective is used with the copular verb *he*, both the verb and the adjective are repeated, here the adjective *eilep* ‘big’ and the loan adjective *patpat* ‘fat’.

- (378) *vatiang ok be eilep be eilep*
 wind PROX 3SG.NFUT.COP big 3SG.NFUT.COP big
 ‘This cyclone was very big.’ [20150410a_h01s121_21]
- (379) *tuvava ak be patpat*
 baby PROX 3SG.NFUT.COP fat
be patpat be patpat
 3SG.NFUT.COP fat 3SG.NFUT.COP fat
 ‘This baby was very very fat.’ 20170331b_n01s034_05

When an adjective is used in a relative clause, the copular verb *he* is no longer required and the adjective can be repeated by itself:

- (380) *liei xil xa te-lep te-lep*
 branch PL REL ADJZR-big ADJZR-big
 ‘very big branches’ [20150226a_n01s098_48]

This strategy can be used with non-stative predicates marked for imperative or prohibitive (including the clitic *ti*), where it intensifies the urgency of the command rather than the lexical semantics of the predicate.

- (381) *Na-ga “to-ha to-ha to-ha!”*
 1SG.NFUT-NFUT.say IMP.PC-go IMP.PC-go IMP.PC-go
 ‘I said “Go go go!”’ [20141107b_n01e020_08]
- (382) *e ona-mmei =ti ona-mmei =ti*
 no 2SG.PROH-come =NEG 2SG.PROH-come =NEG
 ‘no don’t come, don’t come’ [20141107b_n01e020_08]

Although the bare repetition strategies seem to be identical, it is likely that while the durative aspect repetition derives from zero-marked coordination, the intensifier derives from zero-marked subordination, as the other strategy for marking intensification is similar, except that the repeated predicates are

separated by the general subordinator *xa*, *xe*. A similar strategy expresses intensification, emphasis, or excess in many Vanuatu languages.²²

- (383) *mama na-mel mi-sa xa mi-sa*
 mother CL.GEN-1DU.EXCL.POSS 3SG.NFUT-bad SUB 3SG.NFUT-bad
 ‘Our mother is really bad.’ [20141107b_n01e020_08]
- (384) *vo na vueili a-ralu*
 smell CL.GEN pig CL.ED-1DU.INCL.POSS
bos xa bos!
 3SG.NFUT.good SUB 3SG.NFUT.good
 ‘The smell of our pork is really good!’ [20141027a_n01m001_105]
- (385) *vatiang [xa te-mmei] te-xeih*
 cyclone REL 3SG.PRI-come 3SG.PRI-strong
xa te-xeih xa te-xeih
 SUB 3SG.PRI-strong SUB 3SG.PRI-strong
 ‘The cyclone that came was very very powerful.’
 [20150414b_h01e016_10]
- (386) *insaet ig-ak memer xa memer xa memer*
 inside place-PROX 3SG.NFUT.wet SUB 3SG.NFUT.wet SUB 3SG.NFUT.wet
 ‘Inside here was really really wet.’ [20150415a_h02s125_22]

However, a major difference is that the subordination strategy allows a variety of constituents of different sizes to be repeated. Unlike with bare repetition, an entire clause can be repeated as the subordinated clause including subject NPs and auxiliary verbs. Example (102) above shows a subject NP in the repeated subordinate clause with the possessed subject of the psycho-collocation *e xat* ‘be angry’. Example (387) shows an auxiliary verb repeated in the subordinate clause.

- (387) *asu di men xa di men*
 rat CONT.REAL 3SG.NFUT.laugh SUB CONT.REAL 3SG.NFUT.laugh
 ‘The rat was really laughing.’ [20170122b_n01s036_33]

As with the repetition strategy, an adjective can be repeated with the copular verb *he*:

²² e.g. Daakaka (West Ambrym) (von Prince 2015: 400–2); Bierebo (Epi) (Budd 2009: 307–8), Naman (Malekula) (Crowley 2006: 199); Apma (Central Pentecost) (Schneider 2010: 206); Bislama (Crowley 2004b: 185)

(388) *vet* *uhia* *xil* *ak*
 shoot wild_yam PL PROX
be ***te-lep*** ***xa*** ***be*** ***te-lep***
 3SG.NFUT.COP ADJZR-big SUB 3SG.NFUT.COP ADJZR-big
 ‘These wild yam shoots are very big.’ [20141106d_p04e016_13]

But the adjective alone can also be repeated:

(389) *valiu* *na-n* *be* ***val xa val***
 value CL.GEN-3SG.POSS 3SG.NFUT.COP big SUB big
 ‘Its value is very high.’ [20150305f_p01s110_35]

Unlike the unmarked repetition strategy, the repetition with subordinator *xa* can also intensify dynamic verbs, presumably because the presence of the subordinator *xa* disambiguates this from the durative repetition strategy. Example (390) shows the dynamic verb *luk* ‘grow’ repeated with *xa* for intensification, as well as the exclusion of the auxiliary verb *ha* ‘go’ and a locative preposition phrase adjunct from the repetition.

(390) *o-pus* *huras* *sa-m*
 2SG.NFUT-see yam CL.DOM-2SG.POSS
ba ***mi-luk*** *ena*
 NFUT.GO 3SG.NFUT-grow LOC
vatiei *tei* ***xe***
 tree one SUB
mi-luk ***xe*** ***mi-luk***
 3SG.NFUT-grow SUB 3SG.NFUT-grow
 ‘You see your yam go and really really grow up a tree.’
 [20141106d_p10e016_08]

This strategy can also be used with non-verbal predicates such as NPs:

(391) ***us*** ***xa*** ***us,*** ***vatiang*** ***xa*** ***vatiang***
 rain SUB rain wind SUB wind
 ‘There was lots of rain, there was lots of wind.’
 [20150415a_h02s125_56]

There are also examples where only a quantifier is repeated with *xa*:

(392) *Holu* *xil* *la-lel,* ***holu*** ***xa*** ***holu.***
 lots 3PL 3PL.NFUT-die lots SUB lots
 ‘Lots of them died, lots and lots.’ [20141117a_n01m003_27]

4.4 Summary

The Vatlongos TAM categories that are expressed with morphological paradigms can be divided into two broad semantic domains: realis and irrealis.

The realis categories of prior and non-future describe situations that actually have occurred or are occurring in the real or fictive universe, whereas the irrealis categories describe non-actualised situations. Of these irrealis categories, imperative, prohibitive and apprehensive are best described as moods. Imperative and prohibitive are deontic modal categories, while apprehensive is a combination of epistemic and evaluative modality, describing potential and unwanted situations.

However, the four major morphological TAM categories are best distinguished from each other by their temporal extensions, and are therefore described with tense-based labels. Rather than expressing absolute tense, which identifies situations as occurring before, during or after utterance time, these categories express relative tense. This means that they locate the time of situations in relation to a discourse-determined reference point, or topic time, which may or may not be the same as the moment of utterance.

Most typically, the non-future describes situations that occur at topic time, and the prior describes events that precede topic time. However, as they are both also inherently realis categories, neither can be used to express a relationship between the situation and a topic time in the future.

The immediate and distant future are used for events that occur after topic time, including situations that follow a reference point in the past, without making a commitment to the situation being actualised. Both futures can be used for a wide range of irrealis meanings, but they are distinguished from each other primarily by their temporal extensions: while the distant future can be used for all situations that follow topic time, the immediate future entails that the situation occurred soon after it.

In addition to the morphological TAM categories, whose morphological forms are described in the next chapter, Vatlongos has periphrastic strategies which express less central aspectual and modal categories: the use of *vus* 'finish' to express completive aspect; the use of verbal repetition, discourse markers, and *ha* 'go' to express durative aspect; and the use of verbal repetition, and repetition with the subordinator *xa*, for intensification. Verbal repetition in

Vatlongos is unusual in that any object NP is repeated along with the verb. The aspectual and modal functions of SVCs and AVCs are discussed in §6.4.10, §6.4.11, and §8.2.

5 Verbal morphology

Vatlongos's inflectional verbal morphology is complex. It exhibits cumulative exponence resulting from morphological fusion, a great deal of systematic and unsystematic syncretism, and frequent paradigm splitting or overabundance. The concatenative morphology described in §5.1 applies to all verb stems. The affixes mark the TAM categories discussed in Chapter 4 and negative polarity, as well as cross-indexing the person and number of the subject. First or second-person singular and inanimate third-person objects are marked with a pronominal object suffix on transitive verbs, and intransitive verbs can take an object NP when marked with a transitivising suffix, which can also index a third-person object. Non-linear morphological processes apply to closed classes of verbs and include verb initial consonant mutation, vowel elision in the stem in prosodically licensed environments, and reduplication. The derivational morphology includes nominalizing and adjectivizing suffixes.

This chapter begins by discussing inflectional verbal morphology in Vatlongos, before covering two derivational processes that involve verbal roots. Anderson (1992: 73–135, 180–197) discusses cross-linguistic evidence for positing a division between inflectional and derivational morphology, and I follow his characterisation of inflection as the interaction of morphology with syntax, and derivation as the morphological process in the lexicon for creating lexical items from existing ones. One process in Vatlongos that is a challenge for this distinction is the transitivising suffix *-ni*, which could be analysed as deriving a transitive verb from an intransitive one. I follow Bickel and Nichols (2007: 187) in including the transitivising suffix under inflectional morphology on the basis that 'its occurrence is an obligatory response to at least some syntactic environments'. Although the resulting verb form is used in transitive clauses with an object argument, it has a slightly different distribution from morphologically transitive verbs in Vatlongos. It does not take any pronominal object suffixes, but can occur without an NP object constituent in the third-person. Parker's (1970a: viii) dictionary classifies verbs that occur with the transitiviser *-ni* as a separate class of 'optional-object transitive verbs', which are 'almost all derived' from intransitive verb stems. The transitivising suffix is

therefore discussed along with other concatenative inflectional processes in §5.1.3.

The inflectional morphology is further subdivided into concatenative (§5.1) and non-linear (§5.2) processes. Concatenative morphology is segmentable into individual morphemes (Bickel & Nichols 2007: 181–182); the concatenative processes in Vatlongos can be described using the affix template shown in Figure 10. This division is made for descriptive purposes and is not intended as a claim about the mechanics of the morphological processes described.²³ The ease of taking a ‘concatenative’ approach to the data is also heavily dependent on decisions about where to ‘split’ the affix template into multiple slots. For descriptive clarity, I have treated the subject-indexing prefixes as a single affix slot with multiple paradigms for different TAM categories. However, Parker (1968b) posits many more affix slots and requires more interaction between concatenative and non-linear processes requiring portmanteau forms and many morphophonological processes.

Verbal morphology also shows some variation between speaker communities and between individual villages, especially in the frequency of alternative variants of certain subject-indexing prefixes (§5.1.1.2), and patterns of verb-initial consonant mutation (§5.2.1).

5.1 Concatenative inflectional morphology

The primary inflectional morphological processes in Vatlongos that apply to all verb stems are concatenative; they are readily segmentable and arranged in a linear order (Bickel & Nichols 2007: 181–182).

I describe the Vatlongos verbal affix template as having two slots for prefixes and one for suffixes (Figure 10). The initial prefix slot is obligatory if zero morphemes are assumed to ‘fill’ the slot. Zero morphemes are used for the third-person singular non-future in certain phonological environments, for the third-person singular before any of the prefixes in the second slot, and for the second-person singular in imperative mood. The second prefix slot is optional,

²³ Within morphological theory, this distinction has been argued not to reflect any underlying difference in the morphological architecture (Stump 2001: 11).

and the object or transitivity suffix slot is optional depending on the morphological transitivity of the verb and the presence or absence of an object NP constituent (§5.1.3).

Subject indexing	Negative polarity and apprehensive	Verb root	object/transitivity
Multiple TAM paradigms (Table 22)	Negative: Non-future: <i>taa-</i> Elsewhere: <i>naa-</i> Apprehensive: <i>na-</i>		Object pro-index (Table 27) Transitivity: <i>-ni</i>

Figure 10: Vatlongos verbal affix template

I describe the subject-indexing prefixes as separate paradigms for each TAM category; that is, as cumulative formatives expressing both these features (Bickel & Nichols 2007: 188–189). The paradigms for different TAM contexts could be analysed as further segmentable into two or three affix slots, with portmanteau forms in many cells due to phonological processes resulting in morphological fusion. Portmanteau subject and TAM prefixes are common in canonic Oceanic languages (Ross 2004: 501). This is the analysis pursued by Parker (1968b). He identifies seven prefix positions and two suffix positions, deriving the different paradigms of subject-indexing prefixes from combination with tense prefixes via phonological rules, and similarly deriving the dual and paucal paradigms from combination of the plural (or, in his analysis, non-singular) prefixes and specialised dual and paucal prefixes. This analysis is certainly a possible approach to the data, and probably reflects the diachronic origins of the synchronic system.

However, as Crowley (1991: 219) admits, the ‘extreme structuralist model’ adopted in Parker’s article can obscure the patterns in the data, and this is one reason for presenting these prefixes as cumulative formatives instead. In his sketch grammar of the language, Crowley (2002a: 665–666) treats the subject-indexing prefixes as having different paradigms for non-future, immediate future and distant future. However, his description conceals some of the variation that is deducible from Parker’s analysis (1968b), especially the fairly

free variation between CVCV- and CVC- forms in the dual and paucal numbers and prior TAM category.

The sketch grammar also misleadingly suggests that the prior prefix *te-* (described as past tense) combines with the CVCV- forms of the non-future paucal prefixes listed, predicting forms like *rata-te-pat* ‘1PC.INCL-PRI-sleep’ (which is ungrammatical). To the extent that it is segmentable, the prior formative *te-* or *t-* can only combine with the CVC- form of the non-future paucal prefixes, and the resulting geminate consonant is optionally truncated, so that examples of the possible forms in the first-person inclusive paucal are: *rat-te-pat*, *rat-t-pat* ‘1PC.INCL-PRI-sleep’; *rate-pat*, *rat-pat* ‘1PC.INCL.PRI-sleep’. I therefore list the prior as a separate paradigm of subject-indexing prefixes. I have also put the apprehensive prefix *na-* into the same slot as the negative prefixes, as it never cooccurs with these, resulting in only two prefix slots to Crowley’s three (Crowley 1991: 183).

To summarise, I am treating these forms as cumulative formatives in a single prefix slot for reasons of descriptive clarity, to avoid committing to any one system of phonological processing, and to make this data as useful as possible to researchers who want to pursue a more analytic approach.

Oceanic languages typically have pre-verbal subject-indexing markers and post-verbal object-indexing markers (Ross 2004: 495); the Vatlongos affix template matches this pattern. Subject- and object-indexing affixes in Vatlongos differ in their distribution. The subject-indexing prefixes optionally cooccur with subject NPs (§3.1). They are therefore examples of ‘cross-indexes’ in Haspelmath’s (2013) typology of argument-indexing. The object suffixes that occur with transitive verbs cannot cooccur with object NPs in the post-verbal object position. They are ‘pro-indexes’ in Haspelmath’s terms, or they can be described as bound pronominals, or pronominal object suffixes, as they substitute for a NP, unlike the subject-indexing prefixes. Bresnan and Mchombo (1987) describe a similar divide between subject- and object-indexing as a distinction between grammatical and anaphoric agreement in Chichewa. This pattern is also found in other Vanuatu languages, such as Lolovoli (North-East Ambae) (Hyslop 2001: 55–57). However, this is not canonical for Oceanic

languages: if they have object indexes they tend to cross-index any NP in the same way as the subject indexes (Ross 2004: 499).

5.1.1 Subject indexing prefixes

There are four complete TAM paradigms, as well as partial paradigms for the imperative and prohibitive moods. However, there is a great deal of syncretism, where a morphological distinction is not made in some conditions, resulting in identical forms used in different cells of the paradigm (Stump 2001: 212; Baerman, Brown & Corbett 2005; Bickel & Nichols 2007: 207–208). As a result, an individual prefix often does not uniquely identify a specific TAM category, and patterns of syncretism vary in different subject person-number combinations.

The subject prefix paradigms are further complicated by overabundance, the co-existence of different forms for one set of morphosyntactic features in relatively free variation, which have been called ‘cellmates’ (Thornton 2012a). This appears to be the result of processes of phonological reduction, especially elision of the second vowel in a CVCV- sequence. Many complex cells further obscure TAM distinctions, and this means speakers often have a choice between being more or less specific about the TAM category intended. Community membership has an effect on the use of some of these variants.

5.1.1.1 Full paradigms for the relative tenses

There are full paradigms for the four major TAM categories: prior, non-future, immediate future and distant future. Vatlongos subject-indexing prefixes distinguish four persons and four numbers. Like most Oceanic languages (Lynch, Ross & Crowley 2002: 35), Vatlongos makes a distinction between inclusive (including the addressee) and exclusive (excluding the addressee) first person in non-singular numbers (Bickel & Nichols 2007: 220–223). It has four number categories: singular, dual, paucal and plural. These person and number distinctions are shown in the paradigm of South-Vatlongos independent pronouns in Table 21 (Endu-Vatlongos variants are listed in Table 6).

Table 21: Independent pronoun paradigm in South-Vatlongos

	1INCL	1EXCL	2	3
SG		<i>inou</i> <i>nou</i>	<i>xouk</i>	<i>xi</i>
DU	<i>ralu</i>	<i>xamel</i>	<i>xamil</i>	<i>xalu</i>
PC	<i>ratel</i>	<i>xametel</i>	<i>xamitel</i>	<i>xatel</i>
PL	<i>xir</i>	<i>xamem</i>	<i>xamim</i>	<i>xil</i>

Though most Oceanic languages distinguish dual number from plural, the paucal category is less common (Lynch, Ross & Crowley 2002: 35). Vatlongos has a generous paucal category that can be used for fairly large groups of people, with a rhetorical effect of familiarity or minimisation. While in many languages the paucal is used for an absolute number range (for example, between 3 and 7 referents), Vatlongos patterns with other Oceanic languages like Boumaa Fijian and closely related Paamese in having a ‘relative’ meaning component (Dixon 1988: 52; Corbett 2000: 22–26). Crowley (1982: 81) describes how in Paamese the paucal can be used for fairly numerous groups of people when they are being compared to a larger group, for example, the population of Paama as compared to the population of Vanuatu. A stylistic choice to use the paucal for a large group therefore implies a large number as a point of comparison, and perhaps a large audience for the story being recorded.

Using the paucal for a large group also suggests group cohesion and friendliness. A similar connotation has been described for Daakie (South Ambrym), where reference to family members or a group of friends is more important than group size for predicting use of the paucal (Krifka 2011: 6). Krifka (2018) describes this use of number as an ‘affiliative paucal’. Example (393) specifically numbers the subject of the paucal at fourteen: a group of young men from the same village who are running away from the volcano after an eruption. A very frequent use of the paucal is for a whole village, which can be more than a hundred people. In example (394), the subject is the people who originally moved from Maat to Mele Maat: according to another text this was a group of 61 [20141121a_n01m041_66]. The use of the paucal emphasises the unity of the community in this important foundational history.

subset are identified by a lexical NP (Lichtenberk 2000: 1–5). The inclusory use of the Vatlongos prefixes are an example of a split inclusory construction in Lichtenberk’s typology, but phrasal inclusory constructions are also possible with independent pronouns (e.g. in (450) below, the referent of the name *Maison* is a member of the set identified by the following third-person paucal independent pronoun *xatel*). In (397), the verbal prefix marks the full set of subject referents, while the lexical NP identifies one member of that set. Although only Elder Solomon is mentioned in the subject NP, the use of the paucal subject-indexing prefix is not ungrammatical, but instead shows that Elder Solomon was one of a group of teachers.

(397) *Elta* *Solomon* *lat-pispisi* *igak*
 Elder Solomon 3PC.NFUT-teach here
 ‘Elder Solomon [*and his colleagues*] taught here.’ [20150219b_n01m001_54]

Table 22 shows the subject-indexing prefixes in prior, non-future, immediate future and distant future relative tenses. Following Plank’s (1991: 20) suggestion, the TAM categories are arranged in an order that allow cells with the same form to be adjacent as far as possible, giving the following order: prior, non-future, immediate future, distant future. The cell-mates that result from elision of the second vowel in CVCV- (or VCV-) prefixes are shown by splitting the cells, allowing the resulting patterns of syncretism to be seen more clearly. Other cell-mates are shown one below the other. The second /t/ in the paucal prior forms is in brackets to indicate that this is optionally realised as a geminate or simple consonant; it is assumed this is a phonological process.

Table 22: Complete subject-indexing prefixes for major TAM categories

		1 Inclusive		1 Exclusive		2		3			
SG	PRI	n/a		<i>nate-</i>	<i>nat-</i>	<i>ote-</i>	<i>ot-</i>	<i>te-</i>			
	NFUT			<i>na-</i>		<i>o-</i>		<i>a-</i>		<i>mi-</i>	
	IFUT									<i>∅-</i>	
	DFUT			<i>ni-</i>		<i>u-</i>		<i>i-</i>			
DU	PRI	<i>rute-</i>	<i>rut-</i>	<i>mate-</i>	<i>mat-</i>	<i>mute-</i>	<i>mut-</i>	<i>lute-</i>	<i>lut-</i>		
	NFUT	<i>ru-</i>		<i>ma-</i>		<i>mu-</i>		<i>lu-</i>			
	IFUT	<i>ralo-</i>	<i>ral-</i>	<i>malo-</i>	<i>mal-</i>	<i>mulo-</i>	<i>mul-</i>	<i>lalo-</i>	<i>lal-</i>		
	DFUT	<i>rali-</i>		<i>mali-</i>		<i>muli-</i>		<i>lali-</i>			
PC	PRI	<i>rat(t)e-</i>	<i>rat(t)-</i>	<i>mat(t)e-</i>	<i>mat(t)-</i>	<i>mut(t)e-</i>	<i>mut(t)-</i>	<i>lat(t)e-</i>	<i>lat(t)-</i>		
	NFUT	<i>rata-</i>	<i>rat-</i>	<i>mata-</i>	<i>mat-</i>	<i>muta-</i>	<i>mut-</i>	<i>lata-</i>	<i>lat-</i>		
	IFUT	<i>rato-</i>		<i>mato-</i>		<i>muto-</i>		<i>lato-</i>			
	DFUT	<i>rati-</i>		<i>mati-</i>		<i>muti-</i>		<i>lati-</i>			
PL	PRI	<i>rate-</i>	<i>rat-</i>	<i>mate-</i>	<i>mat-</i>	<i>mute-</i>	<i>mut-</i>	<i>late-</i>	<i>lat-</i>		
	NFUT	<i>ra-</i>		<i>ma-</i>		<i>mu-</i>		<i>la-</i>			
	IFUT							<i>li-</i>			
	DFUT	<i>ri-</i>		<i>mu-</i>		<i>li-</i>					

There are some notable patterns of syncretism in this data. Most examples of affix syncretism involve affixes that express person and TAM (Baerman, Brown & Corbett 2005: 100), and Vatlongos fits this cross-linguistic pattern. Firstly, the non-future and immediate future consistently have the same form in singular and plural numbers, except in the third-person. However, the non-future is always uniquely identified in the dual, where the future tenses are optionally not distinguished from each other. In the paucal, all TAM categories can be distinguished with CVCV- forms, but the CVC- forms do not distinguish any TAM categories except the prior, which is optionally distinguished with a geminate consonant.

Sometimes, I have glossed verbs in examples as FUT ‘future’, because in many person-number-polarity environments immediate future and distant future are indistinguishable. However, it is assumed that the speaker intends either immediate or distant future, and sometimes the ‘real’ relative tense is made clear through agreement, for example in SVCs (Chapter 6) and AVCs (Chapter 8). When the intended tense can be identified through independent grammatical evidence the prefix is glossed as the more precise ‘IFUT’ or ‘DFUT’.

The prior form is usually the non-future form with the addition of *te-* or *t-*. However, the paucal paradigm is evidence against treating the prior as the addition of another prefix to the non-future form. The CVCV- cell-mate ending in *a-*, e.g. *rata-* ‘1PC.EXCL.NFUT’, cannot cooccur with a *te-* or *t-* prior prefix: *ratate-* or *ratat-* ‘1PC.EXCL.PRI’ are ungrammatical. This means that even if a *te-* or *t-* prior prefix were assigned to a separate slot in the affix template, a separate paradigm of prior subject prefixes would still be required. The prior paucal variants *ratte-* or *ratt-* ‘1PC.EXCL.PRI’ could be derived from the addition of a prior prefix to the CVC- non-future cell-mate, but in practice speakers do not always produce a geminate consonant. This could be analysed as truncation of extra-syllabic material (Bickel & Nichols 2007: 182) as there seems to be a preferred syllabic template of CVC, but the occasional realisation of the CVCC form even before a consonant shows that this prosodic constraint is not absolute. As word stress is assigned from the end of a word (discussed in §3.6.2, in relation to the negative clitic), it would also be difficult to formulate the correct domain for

such a constraint to apply in many phonological models. For example, in Hayes's (1995) parametric approach, the initial syllable or two syllables may or may not form a foot, depending on the length of the verb root that follows the prefix.

There are some forms that also exhibit unexpected non-adjacent syncretism. *Mu-* appears with second-person plural subjects in all but the prior, as well as being the second-person dual non-future, the first-person plural exclusive distant future and the plural imperative. *Ma-* is the first-person plural non-future and immediate future form, as well as the first-person exclusive dual non-future. *Lu-* is the third-person dual non-future and also the dual imperative (Table 23).

5.1.1.2 Variation in the subject-indexing prefixes

One notable feature of these paradigms is the frequency of overabundance, or complex cells where two or more forms can be used for the same set of grammatical features (Thornton 2012a; 2012b). These are sometimes called 'doublets', though Thornton prefers 'cellmates' to allow for situations where more than two forms appear. This is the case in the prior with a second-person singular subject. Arguably, this is also the case in the prior with paucal subjects in all persons, as the subject prefix can be realised with or without a geminate /t/. This overabundance differs from the Italian examples Thornton discusses, as it occurs in the affixes that apply to all verbs, rather than in certain lexical stems.

There are four main environments with overabundance. Firstly, in the second-person singular non-future and prior cells there is variation between *o-* and *a-*. This phonological alternation is also found in some lexical items, but there are also many examples of minimal pairs requiring a distinction between these phonemes (for example *ga* '3SG.NFUT.eat' and *go* '3SG.NFUT.pass'), so this is not simply underspecification at a phonological level. There is no clear phonological condition (for example, subsequent consonant or vowel) for the occurrence of these forms. The *a-* form seems to be an innovation. Only the *o-* form is mentioned in Parker's (1968b: 28; 1970a: v) descriptions in the late 1960s. Blevins and Lynch (2009: 118) reconstruct **ko-* for Proto-Paamese-Vatlongos.

The occurrence of these two variants was examined in the subcorpus of spontaneous texts. Speakers in Endu, the most conservative speech community, use the *o-* form for 96% of the 239²⁴ tokens of these prefixes in the Endu subcorpus. Speakers in Ase-Taveak use the *o-* form for only 76% of 133 tokens, and speakers in Mele Maat use *o-* for only 65% of 71 tokens. The use of *a-* may have originated in Mele Maat, or at least is being taken up more quickly by speakers there. The effect of community on the choice of these variants was highly significant (Pearson's chi-squared test $p < .005$, $N = 443$). However, the coding of these tokens is based on the perceptions of three transcribers, checked by the researcher. To strengthen this observation, it would be useful to conduct acoustic analysis of the examples.

Similarly, in the third-person singular non-future there is variation between a zero morpheme and *mi-*. There are identifiable conditions on these allomorphs, but they do not account for all the variation. A morphological condition is that the zero morpheme always occurs before the negative non-future prefix *taa-*. This is related to a lexical condition: the *mi-* prefix never occurs before the verb *tavuol*, *tavorel* 'be absent', which is probably derived from a lexicalised negated verb form.

The phonological conditions on the third-person singular zero morpheme are unidirectional. The zero morpheme only occurs before prenasalised consonants ($/^{mb}/$, $/^{nd}/$, $/^{ng}/$), bilabial consonants ($/p/$, $/m/$), and, for some speakers, $/k/$. The prenasalised consonants almost only occur with verb-initial consonant mutation for the non-future (§5.2.1). There is a clear structural justification for using the zero morpheme in these environments, as the stem mutation disambiguates the un-prefixed verb from the other use of bare stem in Vatlongos, for second-person singular imperative (which takes the basic initial consonant). However, the *mi-* variant also occurs in all these environments, though less often than the zero morpheme, e.g. (568).

²⁴ These prefixes are overrepresented in the Endu corpus as some of the speakers tend to use 'you know' as a rhetorical hedge (*o-kil-a* '2SG.NFUT-know-3OBJ').

Although the *mi-* form is possible in more phonological environments, many very frequent verbs undergo verb-initial consonant mutation and usually occur with the zero morpheme. The zero morpheme is therefore more frequent than the *mi-* form, accounting for 75% of 7271 tokens in the subcorpus. Blevins and Lynch (2009: 118, 123–124) reconstruct **mi-* to Proto-Paamese-Vatlongos, and discuss how it has been completely replaced by the zero morpheme in Northern Paamese. The use of a zero morpheme for third-person singular, and present tense or indicative (realis) mood, is also common cross-linguistically (Bybee 1985: 54–55). That it occurs with higher-frequency lexemes in Vatlongos suggests that economy, more than semantic primacy alone, favours the use of the zero morpheme in Vatlongos, two factors discussed by Koch (1995).

Phonological conditions on these variants are a salient marker of Endu-Vatlongos. Speakers from Endu use the *mi-* variant before the prenasalised and, especially, the bilabial consonants, more frequently than other speakers. All the tokens in the corpus where *mi-* is used before /m/ are from Endu speakers. This aligns with the general observation that Endu-Vatlongos is more conservative than South-Vatlongos; Blevins and Lynch (2009: 118) reconstruct **mi-* to Proto-Paamese-Vatlongos. On the other hand, speakers from Endu and Sameo (one of the villages closest to Endu) can use the zero morpheme before /k/. There are only 16 tokens of this in the corpus, but they are all in texts by speakers from Endu or Sameo, or speakers with connections to these communities.²⁵

This observed tendency to use the *mi-* form in more phonological environments is reflected in the overall frequency of these variants in the community-level subcorpora. In Ase-Taveak the zero morpheme accounts for 69% of 3868 tokens, whereas in Endu it only represents 58% of 1544 tokens. In Mele Maat the zero morpheme accounts for 84% of 1859 tokens. This is probably because of the higher frequency of loan verbs in the Mele Maat subcorpus, often introduced by the copular verb which undergoes verb-initial consonant and only ever occurs in the corpus as *be*, with the zero morpheme.

²⁵ For example, speakers in a family from Moru where the grandmother is from Endu.

Again, community has a highly significant effect on the proportion of these two variants in the subcorpus (Pearson's chi-squared test $p < .005$, $N = 7271$).

In the dual future, and paucal non-future and future, as well as all prior cells except the third-person singular, there is variation between a form of the prefix with a syllable shape of CVCV- and a form with CVC- shape (Table 22). The CVC- shape results from phonological reduction, dropping the final vowel from the CVCV- variant. This is an interesting case of overabundance because the CVC- variants have different patterns of syncretism than their CVCV- cell-mates, obscuring the distinction between the two futures in dual numbers, and the distinction between all TAM categories in paucal numbers. In some subject number environments, speakers therefore have a choice to identify a specific TAM category, or not.

This last pattern of overabundance between CVCV- and CVC- variants could be an interim stage of language change moving towards a system with only the phonologically reduced forms. There is some evidence in favour of this prediction. Firstly, the negative prefixes only occur with the CVC- variants of the subject prefixes in the corpus, although speakers usually do accept the CVCV- variants as grammatical in elicitation. The loss of the longer forms in negative polarity could be a first stage towards their loss in other environments. Other evidence from elicitation is that a younger speaker (aged 12) always offered the CVC- variant first, before being prompted to offer the CVCV- version, unlike their parents who usually did the opposite. However, these observations were not tested with other younger speakers and this could simply be an individual difference; this is an important area for further research.

Across all communities, CVC- forms account for more than half of tokens of the prefixes that show this pattern of overabundance. In Ase-Taveak the CVC- forms represent 57% of 632 tokens, and the rate is very similar in Endu at 54% of 312 tokens. Mele Maat uses CVC- forms for 66% of 201 tokens. This could show that Mele Maat speakers are again leading an innovation in Vatlongos, as with the *a*- form of the second-person singular prefix. The grammatical system of Vatlongos could be undergoing simplification in Mele Maat, a possible sign of language attrition (Schmidt 1985: 212–213). If the CVCV- forms are lost, this

could have consequences for the larger TAM system in Vatlongos, as the patterns of syncretism would be so extensive that some distinctions could be lost. For example, the immediate future would only be uniquely identified in the third-person singular, except with verbs that have the most complex patterns of verb-initial consonant mutation (§5.2.1). The effect of community on the variation between CVC- and CVCV- forms is significant, but less significant than on the use of the different forms of the third-person singular non-future and second-person singular non-future and immediate future prefixes (Pearson's chi-squared test $p=.02782$, $N=1145$).

On the other hand, these differences could be a stable pattern of overabundance in the island communities. The same variation existed in the 1960s according to Parker (1968b), although because of his more analytical affix template he lists variants of individual tense prefixes, rather than conditions on the syllable shape of portmanteau prefixes. As a result, the consistency in the pattern is somewhat obscured. This is further obscured by Crowley's (1991; 2002a) accounts which ignore the CVC- variants altogether. Based on data from Italian, Thornton (2012b) argues that overabundance can be very stable diachronically. Perhaps the Vatlongos system is sustainable, especially as the different patterns of syncretism mean that sometimes there is a functional difference between two cell-mates: the phonologically heavier variant is also more semantically specific and might be maintained by the functional need to point to a specific TAM category when it is not specified by other parts of the verbal morphology. It would be interesting to break down the variation between CVC- and CVCV- forms in different cells, to investigate whether the longer forms are more likely when they have a higher functional load. However, this avenue for future research would require a larger corpus, and perhaps some story prompts to encourage the use of some of the rarer cells in the paradigms.

In addition to the cell-mates shown in Table 22, some speakers also seem to have limited vowel harmony. This affects the third-person singular non-future prefix, so that *mi-* '3SG.NFUT' has an allomorph *mu-* when the first vowel of the

verb stem is also /u/ (this is also the case in example (395) above).²⁶ Although there are only seventeen tokens in the corpus, this is another type of variation that is associated with speakers from Endu, or with connections to Endu. This could be an influence of the languages of Northern Ambrym, as in both North Ambrym and Fanbyak third-person singular markers undergo vowel harmony or assimilation (Franjeh 2012: 66–68; 2016).

- (398) *hon buluk mu-kuvun na, Tom*
 horn cow 3SG.NFUT-throw HES Tom
 ‘The cow’s horns threw, um, Tom.’ [20141208b_n01s044_38-9]

Some speakers have vowel harmony in the third-person singular immediate future prefix *va-*, which is realised as *ve-* when the first vowel of the verb stem is /e/. However, this is only found with the verb form *he* ‘go to’ or ‘copula’:

- (399) *mi-sakras moletin ve-he hale*
 3SG.NFUT-unable person 3SG.IFUT-go_to outside
 ‘People couldn’t go outside.’ [20150419d_h01s046_40]

5.1.1.3 Defective paradigms for imperative and prohibitive moods

As well as the full subject-indexing paradigms for the major TAM categories, imperative and prohibitive moods have partial paradigms. As these are used to give orders to the addressee, they could be expected to be restricted to the second person, but in fact both also have a third-person singular form that occurs with psycho-collocations (§3.4.2). The use of the imperative is shown in (400), an elicited example.

- (400) *E-m bovorong-ni xitak!*
 inside-2SG.POSS 3SG.IMP.forget-TR this
 ‘Forget this!’ [20180801b_x01s046]

As discussed in §4.2.4, the distant future can also be used for commands.

Table 23: Paradigm for subject-indexing prefixes in imperative mood

	2	3
SG	∅-	∅-
DU	<i>lu-</i>	n/a
PC	<i>to-</i>	n/a
PL	<i>mu-</i>	n/a

²⁶ These variants are included as forms of *mi-* in the quantitative analysis of *mi-* and zero morpheme forms for the third-person singular non-future above.

The prohibitive subject-indexing prefixes co-occur with the negative clitic *ti* (§3.6.2).

Table 24: Paradigm for subject-indexing prefixes in prohibitive mood

	2	3
SG	<i>ona-</i> <i>ana-</i> <i>on-</i> <i>an-</i>	<i>na-</i>
DU	<i>mulna-</i>	<i>n/a</i>
PC	<i>mutna-</i>	<i>n/a</i>
PL	<i>muna-</i>	<i>n/a</i>

The second-person singular cell has four cell-mates: the /o/ versus /a/ variants reflect the overabundance in the second-person singular non-future forms, and the VCV-, VC- forms reflect the syllabic variants in other TAM categories. The second-person singular forms are the most frequent, with 29 examples in the corpus. The VCV- forms are more frequent (18 compared to 11), and the forms based on /o/ are much more common than forms based on /a/ (24 compared to 4). The higher frequency of this cell might explain why only the second-person singular maintains overabundance.

As with affirmative commands, the distant future in negative polarity can also be used with prohibitive meaning. This could be a calque from Bislama where the future negative is used for prohibitions. Some speakers report that the negative future forms are politer than the prohibitive. Alternatively, it could be for language-internal reasons, because of the similarity between the prohibitive and the apprehensive, which co-occurs with distant future prefixes.

The use of the third-person singular prohibitive form with a psycho-collocation is shown in example (401). The clitic here occurs after the verb, unlike the examples of negated psycho-collocations discussed in §3.6.2.7.

- (401) *e-m* *na-bong-ni* =*ti* *mun*
inside-2SG.POSS 3SG.PROH-forget-TR =NEG again
‘Don’t forget it again’ [20141106d_p05e016_05]

5.1.2 Negative and apprehensive prefixes

The negative prefixes are *taa-* in the non-future, and *naa-* for all other relative tenses (prior, immediate future and distant future). In the third-person singular

non-future, *taa-* appears with the zero-morpheme form of the subject-indexing prefix *mi-* ‘3SG.NFUT’. *Naa-* occurs without the third-person singular immediate future prefix *va-*, while in the distant future *naa-* can optionally cooccur with *i-* ‘3SG.DFUT’. The negative prefixes cannot co-occur with the imperative and prohibitive subject prefixes, or with apprehensive marking. In the corpus, the negative prefixes only occur with the CVC- variants of subject prefixes, although in elicitation some speakers do accept CVCV- variants of future and prior forms with a negative prefix. These possibilities for cooccurrence with the subject prefixes are exemplified for the third person in Table 25. Forms marked in **bold** were produced by speakers either in the corpus, or in elicitation as a translation from Bislama. The other forms were accepted as grammatical, but not volunteered by speakers themselves. The negative prefixes co-occur with the negative clitic *ti* (§3.6.2).

Table 25: Subject-indexing and negative prefixes in the third person

		3
SG	PRI	<i>te-naa-</i>
	NFUT	<i>taa-</i>
	IFUT	<i>naa-</i>
	DFUT	<i>naa-</i> <i>i-naa-</i>
DU	PRI	<i>lut-naa-</i> <i>lute-naa-</i>
	NFUT	<i>lu-taa-</i>
	IFUT	<i>lal-naa-</i> <i>lalo-naa-</i>
	DFUT	<i>lal-naa-</i> <i>lali-naa-</i>
PC	PRI	<i>lat(t)-naa-</i> <i>lat(t)e-naa-</i>
	NFUT	<i>lat-taa-</i>
	IFUT	<i>lat-naa-</i> <i>lato-naa-</i>
	DFUT	<i>lat-naa-</i> <i>lati-naa-</i>
PL	PRI	<i>lat-naa-</i> <i>late-naa-</i>
	NFUT	<i>la-taa-</i>
	IFUT	<i>la-naa-</i>
	DFUT	<i>li-naa-</i>

The apprehensive prefix has the form *na-* and cooccurs with the distant future paradigm of the subject prefixes, except in the third-person singular where it

appears independently. In the second-person singular, as well as the predicted *u-na-* ‘2SG.DFUT-APR-’ form, any of the four prohibitive forms can also be used (*ona-*, *ana-*, *on-*, *an-*). Table 26 shows the possible combinations of prefixes to express the apprehensive. Forms which overlap with the prohibitive paradigm of prefixes are in **bold**. In the dual and paucal numbers, there are CVCV- and CVC- variants of the distant future forms, and the combination of the shorter forms with apprehensive *na-* are identical to the prohibitive prefixes. The main difference between the apprehensive and the prohibitive is that the apprehensive cannot cooccur with the negative clitic *ti*. As such, the prohibitive paradigm could be analysed as a morphologically complex paradigm: a subset of the apprehensive paradigm plus the negative clitic *ti*.

Table 26: Prefixes marking verbs as apprehensive

	1INCL	1EXCL	2	3
SG		<i>ni-na-</i>	<i>ona-</i> <i>ana-</i> <i>on-</i> <i>an-</i> <i>u-na-</i>	<i>na-</i>
DU	<i>ral-na-</i> <i>rali-na-</i>	<i>mal-na-</i> <i>mali-na-</i>	<i>mul-na-</i> <i>muli-na-</i>	<i>lal-na-</i> <i>lali-na-</i>
PC	<i>rat-na-</i> <i>rati-na-</i>	<i>mat-na-</i> <i>mati-na-</i>	<i>mut-na-</i> <i>muti-na-</i>	<i>lat-na-</i> <i>lati-na-</i>
PL	<i>ri-na-</i>	<i>mu-na-</i>	<i>mu-na-</i>	<i>li-na-</i>

5.1.3 Object and transitivity suffixes

Morphological processes are often important to the categorisation of transitivity classes of verbs in Oceanic languages (Hill 2011: 458–459). Verb lexemes in Vatlongos can be sorted into two morphologically-defined transitivity classes based on their cooccurrence with the object and transitivity suffixes, as well as syntactic restrictions on the expression of objects. Transitive verbs can take one of the paradigms of object-indexing suffixes (Table 27), to mark a pronominal object in one of the person-number-animacy combinations that can be indexed by these defective paradigms. When they are not marked with an object suffix, transitive verbs must be followed by an object NP in the post-verbal object position (§3.1). Conversely, intransitive verbs must be marked with the transitivity suffix *-ni* to take an object NP in post-verbal position object, or to refer to a pronominal third-person singular object.

The transitive verb class is more restricted. There are 4304 tokens of transitive verbs in the subcorpus, consisting of 181 different lexemes (of 232 types in total²⁷), just over a quarter of these tokens are marked with an object suffix (1229 tokens). There are 7544 tokens of 307 intransitive verb lexemes in the corpus (of 398 types in total). Morphologically intransitive verbs are marked with the transitivising suffix (and therefore take an object argument) in around a tenth of occurrences (709 tokens).

The object-indexing suffixes only occur with invariantly transitive verbs, and their shape is lexically determined (Table 27). They only occur when there is an object which is not coded with an NP in post-verbal object position (§3.1), and the referent can be understood through context. Although in many Oceanic languages object-marking can co-index an object NP (Ross 2004), there are other Vanuatu languages that have this kind of pronominal agreement (Bickel & Nichols 2007: 232) or pro-indexing (Haspelmath 2013) system, such as North-East Ambae (Hyslop 2001: 260). Not all pronominal objects are coded with an object suffix, they can also be coded with independent pronouns in an object NP. This is especially likely with animate referents, and reflexive or reciprocal structures (§3.4.1).

The object-indexing suffixes have defective paradigms. There are forms for first-person singular, second-person singular and third-person objects. First- and second-person singular referents can also be coded with independent pronouns. The first-person singular independent pronoun *inou* has an alternative form *nou* which is identical to the first-person object suffix, so it is not always possible to distinguish an independent pronoun from an object suffix in this environment. Example (172) shows that *nou* does occur as an independent form, as the negative clitic *ti* cannot intervene between a verb and an object suffix.

The third-person forms can be used with all singular referents, although there is a preference for using an independent pronoun for animate referents.

²⁷ Including the rest of the corpus, entries in Parker's (1970a) dictionary, and other lexemes added to the database during fieldwork.

Inanimate non-singular referents can be expressed with an object suffix. This animacy-based split in agreement patterns mirrors the split in subject-agreement (§5.1.1). As with the subject-indexing of inanimate referents, the third-person object suffix can be used even when an antecedent NP is explicitly marked as plural:

(402) *la-sa* *pepa na-mem* *xil, ma-leh-i*
 3PL.NFUT-give paper CL.GEN-1PL.EXCL.POSS PL 1PL.EXCL.NFUT-take-3OBJ
 ‘They gave us our papers, we took them’
 [20170413e_n01m030_14-15]

First- and second-person non-singular objects can only be expressed with an independent pronoun, see (21). As speakers and addressees are always animate (in so far as they are capable of speaking or being addressed), we can generalise that animate non-singular arguments are expressed with an independent pronoun.

There are five forms of the third-person object suffix, and the second-person singular form is predictable from the third-person form (Table 27). Which series of object suffixes a verb takes is lexically determined, although there are phonological patterns with exceptions. Some verbs such as *leh* ‘take’ are used with different object suffixes depending on speaker background. For example, some speakers from Endu use *leh-e*, while elsewhere speakers use *leh-i*.

Table 27 shows the different forms of the object suffixes and the phonological tendencies in lexical class assignment. The object suffixes appear to be a reflex of proto-Oceanic transitivity **-i* (Pawley 1973: 120–126; Evans 2003: 93–118) which has been reanalysed as a third-person singular object suffix in many modern Oceanic languages (Evans 2003: 100), and also has a zero morpheme alternant with many stems ending in /i/ (Evans 2003: 106; Ross 2004: 508). It also shows how the object suffixes interact with the nominaliser *-en*, which, along with historical evidence, suggests that the lexically specified variants of the *-i* suffix were originally final vowels of the verb stem (Parker 1970a: xi; Crowley 2002a: 662).

Table 27: Object-indexing suffixes

3	1SG	2SG	Nominalizer	Phonological patterns
<i>-i</i>	<i>-nou</i>	<i>-uk</i>	<i>-ien</i>	elsewhere
<i>-e</i>	<i>-nou</i>	<i>-ok</i>	<i>-en</i>	final back vowels or back vowels with /l/ or /r/
<i>-a</i>	<i>-nou</i>	<i>-ak</i>	<i>-aen</i>	final front vowels with /l/ or /v/
<i>-ni</i>	<i>-nou</i>	<i>-nuk</i>	<i>-nien</i>	CV syllable structures ²⁸
<i>-∅</i>	<i>-nou</i>	<i>-uk</i>	<i>-en</i>	final /i/ or /e/

The *-a* and *-ni* series are less frequent, occurring with only eight and seven verb stems respectively, while the *-i* form is by far the most frequent (>100 forms). The *-ni* form also occurs with four prepositions with CV syllable structure (§3.2).

The few verbs which take a zero morpheme for the third-person object can be identified as transitive when they take a second-person singular object, as in (404), or in sentences with a nominal object like (405), as an intransitive verb would require the transitiviser *-ni*.

(403) *tei!*
2SG.IMP.3OBJ.cut
'cut it!' [20150305d_n01s109_09]

(404) *va-tei-uk*
3SG.IFUT-cut-2SG.OBJ
'he'll cut you' [20170217e_n01s142_12]

(405) *ha tei au!*
GO 2SG.IMP.cut rope
'Go cut the rope!' [20141027a_n01m001_58]

Intransitive verbs can take an object when they are marked with the transitivising suffix *-ni*, apparently derived from Proto-Oceanic **akin[i]*, for which very similar functions have been reconstructed (Evans 2003: 119–239).²⁹

²⁸ The seven verbs taking this paradigm of object suffixes also include *seh* 'fill' (e.g. a basket) and *puei* 'climb for' (e.g. fruit). *a* 'eat' has no consonant onset in some TAM environments, but most tokens will have an initial consonant (§5.2.1).

²⁹ Crowley (1983) argues that the Proto-Oceanic transitive suffix **akin[i]* (Evans 2003: 120) has been lost in Paamese, and the transitivising suffix has been reconstituted from preposition *eni*.

Transitiviser *-ni* can be followed by an object NP as in (406) or mark the presence of a pronominal object as in (407). It is analysed as occupying the same slot as the object suffixes rather than preceding them and taking the zero morpheme series of third person object suffixes, because it does not occur with the first or second person singular suffixes.

(406) *ma-pol-ni* *pol-ien* *ak*
 1PL.EXCL.NFUT-do-TR do-NMLZR PROX
 'We do this work.' [20141116a_p02s026_01]

(407) *Lata-pol-ni*
 3PC.NFUT-do-TR
 'They did it.' [20141220g_n01s080_19]

In the Oceanic literature a distinction is often made between stative and dynamic verbs; this distinction correlates with whether the actor or undergoer is the subject of an intransitive verb (see Evans 2003: 23–33). In Vatlongos, one difference between stative and dynamic verbs is how the argument structure of the intransitive verb corresponds to its transitivised equivalent, and the type of event denoted by the transitivised verb. Stative verbs are unaccusative while dynamic verbs are unergative (Foley 2005).

With stative verbs *-ni* has a causative function. The single argument of the intransitive verb maps to the object of the transitivised verb, and the subject maps to an agent or force that causes a change in state. The whole event is equivalent to what Van Valin (2010: 713) describes as a 'causative state'. For example, the stative verb *sale* 'flat, wide' means 'cause to be flat, flatten' when transitivised with *-ni*:

(408) *la-sale-ni*
 3PL.NFUT-flat-TR
 'They flatten it.' [20141028b_p01m003_12]

With dynamic verbs, the transitiviser *-ni* has an applicative function. Dynamic verbs generally denote an activity, and the transitivised equivalent can denote an active accomplishment (Van Valin 2010: 713), where an activity results in a change of state (or creation of something), or a directed activity. The subject of the intransitive verb is also the subject of the transitivised equivalent, the agent performing the activity. The object of dynamic transitivised verbs can express a

- (411) *la-be* *moletin* *xa*
 3PL.NFUT-NFUT.COP person REL
taem *xa* *la-sepin-ni* *venu*,
 time SUB 3PL.NFUT-speak-TR volcano
venu *mi-tteong-ni* *xil*
 volcano 3SG.NFUT-listen-TR 3PL
 ‘They were people who, when they spoke to the volcano, the volcano
 listened to them.’ [20141105f_p01e013_142]
- (412) “*Karagi gel*” *biten-ni* *atuli na-n*.
 crazy.BIS girl.BIS 3SG.NFUT.speak-TR girl CL.GEN-3SG.POSS
 ‘‘Crazy girl!’’ she said to her daughter.’ [20141106f_n01e018]

The transitiviser *-ni* also attaches to intransitive verbs or adverbs that appear as the second element in a complex predicate (§7.1), when the complex predicate as a whole has a transitive argument structure. The second element is often a stative verb modifying the situation predicated by an initial transitive verb (e.g. (566)), so the object is the argument required by the initial transitive verb.

Transitive Bislama loan verbs are usually marked with the transitive suffix when they are incorporated with the copular verb, *he* (§3.5.2). There are also a few examples in the corpus where an adjective used predicatively with the copular verb is marked with the transitive suffix, as in (413). It is not clear whether this is evidence that Bislama verbs are borrowed as adjectives, or whether the frequent use of the copular verb with loan verbs has led to the possibility of using the suffix with native adjectives used predicatively as well.

- (413) *na-be* *imae-ni* *skul* *na-van*
 1SG.NFUT-NFUT.COP smart-TR school CL.GEN-1SG.POSS
 ‘I was smart at my studies.’ [20150219b_n01m001_066]

5.2 Non-linear morphology

Several processes in Vatlongos verbal morphology cannot be modelled as simple concatenation of morphemes as strings of segments. These involve changing individual segments in verb initial-consonant mutation (§5.2.1), copying of segments in reduplication (§5.2.2), deletion of segments in verb-stem alternation (§5.2.3) and other unpredictable changes with an irregular verb stem (§5.2.4).

in negative environments for some verbs (Table 30). Crowley (1991: 212) observes that ‘some verbs in Paamese differ in class membership between dialects’ and that ‘some verbs even within the same dialect vary according to the patterns of more than one class’.

Once these different variants are included, it is necessary to distinguish two more inflectional environments than previously described (Parker 1968b: 33–34; 1970a: vi–vii; Crowley 1991: 185). The third-person singular immediate future and the complex predicate environments are uniquely distinguished by different possibilities when these variants are taken into account. Another difference is that the basic consonant is used with dual and paucal subjects in the immediate future; these are not included in the immediate future environment. The environments are summarised in Table 29.

Table 29: Environments for verb-initial consonant mutation

Reference name	Gloss	Environments
Basic	none	Nominalised form Adjectivized form Affirmative prior Affirmative distant future Affirmative immediate future with dual and paucal subjects Imperative, prohibitive and apprehensive moods
Negative	NEG	Negative polarity
Non-future	NFUT	Affirmative non-future
Immediate future	IFUT	Affirmative immediate future with second-person singular and all plural subjects
First-person singular	1SG.IFUT	Immediate future with first-person singular subject
Third-person singular	3SG.IFUT	Immediate future with third-person singular subject
Complex predicate	v2	Subsequent verb in a complex predicate (§7.1)

Parker and Crowley identify eight classes of irregular verbs by distribution of initial verb mutation. When possible combinations of variants are included, there are 21 different patterns from elicitation, and four more can be posited on the basis of tokens in the corpus. It is therefore perhaps no longer useful to posit classes of verbs identified by this morphological process. Instead, it seems that the individual phonological processes (e.g. /h/ to /g/) and the environments they relate to (e.g. affirmative non-future) are listed for

individual verb lexemes, without reference to lexical classes. Using Round’s (2015) typology of morphemes, we can posit that Vatlongos verb-initial consonant mutation involves meromorphemes (the initial consonant alternations) and metamorphemes (the environments in Table 29), without reference to rhizomorphemes (like Parker’s verb classes).

Rather than trying to list all possible patterns and the verbs that participate in them, Table 30 therefore gives typical examples of common patterns, demonstrating that patterns can differ by possible cell-mates in individual cells. Cell-mates are separated by commas, and capital V indicates an initial vowel, that is, the absence of an initial consonant. The /r/-initial forms of the t initial verbs, and the existence of the k/g alternation have not been previously described, but have parallels in Paamese (Crowley 1991: 190). Some of these patterns are associated with regional dialects. The k/g alternation is only used by speakers connected to Endu, and the two nearest villages Ase and Sameo (Figure 4). In South-Vatlongos the four verb lexemes involved are either regular with invariant initial /k/, or a synonym with initial /h/ is used in one of the patterns including h/g (e.g. *kil* ‘dig’, *hil* ‘dig’). The second t/r/d pattern, which allows the use of initial /r/ in the non-future is a marker of Vatlongos spoken in Toak, and now mainly used by elderly speakers. This is very distinctive as it is the only pattern where a non-prenasalised initial consonant is used for the non-future.

Table 30: Some patterns of verb initial consonant mutation

Onset patterns	Basic	v2	NEG	3SG.IFUT	IFUT	1SG.IFUT	NFUT	Example
h/v/m/b1	h	h	v	h	v	m	b	<i>ha</i> ‘go’
h/v/m/b2	h	h	v	h	v	m, v	b	<i>he</i> ‘go to’
h/v/m/b7	h	h	h, v	h, v	h, v	m, h, v	b	<i>hipas</i> ‘break’
h/g	h	h	h	h	h	h	g	<i>huppas</i> ‘break’
h/v/m/g	h	h	h, v	h, v	h, v	m, h	g	<i>holen</i> ‘smash’
h/v/m/b/g	h	h	h, v	h, v	h, v	m, h, v	b, g	<i>hol</i> ‘dance’
h/V/v/m/g	h	V	h, v	h, v	h, v	m, h	g	<i>husil</i> ‘follow’
v/b	v	v	v	v	v	v	b	<i>vil</i> ‘run’

Onset patterns	Basic	v2	NEG	3SG.IFUT	IFUT	1SG.IFUT	NFUT	Example
v/m/b	v	v	v	v	v	m, v	b	<i>vagos</i> 'breathe'
t/r/d1	t	t, r	t, r	t	t	t	d	<i>ti</i> 'stay'
t/r/d2 (Toak)	t	t, r	t, r	t	t	t	r, d	<i>ti</i> 'stay'
x/g	x	x	x	x	x	x	g	<i>xul</i> 'swim'
x/m/g	x	x	x	x	x	m, x	g	<i>xa</i> 'say'
V/x/v/m/g	V	V	x	v, V, x	v, V	m, V, x	g	<i>a</i> 'eat'
k/g (Endu, Ase, Sameo)	k	k	k	k	k	k	g	<i>kar</i> 'bite'

There are trends within this proliferation of patterns. The consistent use of the prenasalised phonemes (which are relatively rare elsewhere in the language) in the non-future affirmative environment supports Lynch's (1975) argument that the oral/nasal alternation found in many Oceanic languages arose from a preposed particle consisting of a nasal consonant and a vowel (Crowley 1991: 181–183). For all patterns that involve only two initial consonants, this is the environment that is uniquely identified. Similarly, Crowley (1991: 183) posits that Proto-Paamese-Vatlongos had an immediate future prefix *ma-* in the first-person singular and *va-* in all other subject person-number combinations. This explains the restriction of /m/ initial forms to the first-person singular immediate future, and the presence of dedicated immediate future forms in both languages, while the *va-* prefix has been reanalysed as part of either the root or the subject prefixes in different cells of the paradigm.

An important trend for evidence of processes of language change is the tendency for more frequent verbs to have fewer cell-mates, and more specific forms, that is, forms which are restricted to fewer inflectional environments. This is most clear in the /h/ initial verbs. The most frequent verb lexeme in the subcorpus is *ha* 'go' (1254 tokens), and *ha* only allows a single variant in each environment, and does not take the basic consonant in environments where a more specific variant is available. The second and fourth most frequent verbs are the homophonic copular verb *he* (1181 tokens), and the incorporated-preposition verb *he* 'go to' (871 tokens), and both allow variation in only one environment: the first-person singular immediate future can be marked with uniquely-identifying /m/, or /v/ which also occurs in other immediate future

cells and negative polarity, but is not as underspecified as the basic consonant /h/. Less frequent verb lexemes like *hipas*, which does not occur in the subcorpus, allow the use of the basic consonant in all environments except the affirmative non-future. While the patterns do not exactly correlate with token frequency in the corpus for medium-frequent verbs, these examples from either extreme demonstrate the rationale behind this pattern.

The proliferation of cell-mates which makes for a complex system synchronically, could in fact indicate ongoing simplification of this morphological process. The extension of the basic consonant to more environments fits Silva-Corvalán's (1991: 152) model of morphological simplification. Complexity may only be maintained for highly frequent verb lexemes, and if more specific variants of less frequent verbs are lost, eventually inflectional environments could also be lost. The inter-speaker variation in grammaticality judgements for variants in specific environments for some lexemes shows how this system is prone to reanalysis and variation.

This tendency also makes it very difficult to identify inter-speaker or community-level variation in the corpus, as verbs that are frequent enough to occur across the possible inflectional environments are less likely to exhibit any formal variation. Within the time constraints of this project, it was only possible to elicit the pattern for each verb with a single speaker, and the three speakers who took part in this part of the project were all from Ase-Taveak and a single family (Simeon Ben, Madleen Simeon and their son Yaxley). Eliciting these patterns with speakers of different communities and age groups would be an interesting future project. Verb lexemes and inflectional environments where variation is expected on the basis of the description here could be targeted by devising suitable pragmatic contexts, rather than relying solely on traditional elicitation methods.

5.2.2 Reduplication

Reduplication as a morphological process in Vatlongos mainly applies to verbs, though there are also a few reduplicated nouns (e.g. *pang~pang* 'fire') and adjectives (e.g. *tu~tut* 'small') in the corpus, which are probably lexicalised. With stative verbs, reduplication intensifies the gradable component of the

verb’s lexical meaning. This is the same function as the verbal repetition strategy (§4.3.3), which can be a source for partial reduplication cross-linguistically (Bybee, Perkins & Pagliuca 1994: 166). With non-stative verbs, reduplication adds a meaning of random, distributed action, or lack of a specific patient, as in Paamese (Crowley 1982: 152–155). Although not fully productive, reduplication is quite variable, with three different syllabic templates for the reduplicant, which always precedes the rest of the verb root.

The reduplicant can have a CV syllabic structure. Examples of verbs that appear with this form of reduplication are given in Table 31. As many of these reduplicated forms occur only once in the corpus, the meaning of the reduplicated stem is not always transparent. Instead, the context it appears in is given, to identify commonalities in the distribution of different reduplicated forms. Reduplicated forms which only or predominantly occur as the second verb in a complex predicate (§7.1) are marked ‘V2’.

Table 31: Verbs with CV reduplication

Basic verb stem	Gloss	Reduplicated verb stem	Meaning/distribution
<i>ha</i>	‘go’	<i>ha~ha</i>	go off individually in random directions
<i>he</i>	‘go to’	<i>he~he</i>	go to (dual subject) ³¹
<i>sa</i>	‘give’	<i>sa~sa</i>	give things out at random
<i>kea</i>	‘crawl’	<i>ke~kea</i>	crawl (V2)
<i>lau</i>	‘hunt’	<i>la~lau</i>	hunt around, randomly
<i>meao</i>	‘open’ (of a place)	<i>me~meao</i>	very open
<i>kil</i>	‘know, can’	<i>ki~kil</i>	understand (V2)
<i>lel</i>	‘die’	<i>le~lel</i>	lots of people die at different times
<i>lep</i>	‘fat’	<i>le~lep</i>	very fat
<i>mel</i>	‘happy’	<i>me~mel</i>	very happy
<i>pas (V2 only)</i>	‘break’	<i>pa~pas</i>	break at random (V2)
<i>pol</i>	‘full’	<i>po~pol</i>	very full
<i>rat (usually V2)</i>	‘out, take out’	<i>ra~rat</i>	take out many objects at random (V2)
<i>vul</i>	‘break’	<i>vu~vul</i>	randomly break lots of things

³¹ There is only one elicited example of this form in the corpus, from an Endu speaker.

Basic verb stem	Gloss	Reduplicated verb stem	Meaning/distribution
<i>kisi</i>	'soft'	<i>ki~kisi</i>	very soft
<i>bemei</i>	'come'	<i>be~bemei</i>	come (dual subject) ³²
<i>koxol</i>	'block'	<i>ko~koxol</i>	block lots of things at random (V2)
<i>sohuk</i>	'fish' ³³	<i>so~sohuk</i>	fishing (nominalised)
<i>vorong</i>	'forget'	<i>vo~vorong</i>	randomly or casually forget something

There is a coexisting pattern of reduplication where the syllable shape of the reduplicant is CVC. This applies to a different set of lexemes, and participation in a particular pattern of reduplication appears to be lexically determined with some phonological conditioning: CV stems can only undergo CV reduplication. One lexeme, *pol* 'full' occurs with both these patterns. The one token of the CVC form is from a speaker in Endu, but there are not enough tokens to establish whether this is an example of community-level variation.

Table 32: Verbs with CVC reduplication

Basic verb stem	Gloss	Reduplicated verb stem	Meaning/distribution
<i>pol</i>	'full'	<i>pol~pol</i>	very full
<i>pile</i>	'play' (loan verb)	<i>pil~pile</i>	play at random (non-singular subjects)
<i>hiles</i>	'turn'	<i>hil~hiles</i>	random turning and squeezing into mush
<i>husil</i>	'follow'	<i>hus~husil</i>	follow along (of two people, not knowing right direction) ³⁴
<i>kiles</i>	'roll'	<i>kil~kiles</i>	roll lots of things
<i>kilis</i>	'turn'	<i>kil~kilis</i>	turn over at random
<i>pongos</i>	'smell'	<i>pong~pongos</i>	search for by sniffing (e.g. a dog)

Finally, there is one lexeme that takes a CVCV reduplicant.

Table 33: Verbs with CVCV reduplication

Basic verb stem	Gloss	Reduplicated verb stem	Meaning/distribution
<i>sinen</i>	'dry out'	<i>sine~sinen</i>	dry out lots of things

³² There is only one example of this form in the corpus, from an Endu speaker.

³³ a compound of *so* 'put' and loan noun *huk* 'hook'

³⁴ There is only one example of this form in the corpus, from an Endu speaker.

(Crowley 1995b). On the other hand, as Bislama also has a CVC reduplication pattern (Crowley 2004b: 73), either an earlier reduplicated form or the pattern itself could be borrowed from Bislama.

The verbs undergoing reduplication are only a small proportion of all verbs in Vatlongos, and reduplicated forms seem to be prone to lexicalisation. For example, the verb *pis* ‘try, roll’. Reduplicated forms are also more common in restricted grammatical environments, especially nominalised verbs (§5.3.1) and the subsequent verb in a complex predicate (§7.1).

5.2.3 Verb stem alternations

Some verb stems undergo vowel elision when they take a prefix. All but one of these verbs begin with the syllable structure CVCV with identical consonants and vowels. When they are prefixed, the first vowel is elided so the verb stem has a geminate consonant onset CCV. As suggested by Parker (1968a: 34–35), they probably derive from reduplicated forms following the CV- pattern of reduplication (Table 31), though the unreduplicated form is not extant in nearly all cases. Most of the verbs in the tables below are compatible with the meanings of reduplication: either they are gradable stative verbs, or non-stative verbs denoting random action, repeated actions, or multiple participants. For some of these verbs, an unreduplicated form is still extant with a similar or related meaning: *mumun* ‘drunk’ and *mun* ‘drink lots’; *papan* ‘whistle’ and *pan* ‘Golden Whistler bird’; *pipili* ‘red’ and *pili* ‘red’; and *teteli* ‘hold, marry’ and *teli* ‘pull’. However, these forms could be back-derived through reanalysis as a simple consonant, as seems to be happening with the verbs in Table 36.

Table 35 shows the form of the lexeme in contexts without prefixes (the second-person singular imperative, the nominalised form, and the third-person singular non-future form for forms with a bilabial onset), and with a prefix (the third-person plural non-future). None of these verbs undergo verb-initial consonant mutation (§5.2.1).

Table 35: Verb stems undergoing vowel elision with prefixes

2SG.IMP	Gloss	Noun	3SG.NFUT	3PL.NFUT
<i>kakal</i>	‘scratch’	<i>kakal-en</i>	<i>mi-kkal</i>	<i>la-kkal</i>

2SG.IMP	Gloss	Noun	3SG.NFUT	3PL.NFUT
<i>kekereli</i>	'spin'	<i>kekereli-en</i>	<i>mi-kkereli</i>	<i>la-kkereli</i>
<i>kokoei</i>	'close'	<i>kokoei-en</i>	<i>mi-kkoei</i>	<i>la-kkoei</i>
<i>kokora</i>	'crow'	<i>kokora-en</i>	<i>mi-kkora</i>	<i>la-kkora</i>
<i>lilih</i>	'fan'	<i>lilih-ien</i>	<i>mi-llih</i>	<i>la-llih</i>
<i>mamal</i>	'straight'	<i>mamal-en</i>	<i>mamal</i>	<i>la-mmam</i>
<i>memeas</i>	'blue-green'	<i>memeas-en</i>	<i>memeas</i>	<i>la-mmeas</i> <i>la-memeas</i>
<i>memees</i>	'urinate on'	<i>memees-en</i>	<i>memees</i>	<i>la-mmees</i>
<i>memerou</i>	'soft'	<i>memerou-en</i>	<i>memerou</i>	<i>la-mmerou</i>
<i>momoli</i>	'smooth'	<i>momoli-en</i>	<i>momoli</i>	<i>la-mmoli</i>
<i>momoti</i>	'smooth, straight'	<i>momoti-en</i>	<i>momoti</i>	<i>la-mmoti</i>
<i>mumun</i>	'drunk'	<i>mumun-en</i>	<i>mumun</i>	<i>la-mmun</i> <i>la-mumun</i>
<i>mumuu</i>	'dirty'	<i>mumuu-en</i>	<i>mumuu</i>	<i>la-mmuu</i>
<i>nenem</i>	'think'	<i>nenem-ien</i>	<i>mi-nnem</i>	<i>la-nnem</i>
<i>papan</i>	'whistle'	<i>papan-en</i>	<i>papan</i>	<i>la-ppan</i>
<i>pipili</i>	'red'	<i>pipili-en</i>	<i>pipili</i>	<i>la-ppili</i>
<i>pipin</i>	'enough'	⁻³⁵	<i>pipin</i>	<i>la-ppin</i>
<i>pupuluu</i>	'count'	<i>pupuluu-en</i>	<i>pupuluu</i>	<i>la-ppuluu</i>
<i>pupus</i>	'squeeze'	<i>pupus-en</i>	<i>pupus</i>	<i>la-ppus</i>
<i>rorom</i>	'swallow'	<i>rorom-en</i>	<i>mi-rrom</i>	<i>la-rrom</i>
<i>tatang</i>	'insert hand'	<i>tatang-en</i>	<i>mi-ttang</i>	<i>la-ttang</i>
<i>tatanuk</i>	'cloudy'	<i>tatanuk-en</i>	<i>mi-ttanuk</i>	<i>la-ttanuk</i>
<i>teteli</i>	'hold, marry'	<i>teteli-en</i>	<i>mi-tteli</i>	<i>la-tteli</i>
<i>tetenai</i>	'spread leaves'	<i>tetenai-en</i>	<i>mi-ttenai</i>	<i>la-ttenai</i>
<i>teteong</i>	'listen'	<i>teteong-en</i>	<i>mi-tteong</i>	<i>la-tteong</i>
<i>tetep</i>	'swell'	<i>tetep-en</i>	<i>mi-ttep</i>	<i>la-ttep</i>
<i>titi</i>	'share, distribute'	<i>titi-en</i>	<i>mi-tti</i>	<i>la-tti</i>
<i>totoh</i>	'defecate'	<i>totoh-en</i>	<i>mi-ttoh</i>	<i>la-ttoh</i>
<i>totohot</i>	'defecate on'	<i>totohot-en</i>	<i>mi-ttohot</i>	<i>la-ttohot</i>

For some of these verbs the initial CV is optionally lost from the imperative or nominalised form. Sometimes the geminate onset with the prefixed forms is not consistently realised in the corpus ('?' in Table 36), although in elicitation consultants prefer the geminate realisation. Explicit elicitation was only conducted with one consultant per verb stem, and it is likely that, for at least some speakers, these verbs have lost the CV onset to the verb stem altogether,

³⁵ The consultant did not feel that any of the options made sense here: the verb could not be nominalised (§5.3.1).

and are treated as invariant verb stems. It is notable that only two bilabial onset verbs and no /t/ onset verbs behave in this way. The vowel elision pattern is more consistently maintained for bilabial onsets, where the full CVCV form is more likely to occur because it is the preferred third-person singular non-future form, and for /t/ onsets where the t/d or t/d/r pattern of verb-initial consonant mutation is very common, making these forms more marked.

Table 36: Verb stems undergoing vowel elision or losing initial CV syllable

2SG.IMP	Gloss	Noun	3SG.NFUT	3PL.NFUT
<i>kakan</i>	'steal'	<i>kakan-en</i> <i>kan-en</i>	<i>mi-kkan</i>	<i>la-kkan</i>
<i>kekea</i> <i>kea</i>	'crawl'	<i>kekea-en</i>	<i>mi-kkea</i> <i>? mi-kea</i>	<i>la-kkea</i> <i>? la-kea</i>
<i>kekesae</i>	'lost'	<i>kekesae-en</i> <i>kesae-en</i>	<i>mi-kkesae</i> <i>? mi-kesae</i>	<i>la-kkesae</i> <i>? la-kesae</i>
<i>kokot</i> <i>kot</i>	'remove top'	<i>kokot-en</i> <i>kot-en</i>	<i>mi-kkot</i>	<i>la-kkot</i>
<i>memes</i>	'thank'	<i>memes-en</i> <i>mes~mes-en</i>	<i>memes</i>	<i>la-mmes</i>
<i>ngengel</i>	'rest'	<i>ngengel-en</i> <i>ngel-en</i>	<i>mi-ngngel</i> <i>? mi-ngel</i>	<i>la-ngngel</i> <i>? la-ngel</i>
<i>pepen</i>	'shoot'	<i>pepen-en</i> <i>pen-en</i>	<i>pepen</i>	<i>la-ppen</i>
<i>raran</i> <i>ran</i>	'wild'	<i>raran-en</i> <i>ran-en</i>	<i>mi-rran</i>	<i>la-rran</i>
<i>rerei</i> <i>rei</i>	'knock'	<i>rerei-en</i> <i>rei-en</i>	<i>mi-rrei</i>	<i>la-rrei</i>
<i>sisil</i> <i>sil</i>	'send'	<i>sisil-en</i>	<i>mi-ssil</i>	<i>la-ssil</i>

There is only one verb which undergoes vowel elision with different consonants and vowels: *sital* 'come out', which elides to *stal* in prefixed forms like *mi-stal* '3SG.NFUT.come_out'. According to an older consultant, the verb stem *stal* cannot occur with the CVC- form of the CVCV- subject prefixes. However, the twelve-year-old speaker did accept these forms in elicitation, possibly suggesting a change in phonology to tolerate these consonant clusters. In elicitation, these CVC- prefixes are tolerated with the geminate consonant onset verb stems, but the geminate consonant seems to be reduced in length when the form is repeated back. No examples of the geminate onset stems or *stal* occurring with a CVC- prefix are found in the corpus.

5.2.4 Irregular verb *ammei* ‘come’

There is only one uniquely irregular verb in Vatlongos: *ammei* ‘come’. The basic stem for prefixed forms is *mmei*, as in *la-mmei* ‘3PL.NFUT-come’. This stem occurs 551 times in the subcorpus and never occurs with a CVC- form of the CVCV- subject prefixes. It has an irregular third-person singular non-future form *bemei*, which resembles the use of prenasalised /b/ initial forms in non-future affirmative environments, but undergoes CV reduplication as an invariant verb stem (Table 31). The imperative form is *ammei*, which resembles a prefixed second-person singular non-future or immediate future form, with the more innovative *a-* variant of the prefix (§5.1.1.2). To nominalise this verb, the geminate consonant is dropped resulting in *mei-en*. This *mei* form is also used as the derived auxiliary verb *mei* ‘come, inchoative’ (Chapter 8).

There are some irregularities in the ways different speakers use this verb. For some speakers, the form *mei* can also be used as an alternative third-person singular non-future form (13 tokens in the corpus). In most instances this is after the full form *bemei* or another inflected form of *ammei* has been used, but example (420) is the most clear-cut independent use of *mei* for this purpose. Interestingly, it is used with two loan verbs from Bislama; maybe this speaker is influenced by Bislama in reanalysing the auxiliary form as the full verb as there is no formal difference between auxiliary *kam* ‘inchoative’ and lexical *kam* ‘come’ in Bislama (Crowley 2004b: 101).

(420) *lastik* *mei* *aot*
 elastic 3SG.NFUT.come out
 ‘The elastic came off.’ [20150310a_h01s114_09]

One speaker from Endu also uses the *bemei* form for the third-person dual non-future stem form, and it is this same speaker who uses the reduplicated form in the same inflectional environment (Table 31).

(421) *xalu* *ak* *lu-bemei* *sung*
 3DU PROX 3DU.NFUT-NFUT.come first
 ‘These two came first.’ [20141107b_n01e020_19]

5.3 Derivational morphology

There are two main derivational affixes that can be applied to verbs in Vatlongos. The nominaliser *-en* is highly productive and can attach to almost

any verb, supporting Anderson's (1992: 78) observation that derivational morphology need not be any less productive than inflectional morphology. The adjectiviser *te-*, *ta-* can be applied to stative verbs to derive an attributive adjective, but seems to occur most often with a small set of verb lexemes.

5.3.1 Nominaliser *-en*

The nominalising suffix *-en* attaches to a verb stem to denote the state, activity or process predicated by a verb. When it attaches to a transitive verb, it has allomorphs that begin with the lexically specified allomorph of the third-person object suffix (Table 27). The *-ien* allomorph can also be used with intransitive roots ending in a consonant, especially /l/ and /n/, although this seems to be lexically specified by the verb root.

The suffix is very productive and even attaches to some loan verbs such as *skul-ien* 'schooling, education'. However, the nominalised form is prone to further lexicalisation to take on more specific semantic nuances. For example, *sepin* 'speech' nominalises to *sepinien*, which can mean speech in general, a formal speech, a word or a language. Similarly, *kikei-en* 'sing-NMLZR' can denote the activity of singing or a song.

With some verbs, the reduplicated form of the verb stem can be used for nominalisation: for example, *kil~kil-en* 'knowledge', *mes~mes-ien* 'thanks', *pil~pile-en* 'playing, game', *so~sohuk-en* 'fishing'.

This last example, *sohuk-en*, *so~sohuk-en* 'fishing', is a rare example of object incorporation in the nominalised form of the verb, suggesting that this combination of verb and loan noun ('throw' and 'hook') is treated as a verb. Other combinations of verbs and typical objects cannot be nominalised in this way (unlike in Paamese, where nominalisation and verbs with their objects is productive Crowley 1982: 86). In Vatlongos, there is a dispreference for nominalising compounds, especially compounds where the second part is non-verbal. The incorporated-preposition (§3.2.1) and incorporated-adverb verbs (§3.3.1) cannot undergo nominalisation. Similarly, the copular verb *he* cannot be nominalised, and nor can verbs that only appear as the subsequent verb in complex predicates (§7.1). However, the preposition *usil* 'along, about', which

appears to have developed from the form of *husil* ‘follow’ when it occurs as the second verb in a complex predicate, can be nominalised to mean ‘topic’. The stative verb *pipin* ‘enough’ was also judged ungrammatical with the nominaliser, but this might be that it was difficult to interpret semantically.

On the other hand, the bare form of the possessed body-part subject of psycho-collocations is included in the nominalisation. For example, the psycho-collocation meaning ‘like, love’ with possessed subject *e* ‘inside’ and a verb *hei* ‘like’ is nominalised as *e-hei-en* ‘liking, love’. Evidence from reduplication suggests that a morpheme boundary is respected between the possessed subject and the verbal root: *uli-kan-en* ‘fear’ reduplicates to *uli-ka~kan-en*, rather than *ul~ulikan-en* (ungrammatical).

5.3.2 Adjectiviser *te-*, *ta-*

The adjectivizing prefix *te-*, *ta-* derives adjectives from stative verbs. The derived forms have the same distribution as adjectives. They can be used attributively without a relativiser, directly following a head noun within a NP.

(422) *la-kaakau* *ena* *suhon* *ta-melxoluk*
 3PL.NFUT-walk LOC sand ADJZR-dark
 ‘They walk on the black sand.’ [20141105f_p01e013_30]

(423) *an* *te-sa* *mi-kil-a* *bas-i*
 weather ADJZR-bad 3SG.NFUT-can-3OBJ 3SG.NFUT.hit-3OBJ
 ‘Bad weather can damage them [yams].’ [20141106d_p07e016_6]

They can also follow the copular verb *he*(§3.5.2).

(424) *vet* *uhia* *xil* *ak* *be*
 shoot wild_yam PL PROX 3SG.NFUT.COP
te-lep *xa* *be* *te-lep*
 ADJZR-big SUB 3SG.NFUT.COP ADJZR-big
 ‘These wild yam shoots are really big.’ [20141106d_p04e016_13]

(425) *ba* *be* *te-metu* *kuhi*
 GO 3SG.NFUT.COP ADJZR-ready good
 ‘It [*manioc*] gets good and ready.’ [20141030a_p01m004_33]

Like other adjectives, these forms can be used predicatively in relative clauses without the copular verb. The *te-* variant of the prefix in this context is ambiguous with a third person singular subject verb in the prior:

- (426) *tati na-van xa te-tiamu*
 father CL.GEN-1SG.POSS REL ADJZR-first/3SG.PRI-first
 'My father who is first-born.'
 'My father who was first.' [20150305f_p01s110_19]

Sometimes the stem is reduplicated to mark intensity (§5.2.2).

- (427) *lu-be ta-le~lep*
 3DU.NFUT-NFUT.COP ADJZR-RED~big
 'They were very big.' [20141220g_n01s080_03]

This affix is not very productive. Most occurrences are with a few stems: *lep* 'big', *metu* 'ready', *sa* 'bad', *sap* 'different', *tin* 'hot' and *tiamu* 'go ahead, be first'. This strategy is also used for colour terms, especially in translations from Bislama and English: *te-meang* 'yellow', *te-melxoluk* 'black, dark', *te-mieh* 'white, light', *te-mmeas* 'green, blue', and *te-pili* 'red'. Derived adjectives are prone to re-lexicalisation: for example, *te-tiamu* 'first' is also used as an adverb or discourse marker meaning 'before'.

5.4 Multiple exponence of TAMP

Now that all the morphological markers of TAMP have been described, it is worth pointing out that a major feature of Vatlongos morphology is multiple exponence of categories, where a single category is marked by more than one formative in the verb or VP (including the negative clitic *ti*). Multiple exponence is when a feature is realised in more than one position in a domain, and is challenging for many models of morphology (Caballero & Harris 2012).

In Vatlongos, TAMP categories are marked in up to three different positions in the verb. The non-future is uniquely identified by subject-indexing prefixes, the negative prefixes, and the stem of irregular verbs undergoing verb-initial consonant mutation, although the verb-initial consonant does not cooccur with the negative prefixes. Although it is only uniquely identified by the negative prefixes, negative polarity also determines the form of irregular verb stems and the presence of the negative clitic. Multiple exponence of TAMP in Vatlongos is restricted to certain subject person-number environments (due to patterns of syncretism (§5.1.1.1)), or to verb lexemes undergoing verb-initial consonant mutation (§5.2.1).

Example (428) shows non-future tense indicated by the subject prefix and the negative prefix, while negative polarity is indicated by the negative prefix, the initial consonant mutation of the verb stem, and the negative clitic *ti*. Unlike in other examples in the thesis, all possible features for a morpheme are given, separated by a /, following Baerman et al. (2005: 12). This makes patterns of syncretism explicit, and shows how the different parts of the morphology must be analysed together to give the correct analysis. The correct analysis on the basis of the other forms is listed first and emphasised in **bold**.

(428) *na-taa-ve* *sup*
1SG.NFUT/1SG.IFUT/3SG.APR-NFUT.NEG-NEG.COP/IFUT.COP chief
 =*ti*
 =NEG/PART/PROH
 'I'm not the chief.' [20141223b_h01s046_10]

The extensive syncretism in the subject prefix paradigms means that in some environments the intended tense is only clear in negative polarity or with irregular verb stems. The subject prefix *na-* '1SG.NFUT/1SG.IFUT/3SG.APR' exemplifies this. Example (429) is ambiguous for three tense, mood and subject interpretations; they are given on separate gloss lines next to the translations.

(429) *na-mesei* [20141117a_n01m003_8]
 1SG.NFUT-sick 'I am sick'
 1SG.IFUT-sick 'I'm going to get sick'
 3SG.APR-sick 'They might get sick, which is bad'

Negative polarity rules out the apprehensive reading, and distinguishes the non-future and immediate future readings:

(430) *na-taa-sepin* =*ti*
1SG.NFUT/1SG.IFUT/3SG.APR-NFUT.NEG-speak =NEG/PART/PROH
 'I didn't speak.' [20150419e_h01m128_23]

(431) *na-naa-sepin*
1SG.IFUT/1SG.NFUT/3SG.APR-IFUT.NEG/DFUT.NEG/PRI.NEG-speak
 =*ti*
 =NEG/PART/PROH
 'I'm not going to speak'

Irregular verbs like *has* 'hit', an h/v/m/b verb, can also disambiguate the syncretism in the prefix paradigms.

- (432) *na-bas-uk*
 1SG.NFUT/1SG.IFUT/3SG.APR -NFUT.hit-2SG.OBJ
 'I hit you.'
- (433) *na-mas-uk*
 1SG.IFUT/1SG.NFUT/3SG.APR -1SG.IFUT.hit-2SG.OBJ
 'I'm going to hit you.' [20170325f_x01s026]
- (434) *na-has-uk*
 3SG.APR/1SG.NFUT/1SG.IFUT-APR/PROH/IMP/DFUT/PRI.hit³⁶-2SG.OBJ
 'It might hit you, which is bad' [20170126c_n01e025].

Due to patterns of syncretism, the multiple marking of TAMP categories is not redundant in some inflectional and lexical environments, and can therefore be characterised as distributed exponence (Caballero & Harris 2012: 170–171; Carroll 2016: 279–322). The only TAMP category that is marked by distributed exponence in all person-number environments is the prohibitive mood. The prohibitive prefixes are a subset of the apprehensive prefixes, so the presence of the negative clitic is essential for distinguishing the prohibitive from the apprehensive (§5.1.2).

5.5 Summary

The verbal morphology of Vatlongos is complex and prone to variation. Multiple exponence of TAMP categories is an organising principle of the system, and the non-canonical morphological phenomena of syncretism and overabundance are widespread in Vatlongos verbal morphology. Three areas of variation in the subject-indexing prefixes pattern with community membership. Speakers in Endu use the conservative *mi-* form of the third-person singular non-future prefix more often and in more phonological environments than speakers in other communities, as well as the more conservative *o-* form of the second-person singular non-future and immediate-future prefix. Mele Maat speakers use the phonologically reduced CVC- forms of CVCV- subject prefixes, which are also underspecified for many TAM distinctions, resulting in a simpler paradigm. There is also inter- and intra-speaker variation in the use of different verb-initial consonant mutation patterns; some patterns are associated with certain

³⁶ This is the basic stem: not all possibilities are listed but it excludes the non-future and first person singular immediate future, so does disambiguate the subject prefix here.

villages. Variation at a more functional level is seen in the association of the object of transitivity verbs with particular semantic roles: only Endu-Vatlongos allows the interlocutor to occur in object position with certain speech verbs.

6 Serial verb constructions

Serial verb constructions (SVCs) are a much-debated issue cross-linguistically and in Oceanic languages, which employ a variety of serialisation strategies. Crowley's (1987; 2002b) research on SVCs in Paamese has been central to the analysis of these constructions in Oceanic. This chapter gives a detailed description of SVCs in Vatlongos in relation to Crowley's observations.

In Vatlongos there is only one construction which meets the criteria of widely-accepted definitions of SVCs. It is very frequent, occurring on average once every 25 words in the subcorpus, and serves a range of functions.

6.1 Defining and describing SVCs

SVCs have proved difficult to define in a way that can be robustly applied cross-linguistically. The central idea that many linguists agree upon is that more than one verb appears in a single clause (Crowley 2002b; Aikhenvald 2006; Haspelmath 2016). However, theory-specific formulations of clausehood result in different sets of criteria for establishing monoclausality. Similarly, different theoretical stances on the structure of the lexicon make it difficult to prove the verbal status of a lexeme within a construction. As a result, the literature on SVCs proposes many sets of definitional criteria, and sometimes definitions only differ by the inclusion or exclusion of one or two criteria, or by the degree of emphasis they place on the relative importance of the criteria.

The layered-clause model of Role and Reference Grammar (RRG) has been especially influential in the description of SVCs in Oceanic languages (Crowley 1987; Crowley 2002b: 42–43). This model distinguishes three nested layers in the clause: the nucleus, which contains the predicate; the core, which contains the predicate and its core arguments; and the clause, which contains the predicate, its arguments and all non-arguments (which appear in the periphery) (Foley & Olson 1985; Van Valin 1990; Van Valin & LaPolla 1997; Van Valin 2010). Each of these layers is associated with a set of operators; for example, aspect modifies the nucleus; modality, the core; and illocutionary force, the clause. SVCs (and other complex sentences) are described in terms of layer of juncture (which of these units are combined) and nexus relations:

subordination, when one unit is embedded in another; coordination, with no embedding; and cosubordination, where there is no embedding, but one unit is dependent on the other for expression of the relevant operators. The distinction between nuclear-layer and core-layer juncture has become the dominant framework for describing SVCs in Oceanic languages.³⁷ Even research pursuing a generativist analysis of SVCs (following e.g. Baker 1989) employs this distinction to classify SVCs in Oceanic (Cleary-Kemp 2015). In frameworks based on X-bar theory, the distinction between ‘nuclear’ and ‘core’ serialisation may be best formulated as V or VP serialisation (Seiss 2009: 505).

A major problem, beyond different theoretical frameworks, is the applicability of criteria across languages and language families. Not all of the criteria put forward can be equally robustly determined in individual languages. Constraints or variability in constituent order, agreement, case systems, TAM categories and other aspects of a grammatical system mean that some of the criteria cannot be applied to examples in some languages or even language families. This means that SVCs are also challenging to descriptive typologists, because it is difficult to establish robust parameters that can be applied cross-linguistically. Haspelmath’s (2016) response to this problem is to restrict definitions of the ‘comparative concept’ to only those criteria which can be established cross-linguistically, usually functional. However, this can restrict the kinds of cross-linguistic comparisons that can be made, for example about the different functions that SVCs perform in individual languages.

Firstly, this section summarises the morphosyntactic criteria that are discussed in the literature: component lexemes must be verbs, must share arguments, must share polarity and TAM values, and must not be overtly marked for coordination or subordination. I also comment on points of contention in the literature and difficulties that arise in applying these criteria cross-linguistically, and to Vatlongos and other Oceanic languages.

³⁷e.g. Hyslop 2001: 278–279; Budd 2009: 216–248; Schneider 2010: 183–184; Jauncey 2011: 173–177; Franjeh 2012: 173–177.

6.1.1 Morphosyntactic criteria for identifying SVCs

The first major criterion that follows from the definition of SVCs as more than one verb in a single clause is the verbal status of the lexemes involved. The elements participating in an SVC must be verbal: they must also function as a verb in a monoverbal clause. While this sounds straightforward, difficulties arise due to individual language's grammatical systems, or because of diachronic processes typical of SVCs. Establishing the word category of a particular lexeme can be challenging due to flexibility in the grammar. François (2017) observes that this kind of flexibility is typical of Oceanic languages, either where a language allows lexemes to be assigned to multiple word classes, or where a single word class is associated with a wide range of functions and grammatical structures. One example is the use of non-verbal elements as predicates (§3.5.1).

Establishing verbal status through use in monoverbal clauses is also difficult because apparently verbal lexemes often become restricted to SVCs (e.g. in Paamese and Apma, Crowley 2002b: 112; Schneider 2010: 87–88). In such cases, it is sometimes possible to make language-specific arguments for a lexeme's verbal status: for example, if it has morphological properties that are exclusively associated with verbs. However, it is on this basis that I am not calling the constructions in §7.1 an SVC, instead describing them as complex predicates to avoid committing to the verbal status of participating lexemes. There are many lexemes that only occur within these constructions, and some adverbs can also participate, including those that appear to be historically derived from non-verbal sources.

Following an even stricter definition of verbhood, Haspelmath (2016: 303) restricts the comparative concept of 'independent verb' to those which express dynamic events, further limiting the definition of SVCs. Oceanic languages often have a large class of stative verbs that express properties and qualities associated with adjectives cross-linguistically. This is the case in Vatlongos, where the verbal status of these lexemes is immediately apparent morphologically; they participate in all the verbal morphological processes described in Chapter 5. Excluding constructions containing stative verbs from a

definition of SVCs would conceal a major function of SVCs in Oceanic languages and crosslinguistically: those described as ‘ambient’ or ‘event-argument’ SVCs (Crowley 2002b: 41–42; Aikhenvald 2006: 27), where one verb modifies the situation expressed by the other (§6.4.2). Similarly, Haspelmath’s strict definition rules out SVCs where a ‘role-marking’ verb is used to increase valency or merge an argument, which is another common function of SVCs crosslinguistically (Aikhenvald 2006: 25–26) (§6.4.7, §6.4.8).

Even where these criteria are met, and a form can function as a verb independently, it has to be established that the same lexeme is being used in SVCs. This can depend on a theoretical model of lexicon structure (Enfield 2009). Cleary-Kemp (2015: 99–103) stresses the importance of establishing that the same lexical item occurs in a SVC and a monoverbal clause, rather than a homophonous auxiliary, adverb or preposition, diachronically derived from the lexical verb.

The relationship between SVCs and AVCs is contentious. For some authors, encoding aspect and other auxiliary-like categories is a common function of SVCs crosslinguistically (Aikhenvald 2006: 23). Others (Jansen, Koopman & Muysken 1978: 125; Anderson 2006; Anderson 2011; Haspelmath 2016) would exclude any constructions where one verb has aspectual or grammatical meaning; the different meaning is evidence of a separate lexeme. It is widely accepted that AVCs can grammaticalize from SVCs (Seiss 2009: 506), and synchronically it can be difficult to determine whether a verb has lost its lexical meaning and is fully grammaticalized as a new auxiliary. Although these uses should probably be excluded from the comparative concept, descriptions of SVCs within individual languages often list aspectual and other auxiliary functions of SVCs (e.g. Paamese, Crowley 2002b: 77–78). Including these auxiliary functions in language-specific descriptive categories reveals the crosslinguistic tendency to use the same morphosyntactic constructions for prototypical SVC functions as well as grammatical functions like aspect. In Vatlongos, SVCs have some aspectual and modal functions, and can be formally distinguished from AVCs (Chapter 8), which I argue have grammaticalized from SVCs.

A second key criterion for SVCs is that the component verbs must share arguments (Crowley 2002b: 40–42; Aikhenvald 2006: 12–20). The possibilities for argument-sharing differ between languages: most frequently the component verbs share a subject (same-subject SVCs), or the object of the initial verb is the subject of a subsequent verb (switch-function SVCs). Alternatively, the situation expressed by one verb might be the subject of the other verb; one predicate is an argument of the other (event-argument SVCs). Argument sharing is generally held to be a definitional criterion, but there are some counter-examples in Oceanic languages: there is serialisation of causally linked intransitive verbs with different subjects in Mwotlap (François 2006: 231). Possibilities in Vatlongos are exemplified in §6.3.2. Haspelmath (2016: 309) does not treat this as a definitional criterion for SVCs, but finds that it is possible to generalise that the verbs share at least one argument.

A third criterion is sharing of TAMP values (Cleary-Kemp 2015: 123). Both verbs are interpreted as having the same TAMP value, although the surface marking of TAMP values may differ, as they do in Vatlongos (Table 37). Haspelmath (2016: 299) suggests that only polarity is a useful criterion for identifying SVCs cross-linguistically, as TAM categories are not directly comparable between languages. However, he does find that, in the existing literature on SVCs, component verbs always have the same tense and mood values, though certain aspectual categories can be marked separately with different scopes on component verbs (Haspelmath 2016: 307–308). In Avatime (Niger-Congo, Ghana), the recurrent aspect can be marked independently on the verbs in an SVC (Defina 2016b: 659).

Reduplication can be marked separately on component verbs of SVCs in Vanuatu languages Mwotlap (François 2006: 228) and Apma (Ridge 2013: 57), as well as in Toqabaqita, an Oceanic language of the Solomon Islands (Lichtenberk 2006: 263). Reduplication performs similar but not identical aspectual functions in all three languages. In Mwotlap it marks pluractionality (François 2004a: 182–184); in Apma it signifies intensity and repetition (Schneider 2010: 85–86); and in Toqabaqita it marks extended duration or high frequency. Reduplication also serves valency-changing functions in both

Vanuatu languages. In Vatlongos, the optional aspectual category of durativity can be marked by verbal repetition of one component verb of an SVC (§4.3.2, §6.3.1).

The last two criteria aim to rule out a coordination or subordination analysis of possible SVCs. Firstly, there should be no overt markers of coordination, whether by a coordinator or another linking element. Haspelmath (2016: 304) urges caution regarding possible linking elements in constructions: if an element only occurs in multi-verb constructions and does not have another meaning, it is probably a linking element and therefore the construction is not an SVC. In Vatlongos and many other Oceanic languages clausal coordination can be zero-marked, so it is important to establish whether a coordinator can be optionally present without changing the meaning of the construction (Crowley 2002b: 17). However, for any one example it may not be possible to rule out a coordination reading, especially for same-subject SVCs.

There should also be no overt markers of subordination, or the possibility of using them. As well as excluding lexical subordinators, it is important that verbs do not appear in forms dedicated to subordination like participles or infinitives, or take TAM categories like subjunctive that are specialised to subordinate clause environments. These strategies are associated with (non-)finiteness, another concept that may differ in different linguistic frameworks (Nikolaeva 2007), or when applied to different languages and language families. Meyerhoff (2000: 174–181) uses a definition of finiteness based on the specification of reference time (§4.1) to argue that the subsequent verb in Bislama SVCs, and other languages in the region, is non-finite.

Regardless of overt-marking of subordination, a more controversial question is whether a serialisation analysis applies when one verb selects the other. While for some authors, absence of a predicate-argument relationship between component verbs is a definitional property of SVCs (Seuren 1991; Haspelmath 2016), from a generativist perspective Cleary-Kemp (2015: 256) concludes that SVCs must involve complementation between the component verbs, ruling out syntactic adjunction or coordination. Although Haspelmath (2016: 305) excludes complement relationships from serialisation, he does

discuss event-argument or ‘ambient’ SVCs in Oceanic languages – where one verb is the subject of the other – as a possible argument-sharing relationship in SVCs (Haspelmath 2016: 310). Perhaps predicate-argument relationships involving subjects rather than complements are permissible in a definition of SVCs.

To summarise, the main morphosyntactic criteria for defining an SVC are that more than one verb occurs in a single clause. The independent verbal status of the component lexemes can be established by the possibility of appearing in monoverbal clauses, being careful to exclude the possibility of a homophonous lexeme from another word class. The monoclausal status of the construction should be demonstrated by shared polarity, tense and modality between component verbs. Shared aspectual categories and arguments are also important clues to monoclausal status, although there are counter-examples. Any overt markers of clause boundaries, such as coordinators, subordinators or non-finite verb forms, exclude a serialisation analysis.

6.1.2 Other factors that correlate with SVCs

Other factors that are sometimes mentioned in definitions of SVCs are eventhood, intonation and gesture. However, I treat the morphosyntactic features discussed so far as the central criteria of serialisation. The morphosyntactic criteria are easier to observe and test than these other proposed factors. As morphosyntactic structures do not necessarily neatly map onto semantic concepts, I want to avoid using semantic evidence to identify a syntactic phenomenon (Defina 2016c; 2016b: 649). Instead, observations in the literature about how SVCs correlate with semantics, discourse, prosody and gesture are ‘falsifiable generalisations’ (Haspelmath 2016: 306), rather than defining properties.

Many definitions state that SVCs express a single event (Cleary-Kemp 2015:97). Bisang (2009) argues that this is the overarching criterion from which all other criteria for SVCs can be derived. He therefore excludes any examples where the component verbs have different status in the discourse. He rules out narrative SVCs like those described for Kalam (Pawley 2008), and would have difficulties with examples in languages like Avatime, where it is

possible for focus to be marked on either verb and to have scope over only one verb (Defina 2016b: 662–663). Support for the idea that there is a one-to-one relationship between SVCs and events comes from evidence that culturally-determined event types can condition the grammaticality of SVCs (Durie 1997), but it is very difficult to test this by objectively establishing eventhood independently of linguistic evidence (Cleary-Kemp 2015: 120; Defina 2016a: 895; Haspelmath 2016: 306).

Many linguists have observed a relationship between serialisation and intonation, either stating that SVCs should have a ‘single prosodic contour’ or the same intonation patterns as a monoverbal clause (Aikhenvald 2006: 7–8; Cleary-Kemp 2015: 118; Haspelmath 2016: 308). Despite the observed correlation, there are two reasons for being cautious about using prosody as a definitional criterion. Firstly, in under-documented languages like *Vatlongos* there is often very little information available about prosody, especially at the level of intonation. Descriptive accounts of SVCs that assert a relationship between serialisation and prosodic units are usually based on auditory impressions, and field linguists may not be trained in this area of linguistics (Crowley 2002b: 17). Secondly, where more rigorous investigations have been conducted, relationships between prosody and syntax are probabilistic and therefore not useful as a defining criterion. Givón (1991) finds a lower probability of a pause between verbs in SVCs than in clause-combining strategies. Moreover, he finds that there is a continuum of probabilities associated with other non-finite and finite clause combining strategies, so prosody alone would not be able to rule out alternative analyses of a construction. Himmelmann (2013) argues that in natural discourse, prosodic breaks can interrupt any grammatical constituent, limiting the usefulness of this criterion.

Recent research by Defina (2016a; 2016c) has also investigated the relationship between gesture and serialisation. In *Avatime*, she finds that SVCs are more likely to be marked with single iconic co-speech gestures than many complex clausal constructions, except non-finite complement subordinate structures that pattern with SVCs (Defina 2016c: 902). She uses this indirect

evidence from gesture to support the argument that SVCs mark a single event (Defina 2016c: 890), although the cognition of eventhood and its interactions with gesture, intonation and syntax need further research.

6.1.3 Typological description of SVCs

In the description of Vatlongos SVCs below I use Aikhenvald's (2006: 3) typological parameters.

The first parameter is composition: whether an SVC is symmetrical, so any verb can participate in any position in the SVC, or asymmetrical, when one of the verbs is from a restricted class. Cleary-Kemp points out that this is a continuum rather than a binary, and that extreme asymmetry discounts a construction from being an SVC (Cleary-Kemp 2015: 105–108). She argues that if the closed class of verbs occurring in one of the positions in an asymmetrical construction cannot be independently motivated by grammatical or semantic features, then it is likely that the members of the restricted class are grammaticalized and are no longer the same lexeme as their lexical equivalents. If all verbs in the construction are restricted, it is evidence of lexicalised compounding, and the possible combinations do not constitute a productive construction.

The second parameter is contiguity: whether other constituents can intervene between the component verbs. While this is often treated as a binary feature (contiguous or non-contiguous), Cleary-Kemp (2015: 143) points out that it is important to specify what kinds of material or constituents can intervene between the component verbs, as this is important evidence of the syntactic structure of the construction. In some constructions, especially those described as nuclear-layer SVCs in Oceanic languages (e.g. Paamese, Crowley 2002b: 82–106), no morphological formatives at all can intervene the verbs, including affixes. These could be described as 'root-contiguous', like Vatlongos complex predicates (§7.1). SVCs can also be described as contiguous when the inflected verbs must be adjacent (e.g. Tariana (Arawak), Aikhenvald 2006: 2). Non-contiguous SVCs often only allow the object of the first verb to intervene, but Vatlongos is unusual in also tolerating PPs and adverbs (§6.3.1).

This is connected to the third parameter, wordhood. Components verbs of SVCs can be independent grammatical or phonological words, or together form a single word. One-word SVCs are likely to be contiguous, and non-contiguous SVCs are likely to be multi-word; non-contiguous, one-word SVCs have not been attested (Aikhenvald 2006: 39).

Aikhenvald's final parameter is marking of grammatical categories. There is a basic distinction between single-marking, when categories are only marked once in an SVC, and concordant marking, when the category is marked on each component verb. Vatlongos and Paamese are unusual in having non-identical concordant marking. Although the SVC as a whole is understood to have a single TAMP value, a definitional requirement for SVCs, the surface morphological marking on each verb is different in certain TAMP environments (Table 37, Table 38).

6.1.4 Conceptual issues

Before describing Vatlongos SVCs, it is worth commenting on one more contested question: what type of concept is verb serialisation? For some linguists, the concept of SVCs is a constellation of features. This constellation could be pictured as a canonical phenomenon (Brown, Chumakina & Corbett 2013), but is usually described as prototype, the approach taken by Aikhenvald (2006: 3). One problem with the prototype model is that each of the typological parameters is a cline between two extremes. It is not clear which end of each cline should be taken as 'central' to the prototypical concept. This leads to some confusion in the application of the terminology: some authors propose that more tightly-bound constructions are more prototypical (e.g. Cleary-Kemp 2015: 147), while others assume the opposite: for Butt and Seiss (2009) a 'prototypical serial verb' involves several events, in opposition to complex predicates denoting 'a single (albeit complex) event'.

Cleary-Kemp (2015: 128) resolves this paradox by grouping SVCs along the nuclear/core divide, and treating these as separate clusters of prototypical features based on Aikhenvald's parameters. Nuclear-layer SVCs have a fused argument structure, are contiguous, form a single word and have single-marking of verbal categories, while core-layer SVCs have separate sets of

arguments, are potentially non-contiguous, form separate words and have concordant marking of verbal categories. Based on a sample of 36 serialising Oceanic languages, she finds that around half the languages have both types of SVCs, while slightly more languages have only nuclear-layer than only core-layer. Although synchronically Vatlongos appears to only have the core-layer type as a true SVC, the complex predicates discussed in §7.1 are clearly similar to nuclear-layer SVCs in related languages, but are no longer sufficiently productive to meet the definition of SVCs.

For Haspelmath (2016: 313), prototype-based models are not useful for making testable generalisations, and he therefore strives to delineate a cross-linguistic comparative concept that is applicable to all languages (Haspelmath 2010), which consequently mainly appeals to functional criteria. However, strict functional criteria can exclude examples which use an identical morphosyntactic strategy by language-internal criteria: Vatlongos SVCs extend into a wide range of functions (§6.4), some of which would not meet Haspelmath’s cross-linguistic definition.

6.2 Overview of SVCs in Vatlongos

Non-contiguous SVCs in Vatlongos typically involve two verbs which each take a full set of verbal affixes and can be separated by intervening constituents, but cannot contribute differing TAMP values, and are therefore analysed as part of the same clause. Example (435) is a typical example where the subsequent verb contributes path information to the event described by the initial verb, and a goal for its object.

(435) ***mi-leh*** *xil* ***la-be*** *sip*³⁸
 3SG.NFUT-take 3PL 3PL.NFUT-NFUT.go_to ship
 ‘It took them to the ship.’ [20150305h_h01s111_08]

Often Vatlongos SVCs are asymmetrical, with the subsequent verb modifying the initial verb. This is in line with the head-initial constituent order within Vatlongos VPs and NPs, as well as in Oceanic languages more generally (Cleary-

³⁸ In examples in this chapter, serialised verbs are emphasised in **bold**, and any subordinate clauses are demarcated with [square brackets].

Kemp 2015: 150). However, when a basic motion verb is used to indicate the direction of motion prior to the situation expressed by the main verb, the modifying directional verb occurs first, suggesting iconic motivation for the relative order of the component verbs, which is a common feature of SVCs cross-linguistically (Aikhenvald 2006: 22; Haspelmath 2016: 307). A much larger set of lexemes participate in SVCs than in complex predicates, indicative of less constrained asymmetry.

Vatlongos SVCs are very non-contiguous. An object NP is the most frequent constituent to intervene between the two verbs, as in (435), but PPs and adverbs can also intervene, suggesting an unusually loose syntactic relationship between the component verbs.

Each verb is an independent phonological and grammatical word. The syntactic independence of each verb is demonstrated by the possibility of intervening constituents, and morphologically each verb takes a full set of verbal affixes, unlike in complex predicates.

Vatlongos SVCs are typologically unusual in their marking of grammatical categories. Subject agreement is independently marked on each verb and reflects the underlying argument structure of the construction. Subject agreement marking is a good diagnostic of whether an SVC is same-subject, switch-function, or event-argument (in which case the subsequent verb takes third-person singular subject marking).

Morphological TAMP is marked on each component verb, in an unusual pattern of concordant marking (Aikhenvald 2006: 40). The coding on the two verbs does not always match, and neither verb takes an invariant default value. Instead, the encoding on the initial verb determines the TAMP of the SVC, and the morphological marking on the subsequent verb is predictable from the TAMP marking of the first verb. In affirmative polarity the subsequent verb is marked for the same TAMP as the initial verb, as in (435), but the morphological marking on the verbs can differ in negative polarity, as well as the imperative, prohibitive, and apprehensive moods. However, the

interpretation of TAMP values must be identical, despite different morphological marking.

Following Crowley's (1987; 2002b) descriptions of serialisation in Paamese and other Oceanic languages, similar constructions in Oceanic languages are often described as core-layer SVCs. I have opted to refer to these as 'non-contiguous' instead, so as not to tie the analysis to an RRG model of layered clauses. Moreover, some definitions of clause-layer juncture in RRG would not allow anything except core arguments to intervene between the verbs in a core-layer SVC (Foley & Olson 1985: 47), which might make the term inappropriate for Vatlongos SVCs, as adverbs and oblique PPs can intervene, and these would be analysed as part of the peripheral layer in RRG.

It is also important to demonstrate that overt coordinators and subordinators cannot occur between the two verbs without a change in meaning. The example that Crowley (1987; 2002b) uses to exemplify this distinction in Paamese is 'killing a pig dead', and this also works well in Vatlongos. In this case as well as many other examples, the first verb is an 'action' and the second a 'result', both making up a single complex event.

The first verb in these examples is *has* which can mean 'hit' or 'kill', while the second verb, *met* 'die' is an intransitive verb taking the object of the first verb as its subject.

(436) *xi bas ueili met*
 3SG 3SG.NFUT.hit pig 3SG.NFUT.die
 'He killed the pig.' [20150228a_x01s026_01]

Example (436) is a switch-function SVC. It should not be analysed as a reduced coordinate or subordinate multiclausal structure, because forcing a coordinate or subordinate reading results in a very different reading. In example (437), the general coordinator *e* separates the two verbs into separate clauses, and the situation speakers prefer for this sentence is one where a man hits a pig gently and then is surprised that it falls over and dies.

(437) *xi bas ueili e met*
 3SG 3SG.NFUT.hit pig and 3SG.NFUT.die
 ‘He hit the pig and it died.’
 * ‘He killed the pig’ [20170112b_x01s046]

In example (438), the general subordinator *xa* is interpreted as a relative clause marker: speakers prefer the surprising reading of hitting a pig that is already dead.

(438) *xi bas ueili xa met*
 3SG 3SG.NFUT.hit pig REL 3SG.NFUT.die
 ‘He hit the pig that died.’
 * ‘He killed the pig’ [20170112b_x01s046]

These examples demonstrate that an SVC can be used to describe a complex event with a single argument structure frame. However, depending on the semantics of the component verbs it is not always easy to pinpoint such clear differences in interpretation that rule out zero-coordination or subordination analyses.

One syntactic test that distinguishes Vatlongos SVCs from coordinate structures, although not from zero-marked complement clauses, is the possibility of extracting the object of a subsequent verb (Jansen, Koopman & Muysken 1978; Haspelmath 2016: 301). In the relative clause in example (439), the pig is the object of the second verb, co-referenced with an object suffix.

(439) *ueili xa o-ba a-gur-i saotis*
 pig REL 2SG.NFUT-NFUT.go 2SG.NFUT-NFUT.take-3OBJ southeast
 ‘the pig that you went and brought from the south east’
 [20170222d_n01s152_83]

6.3 Formal properties

SVCs can be identified by a combination of morphosyntactic features, but they are sometimes ambiguous with complex predicates, AVCs, zero subordination or juxtaposed independent clauses. Ways to rule out these alternative analyses are pointed out in this section, but it is not always possible for individual examples in certain TAMP, subject person-number, semantic or lexical environments.

6.3.1 Constituent order

The possible constituent orders in an SVC with two component verbs are schematized in Figure 11. In addition, adverbs can intervene between all constituents except a verb and its object (§3.3). This makes SVCs in Vatlongos very non-contiguous compared to other SVCs cross-linguistically. It can be difficult to decide whether a PP is an oblique argument selected by a verb, or an adjunct (§3.2), but it appears that both argument and adjunct PPs can appear in the PP positions. The PPs are marked with a Kleene star to indicate that none to many can appear. This distribution is amenable to an analysis of serialisation as VP adjunction: effectively the second verb in an SVC and its object have the same distribution as any other phrasal adjunct, such as PPs.

(NP) Verb1SUBJ	(Aux)	Verb 1	(NP) Verb1OBJ(=Verb2SUBJ)	PP*	Verb 2	(NP) Verb2OBJ	PP*
-------------------	-------	--------	------------------------------	-----	--------	------------------	-----

Figure 11: Constituent order possibilities in SVCs with two component verbs

Example (435) above shows the cross-linguistically common pattern of the object of the first verb intervening between the two verbs. This is possible even for objects that are encoded with complex, phonologically-heavy NPs. In (440), the object NP between the two verbs contains a quantifier, a locative PP, an apposed coreferential NP and a zero-marked relative clause, as well as being preceded by a hesitation marker. This differs from findings for serialisation in Tuam (Oceanic, Papua New Guinea), where a less-activated object referent expressed by a full NP normally occurs in biclausal structures marked with an overt coordinator (Bugenhagen 2010: 464).

- (440) *mat-tis* *na* *neta,*
 1PC.NFUT-write HES thing
sepin-ien *xil* *ena* *vanuvueiou,*
 speak-NMLZR PL LOC book
ngan *la-kes-i* *ni* *baebol,*
 one 3PL.NFUT-call-3OBJ OBL bible
bemei *ena* *sepin-ien* *xil*
 3SG.NFUT.come LOC speak-NMLZR PL
tahal *e* *Vatlongos*
 from LOC *Vatlongos*
 ‘We translate the um, words in the book, the one they call the Bible, into words from Vatlongos.’ [20141107d_p01s022_05-07]

In (441), only the second verb is transitive and is followed by an object NP.

- (441) *mat-ketteh* *mat-pus* *makon*
 1PC.EXCL.NFUT-look 1PC.EXCL.NFUT-see flying_fox
 ‘We (*looked and*) saw a flying fox.’ [20150225b_n01s097_02]

Intervening PPs are more unusual cross-linguistically. In (442), the proform for locative PPs with a third-person singular object *e-n* (§3.2) occurs between the manner of motion initial verb, and a directional subsequent verb that specifies the path and goal. Example (443) shows the same PP proform in a negative polarity SVC, where the monoclausal status is clear from the scope of negation. The modifying subsequent verb is also negated, although obligatorily marked for affirmative polarity and immediate future (Table 37).

- (442) *na-go* *e-n* *na-be* *Vila*
 1SG.NFUT-NFUT.pass LOC-3SG.OBJ 1SG.NFUT-NFUT.go_to Vila
 ‘I went to Vila in it [*a ship*].’ [20170224a_n01s141_005]
- (443) *la-taa-mmei* =*ti* *e-n* *va-xehik*
 3PL.NFUT-NFUT.NEG-come =NEG LOC-3SG.OBJ 3SG.IFUT-strong
 ‘They don’t come to it [*knowledge*] strongly.’
 [20170413a_h01m169_47]

Full PPs can also intervene between the two verbs. Examples (444) and (445) show PP adjuncts specifying the spatial or temporal circumstances of the first verb, contrasted with a different spatial or temporal PP adjoined to the second verb.

- (444) *kuli mot* *nesau bemei* *e tan*
 dog 3SG.NFUT.fall on_top 3SG.NFUT.come LOC ground
 ‘The dog fell down from up there’ [20150118a_n01m090_17]
- (445) *krup* *ak li-pol* *e maakue*
 group PROX 3PL.DFUT-work LOC morning
i-vus *e wan* *aklok*
 3SG.DFUT-finish LOC one o’clock
 ‘This group will work in the morning until one o’clock.’
 [20150223a_n01s096_32]

The subsequent verbs appear to be hierarchically adjoined in the same way as PP arguments and adjuncts, rather than occupying a slot in a template. This means that where there is both an SVC and a PP in the clause, the PP can precede the subsequent VP or vice versa. Comparing (443) and (446) shows that speakers are free to use either constituent order, even when the PP is an argument of the first verb.

- (446) *na-taa-sepin* =*ti* *va-hik* *mi-ni*
 1SG.NFUT-NFUT.NEG-speak =NEG 3SG.IFUT-strong to-3OBJ
 'I didn't shout at him.' [20150419e_h01m128_24]

Adverbs (and adjuncts of absolute place, (439)) can also intervene between the two verbs, following any object of the first verb.

- (447) *lu-pou* *akis* *sung* *lu-ba*
 3DU.NFUT-carry axe first 3DU.NFUT-NFUT.go
 'First they carried the axe away.' [20141027a_n01m001_007]

- (448) *o-lis* *vuteili* *o-mmei*
 2SG.NFUT-move_over a_bit 2SG.NFUT-come
 'You move over this way a bit.' [20170224a_n01s141_126]

Example (449) shows an oblique PP and an adverb intervening between the two verbs (the first verb is made up of two verbal lexemes in a complex predicate, §7.1):

- (449) *ni-hit mese-ni* *mi ratel* *tang i-mak*
 1SG.DFUT-say+ready-TR to 1PC.INCL just 3SG.DFUT-like_this
 'I'll explain it to us like this.' 20170124b_t01e137_18

Further evidence for analysing these SVCs as adjoined VPs comes from the possibilities for verbal repetition. §4.3 describes how repetition of a verb and its object can be used to express durative aspect or to intensify the gradable component of a stative verb. This process can be applied to the subsequent verb in an SVC. Examples (450) and (451) show the durative repetition strategy applied to a subsequent verb specifying a path. Both examples emphasise the duration of the journey, which is contrasted with shorter journeys in the narratives: in (450) other people run to closer villages while Maison's group run further; in (451), this is the third time the rat has borrowed a burning branch to light his fire but not made it all the way home.

- (450) *Maison* *xatel* *lata-loh* *lata-ba*
 Maison 3PC 3PC.NFUT-run 3PC.NFUT-NFUT.go
lata-ba *lata-ba*
 3PC.NFUT-NFUT.go 3PC.NFUT-NFUT.go
 'Maison and them ran away on and on.' [20150226a_n01s098_23]

- (451) *ngan-ak,* *u-hur-i* *u-ha* *u-ha...*
 one-PROX 2SG.DFUT-take-3OBJ 2SG.DFUT-go 2SG.DFUT-go
 'This one, you'll carry it on and on...' [20170331c_n01s140_29]

Similarly, when a stative verb is used as a subsequent verb to modify the meaning of the initial verb, it can be repeated for intensification:

- (452) *la-kaakau* *la-katten* *la-katten*
 3PL.NFUT-walk 3PL.NFUT-go_fast 3PL.NFUT-go_fast
 ‘They walked really fast.’ [20150225a_x02s046_01]

There are also examples which could be intensification marked by repetition with the subordinator *xa* on the subsequent verb of an SVC, but these examples are all in the third-person singular and are ambiguous with a zero-marked subordination of a sentential subject (§7.2.2):

- (453) *mama* *mi-kaakau* *mi-reru* *xa* *mi-reru*
 mum 3SG.NFUT-walk 3SG.NFUT-slow SUB 3SG.NFUT-slow
 ‘Mum walks very slowly’ / ‘Mum’s walking is very slow.’
 [20150303c_x01e016_63]

- (454) *holesok* *sa-lu* *xil-e* *meul*
 thing CL.DOM-3DU.POSS PL-CONTR 3SG.NFUT.grow
bos *xa* *bos*
 3SG.NFUT.good SUB 3SG.NFUT.good
 ‘Those things of theirs grew really well’ [20141027a_n01m001_29]

There are two examples in the corpus that appear to show coordination of the serialised verbal constituent, further supporting a VP adjunction hypothesis.

- (455) *Xi* *ba* *di* *wes mu di* *Paama*
 3SG 3SG.NFUT.go 3SG.NFUT.stay west or 3SG.NFUT.stay Paama
 ‘They go to West Ambrym or to Paama.’ [20170217j_n01s125_17]

- (456) *i-mas* *i-sumok-ni* *maruana mu i-mun*
 3SG.DFUT-must 3SG.DFUT-smoke-TR marijuana or 3SG.DFUT-drink_lots
 ‘they have to smoke marijuana or drink’ [20170119f_n01s136_36]

However, as both these examples are in affirmative polarity, they have matching TAMP marking and are ambiguous with other analyses. Example (455) could be an AVC (Chapter 8), although discussions with consultants support a directional modification interpretation, with the stay verbs specifying a goal for ‘go’, rather than ‘go’ contributing a prior motion to ‘stay’. Example (456) could be subordination, although the Bislama loan *mas* ‘must’ is otherwise only used in SVCs and AVCs.

Stronger evidence is the elicited example in (457). Here the negative clitic only occurs once after the initial verb, and the final VP constituent would be

ungrammatical if it were an independent clause, as the verb takes a negative prefix but is not followed by a negative clitic.

- (457) *Li-naa-loh* =*ti* *li-naa-ve* *Endu* *mu*
 3PL.DFUT-NEG-run =NEG 3PL.DFUT-NEG-go_to Endu or
li-naa-ve *Sameo?*
 3PL.DFUT-NEG-go_to Sameo
 ‘Won’t they run to Endu or to Sameo?’ [20180801a_x02s046]

Although the examples so far have all had two component verbs, it is also possible to have multiple verbs in an SVC. When the serialised verbs provide path information, the ordering appears to be iconic, as in (458). Otherwise the ordering is quite flexible. In (459), the second verb modifies the direction of gaze, while the third verb describes the result of looking (seeing something). In (460), the second verb is an evaluative description of the first verb, while the third verb modifies the direction.

- (458) *lu-but* *lu-ba* *lu-di* *e-n*
 3DU.NFUT-NFUT.jump 3DU.NFUT-NFUT.go 3DU.NFUT-NFUT.stay LOC-3SG.OBJ
 ‘They jumped out into it [*a canoe*].’ [20170222d_n01s152_91]

- (459) *Mi-ketteh* *be* *nesau* *pus* *asu xal* *vevu*
 3SG.NFUT-look 3SG.NFUT.go_to up 3SG.NFUT.see rat with rail
 ‘He looked up and saw the rat and the rail.’ [20170331c_n01s140_52]

- (460) *anien* *sa-n* *mi-teh* *bos*
 food CL.DOM-3SG.POSS 3SG.NFUT-grow 3SG.NFUT-good
be *nesau*
 3SG.NFUT.go_to up
 ‘Its flesh grows up well.’ [20141030a_p01m004_32]

In examples in affirmative polarity, where matching TAMP marking is expected, it is difficult to rule out a zero-coordination or zero-subordination analysis. In (461), the subject of the first verb and the object of the second verb are together the subject of the third verb: an inclusory argument structure which is common in two-verb SVCs. The auxiliary verb also seems to scope over all the verbs, further evidence that this is true embedding rather than zero coordination of clauses.

- (461) *di* *mi-loh* *del-a* *lu-ba*
 CONT.REAL 3SG.NFUT-run 3SG.NFUT.with-3OBJ 3DU.NFUT-NFUT.go
 ‘it’s running off with him.’ [20141212h_n01s046_62]

There is also one example of multiple serialisation in negative polarity, showing that all non-initial verbs take the same dependent TAMP marking, and that the negative clitic only occurs once per SVC, even with more than two verbs.

- (462) *taa-mmei* =*ti* *va-loh* *va-hos*
 3SG.NFUT-come =NEG 3SG.IFUT-run 3SG.IFUT-good
 'It [*schooling*] didn't start going well' [20170413a_h01m169_60]

The copular verb *he* and its complement can participate in SVCs like any other VP. Predicative adjectives or NPs cannot appear in SVCs without the copular verb. This is one reason for analysing this strategy as an SVC rather than a more general complex predicate or predicate serialisation construction (unlike in Daakaka von Prince 2015: 307). Example (463) shows a copular verb with a proper noun, used to show the result of a name change in a multi-verb SVC with an intervening adverb.

- (463) *ma-ling* *ise-n* *sung* *biles*
 1DU.EXCL.NFUT-put name-3SG.POSS after 3SG.NFUT.turn
be *Pamela*
 3SG.NFUT.COP Pamela
 'We then changed her name to Pamela.' [20150413a_h01s122_10]

6.3.2 Morphological indexing and argument structure

Both verbs in an SVC index their own logical semantic subject. This makes it easier to identify the argument structure of each verb and the SVC than in languages where there is default subject-indexing on the subsequent verb (e.g. Bislama (Meyerhoff 2000: 166), Apma (Schneider 2010: 187)). Vatlongos SVCs exhibit all the argument sharing patterns in SVCs in Oceanic languages discussed by Crowley (2002b: 41), and most major types of argument sharing in Aikhenvald's (2007: 12–20) crosslinguistic survey. Two kinds that are not found in Vatlongos are SVCs where the verbs share an object but have different subjects (Aikhenvald 2006: 20) and resultative SVCs. In resultative SVCs, intransitive verbs with different subjects are serialised and the subsequent verb describes a result of the initial verb, as is found in Mwotlap (Aikhenvald 2006: 19–20; François 2006: 231).

The most common argument sharing pattern found in SVCs crosslinguistically is where verbs share a subject (Aikhenvald 2006: 14). This is often the only argument structure possible in a language, and every language

the first verb, perhaps a speaker is always accessible enough within discourse to be included in this way.

- (474) *la-mmei* *ma-di* *ige*
 3PL.NFUT-come 1PL.EXCL.NFUT-NFUT.stay here
 ‘They came and we stayed here.’ [20150226a_n01s098_69]

Finally, event-argument SVCs are common in Vatlongos. In these constructions, the event or situation predicated by the first verb is the subject of the second verb, which takes third-person singular subject-indexing. This is a common strategy for describing or evaluating a situation, and is the major mechanism for questioning the manner or character of a situation using the interrogative verb *tep* ‘how’, (476).

- (475) *na-pol* *gehik*
 1SG.NFUT-work 3SG.NFUT.strong
 ‘I worked hard.’ [20150219b_n01s001_66]
- (476) *li-pol-ni* *i-tep* *xiak?*
 3PL.DFUT-do-TR 3SG.DFUT-how now
 ‘How are they going to do it then?’ [20170413a_h01m169_110]

6.3.3 Morphological TAMP

The morphological dependency between the two verbs involved in an SVC is complicated. Although the interpretation of TAMP values is always identical in Vatlongos SVCs, the TAMP marking on the two verbs does not always match, and neither does the subsequent verb take default marking. Rather, the marking on the second verb is dependent on the marking on the first verb, which determines the TAMP for the whole construction. This relationship is further complicated by the patterns of syncretism and overabundance that characterise the verbal morphology of Vatlongos (§5.4). The TAMP value marked on the subsequent verb is dependent on the first value, but where there is more than one form associated with that value, any form can be used, even when one of the forms is ambiguous with other TAMP values (§5.1.1).

The morphological TAMP dependencies are summarised in Table 37.

Table 37: Morphological TAMP dependencies in SVCs

Initial verb TAMP	Subsequent verb TAMP
Affirmative polarity	
Non-future +	Non-future +
Immediate future +	Immediate future +
Distant future +	Distant future +
Prior +	Prior +
Negative polarity	
Non-future - (prefix+clitic)	Immediate future +
Immediate future - (prefix+clitic)	Immediate future - (prefix only)
Distant future - (prefix+clitic)	Distant future - (prefix only)
Prior - (prefix+clitic)	Prior - (prefix only) OR Immediate future + OR Immediate future - (prefix only)
Moods	
Apprehensive	Apprehensive
Imperative	Imperative OR Immediate future +
Prohibitive (with <i>ti</i> clitic)	Apprehensive=prohibitive (prefix only) OR Immediate future - (prefix only)

(+ affirmative polarity, - negative polarity)

In affirmative polarity the two verbs are marked for identical TAMP values, but in negative polarity the dependencies are more complicated. In the non-future, the subsequent verb is marked for immediate future and affirmative polarity. However, the TAMP values of the construction as a whole are non-future and negative, so the scope of negation is a good diagnostic of an SVC in this TAMP environment.

(477) *xi taa-vas ueili =ti va-met*
 3SG 3SG.NFUT.NEG-NEG.hit pig =NEG 3SG.IFUT-die
 'He didn't kill the pig.' [20150227a_x01s046_02]

(478) *la-taa-ketteh =ti ve-he nesau*
 3PL.NFUT-NFUT.NEG-look =NEG 3SG.IFUT-go_to up
 'They didn't look up.' [20170217a_n01s038_13]

In the future tenses, both verbs are marked for identical TAM and with a negative prefix, but the negative clitic only appears once in an SVC, after the first verb or its object. The table therefore distinguishes negative polarity marked with 'prefix and clitic' from 'prefix only'. This means it is straightforward to identify SVCs in these TAMP contexts. Outside of an SVC, the second verb would be ungrammatical because the prefix appears without a following clitic (§3.6.2).

- (479) *hura* *xal* ***naa-luk*** =*ti* ***naa-ve*** *igak*
 yam PROX 3SG.FUT.NEG-grow =NEG 3SG.FUT.NEG-go_to here
 ‘This yam won’t grow here.’ [20141106d_p08e016_08]
- (480) ***mu-naa-so*** *vueil* *na* *sa-mim*
 2PL.DFUT-NEG-put pig HES CL.DOM-2PL.POSS
xil =*ti* ***li-naa-ve*** *iat*
 PL =NEG 3PL.DFUT-NEG-go_to.NEG yard
 ‘[If] you don’t put your pigs into pens {we will kill them}’
 [20150219b_n01s001_031]
- (481) ***naa-ketteh*** =*ti* *vari* ***naa-pus*** *xil*
 3SG.FUT.NEG-look =NEG at_once 3SG.FUT.ENG-see 3PL
 ‘They mustn’t look at them.’ [Y3P3_12]

The marking of the subsequent verb in the prior is less predictable. Prior in negative polarity is a fairly rare combination (there are only 14 examples in the subcorpus, and only one of them is an SVC), and the conventions for marking SVCs in this TAM context do not seem to be settled. The two main strategies found in elicitation reflect the prior’s peculiar morphological status: in some ways it patterns with the non-future in expressing actual events, in others it patterns with the future tenses, especially in taking the *naa-* form of the negative prefix. The most frequent strategy is analogical with coding of the future tenses: the subsequent verb is marked for prior and negative polarity, but does not take a negative clitic.

- (482) ***te-naa-vas*** *ueili* =*ti* ***te-naa-met***
 3SG.PRI-NEG-hit pig =NEG 3SG.PRI-NEG-die
 ‘He hadn’t killed the pig.’ [20150227a_x01s046_12]

The second most frequent strategy appears to be based on an analogy with the non-future: the second verb is marked for immediate future in affirmative polarity. This includes the only example of this TAMP environment in an SVC in the subcorpus, shown in (484).

- (483) *xamel* ***mate-naa-kaakau*** =*ti* ***mal-he*** *bien*
 1DU.EXCL 1DU.EXCL.PRI-NEG-walk =NEG 1DU.EXCL.IFUT-go_to sea
 ‘We hadn’t walked to the sea.’ [20150227a_x01s046_49]
- (484) ***lat-naa-pus-i*** =*ti* *tetiamu* ***va-mak***
 3PL.PRI-NEG-see-3OBJ =NEG before 3SG.IFUT-like_this
 ‘They’d never seen one [a place] like this before.’
 [20141105f_p01e013_56]

The final strategy seems to be a mixture of these two strategies, with the second verb marked for immediate future and with the negative prefix. This only occurred a few times with two speakers and in contexts where the prior form was ambiguous with an immediate future. Both speakers were translating equivalent forms in Bislama (with *bin*, Crowley 2004b: 93), and accepted it as a grammatical translation of the intended meaning in checking, but both preferred the other strategies on reflection.

(485) *xamel mat-naa-vas xamil =ti mul-naa-lel*
 1DU.EXCL 1DU.EXCL.PRI-NEG-NEG.hit 2DU =NEG 2DU.EXCL.IFUT-NEG-die
 'We hadn't killed you.' [20150228a_x01s026_60]

(486) *mut-naa-sa mat =ti naa-va-ni xatel*
 2PL.PRI-NEG-give mat =NEG 3SG.IFUT-go-TR 3PC
 'You didn't give them the mat.' [20150303j_x01e016_26]

The apprehensive, imperative and prohibitive moods are not analysed as having affirmative or negative polarity, but their behaviour in SVCs suggests analogies with the other TAM categories in affirmative and negative polarity. Both verbs are marked for the apprehensive, similarly to relative tenses in affirmative polarity.

(487) *na-mot na-mmei*
 3SG.APR-fall 3SG.APR-come
 'It might fall down here' [20170411a_x01s046]

(488) *lali-na-mot lali-na-mmei*
 3DU.DFUT-APR-fall 3DU.DFUT-APR-come
 'They might fall down here' [20170411a_x01s046]

In imperative mood there is a distinction between same-subject SVCs and other argument configurations. In same-subject SVCs, both verbs are marked as imperative:

(489) *suvul ammei e tan!*
 IMP.SG.go_down IMP.SG.come LOC ground
 'Come down to the ground!' [20150305d_n01s109_12]

(490) *xamil lu-ha lu-leh his xa mu-kuk-ni!*
 2DU IMP.DU-go IMP.DU-take banana REL 2PL.NFUT-cook-TR
 'You two go and get the banana you lot cooked!'
 [20150419d_h01s046_38]

- (491) *Ammei mu pus =ti ueili mi-sepin mi-nou*
 IMP.SG.come first IMP.SG.see =PART pig 3SG.NFUT-talk to-1SG.OBJ
 ‘First come and have a look at this pig talking to me’
 [20170222d_n01s152_70]

However, in switch-function SVCs, where the subject of the second verb is the object of the first verb, the second verb is marked for immediate future. This seems to reflect the fact that the object of the first verb is not being given an order, but instead undergoes a change or movement in the condition that the first verb happens: it seems to relate to the irrealis functions of the immediate future (§4.2.3).

- (492) *to-sa mat va-ha-ni xatel*
 IMP.PC-give mat 3SG.IFUT-go-TR 3PC
 ‘Give them the mat.’ [20150303j_x01e016_33]
- (493) *“Lui-ni haromue na-van va-mmei!”*
 IMP.SG.vomit-TR boy CL.GEN-1SG.POSS 3SG.IFUT-come
 “Vomit up my boy!” [20141105e_n01e012_29]
- (494) *lu-kes-i va-hat xai!*
 IMP.DU-call-3OBJ 3SG.IFUT-just_go MED
 ‘Call him away, this one!’ [20170217a_n01s038_24]

Similarly, the subsequent verb takes immediate future marking in event-argument SVCs:

- (495) *Kaakau va-hos/*hos!*
 2SG.IMP-walk 3SG.IFUT-good/*2SG.IMP.good
 ‘Walk well!’ [20180801a_x01s046]

In the prohibitive, there is a distinction between same-subject and other SVCs parallel to that found in the imperative, as imperative and prohibitive have similar illocutionary force: the prohibition is only being made to the addressee, not any objects moved or affected by the actions of the addressee. Verbs in same-subject SVCs are both marked prohibitive, while the subsequent verb in a switch-function SVC can be marked with immediate future.

- (496) *Ona-kaakau =ti an-he bien!*
 2SG.PROH-walk =NEG 2SG.PROH-go_to sea
 ‘Don’t walk to the sea!’ [20161025a_x01m004_26]
- (497) *On-has ueili xil =ti la-naa-met!*
 2SG.PROH-kill pig PL =NEG 3PL.IFUT-NEG-die
 ‘Don’t kill the pigs!’ [20180801a_x01s046]

The subsequent verb in a switch-function SVC can also be marked apprehensive, which seems appropriate: the result of a prohibition is an unwanted possibility. However, the apprehensive is formally identical to the prohibitive in all the person-number combinations possible in the defective paradigm (§5.1.1.3). The negative clitic *ti* only appears after the first verb, so the formal distinction between prohibitive and apprehensive mood marking is neutralised for the subsequent verb in an SVC. In examples like (498), the apprehensive subject prefix is identical to the prohibitive prefix. However, examples like (499) in the third-person plural unambiguously show the apprehensive prefix which follows the distant future prefixes.

(498) *muna-hur sup =ti na-he nesau!*
 PL.PROH-take chief =NEG 3SG.APR-go_to up
 ‘Don’t take the chief up’ 20150305h_h01o111_30

(499) *On-has ueili xil =ti li-na-met*
 2SG.PROH-kill pig PL =NEG 3PL.DFUT-APR-die
 ‘Don’t kill the pigs!’ [20180801a_x01s046]

I also gloss the subsequent verb in event-argument SVCs as apprehensive:

(500) *Muna-xeleial-ni xil =ti na-mak*
 PL.PROH-move-TR PL =NEG 3SG.APR-like_this
 ‘Don’t disturb them like this.’ [20141223a_s01s088_06]

There is also one example of a three-verb SVC in the prohibitive, showing the clitic *ti* appearing only once per construction. The initial verb is a loan from Bislama using the copular incorporation strategy (§3.5.2).

(501) *on-he mestem =ti na-mot na-mmei e tan*
 2SG.PROH-COP mistake =NEG 3SG.APR-fall 3SG.APR-come LOC ground
 ‘Don’t drop them [*breadfruits*] on the ground.’
 [20170331c_n01s140_44]

It is worth restating that while the morphological marking of TAMP on the verbs in an SVC can differ in the ways shown in Table 37, the SVC as a whole has only one value for all of the TAMP distinctions that are morphologically marked in Vatlongos. If a combination of verbs does not match the dependencies laid out in Table 37, they are contributing separate TAMP value and cannot be analysed as an SVC. For example, the collocation of *ketteh* ‘look’ and *pus* ‘see’ is

very common in SVCs ((441), (459), (466), (481)), but if a speaker only wants to negate the result of seeing, then zero coordination has to be used instead:

(502) *lu-ketteh* *lu-taa-pus* =*ti*
 3DU.NFUT-look 3DU.NFUT-NFUT.NEG-see =NEG
 ‘They looked, they didn’t see it.’ [20141208b_n01s044_11]

Similarly, event-argument SVCs cannot be used to express negated qualities of a situation; in such cases zero-subordination of a sentential subject is used (§7.2.2).

6.3.4 Comparison with non-contiguous SVCs in Paamese

The system of serialisation that most closely resembles Vatlongos SVCs is Paamese. Crowley’s (1987; 2002b) description of SVCs in Paamese has been influential in the analysis of SVCs in Oceanic languages, and it is worth summarising their behaviour to allow for detailed comparison with Vatlongos. Crowley (2002b: 50) hypothesises that SVCs in Vatlongos are ‘basically homologous’ to Paamese SVCs, but in fact there are interesting differences in the morphological dependencies between the two verbs, showing that SVCs can vary even between closely-related languages.

One difference between Vatlongos and Paamese non-contiguous SVCs is the possibility for constituents to intervene between the component verbs. In Paamese, the object of the first verb can intervene, and so can a PP at least in the case of locutive serialisation (Crowley 2002b: 79). A variety of clitics can also intervene (Crowley 2002b: 56), but as the clitics attach to the first intransitive verb in the construction, it is not possible to have both a clitic and an object between the component verbs. Many of these clitics are cognate with adverbs that can intervene in Vatlongos (Crowley 1982: 229–237), although in Vatlongos they do not form a single phonological word with the verb, and can appear elsewhere in the clause, such as after transitive verbs, (470), after an object NP, (447), or after a PP, (449). However, if these forms were more phonologically integrated in an earlier form of Vatlongos-Paamese, that could have been a bridging context for other, phonologically heavier, intervening adverbs in SVCs, like *vuteili* ‘a bit’ in example (448).

In Paamese non-contiguous SVCs, as in Vatlongos SVCs, the TAMP marking on the first verb provides the TAMP values for the entire construction, which in turn determines the marking on the subsequent verb which is not necessarily identical. Table 38 (Crowley 2002b: 57) shows that in affirmative polarity, the marking on the subsequent verb is identical to the initial verb, except in the imperative mood where the subsequent verb is immediate or distant future. Vatlongos shows similar patterns except for the distinction between same-subject and switch-function SVCs in the imperative mood: the subsequent verb in a same-subject imperative SVC is imperative rather than immediate future. In Paamese negative realis SVCs, like negative non-future SVCs in Vatlongos, the subsequent verb is marked for immediate future.

The main difference from Vatlongos, however, is the lack of negative marking on the subsequent verb in other negative polarity TAMP environments, neutralising polarity distinctions. There is therefore no basis for comparison with the behaviour of the negative clitic in Vatlongos, both because only the first verb can be marked as negative, and because the cognate form is a suffix rather than a clitic in Paamese (Crowley 1982: 140). Equally, Paamese does not have any equivalent to the Vatlongos category of prior, so no comparisons can be made there. Vatlongos apprehensive is equivalent to Paamese potential, but is incompatible with negative polarity. In fact Crowley's (1982: 140) grammar of Paamese states that Paamese potential is also incompatible with negative polarity, and unfortunately there are no examples of negated potential in his subsequent work on SVCs, so the formal and functional characteristics of the negated potential included in Table 38 are not clear.

Table 38: Mood-polarity dependencies in Paamese non-contiguous SVCs (Crowley 2002b: 57)

Mood of initial verb	Polarity of initial verb	Mood marking of subsequent verb	Polarity marking of subsequent verb
realis	+	realis	+
realis	-	immediate	+
immediate	+	immediate	+
immediate	-	immediate	+
distant	+	distant	+
distant	-	distant	+
potential	+	potential	+
potential	-	potential	+
prohibitive	+	potential	+
imperative	+	immediate/distant	+

6.4 Functions of SVCs

SVCs in Vatlongos are used for a wide range of functions, which could challenge attempts to define SVCs or circumscribe their semantic possibilities. A similarly wide range of functions has been described for non-contiguous SVCs in Paamese (Crowley 1987; 2002b), and many other Vanuatu languages.³⁹ However, in other languages different constructions take on a similarly wide range of functions, for example contiguous SVCs in Mwotlap (François 2004b; 2006) or echo-subject constructions in Southern Vanuatu languages like Erromangan (Crowley 2002b: 206–214).

This challenges the idea that SVCs are an iconic expression of a specific structural relationship between their component verbs. To illustrate this difficulty, this section is roughly organised from functions where there is a loose relationship between the component verbs, moving to functions involving a closer relationship. However, it is difficult to make an objective decision about the closeness of the semantic relationship between two verbs, especially when grammatical meanings are involved.

6.4.1 Discourse: focus and comments

There are some uses of SVCs in Vatlongos that could be analysed as discourse functions. The most clear-cut example of component verbs with different discourse status in Vatlongos is the use of the verbal question word *tep* ‘how’, to

³⁹ e.g. Daakaka (von Prince 2015), Lewo (Early 1994), Bierebo (Budd 2009), Apma (Schneider 2010), Tamambo (Jauncey 2011).

(506)	<i>bemei</i>	<i>be</i>	<i>telep</i>	<i>mi-li</i>
	3SG.NFUT.come	3SG.NFUT.COP	big	3SG.NFUT-beat
	<i>ngan</i>	<i>xa</i>	<i>diamu</i>	
	one	REL	3SG.NFUT.first	
	‘They get bigger than the one before.’ [20141106d_p04e016_08]			

6.4.2 Adjunct-like modification

There are many examples where the second verb is best described as an adjunct to the first verb, as it seems to add extra information or comment on the situation expressed by the first verb, rather than changing its interpretation or argument structure in a more fundamental way. The relationship between the first verb and the subsequent verb is more like that between a verb and an adverb than two components of a complex predicate.

This is very common in the interpretation of event-argument SVCs, and particularly with certain verbs that occur often in this position and have adverb-like semantics. These examples would be excluded from Haspelmath’s (2016) comparative concept of SVCs on the double basis that the subsequent verb is often stative rather than dynamic, and because of the possible predicate-argument relationship inherent in this pattern of argument sharing. However, it is interesting that the descriptive category of SVCs in Vatlongos is used for these functions as well as dynamic events, and that this is the case in so many serialising Oceanic languages (Cleary-Kemp 2015: 131). The typically adjunct-like behaviour of the subsequent verb in these constructions is also a challenge to Cleary-Kemp’s conclusion that SVCs must involve complementation in constituent structure.

In (507) and (508), as well as (475) above, the subsequent stative verb modifies the meaning of the initial verb in a way that resembles an adverb in other languages. Example (508) shows the expected morphological TAMP dependency and wide-scope of negation found in SVCs, whereas examples in affirmative polarity can be ambiguous with zero-marked subordination of a clausal subject.

- (507) *Toak mi-sepin mi-sap vuteili,*
 Toak 3SG.NFUT-speak 3SG.NFUT-different a_little
Edu mi-sepin mi-sap vuteili
 Endu 3SG.NFUT-speak 3SG.NFUT-different a_little
 ‘In Toak they speak a bit differently, in Endu they speak a bit differently.’ [20141107d_p01s022_10]
- (508) *na-taa-sukul =ti va-hos*
 1SG.NFUT-NFUT.NEG-study =NEG 3SG.IFUT-good
 ‘I didn’t do well at school.’ [20170224a_n01s141_72]

SVCs are also used to incorporate information about circumstances of place and time that are otherwise expressed by adjuncts. Location is often specified with verbs meaning ‘stay’ or ‘go’: for example, the locative existential verb *te* ‘be at’ as in (509), *ta* ‘stay’ in (514) below, or *he* ‘go to’ as in (479) above. These are ambiguous between an argument-location reading (where the subject of the locative verb is the subject or object of the first verb) or an event-location reading as an event-argument SVC.

- (509) *tatal xa val pat de vatit vak*
 snake REL big 3SG.NFUT.sleep 3SG.NFUT.be_at tree banyan
 ‘A big snake slept on the banyan tree.’ [20141106f_n01e018_11]

The temporal limits of a situation are often expressed with a subsequent verb *toxoh*, *toxol* ‘reach, touch’:

- (510) *da doxoh xosaliak*
 3SG.NFUT.stay 3SG.NFUT.reach today
 ‘It continues until today.’ [20141223b_h01s046_02]

The demonstrative verb *mak* ‘like this’ is common in this position, and specifies a manner that is either going to be immediately clarified by the discourse, (511), (512), or is provided by the extra-linguistic context, (513). The examples show the expected morphological dependencies in a variety of tenses, as does example (500) in the prohibitive. This demonstrates that this is not a case of juxtaposed independent clauses, zero subordination of clausal subjects, or an invariant adverb.

- (511) *Sisien nen ba mak*
 song of_it 3SG.NFUT.go 3SG.NFUT.like_this
 ‘The song goes like this’ [20141220d_n01s082_28]

(512) *na-miteni* *va-mak*
 1SG.IFUT-IFUT.say 3SG.IFUT-like_this
 ‘I’m going to talk like this.’ [20150419c_h01s004_54]

(513) *mei* *ral-tit* *i-mak*
 COME 1DU.DFUT-just_stay 3SG.DFUT-like_this
 ‘Let’s just come and stay like this.’ [20150223a_n01m096_59]

However, examples like (514) suggest that an invariant adverb form is also emerging, perhaps due to the high frequency of the zero-marked third-person singular non-future form, as in (511). Here the form *mak* is adjoined after the subsequent verb in an SVC specifying location.

(514) *ral-pol-ni* *stoa* *tava* *na-ralu*
 1DU.INCL.DFUT-do-TR store one CL.GEN-1DU.INCL.POSS
i-ta *e-n* *mak*
 3SG.DFUT-stay LOC-3SG like_this
 ‘We’ll make a store for ourselves in it like this.’
 [20170413e_n01m030_32]

The copular verb *he* with an adjective is also a common adjunct-like use of the SVC structure:

(515) *us* *i-mus* *i-he* *eilep*
 rain 3SG.DFUT-rain 3SG.DFUT-COP big
 ‘It will rain a lot.’ [20150419e_h01m128_55]

An SVC can also be used to give a cause, as in (516), when a family is saved by blocking a snake’s mouth with a bushnut. Here SVCs are alternatives to using a subordinate clause or discourse marker of causation.

(516) *rat-be* *laki* *gusil* *tavu* *xie*
 1PC.INCL-NFUT.COP lucky 3SG.NFUT.follow bushnut this
 ‘We’re lucky because of this bushnut.’ [20141220g_n01s080_57]

6.4.3 Prior motion

A very frequent use of SVCs in Vatlongos is to express the direction of motion prior to a situation expressed by the subsequent verb. These examples occur roughly once every 98 words in the subcorpus. The main two verbs used in this way are *ha* ‘go’, expressing direction away from the deictic centre, and *ammei* ‘come’, indicating direction towards the deictic centre. The deictic centre could be the place of speech, (517), fictive speech, (518), (519), or a place that is salient in the discourse, especially in narratives, (520).

- (517) *Inou na-mmei na-pus sista na-van tang igak*
 1SG 1SG.NFUT-come 1SG.NFUT-see sister CL.GEN-1SG.POSS just here
 ‘I’ve just come to see my sister here.’ [20170412a_n01m029_05]
- (518) *muli-ha muli-lau e rin ak*
 2DU.DFUT-go 2DU.DFUT-hunt LOC side PROX
 ‘You must go hunt on this side.’ [20141220g_n01s080_11]
- (519) *ral-ha mu ral-sao-ni nahou*
 1DU.INCL.IFUT-go first 1DU.INCL.IFUT-look-TR garden
sa-ralu
 CL.DOM-1DU.INCL.POSS
 ‘Let’s first go and look for a garden for the two of us.’
 [20141027a_n01m001_25]
- (520) *vevu bemei pus-i*
 rail 3SG.NFUT.come 3SG.NFUT.see-3OBJ
 ‘the rail came and saw it’ [20170331c_n01s140_60]

In Paamese, this type of motion is usually expressed with ‘go’ and ‘come’ verbs as the matrix verb in subordination structures, taking a purposive clause. Crowley (2002b: 64–65) shows that the subordinate clause can be introduced with a subordinator, and that the realis affirmative is followed by the immediate future, which is also preferred for purposive clauses in Vatlongos (§4.2.3).

However, Vatlongos examples like (520) show matching non-future affirmative marking on the two component verbs, and that the SVC expresses prior motion without forcing a purposive interpretation, as the rail sees something unanticipated in the narrative (the rat being pulled out its hole by a dwarf). Similarly, the prior motion SVC in (521) has an inanimate subject, making a purposive interpretation unlikely.

- (521) *vatiang bemei gur rat-i*
 wind 3SG.NFUT.come 3SG.NFUT.take+out-3OBJ
 ‘The cyclone came and took it out.’ [20150419c_h01m004_25]

Example (522) is stronger evidence that these constructions in Vatlongos are true SVCs, showing the morphological TAM dependency and behaviour of the negative clitic expected for the negative future tenses. If the second verb were a separate clause, it would be ungrammatical as it takes a negative prefix without the negative clitic *ti*. A subordinate clause with this intended meaning would be marked for affirmative immediate or distant future, (283), (Table 19), but the SVC or an AVC is strongly preferred with ‘go’ and ‘come’.

- (522) *mut-naa-va* =*ti* *mut-naa-pat?*
 2PC.DFUT-NEG-go =NEG 2PC.DFUT-NEG-sleep
 ‘Aren’t you going to go and sleep?’ [20141028a_c01m002_14]

These examples are evidence for temporal iconicity as an ordering principle in Vatlongos SVCs, in line with cross-linguistic tendencies (Aikhenvald 2006: 21; Cleary-Kemp 2015: 150; Haspelmath 2016: 309). Comparing these prior motion examples with the subsequent motion examples in the next section shows that the order of verbs reflects the temporal order of the events they express, in so far as they are ordered in time, as with the two ‘go’ verbs in the three-verb SVC in (523). When the basic motion verbs are used to specify concurrent direction they occur after the verb, following head-initial constituent-ordering in the Vatlongos VP.

- (523) *lata-ba* *lat-go* *voxoli ak lata-ba*
 3PC.NFUT-NFUT.go 3PC.NFUT-NFUT.pass canoe PROX 3PC.NFUT-NFUT.go
 ‘They went and travelled off in the canoe.’ [20170222f_n01s153_16]

It is difficult to know whether to categorise these as a loose or tightly-bound relationship between the two verbs. On the one hand, they seem more loosely-bound than the basic motion verbs used to add directional information to a preceding verb. They tend to modify activity verbs rather than motion events, as observed by Huang & Tanangkingsing (2005: 318), who exclude such examples from their study of motion events. Because these constructions are translated into English with coordinators ‘go and’ or subordinators ‘go to’ they are often viewed as loosely-bound (Crowley 2002b: 11).

On the other hand, crosslinguistically, constructions used for this meaning are frequently grammaticalized into aspectual or future meanings (for example, English *going to*, *gonna*, *go and*) or seem to violate expectations about clause boundaries (for example, questions in English like *Who are you going to go and visit?* which violate the coordinate structure constraint (Ross 1967: 161)). In Australian languages the category of associated motion, including prior motion, is often marked morphologically, showing extreme grammaticalization (Koch 1984). Cleary-Kemp (2015: 134) argues that SVCs with an initial motion verb in Oceanic languages, which Lynch et. al (2002: 47) describe as ‘sequential’ SVCs, should also be analysed as associated motion, as they always express motion

and when it happened relative to the action. In fact, we could be more specific and call these prior motion SVCs, as in all the examples the motion occurs before the action.

Evidence within Vatlongos suggests that the two verbs have a tight semantic link and that prior motion is a highly grammaticalized concept. Firstly, their very high frequency suggests that in Vatlongos there is a strong convention of specifying any prior motion. Secondly, the verb *ammei* ‘come’ is also used for the aspectual category of inchoative (§6.4.10). Thirdly, ‘go’ and ‘come’ verbs in this construction appear to be the diachronic sources for phonologically reduced auxiliary verbs (Chapter 8), with very similar semantic functions as in these SVCs. Prior motion SVCs are also exceptional in their behaviour with these auxiliaries (§8.1.3).

Another group of verbs which are used in prior motion SVCs are items meaning ‘get up’. Using one of these seems to convey suddenness or effort in performing the action of the subsequent verb, and they are especially common in child-directed narratives (see also (655)).

(524) *vevu mi-temea mi-ga*
 rail 3SG.NFUT-get_up 3SG.NFUT-NFUT.fly
 ‘The rail got up and flew’ [20170222d_n01s152_68]

(525) *lat-tammea lat-go voxoli tei*
 3PC.NFUT-get_up 3PC.NFUT-NFUT.pass canoe one
 ‘They got up and went in a canoe.’ [20170222f_n01s153_15]

(526) *Ma na-tammea na-be aplae*
 so 1SG.NFUT-get_up 1SG.NFUT-NFUT.COP apply.BIS
 ‘So I went and applied’ [20170406a_n01m164_23]

6.4.4 Directional modification

A common function of SVCs cross-linguistically is to add directional or path information to a predicate (Aikhenvald 2006: 22–23), and this is frequent in Vatlongos. The initial verb often expresses manner of motion or causation of motion, while the subsequent verb is usually one of the basic motion verbs, perhaps combined with a preposition to give more information about the path. Example (527) shows an intransitive manner of motion verb modified by a basic motion verb in a same-subject SVC. Examples (468), (469) and (498)

show causation of motion combined with basic motion verbs in switch-function SVCs.

- (527) *mi-kaakau* *bemei*
 3SG.NFUT-walk 3SG.NFUT.come
 'He walked over here.' [20150118a_n01m090_44]

While these examples all express direction of motion concurrent with the main verb, it is also possible for the directed motion to follow the action of the initial verb in switch-function SVCs, as in (528), where the path followed by object of the first verb as a result of the action is specified by the subsequent verb.

Examples like this also suggest an iconic ordering of verbs in Vatlongos SVCs.

- (528) *David mi-sip rat* *futbol bemei* *visel Jemes*
 David 3SG.NFUT-kick+out football 3SG.NFUT.come to Jemes
 'David kicks the football over to Jemes.' [20141215f_k01s026_13]

Again, it is difficult to determine the closeness of the relationship between the two verbs in these constructions. Sometimes the subsequent verb seems to be adding extra information, like an adjunct that could alternatively be expressed by an adjoined PP. On the other hand, it would often be pragmatically infelicitous to use a manner of motion verb alone without the serialised basic motion verb, as Crowley (2002b: 67) observes for Paamese. Cross-linguistically, the co-predicate status of directional verbs in SVCs has been argued for using the ability to translate a whole SVC with a single verb (although the problems with this approach have also been pointed out, Aikhenvald 2006: 4; Defina 2016c: 893). For example, the combination of 'take' and 'come' verbs are translated with 'bring' in examples (469) and (498).

Non-literal path information showing the direction of a non-motion event can be expressed with an event-argument SVC, as seen in the third-person singular agreement marking on the subsequent verbs modifying the verb *ketteh* 'look':

- (529) *La-ketteh* *bemei* *vari* *bien*
 3PL.NFUT-look 3SG.NFUT.come at_once sea
 'They looked over at the sea' [20141117a_n01m003_11]
- (530) *lu-ketteh* *be* *nahou*
 3DU.NFUT-look 3SG.NFUT.go_to garden
 'They looked at the garden.' [20141027a_n01m001_028]

When the locative verb *he* ‘go to’ is used, it is also difficult to decide whether the main function of the SVC is to add path information or to introduce a goal through argument merger, suggesting a closer relationship between the component verbs.

6.4.5 Direct quotations

A very frequent function of SVCs in Vatlongos narratives is to incorporate direct quotations. This is also common in other Vanuatu languages, including Paamese, where Crowley (2002b: 79) describes it as ‘locutive serialisation’.

The most typical use of this structure is a same-subject SVC with two verbs from a set of related lexemes: *hiteni* ‘say’ which takes the listener as an oblique argument in a PP with *mi* ‘to’; and *hit* or *hitene* followed by the quoted speech.

- (531) *lu-biteni* *mi* *tati* *na-lu*
 3DU.NFUT-NFUT.say to dad GEN.CL-3DU.POSS
lu-bit [*“Tati! ...*
 3DU.NFUT-NFUT.say dad
 ‘They said to their Dad “Dad! ...’ [20141220g_n01s080_36]

Although this structure most often appears in narratives in the non-future, where it is ambiguous with a multiclausal analysis, in elicited examples the morphological dependencies shown in Table 37 apply:

- (532) *xi* *taa-viteni* =*ti* *mi-ni* *va-hit...*
 3SG 3SG.NFUT.NEG-NEG.say =NEG to-OBJ 3SG.IFUT-say
 ‘He didn’t say to him that...’ [20170401c]
- (533) *xi* *te-naa-hiteni* =*ti* *mi-ni* *te-naa-hit...*
 3SG 3SG.PRI-NEG-say =NEG to-3OBJ 3SG.PRI-NEG-say...
 ‘He hadn’t said to him...’ [20170401c]

Other verbs specifying a manner of speech can also be the initial verb in these constructions, such as the transitive verb *kes* ‘call’:

- (534) *mi-kes-i* *bit* *“Mael, ...*
 3SG.NFUT-call-3OBJ 3SG.NFUT.say Mael
 ‘He called him saying “Mael, ...’ [20170221e_n01o150_10]

The less frequent verb of speech *xa* ‘say’ is also used as the second verb in these constructions.

- (535) *Ma lu-biteni mi-ni lu-ga,*
 then 3DU.NFUT-tell to-3OBJ 3DU.NFUT-NFUT.say
 “*ma-bit mal-hur nep na-m...*”
 1DU.EXCL.NFUT-NFUT.want 1DU.EXCL.IFUT-take knife CL.GEN-2SG.POSS
 ‘Then they said to her, “We want to take your knife...”
 [20141027a_n01m001_74]

6.4.6 Complement clauses

The ‘say, want’ verbs *hit* and *hitene* can also be used to introduce complement clauses, a function which has probably extended from their role in incorporating direct speech. In this role they seem to act as verbal complementisers within SVCs: they take a zero-marked complement clause, which is the complement of the initial verb in the SVC. In these functions, SVCs can be used instead of alternative complementation strategies such as PPs and overt or zero-marked subordination (§7.2.1).

Examples (223) and (536) show *hit* being used to introduce indirect speech, where the use of the third-person singular for the narrative-internal addressee or speaker shows that the perspective of the narrator is maintained.

- (536) *mama nan biteni bit*
 mum her 3SG.NFUT.say 3SG.NFUT.say
 [*naa-kil =ti naa-ve vatit vak*]
 3SG.FUT.NEG-dig =NEG 3SG.FUT.NEG-NEG.go_to tree banyan
 ‘Her mother told her she shouldn’t dig under the banyan tree.’
 [20141106f_n01e018_19]

They can also be used for thoughts rather than speech. In (538) *hit* is serialised to a verb meaning ‘see’, and the complement clause expresses thoughts deduced from visual evidence, in this case fake blood put in the ocean to allay suspicions when someone has been kidnapped to be eaten.

- (537) *mi-nnem-i bit [kuli san i-met]*
 3SG.NFUT-think-3OBJ 3SG.NFUT.say dog his 3SG.DFUT-die
 ‘He thought his dog would die.’ [20141212h_n01s046_21]
- (538) *li-pus-i li-hit [saak xa ga moletin ok]*
 3PL.DFUT-see-3OBJ 3PL.DFUT-say shark SUB 3SG.NFUT.eat person PROX
 ‘They would see it and think it was a shark that ate the person.’
 [20150228g_n01s100_10]

These verbs can also be used to incorporate complements of verbs of desire, including psycho-collocations (§3.4.2), followed by a complement clause in the

immediate future expressing the (un)desired state of affairs. This retains the polysemy of *hit* and *hitene* as verbs of both speech and desire when used independently in monoverbal clauses.

- (539) *taa-vit* =*ti* *va-hitene*
 3SG.NFUT.NEG-want =NEG 3SG.IFUT-want
[va-pol kuhi-ni *na* *wido* *xil* *ak]*
 3SG.IFUT-do+careful-TR um window PL PROX
 ‘He didn’t want to do these windows properly.’
 [20150419e_h01m128_26]
- (540) *e-n* *bei* *bitene*
 inside-3SG.POSS 3SG.NFUT.like 3SG.NFUT.say
[na-ma *na-mue* *neta,*
 1SG.IFUT-IFUT.go 1SG.IFUT-do_something what
ma *na-me* *tij]*
 GO.1SG.IFUT 1SG.IFUT-COP teach
 ‘She wanted me to um, to go teach’ [20170406a_n01m164_56]

6.4.7 Increasing valency

SVCs can also be used as a valency changing device, adding a new slot to the argument structure of the predicate expressed by the initial verb. The examples in this section show SVCs being used to introduce arguments of various semantic roles. Each semantic role is associated with a specific verb, so it is possible that these verbs have separate lexical entries from their use as main verbs in monoverbal clauses.

Goal arguments which are locations are introduced as the locative object of *he* ‘go to’.

- (541) *xil* *la-loh* *la-be* *Sahuot*
 3PL 3PL.NFUT-run 3PL.NFUT-NFUT.go_to Sahuot
 ‘They ran to Sahuot.’ [20150226a_n01s098_20]
- (542) *mi-til-ni* *vati-n* *vari* *be* *oei*
 3SG.NFUT-stick-TR head-3SG.POSS at_once 3SG.NFUT.go_to water
 ‘He stuck his head into the water at once.’ [20141212h_n01s046_70]

Human goals are instead incorporated with the basic motion verb *go* ‘ha’ marked with the transitivity marker *-ni*:

- (543) *inou* *na-loh* *na-ba-ni* *tuvava*
 1SG 1SG.NFUT-run 1SG.NFUT-NFUT.go-TR baby
 ‘I ran to the baby’ [20150415a_h02s125_04]

Transitivised *ha* ‘go’ is also used for interlocutors, including as a strategy for incorporating interlocutors with novel loan verbs from Bislama such as *teks* ‘text’ in (545).

- (544) *mi-kei* *ba-ni*
 3SG.NFUT-call 3SG.NFUT.go-TR
 ‘He called to him’ [20141220d_n01s082_20]
- (545) *inou ni-he* *teks i-ha-ni* *moletin ak*
 1SG 1SG.DFUT-COP text 3SG.DFUT-go-TR person PROX
 ‘I’ll text this person.’ [20141127c_x01m043_47]

A stimulus of *teteong* ‘listen’ can also be added with transitivised *ha* ‘go’.

- (546) *ma-tteong* *ba-ni* *retio*
 1PL.EXCL.NFUT-listen 3SG.NFUT.go-TR radio
 ‘We listened to the radio.’ [20150226a_n01s098_79]

The verb *tel* ‘accompany, with’ can be used to introduce a concomitant or an instrument. Example (547) is an event-argument SVC describing how the village will eat with the chiefs on Chief’s Day.

- (547) *mati-anien* *i-tel* *xatel*
 1PL.DFUT-eat 3SG.DFUT-with 3PC
 ‘We will eat with them.’ [20150305f_p01s110_30]

In (548) *tel* introduces an instrument into the action of cutting.

- (548) *John dei* *vatiei del* *akis*
 John 3SG.NFUT.cut tree 3SG.NFUT.with axe
 ‘John cut the tree with an axe.’ [20150225a_x01s046_05]

6.4.8 Oblique arguments

As well as increasing the valency of a predicate, the subsequent verb in an SVC can also incorporate oblique arguments already selected by the initial verb, as an alternative to PPs (§3.2). As no verbs in Vatlongos obligatorily select an oblique argument, it can be difficult to determine whether a verb selects an oblique argument, and therefore whether an SVC is increasing valency or merging arguments. The examples in this section all involve verbs that can take a PP argument with the same semantic role, and are common three-place predicates cross-linguistically. Pawley (1973: 141) argues that proto-Oceanic probably used verbs to mark semantic roles (case relations).

The incorporated-preposition verbs (§3.2.1) *he* ‘go to’, *haen* ‘go to it’, *te* ‘stay at’ are used to incorporate goals. This allows for semantic distinctions that are not possible using the general locative preposition *e* they derive from.

- (549) *o-ling-i* *be* *upang*
 2SG.NFUT-put-3OBJ 3SG.NFUT.go_to fire
 ‘You put it on the fire.’ [20141106d_p06e016_06]
- (550) *di* *lu-so* *vueili* *baen*
 CONT.REAL 3DU.NFUT-put pig 3SG.NFUT.go_to_it
 ‘They were putting pork in it.’ [20141027a_n01m001_088]
- (551) *lu-ling-i* *de* *tim* *sa-lu*
 3DU.NFUT-put-3OBJ 3SG.NFUT.stay_at home DOM.CL-3DU.POSS
 ‘They put it in their home.’ [20141027a_n01m001_062]

A recipient can be introduced with transitivised forms of either *ha* ‘go’ or *ammei* ‘come’.

- (552) *Kenet mi-sa-ni* *ba-ni* *tutut Jon*
 Kenet 3SG.NFUT-give-3OBJ 3SG.NFUT.go-TR little Jon
 ‘Kenet passed it to little Jon’ [20141215f_k01s026_03]
- (553) *xamem mate-sa* *anien te-ha-ni* *xamim*
 1PL.EXCL 1PL.EXCL.PRI-give food 3SG.PRI-go-TR 2PL
 ‘We gave the food to you.’ [20141212i_x01s046_32]
- (554) *la-sa* *anien bemei-ni* *xir*
 3PL.NFUT-give food 3SG.NFUT.come-TR 1PL.INCL
 ‘They gave us food.’ [20141212i_x01s046_35]

The option to use either of these verbs allows more path information to be included in the SVC than is possible using the preposition *mi* ‘to’. However, there is no difference in the relationship between the verb *sa* ‘give’ and the recipient in these examples, than in examples like (36) and (37) where the recipient is expressed with a PP.

In (555) a corporate interlocutor referred to with a place term is introduced with locative *he* ‘go to’.

- (555) *ma-sepin* *be* *peis*
 1DU.EXCL.NFUT-speak 3SG.NFUT.go_to base
 ‘We talked to the base.’ [20170224a_n01s141_133]

6.4.9 Sub-events within complex events

The closest lexical semantic relationship between the component verbs in an SVC is when each verb contributes a sub-event to a complex event. This function is more closely associated with the complex predicates described in §7.1, but there are still many examples of this relationship between verbs in an SVC.

Example (556) is a same-subject SVC where the first verb expresses an action and the second verb expresses the result, or intended result, of that action. The frequent collocation of *ketteh* ‘look’ and *pus* ‘see’ is one example of this relationship (e.g. (441), (459), (466), (481)), but (556) demonstrates that SVCs are also used productively for more unusual combinations of action and result.

(556)	<i>tatal</i>	<i>ak</i>	<i>bit</i>
	snake	PROX	3SG.NFUT.want
	[va-ngang	va-a	xamel]
	3SG.IFUT-open_mouth	3SG.IFUT-eat	1DU.EXCL
	‘This snake wanted to open its mouth and eat us.’		
	[20141220g_n01s080_41]		

This relationship is also found in switch-function SVCs. *has* ‘hit, kill’ and *met* ‘die’ are a frequent collocation (e.g. (436)-(438)), but (557) shows a more creative combination of action and result.

(557)	<i>mi-salele-ni</i>	<i>be</i>	<i>eilep</i>
	3SG.NFUT-flat-TR	3SG.NFUT.COP	big
	‘She flattens it out wide.’ [20141212g_p01s046_20]		

6.4.10 Inchoative aspect

One aspectual distinction is marked by SVCs in Vatlongos: the inchoative is expressed by using *ammei* ‘come’ as an initial verb, followed by a stative verb, or often the copular verb *he*.

(558)	<i>hariken</i>	<i>bemei</i>	<i>geih</i>
	hurricane	3SG.NFUT.come	3SG.NFUT.strong
	‘The hurricane got strong.’ [20170222h_h01o154_14]		
(559)	<i>bemei</i>	<i>be</i>	<i>venu</i>
	3SG.NFUT.come	3SG.NFUT.COP	volcano
	‘It turned into a volcano.’ [20170222d_n01s152_102]		

The resulting dynamic predicate can be modified by process adverbs such as ‘quickly’:

(560) *La-mmei* *la-be* *talep* *kamanon* *tavatang*
 3PL.NFUT-come 3PL.NFUT-NFUT.COP big quickly very_much
 ‘They get big very quickly.’ [Y2R4_11.2]

6.4.11 Deontic modality

Finally, a single modal distinction is marked with an SVC: obligation, with the Bislama loan verb *mas*. *mas* appears as the initial verb in the SVC. Cleary-Kemp (2015: 150) observes that imperfective and modal constructions tend to have the minor verb first and suggests this could be inherited from Proto-Oceanic or reflect universal principles.

(561) *li-mas* *li-lang-ni* *vuei*
 3PL.DFUT-must 3PL.DFUT-look_for-TR pig
 ‘They have to look for pigs.’ [20150219a_p01m001_26]

6.5 Discussion

Vatlongos SVCs are an idiosyncratic morphosyntactic strategy characterised by a complex set of dependencies of morphological TAMP marking, and a consistent restriction on the occurrence of the negative clitic. This single strategy is used for a wide variety of functions, involving event and argument structure, grammatical categories, and discourse organisation, and ranging from tightly-bound semantic relationships between the component verbs to much looser semantic configurations. Although similarly wide extensions of particular serialisation strategies have been described for many individual languages, it is important to notice how this challenges dominant theoretical and typological frameworks for analysing these constructions.

The RRG distinction between nuclear and core layer of juncture in SVCs is the prevailing terminology for describing SVCs in Oceanic languages, but this framework involves a specific theory about the interaction of syntactic and semantic structure that makes falsifiable predictions about how operators and semantic relationships relate to the juncture and nexus relations in complex constructions. Descriptively, these terms are usually employed on the basis of how component verbs SVCs relate to the Oceanic verbal complex, especially whether both verbs are marked with the prefixes or proclitics that typically express subject-indexing and TAMP, and the suffixes and enclitics that typically express object-indexing (Ross 2004: 495). However, this morphosyntactic distinction does not necessarily align with the RRG framework’s predictions

about how other syntactic units and operators relate to these terms. In Vatlongos, for example, non-core arguments (PPs) and adjuncts (PPs and adverbs) can intervene between the two verbs, which is not predicted for core-layer cosubordination in the RRG framework, as these constituents belong in the periphery. For this reason, distinctions based on typological parameters such as contiguity, or single-marking versus multiple-marking (e.g. in von Prince’s (2015: 308) description of Daakaka serial predicate constructions) may be a better framework for description.

The RRG interclausal relations hierarchy (Table 39) predicts that the closer semantic relations are expressed by structures with stronger syntactic relations and vice versa (although it should not be interpreted as making any straightforward predictions about how a particular semantic relationship is expressed, rather that within a particular language, stronger syntactic relations are used for closer semantic relations overall).

Table 39: RRG Interclausal Relations Hierarchy (Van Valin & LaPolla 1997: 481)

Syntactic relations	Semantic relations
<i>Strongest</i>	<i>Closest</i>
	Causative
Nuclear cosubordination	Aspectual
Nuclear subordination	Psych-action
Nuclear coordination	Purposive
	Jussive
Core cosubordination	Direct Perception
Core subordination	Propositional attitude
Core coordination	Cognition
	Indirect discourse
Clausal cosubordination	Conditional
Clausal subordination	Simultaneous states of affairs
Clausal coordination	Sequential states of affairs
	Unspecified temporal order
<i>Weakest</i>	<i>Loosest</i>

A separate issue with testing this hypothesis is the difficulty in assessing the ‘closeness’ of semantic relations, especially when comparing areas of grammatical meaning to event semantics. Brill (2007) surveys juncture and nexus relations in complex predicates in Oceanic languages and finds a slightly different cline from the RRG model, shown in Table 40. Most notably, adverbials of time and location are found to be a close semantic relationship, even though

these types of information are predicted to appear in the periphery of the RRG clause.

Table 40: Cline of semantic relations expressed by complex predicates in Oceanic languages (Bril 2007: 302)

Semantic relations
<i>Closest</i>
Modification (manner, direction) or Adverbial (time, location)
Causative
Aktionsart
Purposive
Overlapping actions
Sequential actions
Cause-result
Modality
<i>Loosest</i>

On the other hand, Næss (2012) investigates the use of a different syntactic constructions in Äiwoo in a narrow semantic domain: complex events of cutting and breaking (Majid et al. 2007). Within this domain, semantic integration of event structure can be assessed relatively objectively, allowing for a more robust approach to the relationships between semantic and syntactic integration. She finds a correlation between the morphological integration of constructions and the interdependence of subevents in the complex events they describe.

The relationship between syntactic and semantic integration, and the hypothesis of structural iconicity are interesting empirical questions, and we need to ensure that our terminological framework does not obscure these relationships, nor artificially confirm the hypothesis. This is a problem for typological approaches to defining and describing SVCs. Haspelmath's (2016) definition of SVCs as a comparative concept is necessarily function-based to allow for comparisons across typologically diverse languages. However, if SVCs are excluded from the definition as soon as they extend into new functions, then these categories cannot be used to address the question of structural iconicity without circularity: the syntactic relation of verb serialisation is used to express a narrow range of semantic relations, because anything expressing a closer or looser semantic relation is not an SVC.

Evidence from Vanuatu languages suggests that the morphosyntactic structure of a construction places few constraints on its ability to describe different types of semantic and functional relationships. Many languages, including Vatlongos and Paamese (Crowley 1987; 2002b), use non-contiguous SVCs to express a wide range of semantic relationships between component verbs. On the other hand, Mwotlap (Banks, Northern Vanuatu), uses contiguous SVCs for a similar array of functions (François 2004b; 2006), while languages of Southern Vanuatu instead use echo-subject constructions for similar purposes (Crowley 2002b: 206–214). It seems that a favoured morphosyntactic strategy can extend into new semantic domains regardless of the syntactic relationship between its constituents.

7 Other strategies for combining verbs

This chapter discusses two other strategies for combining verbs in Vatlongos: root-contiguous complex predicates and subordination.

In root-contiguous complex predicates, two or more elements, usually verbal, are combined to form a single syntactic verb with one set of affixes and a single argument structure. Complex predicates could be analysed as a kind of SVC, and similar constructions in other Vanuatu languages have been described as nuclear-layer SVCs (Crowley 1987; Early 1993; Hyslop 2001; Crowley 2002b; Bril 2004; Budd 2009; Schneider 2010; Jauncey 2011). However, I argue that the lack of productivity in possible combinations, and the tendency to reanalyse component verbs as non-verbal or less-verbal elements, or to lexicalise a combination of verbs in a complex predicate, is evidence that this is not (or is no longer) a serialisation strategy. In fact, it may be closer to a system of productive verbal compounding, as has been argued for nuclear-layer serialisation in the region (Meyerhoff 2000: 170).

I use the term ‘complex predicate’ because in many descriptive frameworks this is a more general term that subsumes verb serialisation, but does not make a commitment to the verbal status of the elements involved (Bril & Ozanne-Rivierre 2004; Bril 2007). In other frameworks, ‘complex predicate’ is used to refer to a more specific subset of combinations of predicates, whereby argument structures are merged or fused and individual elements do not take separate TAM marking (Butt & Seiss 2009: 12; Seiss 2009), and the Vatlongos construction also fits this definition. In this tradition, complex predicates can be combined either in the syntax (like SVCs) or in the lexicon (via productive compounding), so this terminology also allows for an analysis as compounding.

In morphosyntactic terms, subordination is usually defined as the embedding of one clause in another. Lehmann (1988: 2) defines subordination as a form of clause linkage whereby clauses are linked in an endocentric construction where one clause is the head of the construction and the other is embedded within it. However, in typological work, functional definitions are preferred, to avoid language-internal criteria that may be restricted to certain

language families or typological profiles. Cristofaro (2005: 2) defines subordination as a cognitive relation between two clauses, one of which is non-autonomous and is construed from the perspective of the other.

As this chapter focuses on combinations of verbs, rather than clause structure more generally, I only discuss subordinate clauses which function as an argument to the verb in the matrix clause. This covers complement clauses, which are selected by the verb, and clausal subjects, which are not restricted by the verb's argument structure. Vatlongos also has many subordinating discourse markers which can mark a range of semantic relationships between a main and subordinate clause. Zero subordination is also used for subordinate clauses which are an adjunct of the main clause, though this can be difficult to distinguish from coordination, which is also frequently zero-marked in Vatlongos.

7.1 Root-contiguous complex predicates

These constructions usually involve two verb roots which are directly contiguous, without any intervening constituents or affixes. Grammatical categories are marked only once, with prefixes occurring on the first verb and suffixes on the final verb. As such, the multiple roots are treated as a single grammatical word in terms of concatenative morphology and syntax: they have the same syntactic distribution as single verbs. However, the component verbs can undergo reduplication and verb-initial consonant mutation individually, suggesting that they remain separate words for non-concatenative morphology. They are also separate prosodic words. Most examples are asymmetrical, with the subsequent verb usually expressing adverbial modifications of the first verb, although complex-event examples could be argued to be symmetrical within the confines of possible semantic and argument structure combinations (Aikhenvald 2006: 29).

- (562) *mi-kas moten-i*⁴⁰
 3SG.NFUT-wash+clean-3OBJ
 ‘They wash it clean.’ [20141216h_p01s046_06]
- (563) *di o-moul moten-i*
 CONT.REAL 2SG.NFUT-weed+clean-3OBJ
 ‘You weed it clear.’ [20141106d_p10e016_03]

Examples (562) and (563) show a result or purpose expressed by the second verb, the transitive *moten* ‘clean’, combined with different action verbs in the first slot. All these verb roots can occur independently as main verbs, although in the semantic domain of purposeful activity it is usually pragmatically felicitous to specify both an action and an intended result by using a complex predicate. They show a single set of subject-indexing and TAMP prefixes occurring on the initial verb, and a single set of object suffixes occurring on the final verb. Example (563) also demonstrates that the construction only takes a single auxiliary verb: both fall under the scope of the continuous auxiliary that precedes the first verb.

Similar constructions occur in many languages of the region and are often described as nuclear-layer SVCs.⁴¹ However, as discussed in §6.1 in relation to SVCs, I avoid committing to an analysis based on layer of juncture in the RRG clause model. I am therefore following a typological approach by describing these constructions as ‘root contiguous’. Other linguists may analyse these forms as compounds, formed in the lexicon and/or morphology rather than in the syntax (Meyerhoff 2000: 170). In North Ambrym (Franjeh 2012: 133), as well as other Vanuatu languages like South Efate (Thieberger 2004: 229–238), a synchronic system of verbal compounding is described, and the possibility that it has arisen from an earlier system of serialisation is discussed. The term ‘complex predicate’ allows for analyses as compounding as well as serialisation.

⁴⁰ Complex predicates are written as separate orthographic words but are aligned as a single unit and glossed with a boundary marked + to reflect their status as single grammatical words. In section §7.1 the complex predicate is marked in **bold**.

⁴¹ e.g. Paamese (Crowley 1987; Crowley 2002b); languages of Epi (Early 1993; Budd 2009: 226); Apma (Schneider 2010: 187–8); North-East Ambae (Hyslop 2001: 282–91); Malo (Jauncey 2011: 345–53).

- (569) 'mi.sa ko.'xo.le
mi-sa koxol-e
 3SG.NFUT-give+block-3OBJ
 'He blocks it [*the football*].' [20141215f_k01s026_36]

There are a few examples where three elements are included in a complex predicate: they remain a single grammatical word, so no constituents can intervene between the component elements and the middle element does not take any affixes. In most of these examples, a complex event or directional meaning is expressed by the first two verbs, which are then in turn modified by an adverb or verb with adverbial semantics, suggesting the hierarchical structure indicated by square brackets in (570).

- (570) *Vatingah* *ba,*
 spear 3SG.NFUT.go
[[*mi-sang xat*] *vari-ni*] *mete-n* *xalu*
 3SG.NFUT-straight+right+at_once-TR eye-3SG.POSS DU
 'The spear went, it went straight into its two eyes.'
 [20170220i_n01e149_29]

7.1.2 Functions of complex predicates

The most productive function of the complex predicate structure is for semantically complex events involving an action and a result. In some cases, like (562) and (563) above, the result is best thought of as a purpose. The result does not have to be successfully effected for the complex predicate to be applied to a situation. In such cases, the complex predicate denotes a purposeful activity, rather than an event with two subevents. The strong tendency to state both the action and result of complex events, without entailing that the result is effected, resembles how a dedicated complex-event strategy is used in Äiwoo (Solomon Islands, Oceanic), where forms that do not function as independent verbs are productively combined to express complex events. Unlike true serialisation in the language, these are used even when 'the expected result' does not occur (Næss 2012: 411). Vatlongos complex predicates are also relatively common in the semantic domain of cutting and breaking events (Majid et al. 2007).

Examples (571) to (574) show how action and result verbs can be productively combined to describe events. The manner verbs *tei*, *ta* 'cut' and *tii* 'bash' are combined with two different result verbs which encode certain effects

on the object: *vul* is used when long objects are broken across their short axis, while *pas* is used to describe splitting a long object along its length, or breaking a more general class of objects, especially anything round.

- (571) *mi-da pas* *vesehi*
 3SG.NFUT-NFUT.cut+break pawpaw
 ‘She cuts the pawpaw.’ [20141127d_x02s043_08]
- (572) *dii pas* *klas e simen*
 3SG.NFUT.bash+break glass LOC cement
 ‘She smashes the glass on the cement.’ [20150303b_x02e016_14]
- (573) *inou na-da vul* *ap ak*
 1SG 1SG.NFUT-NFUT.cut+break_across firewood PROX
 ‘I cut this firewood.’ [20141127c_x01m043_39]
- (574) *mi-dii vul* *liei ni hat*
 3SG.NFUT-NFUT.bash+break_across wood INSTR stone
 ‘She breaks the wood with a stone.’ [20141127d_x02s043_33]

Complex predicates are also used to combine a main verb with a subsequent verb marking direction or position. The intransitive verbs of basic motion, which are very frequent in SVCs, do not appear in this position in complex predicates. Instead, most of the directional verbs which occur here are transitive, and denote the path or position of the object that results from the action expressed by the first verb. As such, this function is semantically adjacent to the action-result combinations. Rather than causing a change of state of the object, the action denoted by the first verb causes a change of location.

Verbs expressing direction in this position seem to be bleached of some of the lexical meaning they have as independent verbs. Many are reduced to expressing only the directional element of their semantics as an independent verb. For example, *husil* ‘follow’ expresses a general directional meaning of ‘along’ and *xol* ‘cover’ has a more general interpretation of ‘over’ or ‘in front of’. Often the directional function in complex predicates becomes the primary use of these lexemes. In complex predicates *xat* meaning ‘on, right on’ appears 108 times in the corpus, but it only occurs twice as an independent verb, meaning ‘affect’ with the impersonal subjects *mae* ‘hunger’ and *mesoh* ‘cold’. Similarly, *rat* meaning ‘out, off, away’ occurs 183 times in complex predicates in the corpus, but only twice as a main verb meaning ‘take out’. These are similar to

what Schutz (1969: 58–59) describes as ‘verbal prepositions’ in Nguna. Crowley (2002b: 173–175) discusses the status and development of similar forms in other Oceanic languages.

Two intransitive directional verbs also fulfil this function: *mamal* ‘straight’ and *tilomun, rilomun, rilamun* ‘return’, which is often used for metaphorical returns, such as rebuilding a house that has been destroyed (576). These verbs must take the transitivising suffix *-ni* when the complex predicate is transitive (§5.1.3).

- (575) *ma-mmei rilamun* *maa...*
 1DU.EXCL.NFUT-come+return.V2 on_and_on
 ‘We keep on coming back...’ [20141105f_p01e013_106]
- (576) *ni-pol rilomun-ni* *nim ak i-tep?*
 1SG.DFUT-make+return.V2-TR house PROX 3SG.DFUT-how
 ‘How would I rebuild this house?’ [20170413a_h01m169_80]

The subsequent verb in a complex predicate can also modify the meaning of the first verb, a function associated with adverbs cross-linguistically (Ernst 2002: 7–8). For example, the verb *sepin* ‘speak’, which has a dedicated reduced form (*sep*) for initial verbs in complex predicates, is modified by *mese* ‘ready’ to mean ‘explain, introduce’, and *pangas* ‘hurt’ to mean ‘tell off, swear’.

- (577) *ni-sep mese-ni* *i-tiamu*
 1SG.DFUT-speak+ready-TR 3SG.DFUT-first
 ‘I’ll explain it first.’ [20141106d_p07e016_14]
- (578) *la-sep pangas* *xamem vena huk*
 3PL.NFUT-speak+hurt 1PL.EXCL for hook
 ‘They told us off about fishing.’ [20170218b_n01s144_18]

While the meaning of the complex predicate is clearly compositional, the precise meaning is not entirely predictable from the main verb uses of the second verb. In some examples, it seems that the combination has been lexicalised, as with the ‘swear’ meaning of *sep pangas*. In others, the subsequent verb seems to be lexicalised separately with a different, though related, sense from the main verb. Often it is difficult to identify precise semantic shifts because for many verbs the use within a complex predicate is much more frequent, or the main verb use has been lost completely. As a main verb, *vuol* ‘finish’ is only used in a psycho-collocation (§3.4.2) meaning ‘tired, exhausted’

- (585) *Edu te-he ripitem rilamun-ni naentin fifti wan*
 Endu 3SG.PRI-COP repeat+return.V2-TR nineteen fifty one
 ‘Endu had repeated 1951 again.’ [20150414b_h01e016_89]
- (586) *la-taa-ve manajem kuhi-ni =ti*
 3PL.NFUT-NFUT.NEG-NEG.COP manage+careful-TR =NEG
 ‘They didn’t manage it carefully.’ [20170413a_h01m169_59]

There is evidence that complex predicates are less productive than SVCs or AVCs. Firstly, a large proportion of the tokens involve a small set of elements in subsequent position. Of the 487 complex predicate tokens in the subcorpus, 133 have *rat* ‘out, away’ and 80 have *xat* ‘on, right on’ as the subsequent element. Along with the directional or adverbial semantics of many of these subsequent verbs, this suggests that the complex predicate construction is extremely asymmetrical, which can itself be a sign of lack of productivity, and is another reason for avoiding an analysis as an SVC (Cleary-Kemp 2015: 105).

Other evidence that complex predication is not very productive in Vatlongos is the tendency of elements occurring as the subsequent element to be relexicalized. Verbs which modify the meaning of the first verb are prone to semantic shift or bleaching, as discussed above. Non-verbal lexemes are frequently derived from the form occurring in complex predicates, which is stronger evidence that they are indeed relexicalized, rather than just being semantically adjusted by a construction-specific process.

Often, an early stage of reanalysis as a different word class is restriction to non-initial position in complex predicates, as has happened in Paamese and Apma (Crowley 2002b: 112; Schneider 2010: 87–88). At this point, the lexeme is no longer a full verb: it cannot function as the main verb in a clause or take the verbal prefixes. Most of the lexemes that can be put in this category have adverb-like meanings, so that their syntactic dependence on the initial verb reflects a semantic requirement for a predicate to modify. Examples are *bin* ‘complete’, *kes* ‘alone’, *pelei* ‘properly’ and *ras* ‘mistake’. It is likely that full verb equivalents of these forms underwent semantic bleaching to derive these adverbial meanings. However, predicate-modifier semantic content is not a prerequisite for restriction to this position: *pas* ‘break, split’ cannot function as a main verb either, despite occurring with the most symmetrical function of

(591) *na-sep+til* *usil-ien* *xa* *tena,*
 1SG.IFUT-speak+stick about-NMLZR REL of
neta, *meul-ien* *te* *vetei*
 what live-NMLZR of breadfruit
 'I will talk about the topic of um, the life of the breadfruit.'
 [20150303d_p01e016_01]

Similarly, *ves* 'for' can function as preposition. Its verbal lexeme is restricted to a subsequent verb in complex predicates. Although the main verb form has been lost, the /v/ initial form would be the appropriate complex predicate form for v/b initial verbs (Table 30). As a subsequent verb it especially correlates with *lang* 'search' (592), but can also be used with other activity verbs denoting a manner of searching. Example (593) shows an unambiguous prepositional use of the form: the preposition phrase *e tan* 'on the ground' intervenes between the verb and the preposition *ves*.

(592) *di* ***lata-lang ves-i***
 CONT.REAL 3PC.NFUT-search+for-3OBJ
 'They were searching for it.' [20170220g_n01s148_22]

(593) *di* *pong~pongos* *e* *tan* *ves* *frog*
 CONT.REAL 3SG.NFUT.RED~sniff LOC ground for frog
 'It's [*the dog*] sniffing on the ground for the frog.'
 [20141212h_n01s046_59]

Pelei 'properly' can also function as an adverb and an adjective. Example (594) shows its use as a verb in a complex predicate: it is a transitive verb that takes the zero-allomorph of the object suffix (Table 27). If it were functioning as an adverb outside of a complex predicate, it could not intervene between the verb and its object (Figure 7).

(594) ***na-sak pelei*** *rong* *ma-metel*
 1SG.NFUT-do+properly bed CL.DR-1PC.EXCL
 'I made up our beds properly.' [20150419d_h01s046_06]

In example (595), *pelei* is an adverb occurring elsewhere in the clause. The intransitive preposition *nesau* 'on top' intervenes between the verb and the adverb.

(595) *naba* *na-e* *be* *nesau* *pelei*
 number CL.GEN-3PL.POSS 3SG.NFUT.go_to on_top properly
 'Their number [*children attending kindergarten*] went up nicely.'
 [20170413a_h01m169_76]

In examples (596) and (597), *pelei* is an adjective meaning ‘proper’, modifying the head noun it follows within an NP.

(596) *uhia, ngan tei xalu be uhia pelei*
 wild_yam one one 3DU 3SG.NFUT.COP wild_yam proper
 ‘wild yams, one kind of them is a proper wild yam’
 [20141106d_p02e016_1]

(597) *Xir moletin pelei xil ra-be holu*
 1PL.INCL person proper PL 1PL.INCL.NFUT-NFUT.COP many
 ‘There are lots of us proper humans.’ [20170217e_n01s142_32]

Apart from relexicalization affecting subsequent verbs in complex predicates, there are also a few verbs which can undergo formal changes when they occur as the initial verb. In the case of *sepin* ‘speak’, which occurs as *sep* in initial position in complex predicates, (577), (578), this may reflect a former morpheme boundary with an affix which has now fossilized as part of the root. The verb *tei* ‘cut’ can occur either as *tei* (or non-future affirmative *dei*) or as *ta* (non-future affirmative *da*) in initial position in complex predicates, (571), (573). In discussions with speakers these forms were described as the same, interchangeable word without any change in meaning: *tei* is a ‘slow’ form and *ta* is ‘quick’. The phonological reduction of verbs in these positions again shows that verbs participating in complex predicates have a different status from when they occur as main verbs.

Lexicalization can also affect combinations of verbs in complex predicates, resulting in lexicalised verbal compounds. Evidence for lexification of verbal compounds includes unpredictable or non-compositional semantic meaning. For example, *kil xat* (lit. ‘know affect’) means ‘have sex’, *loh xat* (lit. ‘run affect’) means ‘rush’, and *met pat* (lit. ‘die sleep’) is a euphemism for dying. For other compounds there is evidence of phonological fusion. The combination *leh rat* ‘take away’ is a common collocation in complex predicates, occurring 29 times in the subcorpus, but it also occurs 14 times with an epenthetic vowel as *leherat*, evidence that it is a lexicalised compound as well. Vowel harmony can also occur in lexicalised compounds: the collocation *sep usil, sep ulis* ‘talk about’, which occurs 18 times in the subcorpus, can also occur as *supulis*.

Nominalisation can be additional evidence for lexicalisation of complex predicates: in elicitation speakers accept some but not all nominalised forms of

complex predicates as grammatical. Although no nominalised complex predicates occur in the subcorpus, there is one example in translations of educational materials: *sep-til-ien* ‘speak-stick-NMLZR’, translating ‘information’.

All these patterns of reanalysis show a marked tendency for verbs participating in complex predicates to lose their verbal character. They may undergo semantic bleaching, be restricted in their distribution, be reanalysed as adverbs or prepositions, or become a component of a lexicalised compound verb.

7.2 Subordination

In Vatlongos subordination, one clause is embedded in another, and the embedded clause does not depend on the matrix clause for the expression of TAMP categories: this is the main distinction between subordination and SVCs (Chapter 6). In an RRG framework, this is the distinction between subordination and cosubordination, where the elements are dependent on each other for expression of operators such as tense and aspect (Van Valin 1990: 216–217). An embedded clause in Vatlongos always takes full verb inflection, and is further distinguished from the subsequent VPs in SVCs by the possibility of including an auxiliary verb and subject NP. There are no dedicated non-finite verb forms.

Rather than presenting a comprehensive account of subordination in Vatlongos, this section is restricted to verbal subordination, when a verb takes a clause as an argument: either a complement clause⁴² that is selected by a particular verb, or a clausal subject, which can occur with any verb.

Complement clauses are frequent in Vatlongos, occurring once every 51 words in the subcorpus. Clausal subjects also occur in Vatlongos, but are ambiguous with a serial verb analysis in most TAMP environments; they are only formally distinguished from event-argument SVCs when the matrix clause

⁴² Sometimes ‘complement clause’ is used to refer to clausal subjects as well, in models where clauses selected by the verb are described as ‘clausal objects’ (Dixon 2006). I reserve the term ‘complement clause’ to clauses selected by the verb, that is, where the clause is a complement argument.

is in negative polarity, or the subordinated clausal subject is negative polarity with prior or future relative tense.

7.2.1 Complement clauses

Complement clauses are selected by many verbs in Vatlongos, especially verbs of desire (*hit*, *hitene* ‘want’; *e hei* ‘want, like’), speech (*hit*, *hitene* ‘say’; *xa* ‘say’), cognition (*nenem* ‘think’), perception (*pus* ‘see’; *long* ‘hear’), feeling (*temanin* ‘be shocked’), ability (*kil* ‘know how to, can’; *sakras*, *saxaras*, *sas* ‘cannot’) and causation (*pol* ‘do, work’; *sak* ‘do, try’). Other more grammaticalized verbs that take complement clauses are *xoni* ‘be like’ and the Bislama loan verb *staat* ‘start’.

It is noticeable that most of these complement-taking verbs are transitive, or are marked with the transitivising suffix when they take a complement clause.⁴³ However, complement clauses have a different distribution than object NPs: they always occur after the negative clitic (§3.6.2.6), and they occur with the third person object-indexing suffix, which never occurs with object NPs (§5.1.3). They therefore do not fit a definition of complement clauses as clausal objects (Dixon 2006: 1–5).

A complement clause can contain all the same elements as a root clause, including a subject NP, (598), (599), (600), a negative clitic, (598), (599) and an auxiliary verb, (600), (601). This is a formal difference between complementation and serialisation (§6.3.1).

(598) ***na-pus-i*** [*nim* *taa-xeih* =*ti*]⁴⁴
 1SG.NFUT-see-3OBJ house 3SG.NFUT.NEG-strong =NEG
 ‘I saw the house wasn’t strong.’ [20150414b_h01e016_09]

(599) *Ma tati na-van* ***bitene*** [*inou ni-naa-skul* =*ti*]
 so Dad CL.GEN-1SG.POSS 3SG.NFUT.say 1SG 1SG.DFUT-NEG-study =NEG
 ‘So my Dad said I wouldn’t go to school.’ [20170406a_n01m164_11]

⁴³ *Temanin* ‘be shocked’, *pol* ‘do, work’ and *sak* ‘do, try’ all take the transitivising suffix when they occur with a complement clause. *Hit* and *hitene* are not morphologically transitive and are not marked with the transitivising suffix.

⁴⁴ In §7.2, the subordinated clause is demarcated with [square brackets] and the verb in the matrix clause that selects the subordinate clause as an argument is marked **bold** for emphasis.

- (600) *mi-long-e* [vat hu di bang]
 3SG.NFUT-hear-3OBJ head hill CONT 3SG.NFUT.make_noise
 ‘He heard the top of the hill was making a noise.’
 [20141107b_n01e020_15]
- (601) *Mut-bit* [ha muto-pat tu xie?
 2PC.NFUT.want GO 2PC.IFUT-sleep already now
 ‘Do you want to go to sleep now?’ [20141028a_c01m002_19]

Another formal difference between subordination and serialisation is that the verb in a complement clause is independently marked for TAMP. In (598) and (599), the verb in the complement clause is negated when the verb in the matrix clause is marked for affirmative polarity, which is not possible in an SVC. Similarly, in (599) and (601) the matrix verbs are marked for non-future, while the subordinate verb is marked for distant or immediate future. This differs from Paamese, where Crowley (2002b: 62) argues that there is a mood dependency between a matrix verb and a subordinate verb.

Instead, the TAMP marking on each verb is determined by the intended semantic interpretation: the choice of marking on a subordinate verb can disambiguate whether there is a commitment to the situation expressed by the complement clause being real. Crowley (2002b: 62) states that in Paamese a negated realis (equivalent of non-future) main verb is always followed by an immediate affirmative verb in the subordinate clause, and this pattern also occurs in Vatlongos:

- (602) *taa-vos* =ti [la-ti tim]
 3SG.NFUT.NEG-NEG.good =NEG 3PL.IFUT-stay home
 ‘It wasn’t good for them [*children*] to stay home.’
 [20170315b_n01m159_07]

In this example, no commitment is made to the reality of the children staying home, and in fact they do not; a kindergarten was built to accommodate them. However, if the speaker wants to make a commitment to the reality of the situation in the complement clause, it can be marked for non-future, regardless of the polarity marking on the matrix verb. This distinguishes subordination in Vatlongos from both serialisation, and a mood-dependency system of subordination like that found in Paamese. In (603), the speaker commits to the reality of the jumping event in the subordinate clause (within the fictive

(606) *inou* *na-pus-i* *tu* [*xa*
 1SG 1SG.NFUT-see-3OBJ already SUB
vatiang *bemei* *di* *mi-geih]*
 wind 3SG.NFUT.come CONT.REAL 3SG.NFUT-NFUT.strong
 'I already saw that the cyclone was getting strong.'
 [20150414b_h01e016_05]

(607) *na-taa-kil* =*ti* *vari* [*tati*
 1SG.NFUT-NFUT.NEG-know =NEG at_once dad
na-van *ba* *xavi]*
 CL.GEN-1SG.POSS 3SG.NFUT.go where
 'I didn't know straightaway where my Dad had gone.'
 [20170222h_h01o154_27]

(608) *u-hiteni* *mi* *xil* [*xa* *xouk* *iaxa...*]
 2SG.DFUT-say to 3PL SUB 2SG so
 'You'll say to them that it's you that {authorises the production of this
 documentary} [20170413a_h01m169_117]

Although complement clauses are usually not marked by any subordinator, an overt subordinator can optionally be used. Zero-subordination of complement clauses is about four times as common as overt subordination, occurring once every 62 words in the subcorpus, compared to only once every 273 words for complement clauses marked with a subordinator. Impressionistically, an overt subordinator is more likely to be used before a topic or subject NP, ruling out a transitive reading where the NP is an object, (608), (609), (613), or before hesitation (610), signalling that the speaker is going to continue with a complement clause. The general subordinator *xa* is the most frequent subordinator for complement clauses in the corpus. Only 6% of the examples of complements with an overt subordinator contain a subordinator other than simple *xa*, including *ma* 'then' (609), *lehit* 'as if' (610), *tena* 'of' (611) and the Bislama loanword *hao* 'how'. Half of these are complex complementisers including *xa*: *tangan xa* 'so that' (612) and the Bislama *hao xa* 'how that'.

(609) *Pol-ni* [*ma* *tatal,* *a,* *volongo-n*
 3SG.NFUT.do-TR then snake HES mouth-3SG.POSS
be *open* *dat* *mak]*
 3SG.NFUT.COP open 3SG.NFUT.just_stay
 'It made it so that the snake, its mouth just stayed open like this.'
 [20141220g_n01s080_55]

- (610) *o-long-e* [lehit na, na, plen
 2SG.NFUT-hear-3OBJ as_if HES HES plane
be aot *be* nesau]
 3SG.NFUT.go_to out 3SG.NFUT.go_to up
 ‘It sounded like, um, a plane taking off’ lit. ‘you heard it as if...’
 [20150303h_n01e106_04-5]
- (611) *be* *tabu* [tena atou va-haen]
 3SG.NFUT.COP taboo of woman 3SG.IFUT-go_in
 ‘It’s taboo for a woman to go in.’ [20141105f_p01e013_015]
- (612) *Be* *akri* [tangan xa haramue na-ne,
 3SG.NFUT.COP agree so_that SUB boy CL.GEN-3SG.POSS
haramue *ak* *i-hur* *atuli* *ak*]
 boy PROX 3SG.DFUT-take girl PROX
 ‘He agrees that his boy, this boy will marry this girl.’
 [20150219a_p01m001_14]

When morphologically transitive verbs (§5.1.3) take a complement clause, they also take an object suffix, encoding a semantically empty third-person singular object (e.g. *pus* ‘see’ and *long* ‘hear’ in (598), (600), (603), (606), (610)). The object suffix does not occur before the negative clitic *ti*, (604), (610), (§3.6.2.6). However, for some speakers, especially from Endu, the object suffix is retained and the negative clitic does not occur in these contexts (e.g. (603), a Mele Maat resident who was born in Endu).

There are a few examples of affirmative transitive verbs being used with a complement clause without taking the object suffix. This suggests that for some speakers these are being reanalysed as formally intransitive verbs when they take a complement clause. The first example is from a Mele Maat speaker, and the second from a speaker who learned Vatlongos as an adult.

- (613) *ma-pus* [xa pupu *tei* na-mel...]
 1PL.EXCL.NFUT-see [SUB grandparent one CL.GEN-1DU.EXCL.POSS
 ‘We saw that a grandparent of ours, {they took him to hospital}’
 [20170412b_n01m167_10]
- (614) *ri-long* [xa ueili ok di tamu
 1PL.EXCL.DFUT-hear SUB pig PROX CONT still
mi-kuh nesau] *xie*
 3SG.NFUT-snore up now
 ‘We’ll hear this pig still snoring up there now.’
 [20170220g_n01s148_50]

- (617) *pol-ni* *[nou* *goni,* *[goni*
 3SG.NFUT.do-TR 1SG 3SG.NFUT.like 3SG.NFUT.LIKE
[nou *na-kila~kil* *be* *eilep*
 1SG 1SG.NFUT-know~RED 3SG.COP big
gusil *metalo* *xil]]]*
 3SG.NFUT.follow white_person PL
 ‘It means that me, it’s like, it’s like me I know a lot about white
 people.’ [20170224a_n01s141_69]

With verbs of perception, it is possible to express the subject of a complement clause as an object NP in the matrix clause: a phenomenon described as raising (Dalrymple 2001: 312–325; Bresnan et al. 2015: 304–306) or matrix coding (Van Valin & LaPolla 1997: 561–572). This is especially clear with transitive verbs, as the absence of an object suffix shows that the NP is an object of the matrix verb syntactically, rather than a subject of the complement clause: otherwise the matrix verb would be ungrammatical.

- (618) *la-long* *tas* *[bang]*
 3PL.NFUT-hear sea 3SG.NFUT.make_noise
 ‘they heard the sea making a noise’ [20150226a_n01s098_31]
- (619) *O-ba* *o-pus* *uhia* *[mi-tteh]*
 2SG.NFUT-NFUT.go 2SG.NFUT-see wild_yam 3SG.NFUT-grow
 ‘You go and see the wild yam growing.’ [20141106d_p10e016_08]

When the subject of the complement clause is coded as an object NP in the matrix clause, it cannot also be expressed by a coreferential NP in the complement clause, including independent pronouns:

- (620) * *o-pus* *Mael* *[xi* *mi-loh]*
 2SG.NFUT-see Mael 3SG 3SG.NFUT-run
 (‘You see Mael running’) [20180801b_x01s046]

However, the complement clause can still be marked with an auxiliary verb (Chapter 8), showing that it is still an independent clause rather than a VP constituent like that found in SVCs (§6.3.1).

- (621) *ti* *ri-pus* *xir* *[ti* *ra-vituei]*
 CONT 1PL.INCL.DFUT-see 1PL.INCL CONT 1PL.IFUT-together
 ‘We’ll see ourselves come together’ [20170124b_t01e137_08]

When an intransitive verb takes the transitivising suffix *-ni* it is more difficult to distinguish this phenomenon from a semantically empty object reading,

because the suffix *-ni* cooccurs with object NPs (§5.1.3). These structures are also sometimes ambiguous with a zero-marked relative clause analysis.

The very frequent complement-clause-taking verb *hit* ‘want, say’, which occurs with zero complementation about once every 120 words, seems to be prone to grammaticalization. §4.2.3 discusses the semantic bleaching of the ‘desire’ sense into one of intention and more general future meanings, as well as evidence that it is undergoing formal reduction, occurring without prefixes. It also functions as a quotative marker and complementiser in SVCs (§6.4). When *hit* is used as a complementiser with *sak* ‘try, cause’, *sak* appears without the transitivity suffix *-ni*, unlike in (615) and (616), where the subordinate clause is a direct complement of *sak*.

- (622) *mat-sak* *mat-bit* *a*, [*pos na-mem...*]
 1PC.NFUT-try 1PC.NFUT-NFUT.want HES boss CL.GEN-1PL.POSS
 ‘We tried to get our boss {to trust us}’ [20170224a_n01s141_49]

7.2.2 Clausal subjects

All the examples of clausal subjects in the corpus are zero-marked: they do not take an explicit subordinator. Impressionistically, they are instead signalled with a rising intonation contour, indicating that the speaker is going to comment on it. In many TAMP environments clausal subjects are ambiguous with event-argument SVCs (§6.4.2). However, there are a few examples in the corpus where the second verb is negated with the clitic *ti*, which is not possible in SVCs. This shows that clausal subjects can take TAMP marking independently from the matrix verb, and are an example of subordination rather than serialisation.

(623) is an especially complex example, where the clausal subject *la-pol* ‘they work’ is inside the complement clause of *pus* ‘see’, and is preceded by a complex NP topic with a zero-marked relative clause (marked with the distinctive prosody associated with fronted constituents (§3.1.1)). It is the poor behaviour of the teachers that the speaker has observed, rather than just their arrival or that they are working. This is therefore evidence of hierarchical structure, rather than a zero-coordination analysis. It also distinguishes a clausal-subject analysis from a topicalised clausal adjunct. In (623), the clausal

8 Auxiliary verb constructions

Vatlongos has an auxiliary verb construction (AVC) that is not found in any of the neighbouring languages including Paamese. AVCs seem to be a recent innovation and I argue that they have emerged from SVCs. They involve a small class of auxiliary verbs that mark continuous aspect and prior motion. The auxiliary verb occurs between any subject NP and the lexical verb. The lexical verb takes full inflection for subject agreement and TAM, while the auxiliary verb does not take any verbal affixes. Some auxiliary verbs can partially inflect to reflect the TAM of the construction, if they are derived from a lexical verb that undergoes verb-initial consonant mutation. This means that Vatlongos AVCs can be analysed as lexically-headed or split-inflectional constructions (Anderson 2006: 21–25; Anderson 2011: 805–806), as the lexical verb is the head of morphological inflection.

Guérin (2016: 271–277) identifies three definitional strategies in the literature on auxiliaries: categorial, where auxiliaries are a separate lexical class, perhaps a subtype of verbs; functional, where auxiliaries are identifiable by the grammatical categories they express; and panchronic, where auxiliaries are intermediary items on a grammatical path from lexical verb to grammatical affix. The panchronic approach is dominant in typology (Heine 1993: 69–70; Kuteva 2001: 5–14; Anderson 2006: 4–5; Anderson 2011: 796–797).

Vatlongos auxiliary verbs are certainly a separate lexical class, on the basis of their morphological behaviour and syntactic distribution. I am analysing them as a subtype of verbs (as does Parker 1970b: viii), on the basis that eligible forms undergo verb-initial consonant mutation (§5.2.1, §8.1.2). The invariant auxiliaries (*mei*, *mas*) are phonological forms that are eligible to function as independent verbs with zero-marked 3SG.NFUT inflection (§5.1.1.2). Cross-linguistically, many invariant forms are described as auxiliary verbs, notably the English modal auxiliaries. Alternatively, Vatlongos auxiliaries could be analysed as a separate class of particles (or perhaps adverbs), although this would obscure the relationship with verbal morphological processes. Under that analysis, these would still be described as auxiliary constructions, but not as AVCs. In the typology of predicative elements that Guérin (2016: 277–283)

proposes, Vatlongos auxiliaries are on the borderline between auxiliary verbs, which resemble defective verbs, and predicate markers, which do not exhibit any verbal morphosyntactic behaviour.

Functional definitions focus on the grammatical categories that auxiliaries typically mark cross-linguistically: aspect, modality, temporality, polarity and voice (Anderson 2006: 5). While the function of the continuous auxiliaries (§8.2.1) fit neatly within this category, the prior motion auxiliaries are less clearly grammatical (§8.2.2). However, auxiliaries with directional meanings are also encountered (Anderson 2006: 35–36, 345–352), and associated motion functions can be fully grammaticalized as affixes (Koch 1984).

The Vatlongos auxiliaries certainly meet panchronic definitions. The relationship with their lexical source verbs is transparent (Table 41), and they are formally reduced elements, which could eventually be incorporated into the verbal affix template (§8.3.2). Moreover, different auxiliaries appear to be at different stages of auxiliiation, rendering the diachronic process relatively clear (§8.3.1), even within the confines of a synchronic descriptive project.

Table 41 lists the auxiliary verbs in Vatlongos, the meaning of their lexical equivalent, their grammatical meaning as an auxiliary, and their token frequency in the subcorpus. A glance at the frequencies shows that the three main auxiliary verbs are *ti* marking continuous aspect; *ha* marking prior motion away from the deictic centre; and *mei* which marks both prior motion towards the deictic centre and inchoative aspect. The three other ‘stay’ verbs also express continuous aspect and are probably used by analogy with *ti*, while *tit* and *tat* contribute a minimising semantic component like their lexical equivalents, which incorporate the adverb *tang* ‘just, only’ (§3.3.1).

The grammatical meanings of *ha* ‘go’, *mei* ‘come’, and the Bislama loan verb *mas* ‘must’ all overlap with the functions of their lexical equivalents as initial verbs in SVCs (§6.4.3, §6.4.10, §6.4.11). *Mei* ‘come’ is the only auxiliary which has a different form from the equivalent lexical root. The lexical root of ‘come’ in prefixed forms is *mmei*, the imperative form is *ammei*, and it takes an irregular third-person singular non-future form *bemei* (§5.2.4). *Mei* does occur as a

variant of *bemei*, but only nine times in the subcorpus, compared to 478 occurrences of lexical *bemei*, and 269 occurrences of auxiliary *mei*. Phonological reduction is not surprising in this highly grammaticalized function and syntactic distribution.

Table 41: Auxiliary verbs in Vatlongos

Auxiliary verb	Lexical meaning	Grammatical meaning	Frequency in subcorpus
<i>ti</i>	'stay'	Continuous	~1/70 words
<i>tit</i>	'just stay'	Continuous	~1/3200 words
<i>ta</i>	'stay'	Continuous and minimiser	~1/11,000 words
<i>tat</i>	'just stay'	Continuous and minimiser	~1/45,000 words
<i>ha</i>	'go'	Prior motion away	~1/100 words
<i>mei</i>	'come'	Prior motion towards; inchoative	~1/170 words
<i>mas</i>	'must' (Bislama loan)	Obligation	~1/15,000 words

On the basis of frequency, AVCs are a very productive construction in Vatlongos, occurring about once every 34 words in the subcorpus. The occurrence of individual auxiliary verbs differs greatly, and productivity is best considered separately for each auxiliary verb. Continuous *ti* is the most productive, followed by the prior motion auxiliaries *ha* and *mei*; this supports an argument that *ti* is the most established, and that *ha* and *mei* have subsequently started to undergo a similar process of formal reduction.

8.1 Formal properties

The formal properties of these constructions support an analysis as AVCs where the auxiliary verb is the syntactic head of a clause. The auxiliary could be analysed as the head of the inflectional phrase in an X-bar analysis of the clause. It appears above the VP in constituent structure: it cannot occur in the subsequent verb in an SVC (§6.3.1), and is not repeated in verbal repetition (§4.3.2, §4.3.3), but can occur in complement clauses (§7.2.1). However, the lexical verb is the head of morphological inflectional, as it is fully inflected, like a verb in a monoverbal clause. The preferred position of prior motion auxiliary verbs in prior motion SVCs is not predicted by the analysis of SVCs and AVCs

presented thus far. The emergence of AVCs may have led to a reanalysis of serialisation in this domain (§8.1.3).

8.1.1 Constituent order

Auxiliary verbs occur after any subject NP and before the lexical verb:

- (626) *xamim di mu-ga-ni taxeak?*⁴⁵
 2PL CONT.REAL 2PL.NFUT-NFUT.eat-3OBJ now
 ‘Are you eating it now?’ [20141127c_x01s043_12]

Adverbs of relative time, restriction and modality can intervene between the subject NP and the auxiliary (§3.3).

- (627) *Asu su ba mi-gol ueili bemei*
 rat after GO.REAL 3SG.NFUT-NFUT.chase pig 3SG.NFUT.come
 ‘Then the rat went and chased the pig over.’
 [20141027a_n01m011_56]

Adverbs of relative time can intervene between the auxiliary and the verb, and the adverb *tamu* ‘still’ is the most frequent adverb to occur in this position after the continuous auxiliary *ti*:

- (628) *xametel di tamu mata-pat*
 1PC.EXCL CONT.REAL still 1PC.EXCL.NFUT-sleep
 ‘We were still sleeping.’ [20141116b_c01m_05]

An argument for analysing auxiliary verbs as the head of the clause, rather than a particle in the verbal complex, is that they are not repeated along with the verb and its object to mark durative aspect (see (361), (362) in §4.3.2). This suggests that they are not part of the constituent formed by a verb and its object. Similarly, the possibility of including an auxiliary is a diagnostic to distinguish subordinate clauses (§7.2) from serialised VPs (§6.3.1). They therefore could be analysed as a correlate of finiteness in Vatlongos (Nikolaeva 2007).

It is possible to have more than one auxiliary verb in a single clause, though there are tight restrictions on their relative ordering. The precedence

⁴⁵ In this section the auxiliary verb, and the lexical verb are both marked with **bold** for emphasis. The prior motion functions of *ha* and *mei* as auxiliary verbs are glossed as GO and COME respectively, as an abbreviation and to distinguish the grammatical status of the restricted prior motion sense from the broader meaning of their lexical equivalents.

being an arbitrary restriction in the grammar, I argue this falls out from the iconic interpretation of the ordering of the prior motion auxiliaries. This explanation necessarily anticipates the discussion of the function of the prior motion auxiliaries (§8.2.2).

The auxiliaries *ha* 'go' and *mei* 'come' are used to express the direction of motion prior to the situation (usually an event or activity) expressed by the main verb. *ha* 'go' expresses that the prior motion is away from the deictic centre, while *mei* expresses that the prior motion is towards the deictic centre, as shown in Figure 12, where the star represents the deictic centre.

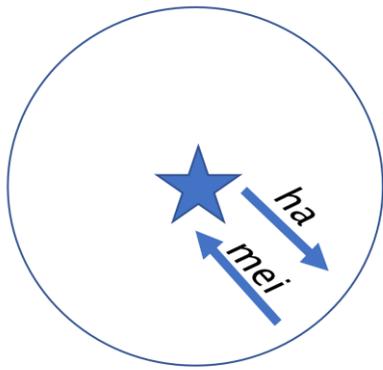


Figure 12: Diagram of prior motion direction expressed by auxiliary verbs *ha* 'go' and *mei* 'come', relative to deictic centre.

The meanings of these auxiliary verbs clearly overlap with the functions of the lexical verbs they are derived from when they occur as initial verbs in SVCs (§6.4.3). As in SVCs, the deictic centre could be the place of speech, a place of fictive speech, or some other location that is important to the discourse, especially the setting of a narrative. As in SVCs, the ordering of directional auxiliary verbs is iconic. In (632), the pigeon is telling the heron to get the shellfish. It is clear in the context of the story⁴⁶ that the pigeon is positioned between the heron and the shellfish, as the pigeon sees the shellfish first, but the heron falsely claims to have seen it first from further away, thanks to its long neck. The pigeon's (fictive) place of speech is the deictic centre, and *mei ha*

⁴⁶ A subtitled video of this story is available online (Ridge 2016)

therefore expresses a coherent path in a single direction that the heron must follow to collect the shellfish, as shown in Figure 13.

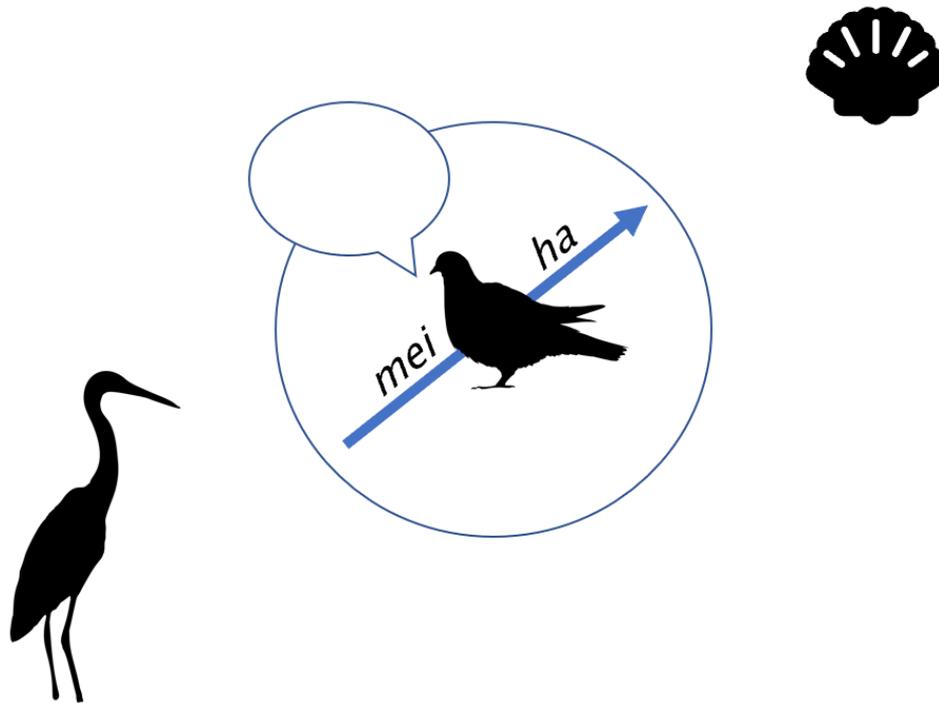


Figure 13: Relative position of speaker, addressee and object in (632)

Similarly, in (635), *mei ha* is used to express a straight path of prior motion. In this case, the deictic centre is the beach where the narrative is currently taking place, and where the subject of the matrix clause is located. The people of Sameo are on the beach and have met a white trader who is in a ship. They want the trader to have a look at a house, which is positioned inland and uphill. The coherent path that the trader must follow in order to reach the house passes through the deictic centre on the beach, so the *mei ha* combination is appropriate.

(635)	<i>la-bit</i>		<i>[metalo</i>	<i>ak</i>	<i>mei</i>	<i>ha</i>
	3PL.NFUT-NFUT.want		white_person	PROX	COME	GO
	<i>va-pus</i>	<i>=ti</i>	<i>nim</i>	<i>tei</i>	<i>e</i>	<i>nesau]</i>
	3SG.IFUT-see	=PART	house	one	LOC	above
	'They wanted this white man to come and go and have a look at a house up the hill.' [20150305h_h01o111_11]					

As the directional auxiliary verbs are arranged iconically, the opposite order *ha mei* is not possible, as it cannot express a coherent path in a single direction that immediately precedes the situation expressed by the main verb. Any path that

began in a direction away from the deictic centre and ended in a movement towards it would entail a change of direction. This is shown schematically in Figure 14.

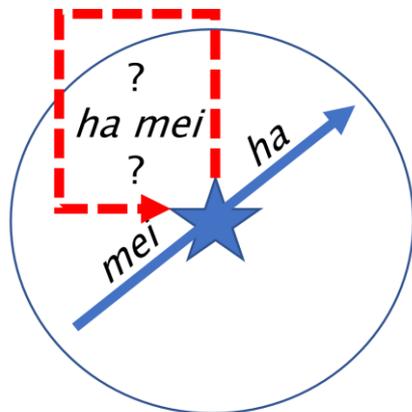


Figure 14: Coherent and incoherent paths expressed by relative ordering of the prior motion auxiliaries

The possibility of using two prior motion auxiliaries together shows that the concept of prior motion is less grammaticalized than the aspectual category, where only one of the continuous aspect auxiliaries can be used per clause. Bybee et al. (1994: 7) describe mutual exclusivity among members of a class developing in line with increasing grammaticalization, along with rigidity of positioning.

8.1.2 Morphological inflection

The main formal difference between AVCs and SVCs is that auxiliary verbs do not take any of the verbal affixes (§5.1). However, the auxiliary verbs that are derived from lexical verbs which undergo verb-initial consonant mutation (§5.2.1) do inflect for some TAM features, in ways that differ slightly from their lexical equivalents. While auxiliary verbs *mei* ‘come’ and *mas* ‘must’ have an invariant form, the continuous auxiliary verbs based on ‘stay’ verbs can appear in both /t/ and /d/ initial forms, and auxiliary *ha* ‘go’ can appear in one of four forms. This is the most important reason for analysing Vatlongos auxiliaries as retaining some of their verbal character, rather than constituting a class of non-verbal particles: *ti* and *ha* inflect for some TAM categories, and *ha* also inflects for subject person-number in some combinations.

The most frequent auxiliary verb *ti* undergoes verb-initial consonant mutation in ways that are slightly different from its lexical equivalent *ti* ‘stay’. Figure 15 shows the root forms of lexical *ti* ‘stay’ in different relative tense and polarity environments, while Figure 16 shows the forms of auxiliary *ti* ‘continuous’. In addition to the relative tenses, the irrealis moods (imperative, prohibitive and apprehensive) also take the basic form *ti*.

Non-future + <i>di</i>	Prior + <i>ti</i>
Immediate Future + <i>ti</i>	Distant Future + <i>ti</i>
Negative <i>ti, ri</i>	

Figure 15: Root forms of lexical *ti* ‘stay’ in different relative tense and polarity environments

Non-future <i>di</i>	Prior <i>di, ti</i>
Immediate Future <i>ti</i>	Distant Future <i>ti</i>

Figure 16: Forms of auxiliary *ti* ‘continuous’ in different relative tense environments

The most immediately obvious different between the inflection patterns of lexical and auxiliary *ti* is that only lexical *ti* inflects for polarity. Auxiliary *ti* is only interested in the relative tense of the clause, as shown in example (636), where the *di* form is used in the non-future, despite the negative polarity of the predicate.

- (636) *di* *la-taa-meul* =*ti* *e* *nahou*
 CONT.REAL 3PL.NFUT-NFUT.NEG-grow =NEG LOC garden
 ‘They don’t grow in gardens’ [20141106d_p03e016_05-06]

The other interesting discrepancy in the verb-initial consonant mutation patterns for lexical and auxiliary *ti* is in the prior. The lexical verb always takes the basic *ti* form as the stem for prior relative tense, and using the *di* form is ungrammatical. However, for the auxiliary verb, *di* is the preferred form for most speakers in prior environments:

(637) *di* *nate-pol* *e* *Santo*
 CONT.REAL 1SG.PRI-work LOC Santo
 ‘I worked in Santo’ [20150219b_n01m001_57]

Ti is also accepted as grammatical in prior environments by most speakers in elicitation, but only occurs twice in the subcorpus, compared to 17 occurrences of *di* in prior relative tense. Both spontaneous occurrences are in a custom story told by Johnson, an older male speaker from Pamal:

(638) *e-ok* *bovorong-ni* *is* *olvala* *nen*
 inside-1SG.POSS forget-TR name old_person of_it
suta *xa* ***ti*** ***lat-hiteni***
 just REL CONT 1PL.PRI-say
 ‘I’ve just forgotten the name that they used to call this ancestor’
 [20170220e_n01s147_19]

The other continuous auxiliaries *ta*, *tit* and *tat* follow the same inflection patterns as *ti*.

As lexical *ha* undergoes one of the most complex patterns of verb-initial consonant mutation, it is not surprising that inflection patterns for auxiliary *ha* are more complicated than for *ti*. Figure 17 shows the root forms of lexical *ha* in different relative tense and polarity environments. Unlike lexical *ti*, the root forms of *ha* used in the immediate future are also conditioned by subject person-number. The *ma* form only occurs with first-person singular subjects in the immediate future, and the *va* form occurs with plural or second-person singular subjects in the immediate future (§5.2.1).

Non-future + <i>ba</i>	Prior + <i>ha</i>
Immediate Future + <i>ha, va; 1SG ma</i>	Distant Future + <i>ha</i>
Negative <i>va</i>	

Figure 17: Root forms of lexical *ha* ‘go’ in different relative tense and polarity environments.

Figure 18 shows the forms of auxiliary *ha* ‘GO’. The other irrealis moods take the basic *ha* form, as shown for the imperative mood in (632).

Non-future <i>ba</i>	Prior <i>ba</i>
Immediate Future <i>ha, va; 1SG ma</i>	Distant Future <i>ha</i>

Figure 18: Forms of auxiliary *ha* 'GO' in different relative tense environments.

Again, the major difference between the inflection patterns of the lexical and auxiliary verb is the lack of auxiliary inflection for polarity. In elicitation, speakers consistently gave the *ba* form for negative polarity clauses in the non-future. However, unlike with the continuous auxiliary, negative polarity main clauses are rarely modified with prior motion auxiliaries in spontaneous texts: there are only two examples in the subcorpus, and only one with *ha*. In (639), the negative clause marked with auxiliary *mei* 'come' contrasts with a clause expressing the ship's intention to bring people to shore.

- (639) *sip mei va-so xil e ut,*
 ship COME 3SG.IFUT-put 3PL LOC shore
mei taa-so xil =ti
 COME 3SG.NFUT.NEG-put 3PL =NEG
 'The ship was coming to put them on the shore, it didn't come and put them [there].' [20150305h_h01o111_35]

In (640) there again seems to be a parallel with a juxtaposed clause that encourages the use of the prior motion auxiliary with a negative clause. Surprisingly, the prior motion auxiliary is modifying a mental state verb *kil* 'know'. The crab does not know that the rat is going to burn the rubbish that he goes to sleep in: perhaps the auxiliary is used to parallel the structure of the following clause. The *ba* form of auxiliary *ha* is used, inflecting to match the non-future relative tense and not the polarity of the clause.

- (640) *ba taa-kil =ti, ba pat dat*
 GO.REAL 3SG.NFUT.NEG-know =NEG GO.REAL 3SG.NFUT.sleep 3SG.NFUT.just_stay
 'He didn't know, he went and slept just there.' [20141119b_n01s036_04]

For auxiliary *ha*, the realignment of inflection in the prior is more consistent than with auxiliary *ti*: there are no examples in the corpus of the basic form *ha* used in this relative tense context, and speakers consistently rejected it in grammaticality judgements. Instead, the *ba* form is used for both the prior

(641) and the non-future, (627). For this reason, the *ba* and *di* forms of auxiliary *ha* and *ti* are glossed as realis rather than non-future like their lexical equivalents, as they are used in both non-future and prior relative tense clauses. This inflection pattern in auxiliary verbs is one piece of formal evidence for the hierarchical TAM structure shown in Figure 8, and suggests that the actuality of the situation expressed might be more important than relative tense to the inflection patterns of auxiliaries.

(641) *ba* *lute-ti* *rute* *tei*
 GO.REAL 3DU.PRI-stay place one
 ‘They’d gone and stayed in one place.’ [20170221e_n01o150_06]

Whereas the basic *ti* form is used for the continuous auxiliary in all other TAM environments, the patterns for auxiliary *ha* are more complicated and depend partly on subject person and number. The *ha* form is always used in the distant future:

(642) *xi* *ha* *i-has-i*
 3SG GO 3SG.DFUT-hit-3OBJ
 ‘he’d go and hit it.’ [20170222d_n01s152_29]

The *ha* form is also used in the immediate future in the same subject person-number combinations as for the lexical verb, namely third-person singular subjects (635), and all dual and paucal subjects:

(643) *Ha* *ralo-xul* =*ti* *e* *tas*
 GO 1DU.INCL.IFUT-swim =PART LOC sea
 ‘Let’s go have a swim in the sea.’ [20141027a_n01m001_95]

With second-person singular subjects and all plural subjects in the immediate future the *va* form is used, as with lexical *ha*:

(644) *moletin* *xil* *va* *la-pat*
 person PL GO.IFUT 3PL.IFUT-sleep
 ‘people were about to sleep’ [20150226a_n01s098_04]

The *ma* form is used in the immediate future with first-person singular subjects:

(645) *na-bit* [*ma* *na-sohuk*]
 1SG.NFUT-NFUT.want GO.IFUT.1SG 1SG.IFUT-fish
 ‘I want to go and fish.’ [20141215a_p01s061_07]

The exact overlap between the inflection patterns of lexical and auxiliary *ha* in different subject person-number combinations in the immediate future shows a very close relationship between the auxiliary and the lexical verb it is derived from, and supports the theory that the Vatlongos directional AVCs derive from prior motion SVCs, where the two verbs always have matching subjects (§6.4.3). On the other hand, the realignment of inflection for prior relative tense to match non-future, resulting in inflection for realis, and the absence of polarity as a determining factor of the inflection of auxiliary verbs, suggests that Vatlongos AVCs maintain a system of inflection that differs from that of lexical verbs. An alternative analysis of these auxiliaries as non-verbal elements would have to posit a system of TAM concord to account for these alternations.

8.1.3 AVCs and prior motion SVCs

The analysis put forward so far suggests that auxiliary verbs head a clause and precede the VP, including SVCs (Figure 11, (461), (513), (550)). However, the distribution of auxiliary verbs in prior motion SVCs is surprising. In this section all verbs in the SVC are marked **bold** for emphasis, because the distribution of the auxiliary verb suggests that the syntactic relationship between the component verbs is not the same as in other SVCs.

The continuous auxiliary verb *ti* precedes the first verb in a prior motion SVC when it is used to express habitual meaning (§8.2.1), a pattern that occurs four times in the subcorpus.

(646) ***di*** ***na-mmei*** ***na-pol***
 CONT.REAL 1SG.NFUT-come 1SG.NFUT-work
 ‘I would come and work.’ [20150219b_n01m001_121]

However, when the continuous auxiliary verb *ti* expresses a progressive meaning, it can follow the first verb in a prior motion SVC:

(647) *bitene* [***lat-ha*** ***ti***
 3SG.NFUT.say 3PC.IFUT-go CONT
lato-pol-ni *nahou* *na-tel]*
 3PC.IFUT-work-TR garden CL.GEN-3PC.POSS
 ‘He said they’d go and be working on their garden.’
 [20170221e_n01o150_23]

I argue for a diachronic explanation of this unexpected word order. As the auxiliary verbs seem to have emerged from SVCs (§8.3.1), they are still subject to time-iconic ordering restrictions like their lexical sources. The progressive use of *ti* is closely tied to the locative meaning of the lexical ‘stay’ verb: this is a very common grammaticalization path cross-linguistically and it has been argued that a locative semantic element is always present in lexical sources for grammatical expression of the progressive (Bybee, Perkins & Pagliuca 1994: 127–137). In examples like (647), the activity of ‘being at work’ follows the prior motion expressed by ‘go’, so the verbs are arranged in this order based on the principle of temporal iconicity that is found in word order in SVCs (Haspelmath 2016: 309).

This word order is also found with the less frequent initial prior motion verb *tammea* ‘get up’:

- (648) *lata-tammea* *di* *lata-lang ves-i*
 3PC.NFUT-get_up CONT.REAL 3PC.NFUT-search+for-3OBJ
 ‘They got up and were looking for it.’ [20170220g_n01s148_22]

However, the auxiliary verbs that most frequently intervene between the verbs in a prior motion SVC are the equivalent prior motion auxiliary verbs. This double expression of the prior motion component through serialisation and auxiliatation is a popular rhetorical strategy in Vatlongos. It is especially frequent with ‘come’, which occurs every 742 words in the subcorpus, but relatively frequent with ‘go’ as well, occurring once every 2100 words. It is possible that the lower figure for ‘go’ is an artefact of how the data has been coded, as the *ba* form of auxiliary *ha* is ambiguous with third-person singular subject non-future lexical *ha*, whereas the form of the ‘come’ auxiliary *mei* is more clearly distinguishable from its lexical equivalent. The high frequency of double expression of this concept demonstrates the importance of prior motion as a category in Vatlongos, showing why it would be prone to grammaticalization via auxiliatation. Multiple expression of grammatical categories is fairly rare cross-linguistically, although Cabellero and Harris (2012) find that it is more common than previously acknowledged. In Vatlongos there is a source construction for this configuration in verbal repetition of basic motion verbs to mark duration (§4.3.2).

Example (649) shows this strategy used for ‘go’ and ‘come’ prior motion in two immediately juxtaposed clauses. The clauses have exactly parallel structures and seem to be contrasting the two events and their prior motion paths.

- (649) *La-ba* *ba* *la-leh* *hura* *a-i*.
 3PL.NFUT-NFUT.go GO.REAL 3PL.NFUT-take *yam* CL.ED-3PL.POSS
La-mmei *mei* *la-din* *hura* *a-i*.
 3PL.NFUT-come COME 3PL.NFUT-NFUT.cook *yam* CL.ED-3PL.POSS
 ‘They go and get their yam. They come and cook their yam.’
 [20141219c_n01s033_03]

It is also possible to find clusters of auxiliary verbs in this position, with the same ordering discussed in §8.1.1.

- (650) *lata-mmei* *mei* *di* *lata-pat* *vongien*
 3PC.NFUT-come COME CONT.REAL 3PC.NFUT-sleep night
 ‘They come and are sleeping at night.’ [20170220g_n01s148_39]

There is one example of inflected *mas* ‘must’, followed by auxiliary *ha* ‘go’ and a main verb. This could be another example of an auxiliary verb intervening between the component verbs of an SVC (§6.4.11), but it is possible that *mas* has also been borrowed as a complementizing verb and this is a subordination structure.

- (651) *asu* *i-mas* *ha* *i-a-ni*
 rat 3SG.DFUT-must GO 3SG.DFUT-eat-3OBJ
 ‘The rat has to go and eat it.’ [20141219c_n01s033_14]

Rather than changing the analysis of auxiliary verbs as heading a clause, I argue that the emergence of these AVCs from prior motion SVCs has led to a reanalysis of their source construction.

One possibility is that the initial motion verb in such examples is treated as a complementizing verb, taking a full subordinate clause that begins with the auxiliary verb. Paamese uses such a subordination strategy to express prior motion (Crowley 2002b: 64–65). However, if *ha* ‘go’ and *ammei* ‘come’ can take full complement clauses, there should also be examples where the complement clause contains a subject NP, or is preceded by the general subordinator *xa*, but these do not occur in the corpus. It is also difficult to explain the interpretation of lexical and auxiliary ‘go’ or ‘come’ when they are used together (649). If the

initial motion verb and the auxiliary motion verb are in different clauses, they should be interpreted as separate motion events, but this interpretation is rejected by speakers.

Instead, it seems more likely that the initial 'go' or 'come' verb in such examples is left adjoined to the clause, and is structurally similar to an adverb appearing between the subject NP and the auxiliary verb (§3.3). Just as there appear to be semantic restrictions on the adverbs that appear in different slots, VPs appearing in this position could be restricted to those expressing prior motion. This pre-auxiliary slot is restricted to adverbs expressing relative time, restriction and modality, so the appearance of inflected *mas* 'must' expressing deontic modality in this position, (651), is not surprising.

8.2 Functions of auxiliary verbs

By definition, auxiliary verbs modify another predicate with a grammatical meaning, rather than contributing a separate predicate like a lexical verb (Anderson 2006: 4–5). However, as auxiliaries typically diachronically emerge from lexical verbs that have undergone semantic bleaching, it can be difficult to draw a clear division between grammatical meanings and highly generalised lexical verbs. Elements of the lexical meaning of verbs in the source construction are often retained by the grammaticised auxiliary verb (Bybee, Perkins & Pagliuca 1994: 15–19).

The functions of AVCs are 'most typically aspectual or modal' (Anderson 2006: 4), and the aspectual functions of *ti* and *ta* fit this cross-linguistic tendency. The inchoative sense of auxiliary *mei* and the change of state function of *ti* and *ha* are aspectual categories concerned with the boundaries of a situation. *Tit* and *tat* combine the aspectual functions of *ti* and *ta* with a minimising semantic element that can be translated with English 'just', and this could be analysed as a form of evaluative modality (Palmer 1986: 12–14, 119–121). The apprehensive mood (§4.2.6) shows that evaluative modality is already a component of a morphological TAM category in Vatlongos, so it is not surprising to find another composite grammatical category that involves the speaker's attitude in the system of auxiliary verbs. *mas* 'must' expresses deontic modality (Timberlake 2007: 329), and has been borrowed from the Bislama

auxiliary verb with the same grammatical meaning (and ultimately from English) (Crowley 2004b: 97).

Although basic motion verbs are a very prone to grammaticization (Bybee, Perkins & Pagliuca 1994: 5), and a common source for auxiliary verbs cross-linguistically (Anderson 2006: 345–352), it is less intuitive to conceive of prior motion as a grammatical category. Semantically, specifying the direction of a motion event preceding a situation expressed by another a verb appears to contribute a separate event, an impression bolstered by the need to express this using coordination or subordination when translating into English (Crowley 2002b: 11). However, prior motion is very frequently expressed in Vatlongos, either in SVCs (once every 98 words in the subcorpus) (§6.4.3) or in AVCs (once every 63 words), and sometimes a combination of the two (once every 550 words). Overall this means that information about a prior motion event is expressed once every 43 words, a similar frequency to the incontrovertibly grammatical category of continuous, which is only expressed every 67 words.

Nor is preoccupation with the deictic direction of motion preceding a situation a quirk of Vatlongos; it is very frequently expressed with SVCs in Oceanic languages. Cleary-Kemp (2015: 134) points out that all examples of ‘sequential’ (Lynch, Ross & Crowley 2002: 47) SVCs in Oceanic languages involve an initial motion verb followed by a verb expressing an event that follows that motion. The frequency of these constructions within specific Oceanic languages in Vanuatu has also been observed, such as in Tamambo (Jauncey 2011: 330–332). The prevalence of expression of prior motion in monoclausal constructions in Oceanic languages, and the morphological expression of prior motion in other language families, both suggest that prior motion is a sufficiently general semantic category to be incorporated into the grammatical system of a language. AVCs that express directional or orientational meanings are also found in other language families: they are common in South East Asian languages and have been shown to be diachronically stable in Turkic languages (Anderson 2006: 333; 2011: 801). This suggests that associated motion is not just an interim stage on a grammaticalization path towards more recognisably grammatical categories.

The Vatlongos prior motion auxiliary verbs are therefore glossed as ‘GO’ and ‘COME’, in order to make the case that they express a more grammatical meaning than their lexical equivalents; this meaning is restricted to prior motion uses of the lexical verb. However, there is a clear identity with the meaning of lexical ‘go’ and ‘come’ as the first verb in SVCs, so the categorical distinction in meaning and grammatical status suggested by the glossing conventions is a simplification.

Whether or not prior motion can be considered a truly grammatical category cross-linguistically, it is clear that in Vatlongos it can be expressed using the same formal construction as is used for the more typical auxiliary categories of aspect and modality. The phonological and morphological reduction of the prior motion verbs is typical of grammaticalization, where it is argued that formal and semantic reduction proceed in tandem (Bybee, Perkins & Pagliuca 1994: 19–21). Aside from the lack of verbal prefixes, *ha* undergoes a simplified pattern of inflection via verb-initial root mutation, and *mei* is phonologically briefer than the lexical verb root *mmei*, and one less syllable than the third-person singular non-future form *bemei*. The prior motion auxiliary verbs demonstrably do participate in Vatlongos AVCs, even if they arguably do not fit strict functional definitions of auxiliaries.

8.2.1 Continuous aspect with ‘stay’ verbs

The auxiliary verbs based on ‘stay’ verbs express progressive and habitual aspects, a combination I am labelling ‘continuous’.

Comrie (1976: 24, 33–35) defines the progressive as a subtype of imperfective aspect: when a situation is viewed from within, making ‘explicit reference to the internal temporal structure of a situation’. Progressive is then defined negatively as imperfective aspect that is neither habitual nor stative. For Dahl (1985: 91) as well, a difference between the progressive and a more general imperfective is the exclusion of stative situations: the progressive marks a dynamic situation that is ongoing at a point in time.

Rather than defining the progressive by the meanings it excludes, Bybee et al. (1994: 136) prefer to focus on the elements of meaning found in sources of

the progressive, some of which are eroded over time, accounting for differences in the extension of progressive in different languages. According to their cross-linguistic survey, strategies for marking the progressive always develop from a lexical source meaning that ‘an agent is located spatially in the midst of an activity at reference time’. As the construction is extended to more contexts, the restrictions to a spatial location, an agent, and an activity can be lost. The Vatlongos auxiliaries have lost these elements, and can be used with non-agentive subjects and developing states as well as dynamic activities. The fact that the Vatlongos continuous auxiliaries have developed from a locative lexical source could be evidence that the progressive meaning is older and has extended into habitual. Timberlake (2007: 302) characterises the progressive as marking an ongoing process that ‘could easily change or cease’ and also observes that progressive-marked situations are ‘often in conflict with (or even interrupted by) other situations’.

Example (652) shows a prototypical progressive situation: the speaker is describing something that is ongoing at the time of speech.

(652) *iaxa rumali ok di lu-dihi xie*
 which 3DU PROX CONT.REAL 3DU.NFUT-3OBJ.NFUT.cut there
 ‘which these two are cutting over there.’ [20150305f_p01s110_08]

In example (653) the continuous auxiliary is used to question a situation that is ongoing at a fictive time of speech:

(653) *Ma oum bitene “e, xouk di*
 then crab 3SG.NFUT.say hey 2SG CONT.REAL
o-pol-ni mi-dep xiak?”
 2SG.NFUT-do-TR 3SG.NFUT-NFUT.how now
 ‘Then the crab said “Hey, what are you doing now?”’
 [20141027a_n01m001_48]

A continuous auxiliary verb can also mark a situation that is ongoing at the topic time in a narrative.

(654) *lii da mavul*
 wood CONT.REAL 3SG.NFUT.break
 ‘The wood was breaking.’ [20170126c_n01e025_16]

The progressive use of the continuous auxiliary verbs is especially common when one situation is interrupted by another:

speaker went to school at home in the village on (most of) the relevant occasions (schooldays) during the stage of his life that he is currently discussing in this autobiographical narrative.

(658) *inou di nat-he sukul e tim Maat tang*
 1SG CONT.REAL 1SG.PRI-go_to school LOC home Maat just
 ‘I just went to school at home in Maat.’ [20150219b_n01m001_004]

On the other hand, in example (659), the speaker is marking the situation as characteristic or typical, in this case equating mat-making with her identity.

(659) *inou na-be atou xa*
 1SG 1SG.NFUT-NFUT.COP woman REL
[di na-pol-ni na, mat.]
 CONT.REAL 1SG.NFUT-do-TR HES mat
 ‘I’m a woman who makes um, mats.’ [20141215c_p01s063_01-02]

This characteristic function of the habitual also works with negated predicates. In example (660), the grouper refuses to take the rat to shore because he does not go to shore, he is a denizen of the reefs.

(660) *inou di na-taa-ve ut. Inou*
 1SG CONT.REAL 1SG.NFUT-NFUT.NEG-NEG.go_to shore 1SG
na-be moletin ta mai xil
 1SG.NFUT-COP.NFUT person of reef PL
 ‘I don’t go to shore. I’m a person of the reefs.’
 [20170122b_n01s036_20]

Using the same grammatical strategy to express progressive and habitual meanings, but not a general imperfective, is common in the region. Paamese uses ‘stay’ verbs as initial verbs in SVCs to mark progressive and habitual aspect (Crowley 2002b: 78) (§8.3). Languages of Epi also show this cogrammaticalization: Bierebo uses a verbal suffix (Budd 2009: 176) and Lewo uses a verbal particle (Early 1994: 244–246) to express the same combination of aspectual meanings. Although not all grammars of Vanuatu languages are explicit about the range of meanings encompassed by terms like ‘continuous’ or ‘durative’, it seems likely that other languages have similar patterns of grammaticalization, as the same pattern is found in Bislama with the auxiliary *stap*, another ‘stay’ verb (Crowley 2004b: 98). Although Dahl (1985: 94) finds that it is rare for a progressive to extend into a general habitual function (as

opposed to marking temporary habits), he does mention some exceptions: Quechua and Fitzroy Crossing.

This pattern may have implications for the modelling of aspectual categories. Comrie (1976: 25) put forward the hierarchical model shown in Figure 19 as a classification of possible aspectual oppositions. The category ‘nonprogressive’ refers to unbounded, non-dynamic situations, that is, states.

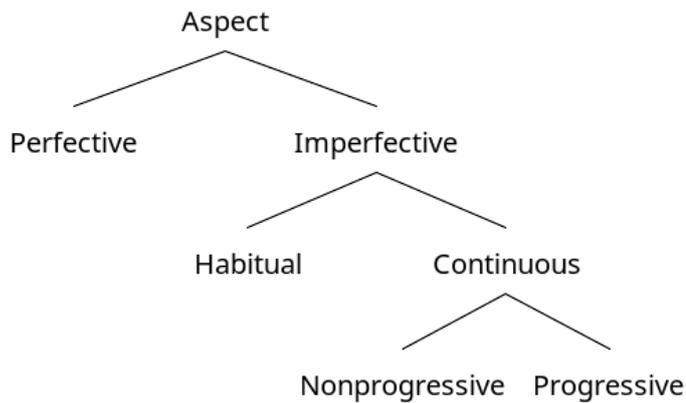


Figure 19: Comrie’s classification of aspectual oppositions (Comrie 1976: 25)

However, neither Dahl (1985) nor Bybee et al. (1994: 139) find evidence of any grammatical strategies in languages included in their surveys that express the ‘continuous’ or ‘nonprogressive’ categories postulated by Comrie. I therefore repurpose the term ‘continuous’ to refer to the combination of habitual and progressive found in Vatlongos and other languages of the region, giving the structure shown in Figure 20, where categories that have been found as marked with a single strategy cross-linguistically are marked in **bold**. As the term ‘continuous’ is not intended to claim any semantic identity between its component meanings, beyond the unbounded status that they also share with states, it is not included in the hierarchy, but is rather an overlay to the hierarchical semantic oppositions.

While the *ti* continuous auxiliary is more frequent than *ta*, it is not always clear what motivates the occasional choice to use *ta*. Parker's (1970a: 28, 33) dictionary does not have separate definitions for the intransitive verb and auxiliary verb senses of these forms, so it is not clear which parts of the definitions apply to the auxiliary or the lexical verb. However, he posits three main distinctions between *ti* and *ta*. Firstly, he states that *ta* has the form *ti* with non-singular subjects. Although there seems to be a preference for singular subjects with *ta*, there are examples of both the auxiliary (664) and the lexical verb (665) taking non-singular subjects, so this no longer seems to be a distinction between *ta* and *ti*.

- (664) *da* *lu-leherat* *na* *neta,* *hahau* *xil*
 CONT.REAL 3DU.NFUT-take_out HES what leaf PL
 'they were removing um what, the leaves.'
 [20141027a_n01m001_103-104]
- (665) *metimel* *xa* [*muta-da* *e-n* *xosali* *xie*]
 village REL 2PC.NFUT-NFUT.stay LOC-3SG.OBJ today then
 'the village that you were in earlier today' [20150226a_n01s098_18]

The second difference described by Parker is that *ti* is restricted to animate subjects. This is definitely no longer true of either the auxiliary (657) or the lexical (352) verb. Now *ti* seems to be the preferred default for all kinds of subjects: it is very frequent and in elicitation some speakers reject the use of *ta* as an auxiliary for any subject type. Auxiliary *tit*, which is derived from a compound of *ti*, also takes both animate and inanimate subjects. Of the 22 occurrences in the corpus, three are inanimate: the tail of a rat, the speaker's blood, and a fish (in the context of fishing). However, there is a tendency for auxiliary *ta* to occur with inanimate or non-human subjects: of the six examples in the corpus, only two examples occur with animate subjects, including (664), which refers to anthropomorphised animals, the crab and the rat. The only non-elicited occurrence of auxiliary *tat* in the corpus also refers to the rat. More typically auxiliary *ta* is used with inanimate subjects, like the wood in (654).

The final difference that Parker suggests between these forms is that *ta* is more definite with regard to 'time, place or other circumstances'. Lexical *ta* does seem to imply a specific location that is accessible to the speaker either because it is previously mentioned in the discourse or because it can be inferred from

the context (§3.5.3). The tendency is less clear for the auxiliary, which only occurs six times in the corpus. In the only example where it applies to an animate subject, it marks a locative predicate. The speaker is recapping plans for after the recording finishes, and the auxiliary *ta* immediately follows a lexical use of *ta* referring to the place of speech. Perhaps the choice of auxiliary is primed by the use of the lexical source.

(666) *ale* *Bell* *i-ta* *ma*
 then Bell 3SG.DFUT-stay then
ta ***i-he*** *tim* *Maat*
 CONT 3SG.DFUT-go_to home Maat
 ‘then Bell will stay on then he’ll be going home to [*Mele*] Maat.’
 [20150223a_n01m096_61]

8.2.2 Direction of prior motion with basic motion verbs

As discussed in §8.1.1, *ha* and *mei* indicate the direction of a motion prior to the situation described by the main verb, along a path away from or towards the deictic centre respectively. Although similar constructions are often analysed as purposive motion cross-linguistically (Huang & Tanangkingsing 2005: 318), examples like (667) show that, in Vatlongos, the auxiliaries mark motion preceding the main verb, regardless of whether the subject has a ‘purpose’: the flowing water is an inanimate force, rather than an agent, and therefore does not have an intention or purpose.

(667) *oei,* *mi-leh* *bemei,*
 water 3SG.NFUT-flow 3SG.NFUT.come
mei ***mi-gur*** *as*
 COME 3SG.NFUT-NFUT.carry ant
 ‘The water, it flowed over and came and carried off the ant.’
 [20170119h_n01s088_07]

The prior motion auxiliaries can modify their lexical equivalents:

(668) *la-nnem-i* *la-bit*
 3PL.NFUT-think-3OBJ 3PL.NFUT-NFUT.say
[atou *ak* ***ba*** ***ba]***
 woman PROX GO.REAL 3SG.NFUT.go
 ‘They thought that this woman had gone and left.’
 [20170331b_n01s034_29]

Auxiliary *mei* can also have an inchoative function when modifying stative predicates, like lexical *ammei* when used as an initial verb in an SVC (§6.4.10). It

can modify stative verbs (669), and adjectival (670) or nominal, (214), (255), predicates with the copular verb *he*.

- (669) *vatiang* *mei* *mi-seten*
 wind INCH 3SG.NFUT-strong
 ‘the cyclone became strong’ [20150419d_h01s046_04]
- (670) *mei* *be* *eilep*
 INCH 3SG.NFUT.COP big
 ‘it got big’ [20170224a_n01s141_180]

There are also a few examples of auxiliary *ha* preceding a stative predicate without a clear directional interpretation, which might be indicating an aspectual category similar to the inchoative or the ‘change of state’ interpretation of the continuous auxiliary:

- (671) *ba* *be* *te-metu* *kuhi*
 GO.REAL 3SG.NFUT.COP ADJZR-ripe properly
 ‘It goes/gets really ripe.’ [20141030a_p01m004_33]

8.2.3 Obligation with Bislama loan word *mas* ‘must’

There are three examples of the Bislama loan word *mas* ‘must’ used as an auxiliary verb in the subcorpus; it also occurs as an initial verb in an SVC taking verbal prefixes with the same meaning (§6.4.11). *Mas* is an auxiliary in Bislama occurring between subject markers and the verb (Crowley 2004b: 97). It conveys an obligation, often to do with social customs, (633), (672).

- (672) *mas la-vas* *vueili tei xa luho-n* *mi-kali*
 must 3PL.IFUT-IFUT.kill pig one REL tooth-3SG.POSS 3SG.NFUT-round
 ‘They have to kill a pig with a round tusk.’ [20150305f_p01s110_37]

8.3 Diachronic developments

AVCs appear to be a recent development in Vatlongos, so it is worth discussing their diachronic origin and defending this analysis. There are also signs that the auxiliary verbs may themselves be incorporated into the verbal prefixes.

8.3.1 Diachronic origin

I argue that Vatlongos AVCs have diachronically emerged from SVCs, especially prior motion SVCs (§6.4.3), and a similar, no longer extant, SVC function with an initial ‘stay’ verb. This argument is based on formal and functional similarities between the two constructions, as well as historical evidence from previous

descriptions of Vatlongos, and parallel functions of serialisation in closely related Paamese.

Apart from the lack of verbal prefixes on the auxiliary verb, which is the rationale for describing AVCs as a separate construction from serialisation, there are several formal parallels between AVCs and SVCs with an intransitive initial verb, following the patterns found with prior motion SVCs. Firstly, no subject NP can intervene between the two verbs in either construction. Intervening subject NPs are possible in subordination of complement clauses (§7.2.1), so this restriction favours serialisation over subordination as a likely source construction. Similarly, the subject-marking on the main verb is one factor determining the inflection of auxiliary *ha* 'GO'. This pattern of verb-initial consonant mutation suggests that both verbs would take identical subject-indexing in the source construction, as is the case in prior motion SVCs.

Another formal argument for identifying SVCs as the source construction of AVCs is the existence of bridging contexts where the two constructions are formally identical. The most important of these in terms of frequency of occurrence is the non-future with a third-person singular subject, where all the lexical sources of the auxiliary verbs can appear without prefixes (because they have a bilabial or prenasalised initial consonant, §5.1.1.2). The imperative mood addressing the second-person singular is another bridging context where all verbs can appear as a bare root (§5.1.1.3).

The main functional evidence for an SVC source for AVCs is the overlap of the function of *ha* and *mei* auxiliaries with their lexical equivalents as initial verbs in SVCs, including both the prior motion meanings (§6.4.3) and the inchoative sense of 'come' (§6.4.10). The participation of the Bislama loan word *mas* 'must' in both constructions is further evidence that speakers consider these two constructions to be functionally equivalent structures (§6.4.11). It suggests that AVCs have emerged fairly recently, as the Vatlongos-speaking community would only have been exposed to Bislama in the last century and a half at most. Some Vatlongos speakers may have been exposed to early forms of Melanesian Pidgin English on plantations in Queensland through the practice of blackbirding starting in the late nineteenth century (Crowley 1990), though

only a few returned. Contact with Bislama would become more significant with the presence of European traders and the progress of missionisation through Vatlongos-speaking communities in the early twentieth century (Tonkinson 1968: 40–42). There is no mention of *mas* in Parker’s (1970a) dictionary, though he does include many other Bislama loanwords. This suggests that *mas* could have been borrowed into Vatlongos in the late twentieth century, and is used both as an inflecting initial verb in SVCs and an invariant auxiliary.

The final piece of functional evidence is the synchronic absence of a ‘continuous’ function of serialisation in Vatlongos, which, to anticipate the next point, might be expected on the basis of its presence in Paamese (Crowley 2002b: 77–78). Cleary-Kemp (2015: 136) finds that ‘continuous’, ‘imperfective’, ‘progressive’ or ‘habitual’ aspect marking is the most common aspectual function of SVCs in Oceanic, usually expressed with a posture or locative verb in initial verb, most often ‘stay’, as in Paamese. The lack of this function of serialisation in Vatlongos suggests that continuous *ti* is the oldest auxiliary verb in Vatlongos and has now fully replaced serialisation as a strategy for expressing this aspectual distinction.

A serialisation strategy in Paamese with a very similar function to the Vatlongos continuous auxiliaries provides evidence of what a Vatlongos serialised source construction might have looked like. Crowley describes ‘aspectual’ (Crowley 1987: 56–58) or ‘auxiliary’ (Crowley 2002b: 77–79) serialisation as a function of core-layer serialisation in Vatlongos. These Paamese SVCs mark the same combination of progressive and habitual aspects as the Vatlongos continuous auxiliaries, using two verbs meaning ‘stay’ in initial position. Example (674) shows the expected morphological TAMP dependency between the negative realis initial verb and affirmative immediate future on the subsequent verb (Table 37, Table 38).

(673) *inau nado: nagu:ken rais*
inau na-doo na-guukeni raisi
 1SG 1SG.REAL-stay 1SG.REAL-cook rice
 ‘I am cooking rice.’ Paamese (Crowley 2002b: 78)

(674) *inau naromultei ma:n veta*
inau na-ro-mule-tei ma-ani vetaa
 1SG 1SG.REAL-NEG1-stay-NEG2 1SG.IFUT-eat breadfruit
 ‘I don’t eat breadfruit’ Paamese (Crowley 2002b: 78)

Paamese SVCs also express inchoative aspect using a verb meaning ‘come’ in initial position, whereas Vatlongos can use either serialisation (§6.4.10) or auxiliary *mei*.

(675) *kosa ro:mai roloŋoloŋo*
kosaa roo-mai ro-loŋoloŋo
 now 1PL.INCL.REAL-come 1PL.INCL.REAL-knowledgeable
 ‘Now we have become knowledgeable.’ (Crowley 2002b: 78)

However, unlike Vatlongos, Paamese does not use serialisation to express prior motion. Instead, the basic motion verbs can take a subordinate clause that Crowley (2002b: 64–65) analyses as a purposive construction.

More concrete evidence for the emergence of auxiliary verbs from SVCs comes from Parker’s research on Vatlongos in the late 1960s, especially his dictionary. Parker (1970a: viii) identifies three auxiliary verbs in Vatlongos, *ti*, *ta* and *tit*, and comments that they are ‘very frequent’ but ‘untranslatable’. Although he does not identify SVCs in Vatlongos, putting together evidence from his descriptions and examples supports the existence of SVCs and their functional overlap with AVCs.

The syntactic distribution of auxiliary verbs in Vatlongos in the 1960s seems to be the same as today. Parker states in his introduction that auxiliary verbs ‘immediately precede finite verbs’ (Parker 1970a: viii), but there is an example of an adverb intervening between the auxiliary and the lexical verb, *tamu* ‘still’ which he finds is ‘always preceded by *ti*’ (Parker 1970a: 30).

(676) *di tamu la-kkora*
 CONT.REAL still 3PL.NFUT-crow
 ‘They’re still crowing.’ (Parker 1970a: 30)

However, the morphological possibilities for auxiliaries at this time suggest that these ‘stay’ verbs were used in both SVCs and AVCs: auxiliary verbs were ‘optionally inflected for person’. At this earlier stage of Vatlongos, the ‘stay, continuous’ verbs resembled prior motion *ha* and *mei* today. In the terms of the

analytical framework put forward here, the ‘stay’ verbs could previously participate in SVCs as an initial verb inflected for person, or participate in AVCs as bare roots without verbal prefixes. When not inflected for person, Parker says they have only two forms an ‘indicative’ (i.e. non-future) initial /d/ form and a ‘non-indicative’ initial /t/ form, which matches the behaviour of the continuous auxiliaries described here. There is only one example of an inflected auxiliary verb in Parker’s dictionary. Unfortunately, it is modifying a verb which is not familiar to speakers today so its exact interpretation and the contribution of the auxiliary is difficult to confirm:

- (677) *o-di* *o-lele-ni* *neh?*
 2SG.NFUT-NFUT.stay 2SG.NFUT-have_premonition-TR what
 ‘What do you have a premonition of?’ (Parker 1970: 33)

The motion verbs are not identified as auxiliary verbs in Parker’s dictionary, but in an earlier article he analyses *ha* as an optional prefix that precedes the other verbal prefixes and ‘indicates that the action is realised in some place other than where it is spoken of, or that some distance must be covered prior to the realisation of the action’ (Parker 1968b: 28). The two examples he gives: *ha-mu-pat* (678) and *ha-lati-pat* ‘GO-3PC.DFUT-sleep’ are both in TAM contexts where the *ha* form of the auxiliary is expected (§8.1.2). Although Parker retracts this analysis by the time of the 1970 dictionary, pre-verbal *ha* must have been frequent enough, and sufficiently semantically bleached, to warrant an analysis as a morphological affix, so this could be evidence of an emerging auxiliary verb. Parker may have been reluctant to analyse the motion verbs as auxiliaries as their function is less obviously grammatical than the continuous auxiliaries. However, there are examples of *ha* included in the dictionary that can be analysed as AVCs, (678), (679), and prior motion SVCs (680).

- (678) *ha* *mu-pat*
 GO 2PL.IMP/DFUT-sleep
 ‘You (plural) go sleep’ (Parker 1968b: 28)
 ‘You (plural) will go sleep’ (Parker 1970: 4)
- (679) *visuvong* *sung* *ha* *ni-pus-i*
 tomorrow after GO 1SG.DFUT-see-3OBJ
 ‘Tomorrow I’ll go see him.’ (Parker 1970: 28)

(680) *u-ha u-hiteni mi Esron kestang*
 2SG.IFUT-go 2SG.IFUT-say to Esron only
 'Go tell it to Esron alone.' (Parker 1970: 10)

There is also an example of lexical *ammei* 'come' in a prior motion SVC, (681), and the *mei* form of 'come' that is now used in AVCs is listed as an alternative third-person singular indicative (non-future) variant of *ammei* (Parker 1970a: 16).

(681) *mi-dep ma o-mmei*
 3SG.NFUT-NFUT.how then 2SG.NFUT-come
o-purun nim ma-van
 2SG.NFUT-dirty house CL.DR-1SG.POSS
 'Why have you come and dirtied my house?' (Parker 1970: 43)

The overall picture suggested by the snapshot of Vatlongos in Parker's dictionary, the use of SVCs in closely-related Paamese, and the synchronic functional overlap between AVCs and SVCs in Vatlongos, is the development of auxiliary verbs as the optional elision of verbal prefixes for some verbs that frequently occur as the initial verb in an SVC with a grammaticalized function.

For the continuous auxiliaries, which were probably the first to follow this grammaticalization path, the unprefixed version then became the only acceptable variant and their use in SVCs was then lost. The basic motion verbs *ha* 'go' and *ammei/mei* 'come' can still participate in both constructions with the same or very similar functions to express prior motion or inchoative aspect. Example (540) above shows a speaker using both in quick succession. The AVC is used to repeat the SVC after a speech disfluency, suggesting that this functional equivalence is evident to speakers. This might suggest that the prior motion verbs entered this process of auxiliation diachronically later than the 'stay' verbs, or that their prior motion meaning is closer to a lexical meaning, maintaining their use as lexical verbs in SVCs. When the Bislama auxiliary *mas* is borrowed into Vatlongos to express obligation, the functional parallel between SVCs and AVCs has sufficient psychological reality for speakers that it can participate in both constructions.

8.3.2 Possible incorporation into the verbal affix template

A further diachronic development of auxiliary verbs could be incorporation into the verbal affix template, being reanalysed as prefixes preceding the other verbal prefixes shown in Figure 10. Reanalysis of auxiliary verbs as affixes is a common grammaticalization path cross-linguistically (Anderson 2006: 334), and Parker’s (1968b: 28) subsequently discarded analysis of *ha* as a verbal prefix demonstrates the potential ambiguities in the distribution of auxiliary verbs. Only the occasional occurrence of adverbs intervening between the auxiliary and lexical verb (§8.1.1) categorically rule out this analysis, so it would seem to be ripe for reanalysis by learners of the language.

There are a few examples suggesting that some speakers in Endu may be reanalysing the continuous auxiliary verb *ti* in this way. For one Endu speaker, the non-future *di* form of the continuous auxiliary consistently undergoes vowel harmony before the third-person dual subject prefix *lu-*, realised as *du*. Although vowel harmony is not used consistently by Vatlongos speakers, it elsewhere occurs within the verbal prefixes (§5.1.1.2), suggesting that this speaker is at least treating the auxiliary as part of the same phonological word as the verb, and perhaps the same grammatical word.

(682)	<i>Du</i>	<i>lu-ba,</i>	<i>du</i>	<i>lu-long,</i>
	CONT.REAL	3DU.NFUT-NFUT.go	CONT.REAL	3DU.NFUT-hear
	<i>na,</i>	<i>ra</i>	<i>haromue</i>	<i>na-lu</i>
	HES	voice	boy	CL.GEN-3DU.POSS
	‘They were going, they were hearing, um, the voice of their boy.’			
	[20141105e_n01e012_26-27]			

An example from another Endu speaker shows even greater phonological integration with the verbal prefixes. The combination of continuous auxiliary verb *ti* (here marking the habitual) with the second-person singular distant future prefix is realised as a single syllable, *tu*, suggesting full integration into the verbal prefix system in this context:

(683)	<i>tu-pol-ni</i>	<i>nenemien</i>
	CONT.2SG.DFUT-do-TR	plan
	‘you make a plan’ [20141106d_p06e016_14]	

8.4 Summary

Vatlongos AVCs appear to have diachronically emerged from SVCs, via semantic bleaching and formal reduction of a grammaticalized initial verb. On the basis of their syntactic distribution, they can be analysed as the syntactic head of the clause, although the lexical verb is the head for morphological inflection. They are analysed as a subclass of verbs, because, depending on the patterns of their lexical source, they undergo verb-initial consonant mutation to inflect for some TAM and subject-agreement combinations. However, these patterns are somewhat reduced, and there are signs that Vatlongos auxiliaries could be further deverbalized, and perhaps incorporated into the verbal morphology.

9 Variation in frequency of verbal constructions

The main aim of this project is to investigate possible differences in how Vatlongos is spoken in Mele Maat compared to communities in Ambrym, looking for early signs of dialect divergence after the separation of two communities (§1.1.2, §1.2.1). It also aims to determine whether any differences are linked to pressures of an urban environment, where there is intense contact with Bislama and English compared to a rural environment (§1.2.2). Endu village is also distinguished from other villages on Ambrym (Ase-Taveak) due to observed differences in the sociolinguistic context, and community recognition of a separate dialect (§1.1.2, Chapter 2). The descriptive chapters of the thesis have identified some formal differences in the verbal morphosyntax in the behaviour of negative clitic (§3.6.2.7, §3.6.2.10) and the frequencies of morphological variants (§5.1.1.2, §5.2.1). This chapter compares the frequency of use of verbal constructions in the subcorpus (henceforth, ‘construction frequency’) across texts by speakers in all three communities. If constructions are used at different rates, this could reflect differential dispersion of variants in separate communities, resulting in different dialectal standards (§1.2.1).

A further research question is whether any observed differences in construction frequency are fairly arbitrary regional features, or better explained by other sociolinguistic factors, such as speaker age and level of education, or text genre. If the latter, different construction frequencies could reflect different speech styles or registers, which are socially meaningful for all speakers of Vatlongos, beyond identifying a speaker’s village or community allegiance. Alternatively, construction frequency differences could indicate simplification or convergence correlating with an individual speaker’s exposure to Bislama and English. In addition to speaker community membership, this chapter therefore examines whether construction frequency varies according to other speaker characteristics (age, gender and years of education) and text genre. Finally, the chapter tries to model the frequency of each construction per text, to see which of these factors are the best predictors of construction frequency.

The quantitative analysis only pertains to the subcorpus of spontaneous texts with full metadata for speakers (§1.3.3). In this chapter I therefore refer to this subcorpus as ‘the corpus’, and use the term ‘subcorpus’ to refer to the subsections of the spontaneous corpus divided by community or other factors, e.g. ‘the Endu subcorpus’.

Before discussing the data and results, I outline the procedures followed for tagging and processing the corpus, and some tentative hypotheses.

The corpus was transcribed and translated into English and Bislama using the SIL Saymore (Moeller 2014) autosegment feature, creating time-aligned ELAN files (MPI Nijmegen 2018) (§1.3.2). While this saved time and ensured it was easy to check audio throughout the workflow process, it had the disadvantage of creating transcription segments that did not directly align with syntactic constituents; these segments tended to be larger than a clause or sentence. The transcriptions were interlinearised using a lexical database and semi-automatic parsing through SIL’s Fieldworks Language Explorer (FLEX Development Team 2018). The FLEX interlinear text template employs the Emeld four-level format (EMELD Project 2000; Bow, Hughes & Bird 2003; Hughes, Bird & Bow 2003), as shown in Figure 21.

56	Word	Asu	su		ba	migol		ueili	bemei	
	Morphemes	asu	su		ba	mi-	gol	ueili	⊙-	bemei
	Lex. Entries	asu	sung + dial. var. of		ha ₂	mi-	xol ₁	ueili	mi-	ammei
	Lex. Gloss	rat	after		go.ind	3sg.nfut	chase.ind	pig	3sg.nfut	come.ind
	Lex. Gram. Info.	n	adv		VAux	v:SubjTAM	vt	n	v:SubjTAM	vi
	Word Gloss	rat	after		go.ind	3sg.nfut-chase.ind		pig	3sg.nfut-come.3sg.ind	
	Word Cat.	n	adv		VAux	vt		n	vi	

Free en Then the rat went and chased the pig over.

bi Rat nao i go ronem pig i kam.

Note b d

Figure 21: Screenshot of a segment of text in the SIL FLEX database showing morpheme-level and word-level glossing, with sentence-level translations and tags [20141027a_n01m001_56].

In addition to morpheme-level and word-level tagging, syntactic constructions were tagged in a separate ‘note’ field that scopes over a time-aligned segment. The tags ‘b’ and ‘d’ in Figure 21 mark the presence of an AVC and an SVC respectively. The constructions discussed in this section are AVCs (Chapter 8), SVCs (Chapter 6), complex predicates (§7.1) and zero-marked complement

clauses (§7.2.1). In addition, SVCs used to express prior motion (§6.4.3) were tagged separately from other SVCs. Originally this was because it was unclear whether they should be analysed as SVCs or as a form of zero-subordination (like similar constructions in Paamese, Crowley 2002b: 64–65). However, it was decided to continue to tag them separately due to their high token frequency, their functional overlap with AVCs, and the unexpected word order when they cooccur with AVCs (§8.3.1). The findings in this chapter show that they do pattern differently from other SVCs.

Some other constructions that were tagged in the corpus were excluded from this analysis due to their low token frequency. All the constructions discussed in this chapter occurred at least once every 100 words (Table 42), whereas the next most frequent construction, complementiser-marked complement clauses, occurred only once every 273 words. Although these overtly marked complement clauses are not discussed separately, they are sometimes added to the zero-marked complement clauses so that patterns in the use of all complement clauses can be separated from the ratio of complementiser-marking to zero-marking, where it elucidates the patterns in frequencies of zero-marked complement clauses (Table 44, Table 47).

The corpus of interlinearised and tagged texts was exported from FLEx as verified generic XML, and then imported into R (R Core Team 2018) as a set of linked dataframes using the *interlineaR* package (Loiseau 2018). In addition, the results of the metadata survey discussed in Chapter 2 were added to the dataframe of texts, as well as text genre, in order to calculate frequencies by community and other text-level factors (speaker characteristics and genre). The data analysis was conducted in the RStudio environment (RStudio Team 2018), in an Rmarkdown notebook (Allaire et al. 2018).⁴⁷

⁴⁷The R packages used in conducting the analysis and generating plots are listed here: COUNT (Hilbe 2016); dplyr (Wickham et al. 2018); ggmosaic (Jeppson, Hofmann & Cook 2018); ggplot2 (Wickham 2016); interlineaR (Loiseau 2018); MASS (Venables & Ripley 2002); msm (Jackson 2011); lattice (Sarkar 2008); msme (Hilbe & Robinson 2018); reshape2 (Wickham 2007); sandwich (Zeileis 2004); stringr (Wickham 2018); tidyverse (Wickham 2017); xml2 (Wickham, Hester & Ooms 2018).

The main hypothesis of this project is that there is a difference in the frequency of use of verbal constructions between the communities. Beyond testing that hypothesis, the quantitative methods employed in this chapter are designed to be explorative, rather than aimed at testing falsifiable hypotheses. However, based on findings in the literature on language change (§1.2), the investigation of the Vatlongos-speaking community contexts described in Chapter 2, and the qualitative analysis of the different constructions presented in the rest of the thesis, it is possible to make some tentative predictions about how and if the frequencies might differ for different constructions.

Comparing Mele Maat to Ase-Taveak, Chapter 8 argues that AVCs were a recent innovation at the time that Mele Maat was established, and have been undergoing change during the time that the speaker communities have been separated, so it seems likely that the communities would diverge in the textual frequency of AVCs. AVCs in Vatlongos express similar categories to auxiliaries in Bislama, but the order of information is different: in Vatlongos the auxiliary precedes subject and negation marking, while in Bislama the auxiliary follows subject or negation marking and is immediately adjacent to the verb (Crowley 2004b: 96). Therefore, it is difficult to predict the effect that a stronger influence of Bislama in Mele Maat could have on the frequency of these constructions.

SVCs are not available in English, so convergence towards English could make them less likely to be used in Mele Maat. However, Vatlongos SVCs have functional overlap with Bislama SVCs (§6.4, Crowley 2004b: 166–172), so exposure to Bislama could favour their use. On the other hand, Vatlongos SVCs are considerably more morphologically complex than their Bislama equivalents (§6.3, Meyerhoff 2000: 160–176; Crowley 2002b: 218–232), especially in that subject-indexing is indexed on the subsequent verb, and their complex morphological TAMP dependencies. Again, it is difficult to make community-level predictions about this construction.

Prior motion SVCs (§6.4.3) on the other hand are less likely to have similar frequencies of use in Mele Maat and on the island, as they are also the only type of SVC to have a parallel, though limited, equivalent in English, as discussed by

Crowley (2002b: 11). They are also less likely to have complex morphology than other SVCs, as they are always same-subject and rarely occur in negative polarity (never in the spontaneous corpus). This means that their surface realisation is more similar to Bislama SVCs.

Complex predicates in Vatlongos have limited productivity and could index more archaic speech styles, so could be expected to be less frequent in Mele Maat. In Bislama the equivalent construction is also not fully productive: a limited set of verbs can function as the subsequent verb in these constructions, and for some groups of speakers they are only available with fixed collocations (Crowley 2002b: 224–225; Crowley 2004b: 85–88). However, Crowley also observes that they are more productive for younger speakers of Bislama.

Zero-marked complement clauses are available in Vatlongos, Bislama and English so are predicted to have the same frequency in all three Vatlongos-speaking communities.

An alternative hypothesis relates to observations about linguistic changes in language shift and attrition (§1.2.2). The vitality of Vatlongos in Mele Maat is lower than in Ase-Taveak (§2.5), and, if language attrition is affecting the Mele Maat community, all these complex constructions could be less frequent in Mele Maat than on the island (Schmidt 1985: 121; 1991: 119; Maher 1991: 68; Mühlhäusler 1996: 291).

In Endu it was observed that Vatlongos was more threatened than in the other Ambrym villages due to competition with North Ambrym for the same domains (§2.5). As a result, Endu might pattern with Mele Maat, despite Endu speakers having more regular communication with speakers from Ase-Taveak. On the other hand, the emphasis on a conservative language ideology in Endu (§2.4) could lead to a higher frequency of complex predicates, which seem to index a more conservative style of speech.

9.1 Frequency of constructions as count data

The frequencies of constructions in individual texts or subcorpora are count data (Hilbe 2014), a kind of discrete (Baayen 2008: 44–57) or nominal (Oakes 1998: 24) data, where discrete events or individuals are counted in periods of

time or regions of space: in this case the number of tokens of constructions in stretches of text. Unlike continuous data, count data cannot have negative or non-integer values, and therefore usually does not meet the assumptions of a normal distribution. It is usually modelled using a Poisson distribution (Baayen 2008: 53–57; Hilbe 2014: 35–72), or modifications of a Poisson distribution (Hilbe 2014: 9–15). These include the negative binomial model (Hilbe 2014: 126–160), used in the regression analyses of construction frequencies by text in §9.5.

Within linguistics, Poisson regression has been advocated for modelling the frequency of syntactic constructions by word count, like the data in this chapter (see Gries 2010 discussing data in Hundt & Smith 2009). This is a logical extension of Poisson as a model of count data: constructions are instantiated in groups of word but are not lexical items themselves, instead word count can be thought of as similar to a measure of units of time in a Poisson model of events over periods of time. Although lexical frequency is normally modelled as binomial data, on the basis that each word in a corpus can be treated as an experimental trial and succeeds and fails on the basis of identity with the lexeme being examined, the Poisson distribution has also been advocated as an efficient model of word frequency when an individual lexical item has a very small frequency relative to the size of a corpus (Baayen 2001: 45; Baayen 2008: 44–57).

Count data is assumed to occur at a certain rate (or mean), with some level of variation from that rate. When I have described construction frequencies so far in the thesis, I have expressed them as the average number of words it takes for a construction to occur in the subcorpus. While this is an intuitive way to think about the rate of very frequent constructions, it is standard practice to express rate of occurrence as a normalized relative frequency over a set number of words, which allows for comparison with other corpora. Table 42 shows the rates as expressed elsewhere in the thesis, and the relative frequency per 10,000 words. 10,000 words was chosen as the base for normalization as it is roughly the size of the smallest community subcorpus (Endu, 9842 words).

Table 42: Rate of occurrence and normalised relative frequency of tagged verbal constructions in the corpus

	Auxiliaries	SVCs	Prior motion SVCs	Complex predicates	0 complement clauses
1 every _ words	34	25	98	98	63
Frequency (per 10,000 words)	293	400	102	102	159

When the rate of occurrence of a construction is high, a Poisson distribution closely resembles a normal distribution (Baayen 2008: 54): a symmetrical bell-shaped density curve that can be derived from the mean and standard deviation of the data (Oakes 1998: 3–6; Baayen 2008: 58–63). Figure 22 is a histogram of the normalised relative frequency of SVCs in the corpus, the most frequent construction. Despite some gaps at the higher frequencies, it fits the normal distribution curve fairly well:

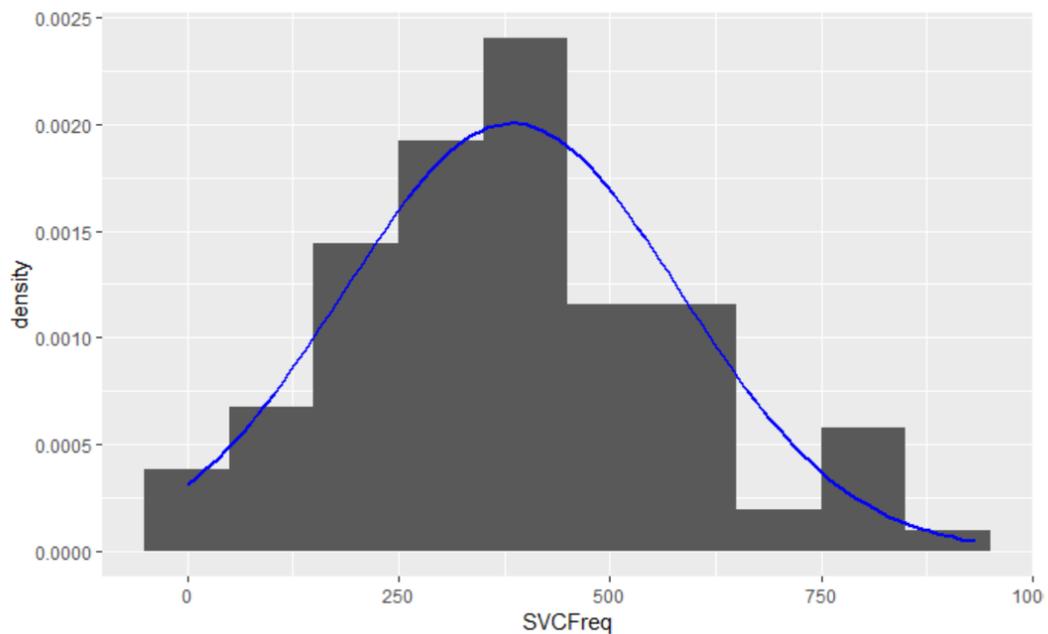


Figure 22: Histogram of normalised relative frequency of SVCs per 10,000 words in each text

On the other hand, at lower average relative frequencies, the Poisson distribution predicts a positively skewed distribution (Oakes 1998: 3–4), approaching negative exponential distributions when the mean is less than 1 (Hilbe 2014: 28). The distribution of complex predicates, one of the least

frequent constructions discussed in this chapter, is heavily skewed with a long tail:

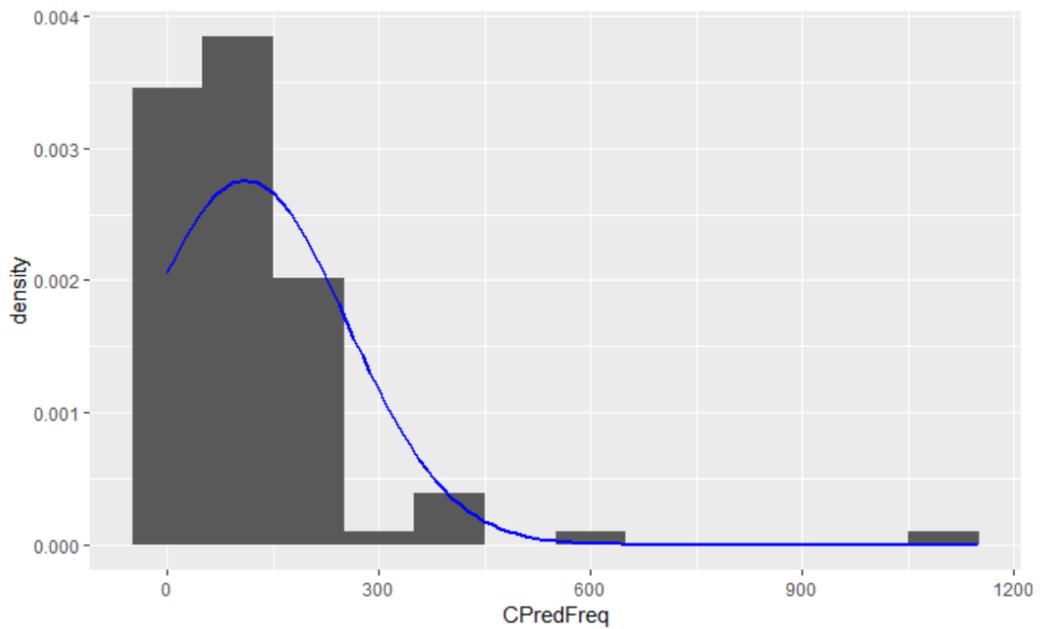


Figure 23: Histogram of normalised relative frequency of SVCs per 10,000 words in each text

One fundamental assumption of the Poisson distribution is that the data is equidispersed: the variation equals the mean (Hilbe 2014: 9). Comparing the boxplots for the normalised frequency of constructions per text (Figure 24), it is clear that the more frequent constructions (AVCs and SVCs) have more variance than the less frequent constructions (e.g. prior motion SVCs and zero-marked complement clauses).

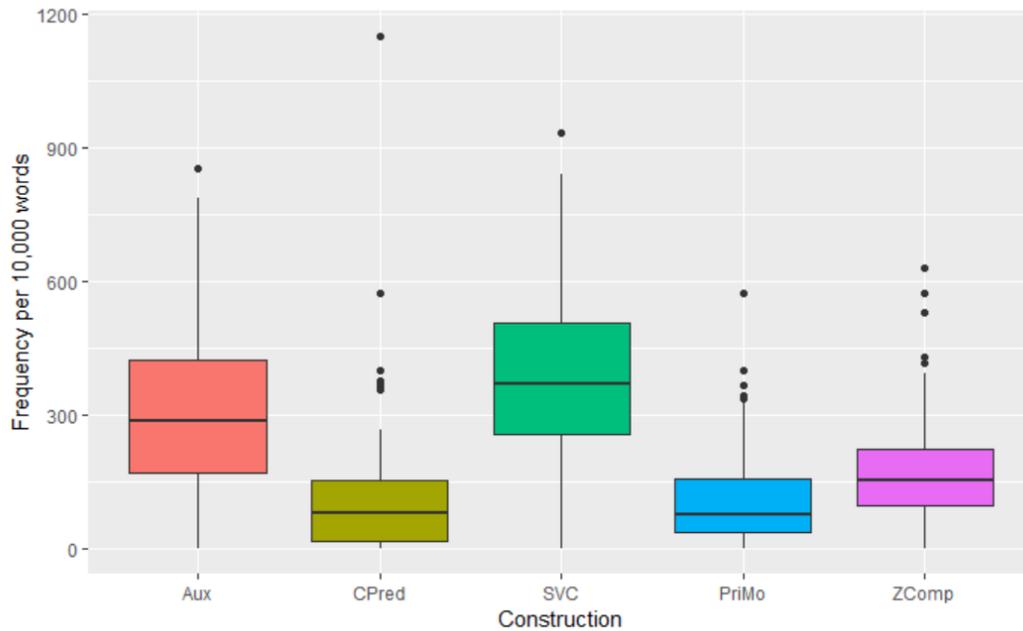


Figure 24: Boxplots of normalised relative frequency per 10,000 words in each text for each construction

However, as Hilbe (2014: 9, 38, 74) stresses, most real-world data is Poisson-overdispersed: the variance is greater than the mean. There are clear linguistic reasons for expecting the construction frequency in a text to be overdispersed. One assumption of the Poisson distribution is that the observed events are independent: one event does not affect the likelihood of another event occurring. However, there are linguistic reasons to expect proneness in the data (Hilbe 2014: 81); the occurrence of a particular construction makes it more likely that the same construction will be used again in the same text. This could be due to exact repetition of a construction with the same lexemes. This is a common feature of spoken language, either due to disfluencies (Foster, Tonkyn & Wigglesworth 2000: 368; Blanche-Benveniste 2006: 60–62; Moreno Sandoval & Guirao 2006: 205), or as an emphatic strategy (Jung & Himmelmann 2011: 211). Some of these verbal constructions can also be repeated through specific repetition strategies that are found in Vatlongos: to mark duration or intensification (§4.3.2, §4.3.3), or in tail-head linkage (see example (355) and discussion there). Proneness of these constructions can also result from syntactic or structural priming, whereby the structure itself, independently of the lexemes involved, is likely to reoccur. This phenomenon has been demonstrated with both experimental and corpus-based approaches (Pickering

& Branigan 1999; Gries 2005; Pickering & Ferreira 2008). It is therefore not surprising that the frequency of constructions in a text is Poisson overdispersed, as there is a positive correlation between individual occurrences of constructions.

Negative binomial regression is a standard approach to modelling overdispersed Poisson data (Hilbe 2014: 33), and this is the model that is used to model construction frequency by text in §9.5. Unlike Poisson regression, it models dispersion as a separate parameter from the mean, allowing for variance to be greater than the mean. This is a more pressing issue in the modelling of frequency per text, but to a lesser extent this effect is also present in the collections of texts that make up community-level and other subcorpora analysed in §9.3 and §9.4.

9.2 Composition of the corpus

The composition of the corpus poses a challenge to addressing the second research question: whether community level differences can be accounted for by other sociolinguistic factors. The content of texts in the corpus has been determined by the priorities of the community and the wishes of individual contributors (§1.3.3). It is an ‘opportunistic’ corpus in the sense of Woodbury (2003: 13), or a ‘convenience’ corpus (Lüpke 2005: 97), rather than a balanced sample of genres and speakers from each community. It therefore fails to meet standards pursued in variationist corpus linguistics for major languages like English, where researchers are advised to carefully control the variables in corpora being compared, in order to isolate differences that result from the factor of interest to a particular study (Oakes 2009: 160). Rather than surrendering the field of variationist and quantitative linguistics to major languages like English, research on endangered languages should prioritise project design and statistical tools that can account for the biases in typical language documentation corpora, without sacrificing the valuable insights that quantitative approaches can bring to the task of describing and analysing understudied languages (Lüpke 2005; Meyerhoff 2017).

Some of the biases in the composition of the corpus reflect the sociolinguistic contexts and attitudes discussed in Chapter 2, while others are

random or brought about by arbitrary logistical constraints during fieldwork. Figure 25 shows the breakdown of the 104 texts in the corpus by age group, community membership, gender, genre and years of education. Younger speakers were more reluctant to contribute to the corpus, especially in Endu and Mele Maat. Ase-Taveak communities contributed more to the corpus due to high enthusiasm for the project and high language vitality. This was despite more researcher time spent in Mele Maat than on the island to try to counteract this tendency. The gender balance in the corpus is reasonable, with slightly more men contributing than women. More than half of contributors only received primary schooling, which is a reasonable reflection of the community as a whole; the national average is 6.8 years of education (UNDP 2016).

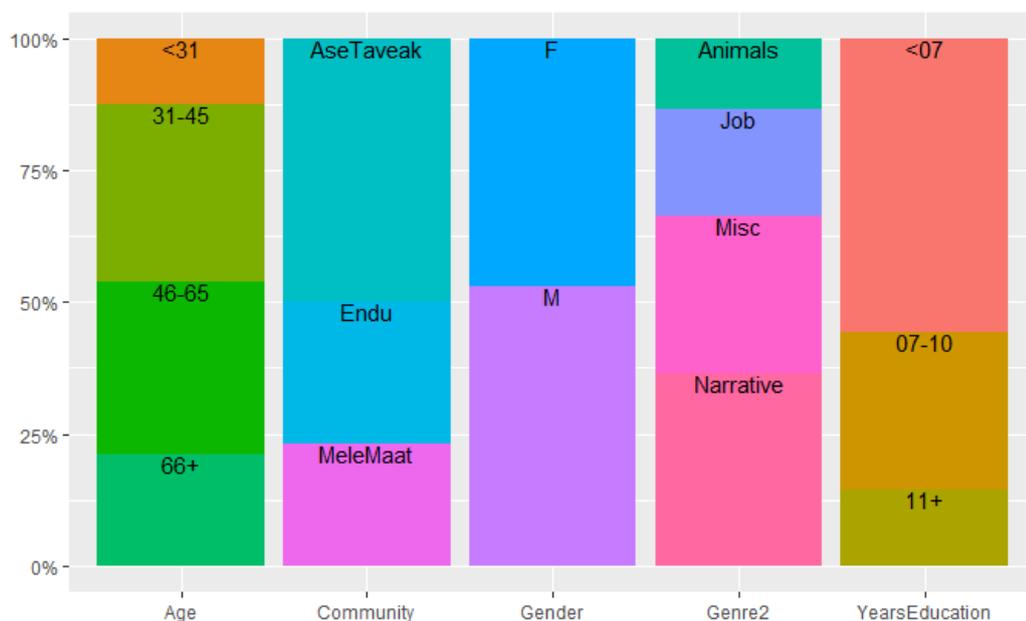


Figure 25: Composition of texts in the corpus by age group, community, gender, genre, and years of education.

However, an additional confounding factor is the length of texts contributed by different speakers in different genres. Comparing Figure 25 with Figure 26, there is a tendency for speakers who are older, from Mele Maat, male, and more educated, to contribute longer individual texts. Similarly, the ‘narrative’ genre, which includes prestigious custom stories and disaster narratives, tends to contain texts that are relatively long, whereas texts in the ‘miscellaneous’ genre tend to be shorter. This is mostly because lots of relatively short procedural texts are included under this category. This meant that despite a lot of interest

in procedural texts, they were underrepresented in the corpus and did not make up a large enough proportion to be grouped separately (for a discussion of these categories and category boundaries see §9.4, Table 48). For the rest of the diagrams in this section, I use word counts rather than text counts, as they are a better reflection of the imbalances in the corpus affecting the chi-squared tests in §9.3 and §9.4.

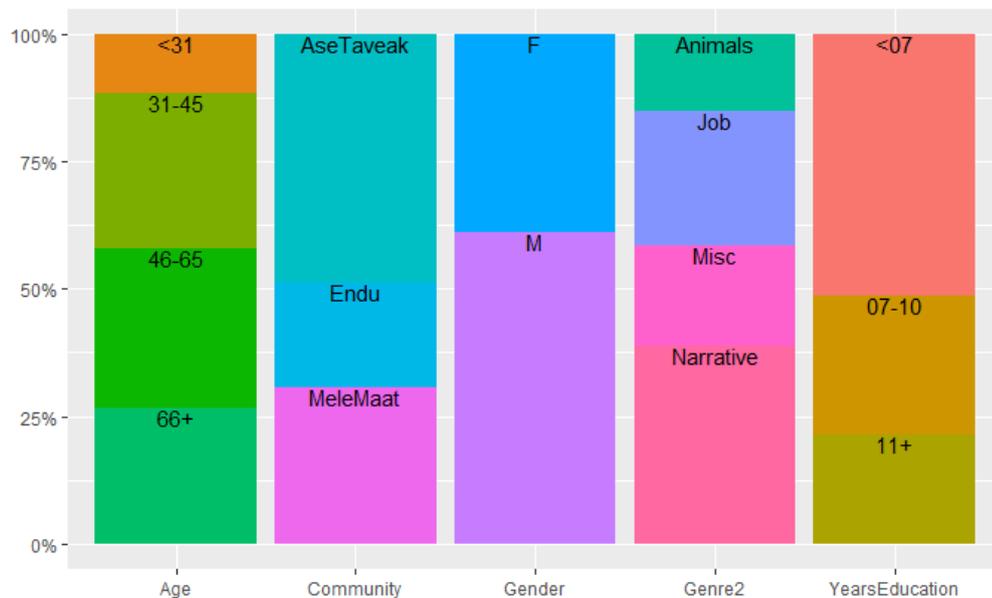


Figure 26: Composition of words in the corpus by age group, community, gender, genre, and years of education.

It is important to examine the representation of these different speaker characteristics and text types by community, as effects at the level of community could be artefacts of differences in the sample composition for each community. Mosaic plots (Baayen 2008: 111–113) visualise these imbalances: the area of each box represents the numbers of words in the corpus produced by speakers in each community in each age group, gender, education level, or genre. The differences in the areas of the boxes reflects the imbalances in the composition of the corpus.

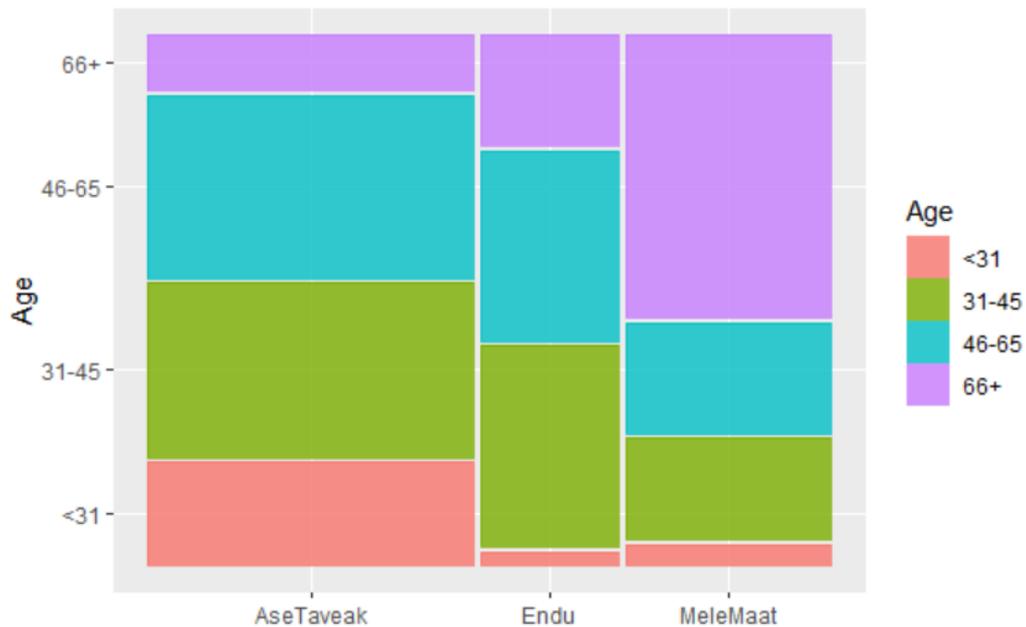


Figure 27: Mosaic plot of words in the corpus by community and age group

First, we examine the interaction between community and speaker age. Figure 27 confirms the observed lack of younger speakers contributing to the corpus in both Endu and Mele Maat (Chapter 2). It further reveals that speakers who are older than 65 make up more than half of the Mele Maat corpus, which makes it difficult to tease apart the effects of age and community membership on construction frequency. The community’s prioritisation of recording elderly speakers in Mele Maat is typical of a situation of interrupted transmission, where older speakers are recognised as having high linguistic and cultural competence, while younger speakers may lack confidence in speaking the language (Schmidt 1985: 38–40; Seifart 2011: 23).

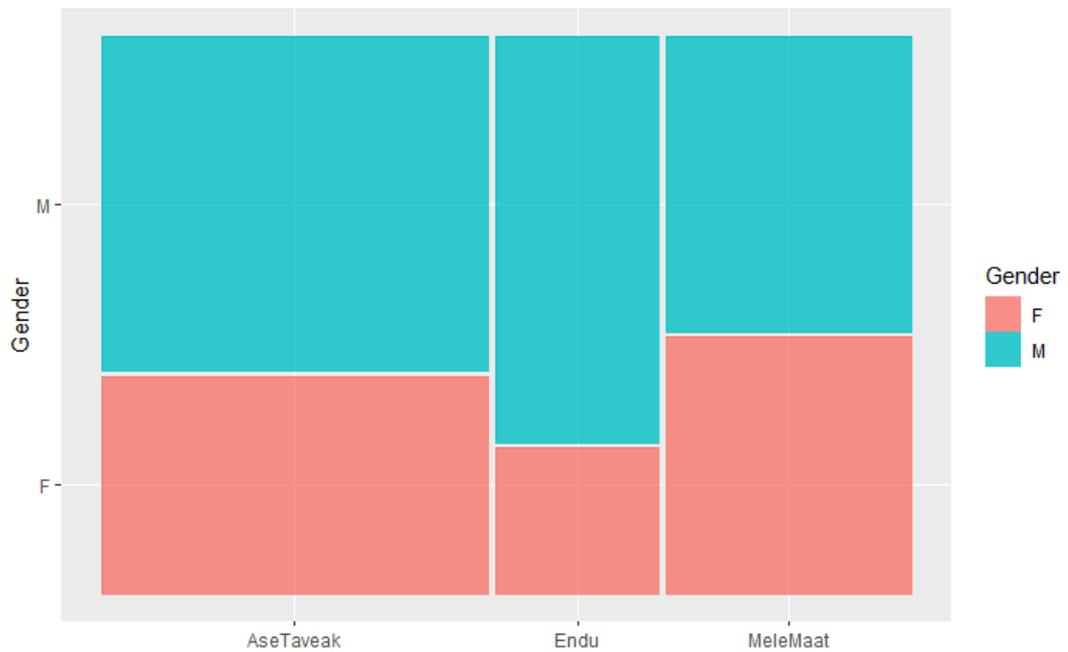


Figure 28: Mosaic plot of words in the corpus by community and gender

Figure 28 shows that the most gender-balanced subsection of the corpus is Mele Maat, with a slightly larger contribution by men, while men also contributed more than half of the Ase-Taveak corpus. However, the greatest imbalance is in Endu, where many women marrying into the community are speakers of North Ambrym (§2.2), and many were not confident about contributing to the Vatlongos corpus. Although the corpus overall appears to be reasonably balanced according to gender, it is useful to recognise this disparity in the Endu subcorpus.

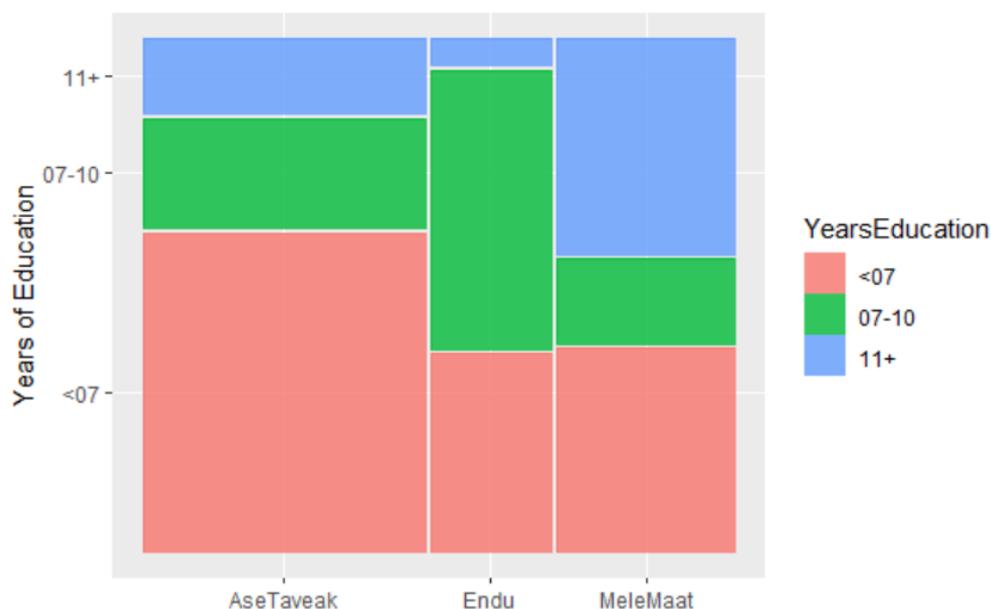


Figure 29: Mosaic plot of words in the corpus by community and years of education

Figure 29 shows that more than half of the Ase-Taveak subcorpus was contributed by speakers who have only had primary-level education, compared to a very similar large minority of the Endu and Mele Maat subcorpora. However, the breakdown of post-primary education in Endu and Mele Maat is strikingly different: most of the Endu subcorpus comes from speakers who have had 7–10 years of education, a level that can be attained at schools in North Ambrym. A large chunk of the Mele Maat corpus on the other hand was contributed by speakers who have completed more than ten years of education, going beyond the national exams in Year 10 and accessing schooling in Port Vila or Santo, if not tertiary education abroad. This makes it difficult to tease apart the effect of community membership and education level on construction frequency.

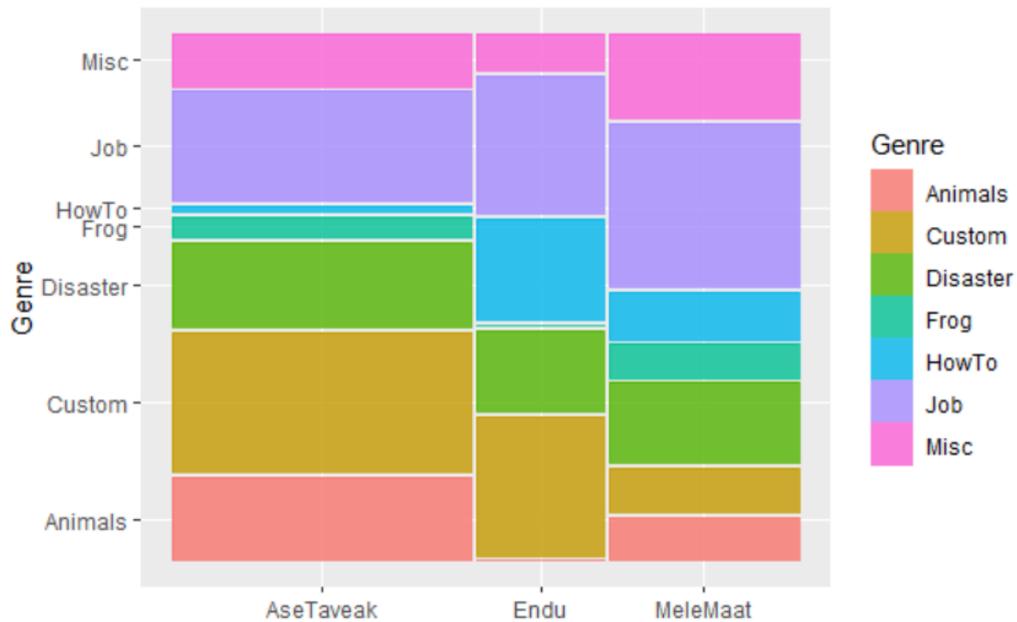


Figure 30: Mosaic plot of words in the corpus by community and genre

The composition of genre by community mostly reflects the different priorities of the urban and island communities. Figure 30 uses more narrowly defined genre categories than the four-way distinction made above (Table 48), because splitting the ‘narrative’ genre into custom stories and disaster narratives reveals an interesting discrepancy between Mele Maat and the two island communities. While all communities have contributed a similar proportion of texts discussing natural disasters – and these included histories of the founding of Mele Maat – custom stories are much less represented in the Mele Maat subcorpus than in either island community. On the other hand, the ‘job’ genre category – a mix of autobiographical and procedural tropes where speakers describe their employment history – unsurprisingly makes up a higher proportion of the Mele Maat subcorpus, where access to the urban cash economy makes for more varied job opportunities.

The other striking difference in Figure 30 is the lack of ‘animals’ texts in the Endu subcorpus. These are narrative texts starring animal characters aimed at children. This reflects observations that Vatlongos has lost out to North Ambrym in this specific domain, which can be an important area for transmission of linguistic and cultural knowledge to children. The lack of animal narratives is also likely connected to the underrepresentation of women in the Endu subcorpus. Figure 31 is a mosaic plot of gender and genre, which shows

that this child-directed genre is associated with female speakers. Conversely, job histories, how-to texts and disaster narratives are more associated with male contributors.

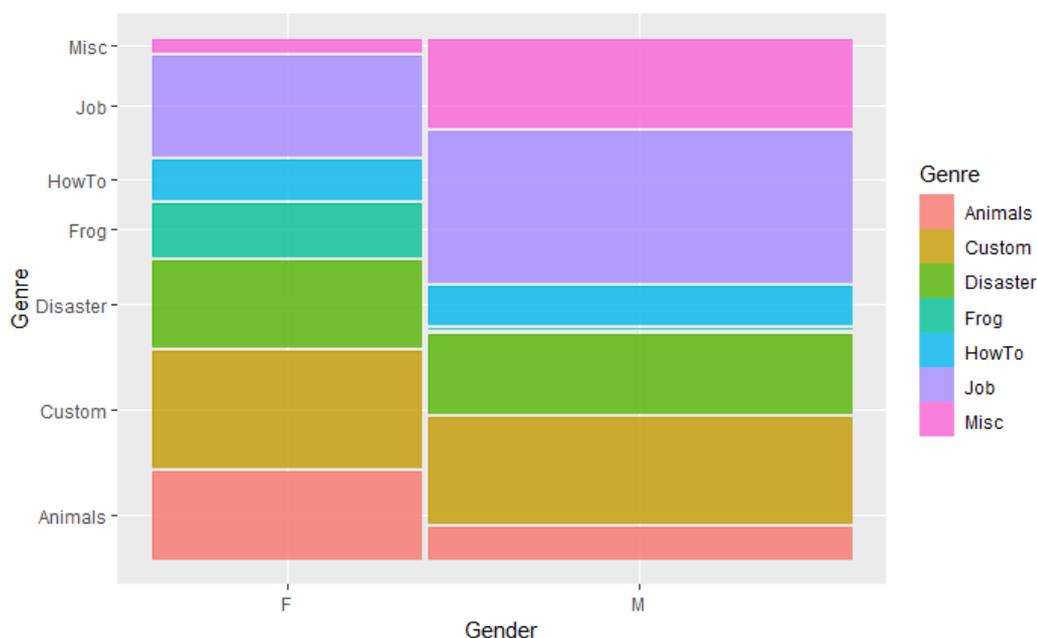


Figure 31: Mosaic plot of words in the corpus by gender and genre

Another possible confounding factor in the composition of the corpus is the overrepresentation of a few enthusiastic contributors, whose personal stylistic preferences could distort apparent evidence of community-level effects. To check for this, word counts were calculated per speaker. Three speakers contributed around 3000 words each to the corpus, while three contributed between 2000 and 3000 words, and 4 contributed 1000–2000 words. The frequency of the constructions in the texts contributed by these speakers was calculated and compared to the overall frequencies in the corpus (Table 42) to look for any distortions that may arise.

The second most prolific speaker, Elder Saksak Ruben, who contributed 3043 words to the corpus, used auxiliaries at a fairly high rate, using 450 every 10,000 words compared to an overall average in the corpus of 292. In fact, calculating a z-score (Oakes 1998: 7–8) for this rate, Elder Saksak is still within one standard deviation of the mean frequency per speaker ($z=0.94$), but this still may cause distortion as he is from Endu, so his speech represents a large chunk of the Endu subcorpus, the smallest of the three community subcorpora.

It could also affect the results for the 7–10 years of education group, and the 45–65 age group.

Two of these speakers from Ase-Taveak use SVCs at a very high rate: around 650 ($z=1.45$) and 750 ($z=2.02$) per 10,000 words, compared to the corpus average of 400. However, this should not have a very large effect on the community level analyses, as they have contributed a total of 2300 words and 1000 words respectively and are both from Ase-Taveak, which is the best-represented community in the corpus (23,273 words). Moreover, the fact that they are two prolific speakers from the same area suggests that this could be a stylistic preference in the community. As one of these speakers is female and the other male, this should not have an effect on the frequency of SVCs by gender.

These imbalances and distortions in the composition of the corpus mean that it is especially important to be cautious about drawing conclusions from the distribution of construction tokens within the corpus, and to choose statistical tools that can account for these distortions. It will be seen in §9.4 that some of apparent effects when factors are examined individually, disappear when other factors are held constant in the regression models. Multiple regression can therefore be an appropriate tool for dealing with these types of imbalances. In a language documentation context, it is important to make the most of the materials available and not miss opportunities to gain insights from quantitative exploration of the data.

9.3 Effect of community on construction frequency

Chi-square tests are often used in linguistics to compare frequencies in different corpora, for example corpora of English from different regions, or from speakers of different genders (Oakes 1998: 24–29; Rayson, Berridge & Francis 2004; Oakes 2009: 163–166; Gries 2010: 16–18). They test the null hypothesis that the observed proportions of category variables are taken from the same overall populations, in this case that the frequency of constructions is the same in each community. The Pearson's chi-square test works by comparing the observed frequencies with expected frequencies, calculated on the basis of the proportions between the cells in the table and the row and column totals. The

significance level or p-value is then calculated by comparing the chi-squared statistic to the degrees of freedom in the contingency table. These tests are especially useful for opportunistic language documentation corpora, because they test raw frequencies in corpora of different sizes (Oakes 2009: 165).

However, chi-square tests are most often used for lexical frequency, which can be organised logically as binomial data: contrasting the count of word tokens which are identical to a particular word form or lemma, and those that are not. When used for syntactic constructions, the chi-square test is usually used to compare frequencies of use of constructions that are analysed as alternatives to each other, for example English present perfect versus simple past (Gries 2010: 17–18 discussing Hundt & Smith 2009), or different English quotative constructions (Desagulier 2017: 178–182 discussing Tagliamonte & Hudson 1999). The understanding of variation as choices between ‘ways of saying “the same thing”’ has underpinned a great deal of sociolinguistic research (Meyerhoff 2014a: 99). However, as Meyerhoff points out, choices about semantic content expressed in language can also constitute variation. One example is the strong tendency in Vatlongos and other Oceanic languages to express directional information about prior motion (§8.2).

Although some of these constructions can be thought of as alternatives to each other in some of their functions (e.g. SVCs and AVCs for prior motion, §6.4.3, §8.2.2), it is the overall frequency of the constructions in each community subcorpus that is being compared here. I therefore treat the construction counts as if they were word counts, subtracting the totals from the overall word count for the subcorpus. This produces contingency tables like the one for AVCs shown in Table 43.

Table 43: Contingency table for AVCs by community

	AVC count	Word count – AVC count	Word count (Totals)
Ase Taveak	745	22528	23273
Endu	310	9532	9842
Mele Maat	347	14392	14739
Totals	1402	46452	47854

For AVCs, it is not unreasonable to count each construction as a word. Each AVC contains an auxiliary verb, which is a lexical class in Vatlongos (Table 41). However, an AVC can contain more than one auxiliary verb in addition to the lexical verb, and SVCs and complex predicates can consist of multiple lexical verbs. In order for the different constructions to be comparable to each other, I treat each construction as a single word for the purpose of these tests.

An alternative strategy, given that these are all verbal constructions, would be to use total verbs in each subcorpus, rather than total words. However, in some of their functions these constructions can be seen as competing with non-verbal strategies. Each individual token of a construction could be conceptualised as a choice between using a simple clause, a different verbal construction, or some other alternative for a particular function, such as an adverb, adjoined PP or even non-verbal cues such as gestures. For example, SVCs could be in competition with PPs for introducing oblique arguments (§6.4.8), or complementisers as a complementation strategy (§6.4.6). As those are alternatives that are available in English, and the influence of English is hypothesised to be a driver of differences between Mele Maat and the island communities, it is important not to lose that information by only examining the proportion of constructions against verbs. Overall, although the data is not technically binomial, examining the proportion of construction count to word count seems to be the best way to access and compare the probability of encountering any given construction in each subcorpus.

Aside from this issue, the data otherwise meets the assumptions of the chi-square test. For each of the speaker-level factors investigated (community, age group, gender and level of education), the data is completely independent. For genre, this is slightly more problematic as individual speakers have sometimes contributed texts within more than one genre. For example, the most prolific contributor to the corpus, Lik Simelum, contributed an animal story, an autobiographical 'job' text, and a description of traditional marriage customs which has been grouped with 'miscellaneous' texts. However, this is only an issue for a few texts, and they are included with texts from other speakers in each genre. Nonetheless, I am cautious about the results of the chi-square tests

for genre, treating them as a justification for inclusion in the regression models in §9.5, rather than conclusive results in themselves. For all of the chi-square tests discussed, there are no cells where the expected frequency is less than 15, meaning the requirements for a chi-square test are comfortably met, and there is no need to use the Yates continuity correction for the two by two contingency tables, such as those for gender (Oakes 1998: 25; Rayson, Berridge & Francis 2004).

To address the overall hypothesis that these constructions are used at different rates in each community subcorpus, it would also be possible to include all the constructions in a single contingency table. Instead, chi-square tests have been performed independently for each construction in order to isolate which constructions have significantly different frequencies across communities, and to decrease the risk of a Type 1 error (Oakes 2009: 166).

For each of the tests in this and the next section, contingency tables have been constructed in the same way as Table 43. The results are reported in the tables below, including the chi-squared value (to three decimal places), the degrees of freedom, and the resultant p-value. As is conventional in much research in linguistics and the social sciences, I am treating p-values of less than .05 as significant, which indicates that the null hypothesis (that the frequency of constructions is the same across the subcorpora) can be rejected with 95% confidence. However, this is an arbitrary cut off point and I also comment on p-values that are close to this level, or less than .1. P-values that are less than .001 are reported as <.001, all other p-values are reported to three decimal places.

In addition to the results of the chi-squared tests, the tables below also include the mean relative token frequency per 10,000 words for each construction. The chi-square test can only indicate whether or not there is a relationship between frequencies and subcorpora, not the direction or magnitude of that relationship. The constructions frequencies can give an initial indication of the trends between communities and other subcorpora, although they are very vulnerable to bias introduced by the imbalances described in §9.2. It is important to clarify that the chi-square tests have been conducted on the raw frequencies, not the normalised frequencies (Oakes 2009: 165).

Table 44: Construction frequency (per 10,000 words) by community and results of chi-square tests

	AVCs	SVCs	PriMo	CPred	0Comp	Comp
Overall	293	400	102	102	159	195
Ase-Taveak	320	457	103	103	147	188
Endu	315	358	94	132	183	223
Mele Maat	235	338	106	79	161	189
X ²	24.864	39.229	0.791	16.357	5.655	4.6881
df	2	2	2	2	2	2
p-value	<.001	<.001	.673	<.001	.059	.096
Significant?	Yes	Yes	-	Yes	<i>Nearly</i>	-

Neither zero-marked complement clauses nor prior motion SVCs show significant differences between communities. This is in line with the hypothesis that these constructions are unlikely to be used less in Mele Maat as both have similar equivalents in both Bislama and English. The mean relative frequencies for prior motion SVCs across communities are very similar, which is reassuring as it suggests that the corpus is large and consistent enough to identify trends in verbal constructions of this frequency, and that the observed differences in other constructions are probably meaningful, and not just noise. The minimum corpus size to investigate different linguistic phenomenon is an empirical question (Lüpke 2005: 96–98; Haig, Schnell & Wegener 2011: 57), and the answer is likely to be different for different languages, and for different types of constructions within a language, depending on their frequency and variability.

However, zero-marked complement clauses do show a tendency to be used more frequently in Endu and Mele Maat than in Ase-Taveak, and the p-value approaches the .05 level of significance. One possibility that could account for this trend is if speakers in different communities are using complementisers and zero-marking in different proportions. I therefore ran another chi-square test comparing the frequency of complementiser-marked and zero-marked complement clauses, and calculated the percentage of all complement clauses that are marked with a complementiser in each community. Speakers in Ase-Taveak used complementisers with 22% of complement clauses, more often

than either Mele Maat (15%) or Endu (18%). This could offset the observed lower frequency of zero-marked complement clauses in Ase-Taveak, but this trend was again just shy of the threshold for significance ($p=.060$). A tendency to use zero-marking rather than complementisers could be a sign of simplification of Vatlongos syntax in Mele Maat. When both zero-marked and complementiser-marked clauses were included, there was almost no difference between the relative frequency of complement clauses in Ase-Taveak (188 per 10,000 words) and Mele Maat (189 per 10,000 words). Although speakers in Endu used more of both kinds of complement clauses (223 per 10,000 words), the effect of community on use of all complement clauses was not significant ($p=.096$), probably because the Endu subcorpus is smaller than the other two community-level subcorpora.

While the tendencies to use slightly different proportions of complementisers or zero-marking with complement clauses in different communities is interesting, the very significant p -values for the effect of community on other constructions suggest that it would be possible to set a more stringent level of significance (for example, Haig et al. (2011: 75) reject effects where $p>.01$ in their study on pronominal coding of patient arguments in four languages). For now, I continue to use a significance threshold of $p<.05$, but it is important to be cautious about trends that only just achieve this level of significance.

The other constructions all showed a significant relationship between frequency and speaker community membership. SVCs seem to be much more frequent in Ase-Taveak. Endu falls between the rates in Ase-Taveak and Mele Maat, but patterns more closely with Mele Maat. AVCs, which I argue are diachronically related to SVCs (§8.3.1), are also more frequent in Ase-Taveak. As the emergence of AVCs seems to have been in progress over the last few generations, this could be a sign that Mele Maat is actually more conservative by this metric. Although Endu appears to pattern with Ase-Taveak on the frequency of AVCs, it is important to remember that the speech of Elder Saksak Ruben, who uses AVCs at a fairly high rate, makes up a large proportion of the

Endu subcorpus, so it could be that the actual rate of AVC use for other speakers in Endu is more similar to the lower rate in Mele Maat.

The pattern for complex predicates is as predicted, on the basis that it is an index of archaic, more conservative speech patterns. It is least frequent in Mele Maat, and more frequent in Endu than in Ase-Taveak, perhaps due to the conservative language ideologies observed in Endu (§2.4). However, the difference between the mean relative frequencies is much less than for SVCs, for example, as complex predicates overall are a less frequent construction.

Mele Maat has lower rates of every construction that does not have very similar formal equivalents in Bislama and English, which could be a sign of language shift in the Mele Maat community, where intergenerational transmission is less robust and Vatlongos is used in fewer domains (§2.3). The higher proportion of zero-marked rather than complementiser-marked complement clauses discussed above could be another sign of language attrition in Vatlongos spoken in Mele Maat.

9.4 Effect of age, gender, education and genre on construction frequency

This section also examines other speaker-level and text-level variables to investigate whether the community-level trends indicate emerging regional dialectal differences, or are more closely related to other sociolinguistic factors. This is especially important given the imbalances in the corpus (§9.2).

In the descriptive chapters of this thesis, I argue that some of the constructions are undergoing change, especially the development of AVCs from SVCs which I argue was in progress in the late twentieth century (§8.3.1). It is therefore hypothesised that age could have a relationship with construction frequency, especially with AVCs. The boundaries for the age groups were determined partly by categories that roughly reflected socially-constructed life stages in the community, and partly to ensure roughly equal-length subcorpora as far as possible. Adults aged 31–45 are likely to be parents to school age children, while those aged more than 65 are likely to be grandparents. The least coherent group, the under 30s, is also the least well represented in the corpus (Figure 26), especially in Endu and Mele Maat (Figure 27). While the other three

age-based subcorpora have roughly similar word counts (14,600; 15022; 12705), the subcorpora of the youngest contributors contains only 5527 words. This means that although it would be preferable to further distinguish children at different developmental stages and young adults, it is not possible to be more specific in this age group within the limitations of the data collected. Recording more texts by speakers in this age group is an important priority for future fieldwork.

Table 45: Construction frequency (per 10,000 words) by age group and results of chi-square tests

	AVCs	SVCs	PriMo	CPred	0Comp
Overall	293	400	102	102	159
<31	293	434	125	101	136
31–45	260	430	92	118	173
46–65	336	400	118	81	166
66+	280	350	84	109	146
X ²	16.308	13.328	12.173	11.037	5.5373
df	3	3	3	3	3
p-value	<.001	0.004	0.007	0.012	0.136
Significant?	Yes	Yes	Yes	Yes	-

Table 45 shows the relative mean construction frequencies and results of chi-square tests across age groups. All constructions except zero-marked complement clauses show a significant relationship between frequency of use and age group. However, it is only SVCs that show a clear trend: younger speakers are using SVCs more than older age groups, and especially more than the oldest age group. This could be a sign that SVCs are growing in popularity in apparent time.

The other constructions show patterns that are harder to explain and may well be artefacts of imbalances in the corpus. Rates of AVCs are fairly consistent across all ages except the 46–65 group who use them more frequently. This could be the influence of Elder Saksak Ruben’s high rate of AVC use, as his texts make up a large chunk of this subcorpus. If AVCs are expanding into new contexts (§8.3.1), this change is not apparent in generational trends. Prior motion SVCs and complex predicates do not show as large a range as the AVCs and SVCs, and no generational trend. A pattern for all constructions except complex predicates, is lower frequency in the speech of the over-65s. This could

reflect a decrease in grammatical complexity in the speech of elderly contributors due to cognitive decline (Kemper, Marquis & Thompson 2001).

The statistical significance of these relationships means it is important to include age in the regression models in §9.4, but the unexpected patterns in the data suggest a need for caution in extrapolating from trends in construction frequency by age, which probably arise from imbalances in the corpus and may be accounted for by other factors.

Vanuatu society is fairly gender-segregated, but previous sociolinguistic work in Vanuatu has found differences in men and women’s use of conversational strategies in Bislama, such as apologies and supporting others’ conversational turns (Meyerhoff 1999; 2003: 308–310), but not at the level of morphosyntactic variables (Meyerhoff 2000: 141). It is worth checking for gender effects given that the nine non-native speakers of Vatlongos in the survey are women who have married into the community (§2.2), and because women are underrepresented in the Endu subcorpus (Figure 28).

Table 46: Construction frequency (per 10,000 words) by gender and results of chi-square tests

	AVCs	SVCs	PriMo	CPred	OComp
Overall	293	400	102	102	159
F	295	433	102	99	162
M	291	379	103	104	157
X ²	0.06	8.42	0.007	0.232	0.333
df	1	1	1	1	1
p-value	0.804	0.004	0.931	0.630	0.684
Significant?	-	Yes	-	-	-

Table 46 shows that the relative frequency of most of these constructions is similar in the speech of men and women. SVCs are the only construction where there is a significant relationship with gender, with women using SVCs at a higher rate. Given the generational increase in use of SVCs apparent in Table 45, this is in keeping with sociolinguistic findings in industrial Western societies, where women often lead language change (Labov 1990; Meyerhoff 2014a: 93–95). However, it could instead be an artefact of imbalances in the corpus, given that SVCs are used at a higher rate in Ase-Taveak, the largest community-level subcorpus, than in Endu, where women are underrepresented (Table 44).

The groupings for level of education are determined by milestones in the Vanuatu education system. Six years of primary education are compulsory in Vanuatu and accessible to most people via government schooling in the past few decades, and previously in mission-run primary schools on most islands (§2.1). The next major milestone is the Year 10 national exams: a lot of students leave secondary education after this point, and many secondary schools on the outer islands only offer education to this level. Ranon High School in North Ambrym only offers education to this level, and education in Southeast Ambrym is only available up to Year 8. Therefore, receiving more than 10 years of education usually necessitates living on another island, usually in an urban centre, if not attending tertiary education abroad, and indicates greater exposure to English and highly multi-lingual contexts.

Table 47: Construction frequency (per 10,000 words) by years of education and results of chi-square tests

	AVCs	SVCs	PriMo	CPred	0Comp	Comp
Overall	293	400	102	102	159	195
<7	291	451	104	105	151	186
7-10	313	361	115	90	187	233
11+	273	326	81	109	143	170
X ²	3.403	36.72	6.982	2.563	9.189	14.075
df	2	2	2	2	2	2
p-value	0.182	<.001	0.03	0.278	0.01	<.001
Significant?	-	Yes	Yes	-	Yes	Yes

The only very clear effect is again with SVCs, which more educated speakers use at a lower rate. This is perhaps due to exposure to English which lacks verb serialisation. There is a similar, but less significant, trend for prior motion SVCs, a function for which English does have a broadly similar strategy in this functional domain involving reduced verbal marking without overt linking devices (Crowley 2002b: 11).

The significant relationship between education level and zero-complement clauses is harder to explain, as there is no clear trend with increasing years of education. However, further chi-square tests including complementiser-marked complement clauses reveals an interaction of two patterns. When the frequency of all complement clauses is compared, there is a very significant effect of years

of education ($p < .001$). Speakers who have 7–10 years of education use more of these complex sentences than speakers who have less than 7 years of education. Although secondary education is not conducted in Vatlongos, perhaps the emphasis on syntactic complexity in formal schooling impacts on their use of Vatlongos, especially if a recording event is viewed as a performance. However, when speakers have received more than ten years of education, they are less likely to use these complex sentences. It is likely that this group are less proficient speakers of Vatlongos, having necessarily spent a lot of time in English-dominant linguistic contexts and away from Vatlongos-speaking communities.

This tendency for the most highly-educated speakers to use simpler structures is also evident in the ratio of zero-marked complement clauses to overt complementisers: highly educated speakers are slightly less likely to use overt complementisers, using them for only 16% of complement clauses, compared to 19% and 20% for those with less than 7 or 7–10 years of education. Although this effect is likely due to random factors in this corpus ($p = .575$), it offsets the tendency for highly educated speakers to use fewer complement clauses overall, and explains why the effect of education on zero-marked complement clauses is so much less significant than the effect on all complement clauses.

Rather than eliciting texts in particular genres, community and contributor preferences determined the contents of the corpus as far as possible (§1.3.3). A particular genre can be defined by the features of its content (e.g. fiction versus non-fiction, narrative), or purpose and audience (e.g. education, humour, oratory). Rather than restricting the categories to a particular theoretical view of genre, I have tried to group texts into genres based on how Vatlongos speakers describe and categorise them, especially in planning sessions when discussing what to record. Initially I grouped the texts into seven narrow genres, shown in Table 48.

Disaster narratives were a very popular topic for speakers, including accounts of the volcanic explosion, cyclone and tsunami of 1951 that led to the formation of Mele Maat, Cyclone Pam in 2015 and other cyclones, shipwrecks,

earthquakes and volcanic explosions. Paviour Smith (2015) discusses the importance of similar disaster narratives for individual and community identity of speakers of Aulua (Malekula). Custom (Bislama *kastom*) stories are a widely recognised genre in Vanuatu (Budd 2011: 190; Meyerhoff 2015: 92), consisting of traditional stories and myths that often include a song. The ‘animal’ genre also consists of traditional narratives, but these are myths and fables starring animal protagonists aimed at children. Von Prince (2015: 406–410) makes a similar distinction between factual accounts, myths and children’s stories for Daakaka narratives. Frog stories were distinguished from other animal texts because they were the only prompted genre included in the corpus, in which a speaker told a story based on pictures in *Frog Where Are You?* (Mayer 1969).

Autobiographical texts recounting employment histories were another popular genre, which I am calling ‘job’ texts. These are a mixture of personal narrative with some procedural interludes, with detailed descriptions of how to perform particular tasks within a job. They are the most modern genre, involving jobs, technologies and concepts that did not exist in traditional pre-contact Vanuatu. How-to texts are procedural, explaining how to perform tasks such as building a house, planting crops, or preparing food. This genre was very popular but the individual texts are often very short. Despite the low total word count, there are 21 how-to texts in the corpus, equal to the number of ‘job’ texts and second only to custom stories (22). The miscellaneous category is a mix of other genres, each of which was represented by less than five texts. These include speeches at ceremonies, evangelical histories of individual villages, descriptions of current activities and humorous anecdotes.

Table 48: Narrow and broad definitions of genre, with word counts

Narrow genre	Word count	Broad genre	Word count
Disaster	7942	Narrative	19585
Custom	10643		
Animals	5180	Animals	7185
Frog	2005		
Job	12687	Job	12687
How-to	3707	Miscellaneous	9397
Miscellaneous	5690		

Some of these genres have such low word counts that they are unlikely to be useful for generalisations, and seven is too many category levels to be useful for the regression models for frequency per text in §9.5. In order to decide how to organise them into four broader categories, construction frequencies in each genre subcorpus were calculated and used as the basis for regrouping, in order not to lose the boundaries that proved most relevant to construction frequency. The results are shown in Table 49, but should be treated with caution.

Table 49: Construction frequency (per 10,000 words) by narrowly-defined genre and results of chi-square tests

	AVCs	SVCs	PriMo	CPred	0Comp
Overall	293	400	102	102	159
Custom	365	420	96	142	156
Disaster	271	473	92	94	166
Animals	429	475	133	110	120
Frog	264	753	85	55	85
Job	198	300	111	80	182
How to	340	337	97	119	170
Misc.	258	330	90	83	158
X ²	100.27	128.77	8.766	31.034	17.1
df	6	6	6	6	6
p-value	<.001	<.001	0.186	<.001	0.009
Significant?	Yes	Yes	-	Yes	Yes

AVCs, SVCs and complex predicates all showed significant interactions with narrowly defined genre. SVCs and AVCs were more frequent in purely narrative genres than in the job, how-to and miscellaneous texts. They were even more frequent in animal stories, which, conversely, had fewer zero-marked complement clauses.

It was particularly important to check the distribution of constructions in frog stories, as the only prompted texts in the corpus. Foley (2003) compares frog stories to traditional oral narratives, and finds major differences in discourse organisation as syntactic integration. Frog stories in Vatlongos had a very high rate of SVC tokens at 753 per 10,000 words ($z=1.86$). However, individual texts in other genres, including disaster narratives, had even higher rates of SVCs. On the other hand, the frog stories had the lowest rates of prior motion SVCs, complex predicates and zero-marked complement clauses, so did

not reproduce Foley's finding of higher syntactic integration in the prompted texts, a tendency he linked to conventions of Western literacy. Overall the construction frequencies for frog stories seem to be an exaggeration of patterns for other child-directed animal stories, so these two genres are grouped together into a single broad genre, contrasted with the other narrative genres.

One distinction that is unavoidably lost by grouping custom and disaster narratives is the higher rate of complex predicates in custom stories compared to all other genres, confirming the impression that these index conservative speech patterns in these texts dealing with traditional settings. They are also fairly rare in the modern 'job' genre. The procedural how-to texts had lower rates of SVCs than the narrative genres, and higher rates of complex predicates, probably largely due to use of *rat* 'out, away' as a subsequent verb in physical instructions (§7.1.3). Given the low word count for this genre, and their differences from the narrative genres, they were absorbed into the 'miscellaneous' category.

The results for these broader genre categories are shown in Table 50. The effect of genre is significant for all constructions except prior motion SVCs, but especially significant for AVCs and SVCs. AVCs and SVCs are more frequent in purely narrative genres, and especially in child-directed animal narratives. SVCs are less frequent in 'job' texts, perhaps because these tend to describe more abstract events from a broader temporal and locational perspective (e.g. 'I worked there for five years'), rather than physical events from a narrower perspective, where directional and evaluative modification of verbs would be more relevant (e.g. 'it ran away quickly'). Complex predicates are more frequent in narrative texts (probably because they include custom stories) than in other genres, especially 'job' texts which are in a modern setting. The significance of this effect is less than when custom stories were separated from disaster narratives. The zero-marked complement clauses are less common in animal texts, perhaps due to a simpler sentence structure in child-directed texts.

Table 50: Construction frequency (per 10,000 words) by broadly-defined genre and results of chi-square tests

	AVCs	SVCs	PriMo	CPred	OComp
Overall	293	400	102	102	159
Narrative	324	443	94	122	160
Animals	383	553	120	95	110
Job	198	300	111	80	182
Misc.	291	333	93	97	163
X ²	67.23	96.217	5.223	13.6	15.471
df	3	3	3	3	3
p-value	<.001	<.001	0.156	0.004	0.001
Significant?	Yes	Yes	-	Yes	Yes

Given the imbalances described in §9.2, it is important to be cautious about making claims on the basis of the trends indicated by mean relative token frequencies in subcorpora based on these sociolinguistic variables. However, the factors that have a significant effect in these subcorpus-level chi-squared tests are included in the regression models of frequency per text, which help to rule out effects that are artefacts of sample imbalances.

9.5 Modelling frequency of each construction in a text

The sociolinguistic factors discussed in the previous two sections are characteristics of individual texts. Rather than cutting the corpus into subsections on the basis of individual factors, this section aims to model the frequency of constructions within each text, in order to investigate how each factor contributes to the construction frequencies when all factors are considered. The chi-square tests used so far cannot handle multiple independent variables or interactions between them (Oakes 1998: 37). Using regression techniques to model frequency per text also helps to make sure that the overall trends described so far are not simply artefacts of the imbalances in the corpus.

Generalised linear models have been used to address a wide range of linguistic questions relating to discrete data (Oakes 1998: 33–46; Baayen 2008: 195–214). In sociolinguistic studies, this has normally pertained to binomial data structured as the frequencies of alternative variants, modelled as a variable rule and often implemented using VARBRUL and subsequent software

implementations of variable rule programs, including within R (Sankoff & Labov 1979; Sankoff, Tagliamonte & Smith 2005; Johnson 2009). Although I have treated construction frequency as equivalent to lexical frequency for the purposes of the chi-square tests above, it is preferable to choose a probability distribution function that is closer to the real-world relationship between number of constructions and number of words in a text in order to create models that best fit the data.

As argued in §9.1, the frequency of constructions in texts is not binomial, but overdispersed-Poisson. Constructions are therefore assumed to occur at a certain rate over lengths of speech that can be measured by word count. There is expected to be a level of variance which is greater than the mean rate, and this variance is modelled as a separate parameter using negative binomial regression (Hilbe 2014: 126–160). I therefore use the raw frequency count of each construction per text as the dependent variable, and use the word count per text as an offset within the model, as a measure of exposure (Hilbe 2014: 62–66). Typically, exposure is a measure of time, area or space, but other ways to enumerate the opportunities for an event to occur are also modelled in this way, such as population of a county when modelling death rates. Word count is a good proxy for exposure in modelling construction frequency, because the basic Poisson distribution assumes that counts are made in very small intervals of time, such that only a single count can be made (Hilbe 2014: 63). While constructions are not words in themselves, a construction can be counted with the addition of a single word, for example the second verb in an SVC. The incorporation of an offset for exposure makes this a proportional intensity model.

Very short texts could be included in the chi-square tests in §9.3 and §9.4 because they were amalgamated with other texts and analysed as entire subcorpora. However, for the purpose of modelling frequency for individual texts, texts that are very short relative to the rate of occurrence of a particular construction are not informative. For example, complex predicates occur roughly once every 98 words (Table 42). If a text is shorter than 98 words and does not contain a complex predicate, it cannot tell us whether the speaker,

speaking in that genre, tends to use complex predicates at a lower or higher rate than is usual in Vatlongos. Conversely, if a complex predicate does occur in text shorter than 98 words, the relative frequency of complex predicates in that text will appear to be very high, even though if they had continued speaking, they may have used complex predicates at the usual rate. This phenomenon can account for the extreme outliers that can be seen in Figure 24. These short texts are therefore not informative and may actively distort the model, so texts that were shorter than the number of words a particular construction tends to occur in (Table 42), are excluded from the model of that construction. The number of texts included for each model is shown as N in the tables below.

All of the factors that were found to be significant in the chi-square tests above were included in the model for a construction. When constructing a regression model for a balanced sample, it is usual practice in sociolinguistics to choose which predictors to include by repeating the analysis, adding significant predictors and removing non-significant predictors in an iterative, stepwise procedure. This is the procedure used in the GoldVarb implementation of VARBRUL, for example, but its effectiveness has been questioned (Johnson 2009: 22). Given the clear imbalances in the Vatlongos corpus (§9.2), it is important to include all the factors that have been shown to have an effect on the overall construction frequencies in the corpus. For example, both community and age group have been shown to have an effect on the frequency of SVCs, but Mele Maat speakers are disproportionately in the oldest age group (Figure 27). If only one of these factors is included, it becomes impossible to tease apart the potentially separate effects: whichever is included appears to have higher significance than if both are included so that the effect of one is calculated while holding the other constant. Partly this decision comes down to the goals of this project: the aim of the models in this section is not to be a maximally efficient model of the corpus, but rather to identify robust trends that can be extrapolated to the wider use of Vatlongos as far as possible. The factors that have been identified as significant are shown in Table 51.

Table 51: Significant factors for each construction according to chi-squared tests

	Community	Age	Gender	Years of Education	Genre
Auxiliaries	✓	✓	-	-	✓
SVCs	✓	✓	✓	✓	✓
Prior motion SVCs	-	✓	-	✓	-
Complex predicates	✓	✓	-	-	✓
0 Complements	-	-	-	✓	✓

All of the factors except gender have more than two levels, so it is important to explain how multi-level parameters are calculated in regression models. For each parameter, one level is used as a reference level against which the effect of the other levels is measured. These are shown in **bold** next to the parameter name in the summary tables below. For genre, miscellaneous is the obvious choice for comparison with the more internally coherent categories. The Ase-Taveak community and primary level education group both represent the greatest proportion of speakers, both in the speaker population and the corpus, so are sensible choices for reference levels. For age, the 31–45 group is used as a reference level. The youngest age group has a very small word count and is expected to have high variance as it includes both children and young adults, so is not a suitable choice for a reference level. Looking at the mean frequencies in the tables above, the oldest group and, to a lesser extent, the 46–65 age group are exceptional for several constructions, while the 31–45 group patterns most consistently with the construction frequencies in the whole corpus.

The models for each construction are summarised in Table 52–Table 56. I explain the summary statistics in relation to the model for SVCs, which proved to be the most revealing, and also illustrates some important considerations in model construction and criticism. One widely-used measure of model fit is the Akaike Information Criterion (AIC) (Hilbe 2014: 115–118), which can be used to compare non-nested models of the same variable. Although negative-binomial regression is the most common method for modelling overdispersed Poisson data (Hilbe 2014: 33), I also ran the same models with a Poisson distribution, and in each case the AIC for the negative-binomial model was significantly lower, confirming the choice of distribution function. AIC_N is a measure of AIC adjusted for N, and again a lower value indicates a better fit.

Alpha (α)⁴⁸ is the dispersion parameter of the negative binomial model: the inclusion of this parameter is the difference between the negative binomial model and Poisson regression. The closer alpha is to zero, the closer the model is to a Poisson distribution, and the less overdispersion there is in the data. Alpha is lowest in the model for SVCs (.065), suggesting this is the construction that most closely resembles a Poisson distribution. However, the AIC for the negative binomial model of SVCs was significantly lower than the AIC for the Poisson regression model (675), according to Hilbe's (2014: 118) significance levels. The value of alpha was highest for complex predicates (.252), showing that this construction is the most overdispersed and auto-correlated. This supports the impression that complex predicates have low productivity and are the most lexicalised of these constructions. Greater overdispersion is expected for more lexicalised constructions, and greater priming effects (which cause overdispersion) also occur when there is a stronger association between individual lexemes and constructions (Gries 2005: 391). Alpha is shown to three decimal places in the summary tables.

Table 52: Summary of model of frequency of SVCs per text

SVCs	Parameters	Levels	IRR	95% CI		Sig?	P	Sig?
N=104 AIC: 642 AICn: 6.18 α : .065	Community: Ase-Taveak	Endu	0.94	0.767	1.145	-	.525	-
		Mele	0.86	0.700	1.051	<i>Nearly</i>	.140	-
		Maat						
	Age: 31-45	<31	1.01	0.783	1.294	-	.958	-
		46-65	0.98	0.812	1.185	-	.844	-
		66+	0.93	0.742	1.161	-	.517	-
	Gender: F	M	1.14	0.960	1.365	-	.132	-
	Education: <7	7-10	0.77	0.636	0.921	Yes	.005	Yes
		11+	0.70	0.554	0.891	Yes	.004	Yes
	Genre: Miscellaneous	Animals	1.51	1.160	1.977	-	.002	Yes
		Narrative	1.24	1.018	1.521	-	.330	-
		Job	0.89	0.703	1.521	-	.033	Yes

Coefficients based on the Poisson distribution can be difficult to interpret, so I have instead calculated incidence rate ratios (IRRs) by exponentiating the coefficient (Hilbe 2014: 55-61). IRRs are a more transparent expression of a predictor's effect on the likelihood of an occurrence. The likelihood is compared

⁴⁸ In fact, the glm package in R uses the inverse of alpha, theta. I have converted the theta values to alpha for easier interpretation of the dispersion parameter (Hilbe 2014: 131).

to the reference level (shown in **bold**) for that parameter, when all other parameters are held constant.

An IRR of 1 shows no effect on the likelihood of a construction. Interestingly for SVCs, the IRR for all age groups is nearly equal to 1, showing that age has no effect on the use of SVCs when all other parameters are held constant. This is surprising given the apparently significant effect of age on the frequency of SVCs in the chi-square test above (Table 47), where the oldest age group have a much lower mean relative frequency. Instead, it seems likely that this effect is a result of the high proportion of contributors of this age group who are from Mele Maat (Figure 27). The overall difference in mean relative token frequency for the oldest age group is nearly entirely due to the nearly significant effect of community on SVC use: speakers from Mele Maat are only 86% as likely to use SVCs as speakers from Ase-Taveak (IRR=0.86).

An IRR of less than 1 indicates a factor that lowers the likelihood of the construction occurring, while a value higher than 1 shows higher probability, relative to the reference level. For example, texts in the animal narrative genre are one and half times as likely to contain SVCs as miscellaneous texts (IRR=1.51). Surprisingly, given the apparent trend in the relative mean construction frequencies in (Table 46), the model suggests men are actually 14% more likely to use SVCs than women once other effects were held constant (IRR=1.14). IRRs are reported to two decimal places in the summary tables.

However, the significance of these effects must be evaluated. There are multiple ways of determining the significance of factors in a negative binomial model, and they are more or less conservative. The simplest, but least conservative, method is to look at the p-value for the z-value: the ratio of the raw coefficient and the model standard errors (Hilbe 2014: 57–58). For this measure I again consider values less than .05 as significant, though I also comment on effects with a p-value less than .1. A more robust method is to look at the exponentiated confidence intervals based on profile likelihood errors (Hilbe 2014: 59), rather than model standard errors as used to calculate the z-values. If both the upper and lower confidence interval is less than one, the effect is significant. As can be seen in Table 52, this measure is more

conservative. On the basis of z-values, education and genre each contain significant predictors of SVC frequency. Looking at confidence intervals, genre is not a significant parameter. These indicators of significance are reported to three decimal places in the summary tables.

To summarise the significant effects in Table 52, level of education seems to be the most significant predictor of SVC frequency. Compared to speakers who have only primary level education, speakers with 7–10 years of education are only 77% as likely to use SVCs, and this effect is even greater for speakers with more than ten years of education, who use SVCs 70% as much as primary-educated speakers. Community, which appeared to be a significant factor in chi-square tests (§9.3), did not reach significance in this model. There was a nearly significant effect on SVC use, with speakers from Mele Maat 14% less likely to use SVCs than speakers from Ase-Taveak. Endu falls between Mele Maat and Ase-Taveak, showing 6% less use of SVCs than speakers from Ase-Taveak, but this effect is not significant. The effect of genre is significant when looking at z-values, but not likelihood-based confidence intervals. Speakers are 51% more likely to use SVCs when telling an animal story than a miscellaneous text, but 11% less likely to use them when recounting their employment history. Narrative texts are also 24% more likely than miscellaneous texts to use SVCs, but this effect was not significant by either measure.

Table 53: Summary of model of frequency of AVCs per text

AVCs	Parameters	Levels	IRR	95% CI	Sig?	P	Sig?	
N=103 AIC: 645 AICn: 6.27 α : .179	Community: Ase-Taveak	Endu	1.08	0.805	1.443	-	.604	-
		Mele Maat	0.76	0.570	1.013	<i>Nearly</i>	.054	<i>Nearly</i>
	Age: 31–45	<31	0.96	0.670	1.388	-	.835	-
		46–65	1.31	0.997	1.714	-	.049	Yes
		66+	1.16	0.855	1.959	-	.342	-
	Genre: Miscellaneous	Animals	1.34	0.922	1.959	-	.123	-
		Narrative	1.04	0.783	1.385	-	.778	-
		Job	0.64	0.458	0.909	Yes	.010	Yes

Table 53 summarises the model for AVCs. Here there were fewer significant effects. The only significant effect by both measures was the lower likelihood of using AVCs in employment histories: speakers were more than a third less likely to use AVCs in these texts than in miscellaneous texts. Narrative texts and miscellaneous texts were very similar, while AVCs were a third more likely to

occur in animal stories, although this effect was not significant. Mele Maat speaker community membership nearly had a significant effect by both measures: speakers from Mele Maat were nearly a quarter less likely to use AVCs. Surprisingly, the 46–65 age group had a significant effect on AVCs, with speakers in this age group being nearly a third more likely to use AVCs than the 31–45 age group. However, this effect is not significant when considering confidence intervals, and is probably related to the many texts contributed by Elder Saksak Ruben, who uses AVCs at a high rate (§9.2). The oldest age group also used AVCs at a high rate, being 16% more likely to AVCs than the 31–45 age group, but this was not significant.

Table 54: Summary of model of frequency of zero-marked complement clauses per text⁴⁹

OComp	Parameters	Levels	IRR	95% CI		Sig?	P	Sig?	
N=98 AIC: 515 AICn: 5.26 α: .121	Education:	7–10	1.30	1.019	1.673	-	.032	Yes	
		11+	1.10	0.813	1.488	-	.526	-	
	Genre:	Animals	0.72	0.492	1.059	<i>Nearly</i>	.086	<i>Nearly</i>	
		Miscellaneous	Narrative	1.08	0.785	1.484	-	.638	-
		Job	0.98	0.737	1.210	-	.875	-	

Zero-marked complement clauses were the only other construction to show borderline significant effects of the factors discussed in this chapter. For genre, both narratives and jobs had quite similar rates of zero-marked complement clauses to the miscellaneous texts, but they were more than a quarter less likely to occur in animal stories, and this effect was nearly significant by both measures. This could reflect a tendency to use less complex sentence structures in texts aimed at children. For education, having more education increased the likelihood of using zero-marked complement clauses, but this effect was only significant for the 7–10 years of education group, who were a third more likely to use zero-marked complement clauses. However, this effect was not significant by the likelihood-based confidence intervals measure.

⁴⁹ As community had a nearly significant effect on zero-marked complement clause construction frequency (Table 44), another model was run including community. However, this model was no more efficient (AIC: 519), community was significant in the model, and the trends for education and genre were very similar to the model excluding community, so only the simpler model is shown here.

None of the effects of these factors on complex predicates or prior motion SVCs proved to be significant. However, the IRRs for the communities in the complex predicate model (Table 55) do reflect the overall pattern that complex predicates are associated with more conservative speech: they are a quarter more likely in Endu than in Ase-Taveak, and a quarter less likely in Mele Maat.

Table 55: Summary of model of frequency of complex predicates per text

CPred	Parameters	Levels	IRR	95% CI		Sig?	P	Sig?
N=92 AIC: 447 AICn: 4.86 α : .252	Community: Ase-Taveak	Endu	1.28	0.871	1.883	-	.209	-
		Mele	0.74	0.501	1.105	-	.143	-
		Maat						
	Age: 31-45	<31	0.72	0.431	1.216	-	.222	-
		46-65	0.74	0.504	1.091	-	.129	-
		66+	1.09	0.729	1.635	-	.670	-
	Genre: Miscellaneous	Animals	1.13	0.647	1.966	-	.672	-
		Narrative	0.89	0.551	1.443	-	.641	-
		Job	1.34	0.893	2.019	-	.156	-

Prior motion SVCs are strikingly stable across the various sociolinguistic variables we have discussed, but Table 56 shows that they are more likely to occur in the speech of under-31-year-olds, and less likely to occur in the speech of more highly educated speakers, but neither of these effects were significant.

Table 56: Summary of model of frequency of prior motion SVCs per text

PriMo	Parameters	Levels	IRR	95% CI		Sig?	P	Sig?
N=92 AIC: 442 AICn: 4.80 α : .230	Age: 31-45	<31	1.34	0.837	2.161	-	.220	-
		46-65	1.07	0.749	1.543	-	.696	-
		66+	0.96	0.637	1.460	-	.861	-
	Education: <7	7-10	1.16	0.835	1.614	-	.380	-
		11+	0.79	0.513	1.228	-	.292	-

All the significant and near-significant effects in the models for these constructions are summarised in Table 57. The summary of significant effects is highlighted in **bold**. Effects that were nearly significant by both measures, or only significant by the z-value based on standard model errors, are marked in *italics*.

Table 57: Summary of significant and near-significant trends in models of construction frequency per text

Constr	Parameter	Level	IRR	Summary
SVCs	<i>Community: Ase-Taveak</i>	<i>Mele Maat</i>	0.86	<i>Speakers in Mele Maat are 14% less likely to use SVCs than speakers in Ase-Taveak</i>
	Education: <7	7–10	0.77	More educated speakers are around a quarter less likely to use SVCs than speakers who have only had primary-level education.
		11+	0.70	
	<i>Genre: Miscellaneous</i>	<i>Animals</i>	1.51	<i>Speakers are one and a half times as likely to use SVCs when telling animal stories.</i>
<i>Job</i>		0.89	<i>Speakers are 10% less likely to use SVCs when telling their job history.</i>	
AVCs	<i>Genre: Miscellaneous</i>	<i>Job</i>	0.64	Speakers are a third less likely to use AVCs when telling their job history.
	<i>Community: Ase-Taveak</i>	<i>Mele-Maat</i>	0.76	<i>Speakers in Mele Maat are nearly a quarter less likely to use AVCs than speakers in Ase-Taveak.</i>
	<i>Age: 31–45</i>	46–65	1.31	<i>Speakers aged 46–65 are nearly a third more likely to use AVCs than speakers aged 31–45.</i>
0Comp	<i>Genre: Miscellaneous</i>	<i>Animals</i>	0.72	<i>Speakers are more than a quarter less likely to use zero-marked complement clauses when telling animal stories.</i>
	<i>Education: <7</i>	7–10	1.30	<i>Speakers who have had 7–10 years of education are nearly a third more likely to use zero-marked complement clauses than speakers who have only had primary level education.</i>

9.6 Summary

To conclude, this chapter has tested the hypothesis that verbal constructions are used at different rates in the subcorpora collected from different speaker communities, and tried to assess the contribution of various sociolinguistic factors to variation in construction frequency. While many trends in community membership and frequency of constructions were observed in §9.3, community was a not significant parameter once other sociolinguistic factors were included in models. However, there was a lower rate of use of SVCs and AVCs in Mele Maat than in Ase-Taveak, which approached significance by both measures.

Both of these trends tend to confirm the hypothesis that speakers in Mele Maat would be less likely to use the constructions that are most dissimilar to English, and, to a lesser extent, Bislama. The significant relationship between education and frequency of SVCs seems to substantiate the hypothesis that greater exposure to English could be a possible cause of community level differences. More educated speakers were up to a quarter less likely to use SVCs, and this effect was slightly higher for the highest education group, who are likely to have been most exposed to English and most removed from Vatlongos-speaking communities. The fact that the most significant effects were observed for SVCs rather than the other constructions, confirms the cross-linguistic impression that SVCs are prone to contact effects, including reduction in use leading to their loss (Aikhenvald 2006: 53).

It was also interesting that the prior motion SVCs did not show any of the same conditioning factors as the other SVCs. It would be interesting in further research to break down SVCs for different functions even further, to see if variation could be isolated to a few key functions, while other functions are more stable. A hypothesis would be that functions that are not available in Bislama, or where formal differences from Bislama SVCs are more evident, would show more variation by speaker education level and community membership.

Education level also had an effect on the frequency of zero-marked complement clauses. Education beyond primary level increased the likelihood of the use of these subordinate structures. Speakers who had 7–10 years of education were a third more likely to use zero-marked complement clauses, although this effect was only significant by the less robust measure. However, this effect tailed off with education beyond 10 years, suggesting that competence in Vatlongos may decrease due to more extreme exposure to English at this level.

Two genres had significant or near-significant effects on frequency of constructions. Texts in the job genre were significantly less likely to contain AVCs (only 64% as often as miscellaneous texts), and SVCs (89% as often as miscellaneous texts), although the effect on SVCs was only significant by the less

robust measure. This seems to suggest that SVCs are not used when discussing more modern or Western concepts, which could be another explanatory factor in the lower construction frequency for SVCs in Mele Maat texts. Conversely, SVCs were more common in animal stories, where SVCs were nearly one and half times as likely to occur. Despite their apparent decline in Mele Maat, their presence in these child-directed narratives could support the maintenance of these morphologically complex structures. On the other hand, zero-marked complement clauses were a quarter less likely to occur in animal stories, suggesting that multi-clause structures might be avoided in child-directed speech, in contrast to the monoclausal SVCs.

Interestingly, age did not have a significant effect on the use of either AVCs or SVCs, despite the hypothesis that these structures are involved in diachronic changes resulting in the emergence of AVCs from SVCs (§8.3). In fact, for SVCs, the relationship with age that was significant according to the chi-square tests (§9.4) disappeared completely once other factors were included in the model (§9.5). This shows the importance of modelling frequencies in a way that includes multiple factors, especially with unbalanced corpora (§9.2). This suggests that any changes in these structures are affecting all age groups similarly, rather than changing in apparent time. Sankoff (2006) argues that while observed trends in apparent time are a robust indicator of diachronic change, apparent time often underestimates the rate of change because older speakers adjust to new community norms.

Although there are no significant effects for complex predicates, the trends are in keeping with the hypothesis that this less productive construction indexes conservative speech; they are more frequent in Endu and less frequent in Mele Maat. Complex predicates also had the most overdispersed frequencies by text, suggesting that they are more lexicalised than the other grammatical constructions described in this section.

Overall, this chapter has given a snapshot of variation in the token frequency of constructions between different communities. It has demonstrated relationships between this variation and sociolinguistic factors, and found that community-level differences can be partly explained by different sociolinguistic

environments in these three communities, especially in the relocated, peri-urban Mele Maat community. This chapter has also demonstrated that appropriate statistical tools can allow this variation to be observed in a relatively small, imperfectly balanced corpus, of the kind which is typical in language documentation projects.

10 Conclusion

Verbal morphosyntax in Vatlongos exhibits patterns of variation in the frequency of use of morphological variants and morphosyntactic constructions that can be linked to sociolinguistic factors across three speaker communities, especially the relocated peri-urban community of Mele Maat. After documenting the sociolinguistic environment of each community (Chapter 2), the thesis describes Vatlongos verbal morphosyntax, highlighting areas of variation (Chapters 3–8), before investigating variation in the token frequency of verbal constructions in a corpus of spontaneous texts from across the speaker communities (Chapter 9).

In any documentation of an understudied language, it is likely that unanticipated linguistic features prove interesting or puzzling. The first section of the conclusion therefore summarises some of the most notable features of Vatlongos morphosyntax. Their interest lies either in the challenge of accounting for them in major theoretical frameworks, or their rarity in the region and language family. §10.2 draws together the different areas of variation in Vatlongos verbal morphosyntax that have been uncovered through qualitative (Chapters 3–8) and quantitative analysis (§5.1.1.2, Chapter 9), and comments on their implications for the study of language variation and change (§1.2). The conclusion then summarises the documentary, descriptive, theoretical and methodological contribution of the project (§10.3), and points out various avenues for future research (§10.4).

10.1 Notable features of Vatlongos verbal grammar

Simple clauses in Vatlongos exhibit many canonical features of Oceanic languages, including nominative-accusative alignment, SVO constituent order, head-marking, and optional syntactic expression of subjects and objects (Chapter 3). Vatlongos syntax interacts with its morphological paradigms and verb classes in ways we can now consider more fully, in light of the description of the verbal morphology in Chapter 5. Morphologically transitive verbs obligatorily occur with either an object suffix or an immediately post-verbal object NP. The object suffix paradigm is defective, with no morphological means for expressing first or second-person non-singular objects, or animate third-

person non-singular objects. This means that if an object of a transitive verb with one of these person-number-animacy combinations is topicalised in clause-initial position, a co-referential independent pronoun must also appear in post-verbal object position, to avoid violating the constraint that transitive verbs must have a post-verbal object NP when no object suffix is present (§3.1.1).

The interaction of syntax and morphological processes also results in the existence of small classes of verbs diachronically derived from incorporation of prepositions and adverbs (§3.2.1, §3.3.1), which are the only verbs that cannot be nominalized with the, otherwise, very productive *-en* suffix (§5.3.1). The status of the NP complement of the incorporated-preposition verbs is ambiguous, as shown by the variable acceptability and position of the negative clitic (§3.6.2). For some speakers these verbs cannot be negated, while for others the negative clitic can follow this NP, a position that is otherwise only possible following an object NP or the NP complement of the copular verb. However, no speakers allow the negative clitic to occur between the incorporated-preposition verb and its complement NP, a position that is available with object NPs, including object NPs instantiating a goal argument (§3.6.2.1).

While double marking of negation is common in Oceanic languages (Vossen & van der Auwera 2014), the variable positioning of the negative clitic in Vatlongos is unusual, and resists an analysis as a bipartite morpheme. The position of the negative clitic is partially conditioned by verb class and construction type, as well as by the semantic role of the object in transitive VPs and scope of negation. The pre-verbal positions mark existential negation of the subject, which is focussed when the clitic appears before the subject NP. Despite the number of different processes that appear to be involved in determining the position of the negative clitic, negative clauses nearly always require a negative clitic to be present, and never allow more than one negative clitic, even when they contain more than one verb (in SVCs, complex predicates and arguably AVCs), or even more than one negative prefix, notably in SVCs marked for TAM categories other than the non-future.

The extension of morphological TAM categories in Vatlongos into different TAM contexts differs from what has been described for both Paamese and other languages of Ambrym. These extensions also illustrate the difficulty in drawing a sharp line between mood (the realis/irrealis distinction that is reconstructed for proto-Oceanic) and temporality as organisational principles of TAM in Vanuatu languages (cf. Barbour 2011). The four main TAM categories are best differentiated from each other in their temporal extensions, but both these relative tenses and the three mood paradigms (imperative, prohibitive and apprehensive) can also be grouped according to the realis–irrealis divide that is prominent in the region and in Oceanic languages. The likely diachronic consequences of this divide can account for otherwise puzzling modal restrictions (of the prior to realis contexts) and overlaps in modal extensions (between immediate and distant future for many irrealis functions). The parallels of the prior category in the languages of West Ambrym could also shed light on contact between Ambrym languages. Another notable feature of the Vatlongos TAM system is the inclusion of the object NP in verbal repetition expressing duration or intensification, something that has not been attested in other Vanuatu languages, despite the very common use of repetition to express these categories (§4.3.2, §4.3.3).

Vatlongos exhibits multiple exponence of TAM and polarity (TAMP) features in the subject-indexing prefix paradigms, the negative prefixes and verb-initial consonant mutation, as well as the negative clitic. Each of these morphological processes divides TAMP categories in slightly different ways. The subject-indexing prefixes distinguish the four main TAM categories without indicating polarity, but there are different patterns of syncretism in different subject person-number combinations, sometimes grouping the two future tenses together and sometimes grouping immediate future with non-future. Aside from marking negation, the negative prefixes further distinguish non-future from all other TAM categories. The simplest pattern of verb-initial consonant mutation distinguishes non-future in affirmative polarity from all other TAMP environments, and more complex patterns additionally distinguish negative polarity, and affirmative immediate future in some subject person-number combinations. In light of the description of AVCs in Chapter 8, we can

identify two more morphologically marked divisions of TAMP contexts, distinguishing non-future and prior from the future contexts without regard for polarity, and additionally distinguishing immediate future in some subject person-number combinations. These many different patterns in different areas of morphology make it difficult to use formal evidence for the relationship between the four main TAM categories and polarity. Prohibitive mood is marked by the combination of apprehensive prefixes and the negative clitic, an example of distributed exponence.

The data on verb-initial consonant mutation described in §5.2.1 complicates the lexical class based analysis described by Parker (1968b; 1970a), and in Crowley's (1991) comprehensive overview of this process in Central Vanuatu languages. Instead of organising verbs into lexical classes that undergo the same patterns of mutation, it seems that speakers produce and accept many different combinations of mutations with different lexemes. It is not clear to what extent this amount of variation is an established feature of the Vatlongos system, or whether this is a recently regular system that is rapidly being lost. An overall characteristic of Vatlongos morphology, apparent both in verb-initial consonant mutation and in the subject-indexing prefixes, is the proliferation of variants that are not straightforwardly conditioned by morphological or phonological environments. These areas of variation are summarised in the next section, but it is worth observing here that these are a challenge to many models of morphology (Thornton 2012a).

The SVCs described in Chapter 6 exhibit interesting parallels and differences from Paamese SVCs, which have been very influential in the description and analysis of SVCs in Oceanic languages (Crowley 1987; 2002b). Like Paamese SVCs, Vatlongos SVCs are structured by dependencies in the marking of morphological TAMP on the component verbs. However, Vatlongos SVCs allow more variation in some TAMP environments, especially the prior, and the subsequent verb can be marked with negative prefixes, although the negative clitic only appears once in an SVC. Vatlongos SVCs also allow a range of constituents to intervene between the component verbs, including PP arguments and adjuncts, and adverbs. Like SVCs in many other Oceanic

languages, Vatlongos SVCs have a broad functional extension, and this itself can be a challenge for theories of verb serialisation that posit strict parallels between syntactic and semantic integration (Van Valin & LaPolla 1997: 481). Attempts to define SVCs that rely on exclusively functional criteria in order to enable cross-linguistic comparison (Haspelmath 2016) make it difficult to test these kinds of hypotheses, and risk artificially confirming them by excluding examples of serialisation in less common functions. It is interesting that a single morphosyntactic strategy that can be rigorously identified by language internal criteria is used for such a broad range of functions, and it is important to have analytical tools that do not obscure this observation.

In contrast, the complex predicates described in the first part of Chapter 7 are much more restricted, both functionally and lexically. While similar strategies have been described as serialisation in Paamese and other Vanuatu languages (Crowley 2002b), in Vatlongos these appear to be less productive than an analysis as serialisation would imply, and may be better analysed as a fairly productive verbal compounding strategy. However they are best labelled, they show an interesting mismatch between phonological and grammatical wordhood. The component lexemes appear to constitute a single grammatical word in terms of syntax and affix-based morphology, but the second verb can undergo separate non-linear morphological processes (verb initial consonant mutation and reduplication), and take separate word stress.

The processes of subordination described in §7.2 are very common cross-linguistically, but still reveal some interesting properties of Vatlongos morphosyntax. The possibility of raising the subject of the subordinate clause to the object of the matrix clause, and the requirement for a transitive matrix clause to take default third-person object agreement, both demonstrate the consequences of morphological transitivity, which always requires object marking with a suffix or NP object even without a semantic referent. This section has also demonstrated that the TAMP marking in a complement clause contributes a meaningful semantic distinction, rather than being predictable from the semantic properties of the matrix verb, as is described for Paamese (Crowley 2002b: 62).

Chapter 8 describes a distinctive AVC that is not found in Paamese or other related languages, and argues that it has developed from SVCs in recent generations, by erosion of verbal prefixes on initial verbs with functional semantics. Auxiliation is most advanced with the use of *ti* ‘stay’, which is the only strategy for expressing continuous aspect (defined as a combination of progressive and habitual). By analogy, this is also used with other ‘stay’ verbs, including verbs derived from compounding with the adverb *tang* ‘just’, to express continuous aspect with a minimising evaluative component. Since Parker’s (1970a) description, the verbs *ha* ‘go’ and *ammei* ‘come’ have been added to the inventory of auxiliary verbs; the latter is reduced to an invariant *mei*, while *ha* undergoes complex patterns of verb initial consonant mutation. These express the direction of prior motion, which can alternatively be expressed with an SVC (§6.4.3). This category is less obviously grammatical than continuous aspect, but seems prone to frequent expression and grammaticalization in this region. A Bislama auxiliary *mas* has also been borrowed as an auxiliary verb to express deontic modality. These auxiliary verbs could be analysed as the head of the Vatlongos clause: they occupy a dedicated syntactic position between the subject and verb, and can mark verbs in subordinate clauses but not the subsequent verb in an SVC (except with the double marking of prior motion discussed in §8.1.3).

10.2 Variation in Vatlongos verbal morphosyntax

Variation in the syntax of simple clauses (Chapter 3) is mainly the position of the negative clitic *ti*. An overarching community-level difference is the optionality of the negative clitic in Endu-Vatlongos. There are several examples in the corpus⁵⁰ where speakers from Endu, or with family connections to Endu, use a negative prefix without the clitic, which is ungrammatical for most speakers (except subsequent verbs in SVCs). As Endu-Vatlongos is otherwise a conservative dialect, this may reflect an earlier stage in Jespersen Cycles, where the partitive clitic is used to reinforce emphatic negation, rather than marking

⁵⁰ In this chapter I am using ‘corpus’ and ‘subcorpus’ as they are used in Chapters 3–8, to refer to the corpus of interlinearised texts, and the corpus of spontaneous interlinearised texts that are tagged for quantitative analysis respectively.

standard negation. However, speakers do not report a difference in emphasis. Another community-level difference occurs with psycho-collocations: whereas most speakers require a negative clitic between the subject and the verb in these constructions, speakers in Mele Maat, or connected to Mele Maat, can again choose not to use a negative clitic at all. The other main area of variation is in transitive clauses with an NP object. Here the variation is partially conditioned by the semantic role of the object, although in elicitation speakers accept both word orders with all verbs. This does not seem to correlate with speaker-level variables: no clear patterns emerge in the 80 examples in the subcorpus.

There is a choice between different morphological TAM categories in some contexts, especially between the prior and the non-future for events preceding topic time, the immediate future and distant future for events immediately following topic time, and between the two future categories in many irrealis contexts. No clear patterns in TAM choices emerge in the subcorpus, which may be too small to address this kind of question. The various patterns of syncretism in different subject person-number environments described in Chapter 5 also mean that many potential examples in the corpus are ambiguous between the relevant TAM categories.

There is a great deal of variation in the verbal morphology of Vatlongos, with multiple variants in single cells of the subject-indexing prefix and verb-initial consonant mutation paradigms. Firstly, there is variation between the *o-* and *a-* variants of the second-person singular non-future prefix. Whereas speakers in Endu overwhelming use the *o-* variant, speakers in Ase-Taveak and Mele Maat only use it about three-quarters of the time. The *a-* form seems to be more innovative, as it is not observed by Parker in the 1960s, and this corroborates speaker impressions that Endu-Vatlongos is more conservative. The variation between *mi-* and zero in the third-person singular non-future is partially phonologically conditioned, but is another area where Endu-Vatlongos differs from Vatlongos spoken elsewhere. Speakers in Endu are more likely to use the *mi-* form before prenasalised or bilabial consonants than speakers elsewhere, and again this seems to be a conservative feature. Finally, there is

variation to do with syllable shape: there is a choice between a CVCV- and a CVC- variant of the subject-indexing prefixes in many person-number-TAM combinations. The use of the CVC- forms results in patterns of syncretism that obscure the distinction between TAM categories in many person-number combinations. This is an area of variation where the island communities of Ase-Taveak and Endu pattern together, using the CVC- variants for about half of the tokens in the corpus, whereas speakers in Mele Maat use the shorter, less specific variants for nearly two-thirds of tokens. This could be a sign of simplification of the verbal morphology in Mele Maat. Community-membership was statistically significant in chi-squared tests for all these areas of variation in the subject-indexing prefixes.

Some additional patterns of verb-initial consonant mutation in the corpus were associated with particular villages in Southeast Ambrym. The k/g alternation was associated with Endu, Ase and Sameo, while the use of /r/ with /t/ initial verbs in the affirmative non-future is a salient feature of traditional speech in Toak. However, rather than relying on the corpus, overall variation in verb-initial consonant mutation was addressed through elicitation and grammaticality judgements with all verb lexemes found in Parker's (1970a) dictionary, and in the growing interlinearised corpus, during the second period of fieldwork. There were found to be many more patterns of acceptable variants in the different environments than could be accounted for with the nine verb classes proposed by Parker. Although these could not be repeated with speakers from each community, when speakers were asked about the same verb lexeme they did not always agree, and further options were sometimes found in the corpus. More frequent verb lexemes, like the copular verb *he* and *ha* 'go', show less variation in each environment, and tend to use the most specific variant to the exclusion of others, whereas lower frequency verbs have more variants in each environment, with the basic consonant extending into more contexts. If the more specific variants are then lost this could to simplification of the system, with irregularity maintained only for the most frequent verb lexemes.

Both the variation in verb-initial consonant mutation and between CVCV- and CVC- subject-indexing prefixes illustrate a paradox of simplification in process. Simplification has been characterised as both the loss of allomorphs or forms, and the extension of a form into more contexts, but in the short term this extension can lead to more variants being available to speakers (Maher 1991: 68; Silva-Corvalán 1991: 152). If the more specific forms (the CVCV- subject-indexing prefixes and the verb stems that identify more specific TAM environments) do indeed go on to be lost altogether, these developments would lead to a simpler verbal morphology in Vatlongos. However, the synchronic situation described in this project is highly complex and involves a proliferation of variants within paradigmatic environments. Longitudinal research in the future may resolve the consequences of these changes, but if the current system is maintained it could be argued to constitute complexification rather than simplification.

Chapter 9 describes variation in the frequency of verbal constructions, and investigated connections with community membership, other speaker characteristics and genre. The most robust effects on frequency of verbal constructions were the effects of community membership and education level on SVCs: when other factors are held constant, more educated speakers are a quarter less likely to use SVCs than speakers who are educated to a primary level, and there was also a nearly significant effect whereby speakers in Mele Maat are 14% less likely to use SVCs than speakers in Ase-Taveak. The negative correlation of SVC use with urban lifestyles and higher education was also reflected in the borderline significant genre effects: speakers are a tenth less likely to use SVCs in employment history texts, and one and a half times more likely to use them in traditional child-directed animal narratives. AVCs show a broadly similar pattern. There was a significant effect of genre on the use of AVCs, with speakers a third less likely to use AVCs in employment histories than in miscellaneous texts. There was also an effect of community membership that approached significance, whereby speakers in Mele Maat are a quarter less likely to use AVCs than speakers in Ase-Taveak. Overall, these patterns of variation in the frequency of constructions confirm that new standards can emerge in early stages of dialect differentiation, and that heavy exposure to

Bislama and English can lead to a lower rate of use of the constructions that are most characteristic of Vatlongos, especially in domains like employment that are most closely associated with the national languages.

10.3 Significance of the project

This thesis contributes to the documentation of an under-described Oceanic language of Vanuatu, and to the study of language variation and change in contexts of migration and urbanisation. It has also trialled a workflow for documentation involving several innovations that could be utilised in the study of other endangered languages, with the potential to improve the efficiency, accessibility and comparability of language documentation projects, and opening them up to the study of new kinds of research questions.

The documentary and descriptive goals of this project were chosen to fill gaps in the previously available materials on Vatlongos. The project has produced the first corpus of Vatlongos texts and recordings, comprising a wide range of text-types and contributions from male and female speakers in all generations, from every Vatlongos-speaking village. Time-aligned transcriptions and translations in Bislama have been produced for the majority of recording events, and a third of the recordings have been interlinearised in a lexical database with English translations, resulting in a corpus of >65,000 words that is maximally accessible for linguists, anthropologists, other researchers, and semi-speakers in the community (especially in the Mele Maat community). Steps are being taken to archive the corpus, making as much of the material as possible available online.

The project has also facilitated the production of a range of community-directed materials, including videos, subtitles, e-books, and a dictionary app, all of which have been distributed widely in formats that can continue to be shared from speaker to speaker. These materials, improved by the consistency of analysis and orthography facilitated by lexical database software, can support the vernacular literacy and education goals of both the Vatlongos-speaking community and the national government (Vanuatu Ministry of Education 2012).

The project has updated and supplemented earlier research on Vatlongos and its speakers. The sociolinguistic survey described in Chapter 2, reveals developments in the Mele Maat community of interest to the anthropological literature (Tonkinson 1968; 1979; 1982; Johansen 2012), as well as providing more detail on the language ecology of all Vatlongos-speaking communities, with implications for language vitality. Previous work on the language has concentrated on phonology, morphophonemics and lexicography (Parker 1968a; 1968b; 1970b; 1970a), so the descriptive element of this project has centred instead on verbal morphosyntax. It especially focuses on constructions involving multiple verbs, an area that has been well-documented for closely-related Paamese (Crowley 1987; 2002b). This sheds new light on the similarities and differences between these two languages, and adds to a rich and growing literature on verb-serialisation and adjacent constructions in Oceanic languages.

This thesis explores questions about language variation, change and divergence by taking the sociolinguistic context of Vatlongos-speaking communities as an opportunity to investigate evidence of incipient language divergence. While work on language variation and change often takes signs of linguistic divergence as a starting point for building hypotheses about the history and circumstances of social divisions and language contact, this project examines language use in a speech community which has undergone a recent and well-documented split, to look for signs of differentiation in the speech patterns of each group. Thanks to the anthropological research of Tonkinson (1968; 1979; 1981; 1982; 1985; 2011), we have access to a wealth of information about the circumstances of Mele Maat's relocation, and the social context of both the relocated and original communities since the initial displacement. This longitudinal historical detail, based on contemporary observations, removes some of the conjecture that is usually inherent in studies of language divergence. The Mele Maat context has also allowed this project to investigate linguistic divergence at a very early stage, just a few generations after the separation of the communities. Overall, the results of the project seem to corroborate a model of language divergence whereby innovative or pre-existing variants spread at differential rates in separated communities, at least

partially conditioned by differences in social factors and patterns of language contact.

The thesis makes two types of contribution to the methodology of language documentation and description. Firstly, innovations in the combination of software tools used in this project have demonstrated potential for gains in efficiency of language documentation workflows, and the potential to use the corpora arising from language documentation projects to address new kinds of research questions.

The production of time-aligned transcriptions and translations was considerably accelerated by the use of SayMore's (Moeller 2014) autosegment tool, in conjunction with Audacity (2017), to produce mp3 files of individual segments that can be played on cheap mobile phones or mp3-players. This meant that independent transcription could proceed quickly, without the need for expensive hardware like laptops, or training in specialised software like Elan, which are not especially useful to participants beyond the goals of linguistic projects. As mobile phone coverage and access to small-scale solar power becomes more and more common in remote areas worldwide, this workflow will be suitable for an increasing number of endangered language contexts. It is ideal for situations where literacy and access to mobile phones are widespread, but not computer literacy, computing hardware, or reliable and powerful sources of electricity. However, where there is interest in developing IT skills and documentation-specific skills in the community, and the infrastructure to use and maintain hardware, budgeting for laptops and IT training should be the preferred approach.

The second innovation in combining software tools is importing a corpus of interlinearised texts in a FLEx (2018) database into the statistical environment R (2018) for quantitative analysis, using the *interlineaR* package (Loiseau 2018). While FLEx and its precursor Toolbox (Buseman 2018) are commonly used in language documentation projects, importing the resulting corpora into R greatly increases the possibilities for quantitative corpus analysis, beyond the concordancing tools available within the linguistic database softwares.

This project has made use of only two of the functions made possible by accessing the corpus in R. Firstly, in R it is possible to cross-reference the corpus with any other sources of data, in this case the results of the sociolinguistic speaker survey, and text genres. While some of these fields could have been included in the FLEx database, there are many situations where it is useful to introduce additional sources of information, or where it is more useful to store this information in a different format. Geographical information, speaker family trees, social network analysis and any other data relevant to a particular project could be incorporated in the same way. Secondly, using R allows tags in the note field to be accessed, counted and manipulated, a capacity that is very limited within FLEx. In this case, a simple system of construction tags was used, but more complex systems making full use of R's capacity for handling character strings could have great potential for exploring complex interactions and research questions, in such a way that annotated information continues to be available as the FLEx database grows and is updated.

A functionality that was not exploited in this project, but is also made possible by the *interlinearR* package, is cross-referencing the interlinearised texts with the FLEx lexicon, including custom fields. This could greatly expand the possibilities for researching lexical effects beyond the basic lexeme reference, gloss, and morpheme category information available in the EMELD interlinear format.

Once these different kinds of information have been imported into R, all of R's very powerful and flexible tools for statistical analysis are available in support of the research goals. This project has employed tools for chi-square tests and negative binomial regression, but as R is an open-source software and a programming language that is widely used by statisticians and developers, there is no practical limit to the kinds of statistical analysis that can be implemented.

At the level of data structure, this thesis contributes to methodological approaches to syntactic variation by tagging constructions per time-aligned segment within a corpus. While tagging at morpheme-level and word-level, including for part of speech, is common in corpus linguistics and especially in

the construction of corpora for language documentation, linguistic features beyond the word-level have received less attention. One reason for this is the difficulty of modelling the potential for recursive hierarchies in xml data structures in a way that is robustly retrievable for further analysis. FLE_x and ELAN are both structured around xml files that are accessible because of the consistency of the data structure: this consistency makes it possible for a certain path through the nodes of the structure to reliably return data of a certain type. Once syntactic recursion is included in these models, the relationship between node-paths and information type is lost. For example, it might be possible to retrieve NPs in a main clause, but not subordinate clauses. FLE_x allows tagging of larger constituents, but without any nesting, and models clause structure with a text-frame approach from RRG (Van Valin & LaPolla 1997; Van Valin 2010). Both of these tools are of limited value as they cannot be exported to other formats.

One exception is the Grammatical Relations and Animacy in Discourse (GRAID) annotation system (Haig & Schnell 2014), which particularly focuses on referential expressions in the clause. While it includes means for glossing certain syntactic constructions, such as complement clauses, relative clauses, and verb complexes, these force the annotator to make explicit decisions about the boundaries of constituents within these constructions. While this is undoubtedly important for the rigorous cross-linguistic comparative work that the GRAID system enables, it is often not practical in the early stages of a documentation project, and relies on an already complete descriptive analysis of challenging constructions.

The advantage of the system of tagging used in this project, where just the presence of a construction is tagged, is that it can be used before an analysis of the constituency of the construction is complete, and then used to conduct concordances which can inform that analysis. It can also be used to accelerate a subsequent, more fine-grained annotation once the analysis is secure. While it does not attempt to address the problem of hierarchical structure within syntax, this approach allows the presence of constructions to be registered within a corpus that can continue to be updated, and is accessible for future

projects and other researchers. In this respect it differs from approaches where examples are extracted into a separate database and then coded for features that are relevant to the goals of an individual project.

The value of this simple annotation strategy is made clear by the final methodological contribution of this thesis. Chapter 9 investigates intra-language variation in the token frequency of constructions (measured by word count per text) and how it correlates with sociolinguistic variables. The speaker-level variables of community membership and level of education, and text-level variable of genre were shown to be relevant to the distribution of SVCs, AVCs and zero-marked complement clauses in Vatlongos. This demonstrates that different community standards can emerge in relation to construction frequency, and that these morphosyntactic constructions are vulnerable to shifts in sociolinguistic environment.

10.4 Directions for future research

Many possible avenues for future research have been flagged in the main body of the thesis, but some of the main priorities are summarised here.

The sociolinguistic survey described in Chapter 2 provided a broad outline of the social and linguistic context of the Vatlongos-speaking communities, and could form the basis of more detailed documentation of this setting. Firstly, responses to the open questions about language attitudes could inform a more systematic survey of language ideology, using frequent responses as example views, and asking respondents to rate how far they agree with each statement. This would make it possible to establish how widely-held the views are in each community, and investigate how different views correlate with each other. Using these unprompted responses from Vatlongos speakers will minimise the risk of introducing value systems from external frameworks. Two features of Vatlongos contexts that would be interesting to investigate in more targeted ways are delayed acquisition of Vatlongos in Mele Maat, and patterns of multilingualism with North Ambrym languages in Endu. Finally, the sociolinguistic contexts and speech patterns of Vatlongos speakers in other urban communities in Port Vila and Luganville (Santo) would make for an illuminating comparison with Mele Maat.

A limitation of the descriptive aspects of this thesis is the neglect of phonetic and phonological features of Vatlongos. This was a considered choice given the time available and the detailed previous research on Vatlongos phonology and morphophonology that was the main focus of Parker's (1968b; 1968a; 1970b) articles. However, more detailed acoustic and articulatory research could illuminate patterns of variation in morphology, and the nature of some constructions. For example, acoustic analysis of the realisation of /o/ and /a/ in different communities could reveal whether this is part of a restructuring of Vatlongos phonology, or a pattern of variation restricted to particular lexical and morphological items. Similarly, the few examples of vowel harmony in the corpus have been identified impressionistically, and would benefit from closer acoustic investigation. Establishing the phonetic correlates of word stress in Vatlongos would strengthen the analysis of prosodic wordhood in relation to the clausal clitics and complex predicates.

The description of Vatlongos verbal morphosyntax outlined here is also limited at the boundary between syntax and discourse. The information status of discourse referents has been mentioned in relation to non-basic word orders, but a more thorough analysis of the tracking of discourse referents, prosodic marking of topic or focus status and other discourse cohesion strategies like tail-head linkage would be a fruitful avenue for future research. Tagging for information status of referents could add value to the corpus constructed for this project, and help to shed light on how variants of the subject-indexing prefixes and the constructions tagged for this project interact with discourse factors, in addition to the sociolinguistic factors considered in Chapter 9.

Some clear directions for investigating the observed patterns of variation would involve separation of paradigmatic, functional or lexical environments. Breaking down the variation between CVC- and CVCV- prefixes into different TAM, subject person-number and verb stems undergoing verb-initial consonant mutation could reveal whether the CVCV- forms are more likely to be used when they have a higher functional load, that is, when morphological TAM is not uniquely identified if the CVC- form is used. Dividing tokens of SVCs by the different functions outlined in Chapter 6 could help isolate the exact functions

where the observed variation in frequency occurs. Similarly, looking at the frequency of each auxiliary verb, rather than grouping all AVCs together, could reveal differences in distribution, addressing the hypothesis that more variation will be observed for the less grammaticalized prior-motion auxiliaries than the most grammaticalized continuous auxiliary. Now that the functions of each construction have been identified, it is possible to isolate semantic and grammatical functions that can be expressed with more than one strategy, allowing a more traditional variationist approach. These include the use of SVCs and prepositions to incorporate goal arguments; the use of complementisers, SVCs, or both to introduce complement clauses and quotations; and the use of SVCs and AVCs to express prior motion.

Some of the contexts for variation that have been identified are unlikely to occur often enough in the corpus to produce enough tokens for quantitative analysis, especially those restricted to less frequent TAM-person-number or lexical environments. It would therefore be helpful to design elicitation materials, especially narrative contexts, to target these environments with speakers in each community. For example, parallel elicitation of variable TAM contexts with storyboards targeting unambiguous subject person-number combinations, or verbs that undergo disambiguating verb-initial consonant mutation, could address the question of community level differences in TAM extensions. It would be particularly interesting to see if the TAM system is being simplified in the direction of Bislama's three-way distinction between prior, non-future and future in the Mele Maat community, perhaps through the loss of the immediate future.

In terms of methodology, there are clear routes for developing the system of construction tagging. A pragmatic extension of the system would be to exploit the orthographic conventions of punctuation, so as to link the presence of constructions to a more syntactically relevant unit than the intonation-based temporal segments automatically generated in SayMore. The use of full stops and speech marks in the transcriptions is independently motivated by community education goals, where it is hoped that early literacy in Vatlongos will support literacy in English. English punctuation conventions have therefore

been followed as closely as possible so that the corpus can be easily used to create community materials. These conventions also convey important syntactic information about the boundaries of sentence units made up of one or more clauses, and the difference between quotations and complement clauses, which is often syntactically relevant cross-linguistically. If these were also included in the phrase-level note field in FLE_x, they could be used to regroup both words and construction tags into sentences rather than temporal segments in the dataframes in R. This would make it possible to address questions about the cooccurrence of constructions at the sentence level rather than only at text level. It would also allow for comparisons of measures of syntactic complexity across communities, age groups and education levels, a line of enquiry that follows on from the conclusions of Chapter 9.

The tagging system could also be augmented in numerous ways to register the more detailed categories that have been revealed in the course of the qualitative analysis described in this thesis, for example to reflect the different functions of SVCs. The goal in developing the tagging system further will be to ensure that all new kinds of information being included can be cross-referenced with the information that has already been annotated.

10.5 Summary

This thesis has investigated variation in an unusual sociolinguistic environment, using the recent and well-documented relocation of the Vatlongos-speaking Mele Maat community as a springboard for investigating early stages of language divergence under the pressures of urbanisation and globalisation. It has also examined more established dialectal differences in Endu, the northernmost Vatlongos-speaking village. It has focused on verbal morphosyntax as an area of grammar that exhibits a great deal of variety in languages of Vanuatu, and has described the syntax of simple clauses, the TAM categories, the verbal morphology, and several complex verbal constructions in the language, identifying many features of regional and cross-linguistic interest. Throughout the descriptive portions of this thesis, careful attention has been paid to formal variation, and inter-community differences in the frequency of morphological variants have been identified. Through quantitative analysis of a

tagged corpus of spontaneous texts, the thesis has demonstrated some robust differences in the frequency of SVCs and AVCs connected to community membership, level of education, and genre.

References

- Acciaoli, Greg. 2013. The Dialectic of Robert Tonkinson's Dynamic Mode of Argumentation: Paradoxes, Adaptive Strategies, and Accommodations. *Anthropological Forum* 23(4). 321–341.
- Aikhenvald, Alexandra Y. 2006. Serial verbs constructions in a typological perspective. In Alexandra Y. Aikhenvald & Robert M. W. Dixon (eds.), *Serial verb constructions: A cross-linguistic typology*, 1–68. Oxford University Press.
- Aikhenvald, Alexandra Y. 2007. Typological distinctions in word-formation. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 3, 1–65. 2nd ed. Cambridge: Cambridge University Press.
- Albury, Nathan John & Lyn Carter. 2018. "An unrealistic expectation": Māori youth on indigenous language purism. *International Journal of the Sociology of Language* 2018(254). 121–138. doi:10.1515/ijsl-2018-0036.
- Allaire, JJ, Yihui Xie, Jonathan McPherson, Javier Luraschi, Kevin Ushey, Aron Atkins, Hadley Wickham, Joe Cheng & Winston Chang. 2018. *Rmarkdown: Dynamic documents for R*. <https://CRAN.R-project.org/package=rmarkdown>.
- Anderson, Gregory D. S. 2006. *Auxiliary verb constructions*. Oxford University Press.
- Anderson, Gregory D. S. 2011. Auxiliary Verb Constructions (and Other Complex Predicate Types): A Functional–Constructional Overview. *Language and Linguistics Compass* 5(11). 795–828. doi:<https://doi.org/10.1111/j.1749-818X.2011.00311.x>.
- Anderson, Stephen R. 1992. *A-morphous morphology* (Cambridge Studies in Linguistics 62). Vol. 62. Cambridge University Press.
- Andrews, Avery D. 1997. Complex predicates and nuclear serial verbs. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG97 Conference*. CSLI Publications, Stanford University University of California, San Diego.
- Andrews, Avery D. 2007. The major functions of the noun phrase. In Timothy Shopen (ed.), *Language typology and syntactic description: Clause structure*, vol. 1, 62–154. 2nd ed. Cambridge: Cambridge University Press.
- Audacity Team. 2017. *Audacity: Free, open source, cross-platform audio software for multi-track recording and editing*. <https://www.audacityteam.org/>.
- Austin, Peter & Joan Bresnan. 1996. Non-Configurationality in Australian Aboriginal Languages. *Natural Language & Linguistic Theory* 14(2). 215–268.
- Baayen, R. Harald. 2001. *Word frequency distributions* (Text, Speech and Language Technology 18). Kluwer Academic Publishers.
- Baayen, R. Harald. 2008. *Analyzing linguistic data: a practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Baerman, Matthew, Dunstan Brown & Greville G. Corbett. 2005. *The syntax-morphology interface: A study of syncretism*. Cambridge University Press.
- Baker, Mark C. 1989. Object sharing and projection in serial verb constructions. *Linguistic inquiry* 513–553.
- Baker, Mark C. 1996. *The polysynthesis parameter*. Oxford University Press.

- Baker, Mark C. 2001. Configurationality and polysynthesis. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals: An international handbook*, vol. 2, 1433–1441.
- Barbour, Julie. 2011. Exploring Mood in Neverver. *Oceanic Linguistics* 50(1). 198–220.
- Barbour, Julie. 2015. Jespersen Cycles on Malekula. *Linguistic Typology* 19(3). 425–462.
- Bhat, D. N. Shakara. 1999. *The prominence of tense, aspect and mood* (Studies in Language Companion Series 49). John Benjamins Publishing.
- Bickel, Balthasar, Bernard Comrie & Martin Haspelmath. 2015. The Leipzig Glossing Rules: Conventions for interlinear morpheme-by-morpheme glosses. <https://www.eva.mpg.de/lingua/resources/glossing-rules.php>.
- Bickel, Balthasar & Johanna Nichols. 2007. Inflectional morphology. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 3, 169–240. 2nd ed. Cambridge: Cambridge University Press.
- Bisang, Walter. 2009. Serial Verb Constructions. *Language and Linguistics Compass* 3(3). 792–814. doi:10.1111/j.1749-818X.2009.00128.x.
- Blanche-Benveniste, Claire. 2006. Linguistic analysis of spoken language: the case of French language. In Yuji Kawaguchi, Susumu Zaima & Toshihiro Takagaki (eds.), *Spoken language corpus and linguistic informatics*, 35–66. John Benjamins Publishing Company.
- Blevins, Juliette & John Lynch. 2009. Morphological Conditions on Regular Sound Change?: A Reanalysis of* l-loss in Paamese and Southeast Ambrym. *Oceanic Linguistics* 48(1). 111–129.
- Bohnemeyer, Jürgen. 2014. Aspect vs. relative tense: The case reopened. *Natural Language & Linguistic Theory* 32(3). 917–954.
- Bohnemeyer, Jürgen, Melissa Bowerman & Penelope Brown. 2001. Cut and break clips. In Stephen C. Levinson & Nick J. Enfield (eds.), *Manual for the field season 2001*, 90–96. Nijmegen: Max Planck Institute for Psycholinguistics.
- Bow, Catherine, Baden Hughes & Steven Bird. 2003. A Four-Level Model for Interlinear Text. <http://minerva-access.unimelb.edu.au/handle/11343/33786>.
- Brenzinger, Matthias, Akira Yamamoto, Noriko Aikawa, Dmitri Koundioubá, Anahit Minasyan, Arienne Dwyer, Colette Grinevald, et al. 2003. Language vitality and endangerment. *Paris: UNESCO Intangible Cultural Unit, Safeguarding Endangered Languages*.
- Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2015. *Lexical-functional syntax* (Blackwell Textbooks in Linguistics 16). 2nd edn. John Wiley & Sons.
- Bresnan, Joan & Sam A. Mchombo. 1987. Topic, Pronoun, and Agreement in Chicheŵa. *Language* 63(4). 741–782. doi:10.2307/415717.
- Bril, Isabelle. 2004. Complex nuclei in Oceanic languages: contribution to an areal typology. In Isabelle Bril & Françoise Ozanne-Rivierre (eds.), *Complex predicates in Oceanic languages: Studies in the dynamics of binding and boundness*, 1–48.

- Bril, Isabelle. 2007. Nexus and juncture types of complex predicates in Oceanic languages: Functions and semantics. *Language and linguistics (Academia Sinica)* 8(1). 267–310.
- Bril, Isabelle & Françoise Ozanne-Rivierre. 2004. *Complex predicates in Oceanic languages: Studies in the dynamics of binding and boundness*. Walter de Gruyter.
- Brown, Dunstan, Marina Chumakina & Greville G. Corbett. 2013. *Canonical morphology and syntax*. Oxford University Press.
- Budd, Peter. 2009. *Topics in the grammar of Bierebo, central Vanuatu, with a focus on the realis/irrealis categories*. School of Oriental and African Studies (University of London) PhD Thesis.
- Budd, Peter. 2010. Negation in Bierebo and the Other Languages of Epi, Central Vanuatu. *Oceanic Linguistics* 49(2). 511–542.
- Budd, Peter. 2011. On borrowed time? The increase of Bislama loanwords in Bierebo. (Ed.) Julia Sallabank. *Language Documentation and Description* 9. 169–198.
- Bugenhagen, Robert D. 2010. Event integrity and continuity in the Tuam language. In John Bowden, Nikolaus P. Himmelmann & Malcolm D. Ross (eds.), *A journey through Austronesian and Papuan linguistic and cultural space: Papers in honour of Andrew K. Pawley*, 455–474.
- Buseman, Alan. 2018. *Field Linguist's Toolbox*. SIL International. <https://software.sil.org/toolbox/>.
- Butt, Miriam & Melanie Seiss. 2009. Complex predicates: an overview. University of Konstanz, ms.
- Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form* (Typological Studies in Language 9). John Benjamins Publishing.
- Bybee, Joan L., Revere Dale Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. University of Chicago Press.
- Caballero, Gabriela & Alice C. Harris. 2012. A working typology of multiple exponence. In Ferenc Kiefer, Mária Ladányi & Péter Siptár (eds.), *Current issues in morphological theory:(Ir) regularity, analogy, and frequency*, 163–188.
- Carroll, Matthew J. 2016. *The Ngkolmpu language with special reference to distributed exponence*. Australian National University PhD Thesis.
- Clark, David. 2009. **Leo tuai: A comparative lexical study of North and Central Vanuatu languages*. Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University.
- Cleary-Kemp, Jessica. 2015. *Serial verb constructions revisited: A case study from Koro*. University of California, Berkeley PhD Thesis.
- Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge University Press.
- Comrie, Bernard. 1985. *Tense*. Cambridge University Press.
- Corbett, Greville G. 2000. *Number*. Cambridge University Press.
- Cristofaro, Sonia. 2005. *Subordination*. Oxford University Press.
- Croft, William. 2000. *Explaining language change: An evolutionary approach*. Pearson Education.
- Crowley, Terry. 1982. *The Paamese language of Vanuatu* (Pacific Linguistics Series B 87). Australian National University.

- Crowley, Terry. 1983. Development of a Paamese transitive suffix. *Papers from the Third International Conference on Austronesian Linguistics* (Pacific Linguistics), vol. 4: Thematic variation, 269–283. Australian National University.
- Crowley, Terry. 1987. Serial verbs in Paamese. *Studies in language* 11(1). 35–84.
- Crowley, Terry. 1990. *Beach-la-Mar to Bislama: The emergence of a national language in Vanuatu*. Oxford: Clarendon Press.
- Crowley, Terry. 1991. Parallel development and shared innovation: Some developments in Central Vanuatu inflectional morphology. *Oceanic Linguistics* 30(2). 179–222.
- Crowley, Terry. 1992a. *An introduction to historical linguistics*. Oxford University Press.
- Crowley, Terry. 1992b. *A dictionary of Paamese* (Pacific Linguistics Series C 121). Australian National University.
- Crowley, Terry. 1995a. Inalienable possession in Paamese grammar. In Hilary Chappell & William McGregor (eds.), *The Grammar of Inalienability: a typological perspective on body part terms and the part-whole relation* (Empirical Approaches to Language Typology 14), 383–432. Walter de Gruyter.
- Crowley, Terry. 1995b. *A new Bislama dictionary*. Institute of Pacific Studies; Pacific Languages Unit.
- Crowley, Terry. 2000. The Language Situation in Vanuatu. *Current Issues in Language Planning* 1(1). 47–132. doi:10.1080/14664200008668005.
- Crowley, Terry. 2002a. Southeast Ambrym. In John Lynch, Malcolm D. Ross & Terry Crowley (eds.), *The Oceanic languages* (Curzon Language Family Series), 660–670. Richmond, Surrey: Curzon Press.
- Crowley, Terry. 2002b. *Serial verbs in Oceanic: A descriptive typology*. Oxford University Press.
- Crowley, Terry. 2004a. The question of dialect and language in Oceania. In John Bowden & Nikolaus P. Himmelmann (eds.), *Papers in Austronesian subgrouping and dialectology* (Pacific Linguistics), 3–20. Canberra: Australian National University.
- Crowley, Terry. 2004b. *Bislama reference grammar* (Oceanic Linguistics Special Publication 31). University of Hawai'i Press.
- Crowley, Terry. 2004c. Borrowing into Pacific languages: Language enrichment or language threat? In Jan Tent & Paul Geraghty (eds.), *Borrowing: a Pacific Perspective* (Pacific Linguistics), 41–53. Canberra: Australian National University.
- Crowley, Terry. 2005. Competing agendas in indigenous-language renewal: Initial vernacular education in Vanuatu. *International journal of the sociology of language* 2005(172). 31–49.
- Crowley, Terry. 2006. *Naman: a vanishing language of Malakula (Vanuatu)* (Pacific Linguistics). Australian National University.
- Cru, Josep. 2016. Shifting Language Ideologies among Young Maya Professionals: Overcoming Purism in Yucatán. *Critical Multilingualism Studies* 4(2). 111–132.
- Crystal, David. 2000. *Language Death*. Cambridge University Press.
- Dahl, Östen. 1985. *Tense and aspect systems*. Oxford: Basil Blackwell.
- Dalrymple, Mary. 2001. *Lexical functional grammar*. Brill.

- Dayal, Veneeta. 1998. Any as inherently modal. *Linguistics and philosophy* 21(5). 433–476.
- Dayal, Veneeta. 2004. The universal force of free choice any. *Linguistic variation yearbook* 4(1). 5–40.
- Decker, Donald M. 1999. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge University Press.
- Defina, Rebecca. 2016a. *Events in language and thought: The case of serial verb constructions in Avatime*. Nijmegen: Radboud University PhD Thesis.
- Defina, Rebecca. 2016b. Serial verb constructions and their subtypes in Avatime. *Studies in Language* 40(3). 648–680. doi:10.1075/sl.40.3.07def.
- Defina, Rebecca. 2016c. Do serial verb constructions describe single events?: A study of co-speech gestures in Avatime. *Language* 92(4). 890–910. doi:10.1353/lan.2016.0076.
- Desagulier, Guillaume. 2017. *Corpus Linguistics and Statistics with R: Introduction to quantitative methods in linguistics* (Quantitative Methods in the Humanities and Social Sciences). Springer.
- Dixon, Robert M. W. 1988. *A grammar of Boumaa Fijian*. Chicago: University of Chicago Press.
- Dixon, Robert M. W. 2006. Complement clauses and complementation strategies in typological perspective. In Robert M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Complementation: A cross-linguistic typology*, 1–48.
- Donaldson, Tamsin. 1985. From speaking Ngiyampaa to speaking English. *Aboriginal History* 9(1/2). 126–147.
- Donohue, Mark & Simon Musgrave. 2007. Typology and the linguistic macrohistory of Island Melanesia. *Oceanic Linguistics* 46(2). 348–87.
- Dorian, Nancy C. 1994. Purism vs. compromise in language revitalization and language revival. *Language in Society* 23(4). 479–494.
- Dryer, Matthew S. 2007. Clause types. In Timothy Shopen (ed.), *Language typology and syntactic description: Clause Structure*, vol. 1, 224–275. 2nd ed. Cambridge: Cambridge University Press.
- Dunn, Michael, Robert A. Foley, Stephen C. Levinson, Ger Reesink & Angela Terrill. 2007. Statistical Reasoning in the Evaluation of Typological Diversity in Island Melanesia. *Oceanic Linguistics* 46(2). 388–403.
- Dunn, Michael, Stephen C. Levinson, Eva Lindström, Ger Reesink & Angela Terrill. 2008. Structural Phylogeny in Historical Linguistics: Methodological Explorations Applied in Island Melanesia. *Language* 84(4). 710–759.
- Dunn, Michael, Angela Terrill, Ger Reesink, Robert A. Foley & Stephen C. Levinson. 2005. Structural Phylogenetics and the Reconstruction of Ancient Language History. *Science* 309(5743). 2072–2075. doi:10.1126/science.1114615.
- Durie, Mark. 1988. Verb Serialization and “Verbal-Prepositions” in Oceanic Languages. *Oceanic Linguistics* 27(1/2). 1–23. doi:10.2307/3623147.
- Durie, Mark. 1997. Grammatical structures in verb serialization. In Alex Alsina, Joan Bresnan & Peter Sells (eds.), *Complex predicates*, 289–354.
- Early, Robert J. 1993. Nuclear Layer Serialization in Lewo. *Oceanic Linguistics* 32(1). 65–93. doi:10.2307/3623097.

- Early, Robert J. 1994. *A grammar of Lewo, Vanuatu*. Australian National University, Canberra PhD Thesis.
- Early, Robert J. 2002. Lamén. In John Lynch, Malcolm D. Ross & Terry Crowley (eds.), *The Oceanic Languages* (Routledge Language Family Series), 671–80. Curzon Press.
- Early, Robert J. 2004. Periphrasis as a verbal borrowing strategy in Epi languages. In Jan Tent & Paul Geraghty (eds.), *Borrowing: A Pacific perspective* (Pacific Linguistics), 55–63. Canberra: Australian National University.
- Eckert, Penelope. 1989. *Jocks and burnouts: Social categories and identity in the high school*. Teachers College Press.
- EMELD Project. 2000. Electronic Metastructures for Endangered Language Documentation. <http://emeld.org/>.
- Enfield, Nick J. 2003. *Linguistic epidemiology: Semantics and grammar of language contact in mainland Southeast Asia*. Routledge.
- Enfield, Nick J. 2009. Serial verb constructions: A cross-linguistic typology (review). *Language* 85(2). 445–451.
- Ernst, Thomas. 2000. Semantic features and the distribution of adverbs. *Approaching the Grammar of Adjuncts* (ZAS Papers in Linguistics), vol. 17, 79–98. Oslo.
- Ernst, Thomas. 2002. *Syntax of Adjuncts* (Cambridge Studies in Linguistics 96). Cambridge University Press.
- Erteschik-Shir, Nomi. 2007. *Information Structure: The Syntax-Discourse Interface*. Oxford: Oxford University Press. (22 September, 2018).
- Ethnologue. 2018. Ambrym, Southeast. *Ethnologue*. <http://www.ethnologue.com/language/tvk>.
- Evans, Bethwyn. 2003. *A study of valency-changing devices in Proto Oceanic*. Pacific Linguistics.
- Evans, Bethwyn & Malcolm D. Ross. 2001. The History of Proto-Oceanic *ma-. *Oceanic Linguistics* 40(2). 269–290. doi:10.1353/ol.2001.0018.
- Evans, Nicholas. 2016. Typology and coevolutionary linguistics. *Linguistic Typology* 20(3). 505–520. doi:10.1515/lingty-2016-0023.
- Fernández, Eva M., Ricardo Augusto De Souza & Agustina Carando. 2017. Bilingual innovations: Experimental evidence offers clues regarding the psycholinguistics of language change. *Bilingualism: Language and Cognition* 20(2). 251–268. doi:10.1017/S1366728916000924.
- Fishman, Joshua A. 1991. *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages* (Multilingual Matters 76). Clevedon.
- FLEX Development Team. 2018. *FieldWorks Language Explorer (FLEX)*. SIL International. <https://software.sil.org/fieldworks/>.
- Florey, Margaret. 2004. Countering purism: confronting the emergence of new varieties in a training program for community language workers. *Language Documentation and Description* 2. 9–27.
- Foley, William A. 2003. Genre, register and language documentation in literate and preliterate communities. *Language Documentation and Description* 1. 85–98.

- Foley, William A. 2005. Semantic parameters and the unaccusative split in the Austronesian language family. *Studies in Language* 29(2). 385–430. doi:10.1075/sl.29.2.05fol.
- Foley, William A. 2007. A typology of information packaging in the clause. In Timothy Shopen (ed.), *Language typology and syntactic description: Clause structure*, vol. 1, 362–446. 2nd ed. Cambridge: Cambridge University Press.
- Foley, William A. & Mike Olson. 1985. Clausehood and verb serialization. *Grammar inside and outside the clause* 17–60.
- Foster, Pauline, Alan Tonkyn & Gillian Wigglesworth. 2000. Measuring spoken language: A unit for all reasons. *Applied linguistics* 21(3). 354–375.
- François, Alexandre. 2001. *Contraintes de structures et liberté dans l'organisation du discours. Une description du mwotlap, langue océanienne du Vanuatu*. Université Paris-Sorbonne-Paris IV PhD Thesis.
- François, Alexandre. 2004a. La réduplication en mwotlap: les paradoxes du fractionnement. *Les langues austronésiennes. Faits de langues* (23–24). 177–194.
- François, Alexandre. 2004b. Chains of freedom: Constraints and creativity in the macro-verb strategies of Mwotlap. In Isabelle Bril & Françoise Ozanne-Rivierre (eds.), *Complex predicates in Oceanic languages: studies in the dynamics of binding and boundness*, 107–143. Walter de Gruyter.
- François, Alexandre. 2006. Serial verb constructions in Mwotlap. In Alexandra Y. Aikhenvald & Robert M. W. Dixon (eds.), *Serial verb constructions: A cross-linguistic typology*, 223–238. Oxford: Oxford University Press.
- François, Alexandre. 2009. The languages of Vanikoro: three lexicons and one grammar. In Bethwyn Evans (ed.), *Discovering history through language: Papers in honour of Malcolm Ross* (Pacific Linguistics 605), 103–126. Canberra: Australian National University.
- François, Alexandre. 2011. Social ecology and language history in the northern Vanuatu linkage: A tale of divergence and convergence. *Journal of Historical Linguistics* 1(2). 175–246.
- François, Alexandre. 2012. The dynamics of linguistic diversity: egalitarian multilingualism and power imbalance among northern Vanuatu languages. *International Journal of the Sociology of Language* 2012(214). 85–110. doi:10.1515/ijsl-2012-0022.
- François, Alexandre. 2014. Trees, waves and linkages: models of language diversification. *The Routledge handbook of historical linguistics*, 179–207. Routledge.
- François, Alexandre. 2017. The economy of word classes in Hiw, Vanuatu. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"* 41(2). 294–357.
- François, Alexandre, Michael Franjeh, Sébastien Lacrampe & Stefan Schnell. 2015. The exceptional linguistic density of Vanuatu. *The Languages of Vanuatu: Unity and Diversity* (Studies in the Languages of Island Melanesia (SLIM) 5). Asia-Pacific Linguistics Open Access.
- Franjeh, Michael J. 2012. *Possessive classifiers in North Ambrym, a language of Vanuatu: explorations in semantic classification*. London: SOAS, University of London PhD Thesis.
- Franjeh, Michael J. 2014. *Rral: North Ambrym Dictionary*.

- Franjeh, Michael J. 2016. Vowel harmony in Fanbyak: An Oceanic language of Vanuatu. Oral Presentation presented at the Australian Linguistics Society annual conference, Monash University, Melbourne.
- Frater, M. 1922. *Midst volcanic fires: an account of missionary tours among the volcanic islands of the New Hebrides*. London: James Clarke & Co. Ltd.
- Garrett, Peter. 2000. *Language attitudes: methods and interpretation in sociolinguistic research*. Cardiff University PhD Thesis.
- Garrett, Peter, Nikolas Coupland & Angie Williams. 2003. *Investigating language attitudes: Social meanings of dialect, ethnicity and performance*. University of Wales Press.
- Gippert, Jost, Nikolaus P. Himmelmann & Ulrike Mosel. 2006. *Essentials of language documentation* (Trends in Linguistics: Studies and Monographs 178). Walter de Gruyter.
- Givón, Talmy. 1983. *Topic Continuity in Discourse: A quantitative cross-language study*. John Benjamins Publishing.
- Givón, Talmy. 1991. Some substantive issues concerning verb serialization: Grammatical vs. cognitive packaging. In Claire Lefebvre (ed.), *Serial verbs: grammatical, comparative and cognitive approaches*, 137–184.
- Givón, Talmy. 1994. Irrealis and the subjunctive. *Studies in Language* 18(2). 265–337.
- Gooskens, Charlotte & Cindy Schneider. 2016. Testing mutual intelligibility between closely related languages in an oral society. *Language Documentation* 10. 278–305.
- Goyal, Kovid. 2018. *Calibre: E-book management*. <https://calibre-ebook.com/> (2 April, 2018).
- Greenhill, Simon J., Chieh-Hsi Wu, Xia Hua, Michael Dunn, Stephen C. Levinson & Russell D. Gray. 2017. Evolutionary dynamics of language systems. *Proceedings of the National Academy of Sciences* 114(42). E8822–E8829. doi:10.1073/pnas.1700388114.
- Gries, Stefan Th. 2003a. Grammatical variation in English: A question of “structure vs. function”? *Determinants of grammatical variation in English*, 155–73.
- Gries, Stefan Th. 2003b. Towards a corpus-based identification of prototypical instances of constructions. *Annual Review of Cognitive Linguistics* 1. 1–27. doi:10.1075/arcl.1.02gri.
- Gries, Stefan Th. 2005. Syntactic Priming: A Corpus-based Approach. *Journal of Psycholinguistic Research* 34(4). 365–399. doi:10.1007/s10936-005-6139-3.
- Gries, Stefan Th. 2010. Useful statistics for corpus linguistics. *A mosaic of corpus linguistics: Selected approaches* 66. 269–291.
- Guérin, Maximilien. 2016. *Les constructions verbales en wolof: vers une typologie de la prédication, de l’auxiliation et des périphrases*. Université Paris Nouvelle - Paris 3 PhD Thesis.
- Haig, Geoffrey L. J. & Stefan Schnell. 2014. Annotations using GRAID: Manual version 7.0.
- Haig, Geoffrey L. J., Stefan Schnell & Claudia Wegener. 2011. Comparing corpora from endangered language projects: Explorations in language typology based on original texts. In Geoffrey L. J. Haig, Nicole Nau, Stefan Schnell & Claudia Wegener (eds.), *Documenting endangered languages:*

- Achievements and perspectives* (Trends in Linguistics: Studies and Monographs 240), 55–86. Berlin/Boston: De Gruyter Mouton.
- HandBrake Team. 2017. *HandBrake: Open Source Video Transcoder*. <https://handbrake.fr/>.
- Harrison, Sheldon P. 2003. On the limits of the comparative method. *The handbook of historical linguistics*, 213–243. Routledge.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687.
- Haspelmath, Martin. 2013. Argument indexing: A conceptual framework for the syntactic status of bound person forms. In Dik Bakker & Martin Haspelmath (eds.), *Languages across boundaries: Studies in memory of Anna Siewierska*, 197–226.
- Haspelmath, Martin. 2016. The Serial Verb Construction: Comparative Concept and Cross-linguistic Generalizations. *Language and Linguistics* 17(3). 291–319. doi:10.1177/2397002215626895.
- Haugen, E. I. 1972. *The ecology of language* (Language Science and National Development). Stanford, California: Stanford University Press.
- Hayes, Bruce. 1995. *Metrical stress theory: Principles and case studies*. University of Chicago Press.
- Heine, Bernd. 1993. *Auxiliaries: Cognitive Forces and Grammaticalization*. Oxford University Press.
- Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge University Press.
- Hilbe, Joseph M. 2014. *Modeling Count Data* (Cambridge Core: Statistical Theory and Methods). doi:10.1017/CBO9781139236065.
- Hilbe, Joseph M. 2016. *COUNT: Functions, data and code for count data*. <https://CRAN.R-project.org/package=COUNT>.
- Hilbe, Joseph M. & Andrew Robinson. 2018. *msme: Functions and datasets for "Methods of statistical model estimation."* <https://CRAN.R-project.org/package=msme>.
- Hill, Deborah. 2011. Transitivity in Longgu: The interdependence of verb classes and valency-changing derivations. *Oceanic Linguistics* 50(2). 458–482.
- Hill, Jane H. & Kenneth C. Hill. 1980. Mixed grammar, purist grammar, and language attitudes in modern Nahuatl. *Language in Society* 9(03). 321–348. doi:10.1017/S0047404500008241.
- Himmelmann, Nikolaus P. 2004. Lexicalization and grammaticization: Opposite or orthogonal. In Nikolaus P. Himmelmann, Walter Bisang & Björn Wiemer (eds.), *What makes grammaticalization?: A Look from Its Fringes and Its Components*, 21–42.
- Himmelmann, Nikolaus P. 2006. Language documentation: What is it and what is it good for. In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation* (Trends in Linguistics: Studies and Monographs 178), 1–30.
- Himmelmann, Nikolaus P. 2013. Some observations on clausehood and event segmentation in Austronesian and Papuan languages. Oral Presentation presented at the Sixth Austronesian and Papuan languages and linguistics (APLL) conference, SOAS, University of London.
- Holm, John A. 1988. *Pidgins and creoles*. (Cambridge Language Surveys). Vol. 1. Cambridge University Press.

- Hopperdietzel, Jens. 2018. The morphosyntactic status of the subject (marker) in Daakaka (West Ambrym). Oral Presentation presented at the Austronesian and Papuan Languages and Linguistics 10, University of Surrey.
- Huang, Xuanfan & Michael Tanangkingsing. 2005. Reference to motion events in six western Austronesian languages: Toward a semantic typology. *Oceanic Linguistics* 44(2). 307–340.
- Hudson, Alan. 2002. Outline of a theory of diglossia. *International Journal of the Sociology of Language* 1–48.
- Hughes, Baden, Steven Bird & Catherine Bow. 2003. Encoding and presenting interlinear text using XML technologies. *Proceedings of the Australasian Language Technology Workshop 2003*, 61–69. Melbourne, Australia. <http://www.aclweb.org/anthology/U/U03/U03-1008>.
- Hundt, Marianne & Nicholas Smith. 2009. The present perfect in British and American English: Has there been any change, recently. *ICAME journal* 33(1). 45–64.
- Hyslop, Catriona. 2001. *The Lolovoli Dialect of the North-East Ambae Language: Vanuatu* (Pacific Linguistics). Vol. 515. Australian National University.
- Inkelas, Sharon & Cheryl Zoll. 2005. *Reduplication: Doubling in morphology* (Cambridge Studies in Linguistics 106). Cambridge University Press.
- Jackendoff, Ray S. 1972. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT press.
- Jackson, Christopher H. 2011. Multi-state models for panel data: the msm package for R. *Journal of Statistical Software* 38(8). 1–29.
- Jansen, B., H. Koopman & Pieter Muysken. 1978. Serial verbs in the creole languages. *Amsterdam Creole Studies* 2. 125–159.
- Jauncey, Dorothy. 2011. *Tamambo, the language of west Malo, Vanuatu* (Pacific Linguistics 622). Australian National University.
- Jelinek, Eloise. 1984. Empty categories, case, and configurationality. *Natural Language & Linguistic Theory* 2(1). 39–76.
- Jenkins, Rebecca Sue. 2005. *Language contact and composite structures in New Ireland*. SIL International.
- Jeppson, Haley, Heike Hofmann & Di Cook. 2018. *ggmosaic: Mosaic plots in the "ggplot2" framework*. <https://CRAN.R-project.org/package=ggmosaic>.
- Jespersen, Otto. 1917. *Negation in English and other languages*. AF Høst.
- Johansen, Gry Krath. 2012. *Yu Blong Wea? An anthropological study of Vanuatu land struggle*. University of Copenhagen Master's thesis.
- Johnson, Daniel Ezra. 2009. Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and linguistics compass* 3(1). 359–383.
- Johnston, Raymond Leslie. 1980. *Grammar and basic vocabulary in Oceanic Austronesian languages: A standard elicitation schedule*. Summer Institute of Linguistics.
- Jung, Dagmar & Nikolaus P. Himmelmann. 2011. Retelling data: working on transcription. In Geoffrey L. J. Haig, Nicole Nau, Stefan Schnell & Claudia Wegener (eds.), *Documenting Languages: Achievements and Perspectives*, 201–20. Berlin: Mouton De Gruyter.
- Kahle, David & Hadley Wickham. 2013. ggmap: Spatial visualizaiton with ggplot2. *The R Journal* 5(1). 144–161.

- Kemper, Susan, J. Marquis & M. Thompson. 2001. Longitudinal change in language production: effects of aging and dementia on grammatical complexity and propositional content. *Psychology and Aging* 16(4). 600–614.
- Kiesling, Scott F. 2011. *Linguistic variation and change*. Edinburgh University Press.
- Kirkness, Alan. 1994. Aero-lexicography: Observations on the treatment of combinemes and neoclassical combinations in historical and scholarly european dictionaries. *Euralex 1994 Proceedings*, 530–535. http://www.euralex.org/elx_proceedings/Euralex1994/.
- Klamer, Marian. 2002. Typical Features of Austronesian Languages in Central/Eastern Indonesia. *Oceanic Linguistics* 41(2). 363–383. doi:10.2307/3623314.
- Klein, Wolfgang. 1992. The Present Perfect Puzzle. *Language* 68(3). 525–552. doi:10.2307/415793.
- Klein, Wolfgang. 1994. *Time in language* (Germanic Linguistics). London: Routledge.
- Koch, Harold. 1984. The category of ‘associated motion’ in Kaytej. *Language in central Australia* 1(1). 23–34.
- Koch, Harold. 1995. The creation of morphological zeroes. *Yearbook of Morphology 1994*, 31–71. Kluwer Academic Publishers.
- Krifka, Manfred. 1995. The semantics and pragmatics of polarity items. *Linguistic analysis* 25(3–4). 209–257.
- Krifka, Manfred. 2011. Notes on Daakie (Port Vato): sounds and modality. *Proceedings of the Eighteenth Meeting of the Austronesian Formal Linguistics Association (AFLA)*, 46–65. Harvard.
- Krifka, Manfred. 2016. Realis and non-realis modalities in Daakie (Ambrym, Vanuatu). *Semantics and Linguistic Theory*, vol. 26, 566–583.
- Krifka, Manfred. 2018. Social Meaning of Number: Honorific Duals and Affiliative Paucals. Oral Presentation presented at the Vanuatu Languages Conference, USP Emalus Campus, Port Vila; Vanuatu.
- Krifka, Manfred, Kilu von Prince, Ana Krajinovic & Stephan Druskat. 2018. MelaTAMP: Project description. <https://www.projekte.hu-berlin.de/en/melatamp> (21 November, 2018).
- Kroskrity, Paul V. 2010. Language ideologies: Evolving perspectives. In Jürgen Jaspers, Jef Verschueren & Jan-Ola Östman (eds.), *Handbook of Pragmatics Highlights*, vol. 7: Society and language use, 192–205. John Benjamins Publishing.
- Kuteva, Tania. 2001. *Auxiliation: An Enquiry Into the Nature of Grammaticalization*. Oxford University Press.
- Labov, William. 1963. The social motivation of a sound change. *Word* 19(3). 273–309.
- Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2(2). 205–254. doi:10.1017/S0954394500000338.
- Lambrecht, Knud. 1996. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents* (Cambridge Studies in Linguistics 71). Cambridge University Press.

- Laycock, Donald C. 1982. Melanesian linguistic diversity: a Melanesian choice? In Ronald James May & Hank Nelson (eds.), *Melanesia: beyond diversity*, 33–38.
- Lehmann, Christian. 1988. Towards a typology of clause linkage. In John Haiman & Sandra A. Thompson (eds.), *Clause combining in grammar and discourse* (Typological Studies in Language 18), 181–225. John Benjamins Publishing Company.
- Lewis, M. Paul & Gary F. Simons. 2010. Assessing endangerment: expanding Fishman's GIDS. *Revue roumaine de linguistique* 55(2). 103–120.
- Lewis, M. Paul & Gary F. Simons. 2016. *Sustaining language use: Perspectives on community-based language development*. Dallas, TX: SIL International.
- Lichtenberk, Frantisek. 2000. Inclusory Pronominals. *Oceanic Linguistics* 39(1). 1–32. doi:10.1353/ol.2000.0006.
- Lichtenberk, Frantisek. 2006. Serial verb constructions in Toqabaqita. In Alexandra Y. Aikhenvald & Robert M. W. Dixon (eds.), *Serial verb constructions: A cross-linguistic typology*, 254–272.
- Lichtenberk, Frantisek. 2010. Subject incorporation in Toqabaqita and other Oceanic languages. In John Bowden, Nikolaus P. Himmelmann & Malcolm D. Ross (eds.), *A journey through Austronesian and Papuan linguistic and cultural space: Papers in honour of Andrew Pawley*, 375–90. Canberra: Pacific Linguistics.
- Lindstrom, Lamont. 1990. *Knowledge and power in a South Pacific society*. (Smithsonian Series in Ethnographic Inquiry). Washington, D.C.: Smithsonian Institution Press.
- Lindstrom, Lamont. 2008. Melanesian kastom and its transformations. *Anthropological Forum* 18(2). 161–178.
- Loiseau, Sylvain. 2018. *interlineaR: Importing interlinearized corpora and dictionaries as produced by descriptive linguistics software*. R. Paris 13 University / UMR 7107 Lacito. <https://CRAN.R-project.org/package=interlineaR>.
- Lüpke, Friederike. 2005. Small is beautiful: contributions of field-based corpora to different linguistic disciplines, illustrated by Jalonke. *Language documentation and description* 3. 75–105.
- Lynch, John. 1975. Oral/Nasal Alternation and the Realis/Irrealis Distinction in Oceanic Languages. *Oceanic Linguistics* 14(2). 87–99.
- Lynch, John. 2001. Article Accretion and Article Creation in Southern Oceanic. *Oceanic Linguistics* 40(2). 224–246. doi:10.2307/3623439.
- Lynch, John. 2008. Liquid vocalization and loss in Central Vanuatu. *Oceanic Linguistics* 47(2). 294–315.
- Lynch, John, Malcolm D. Ross & Terry Crowley (eds.). 2002. *The Oceanic Languages* (Curzon Language Family Series). Vol. 1. Richmond, Surrey: Curzon Press.
- Maher, Julianne. 1991. A crosslinguistic study of language contact and language attrition. In Herbert W. Seliger & Robert Michael Vago (eds.), *First language attrition*, 67–84.
- Majid, Asifa, Melissa Bowerman, Miriam van Staden & James S. Boster. 2007. The semantic categories of cutting and breaking events: A crosslinguistic perspective. *Cognitive Linguistics* 18(2). 133–152. doi:10.1515/COG.2007.005.

- Malau, Catriona Hyslop. 2011. Sustaining Vurës: Making products of language documentation accessible to multiple audiences. In Geoffrey L. J. Haig, Nicole Nau, Stefan Schnell & Claudia Wegener (eds.), *Documenting Endangered Languages: Achievements and Perspectives*, 305–19. De Gruyter Mouton.
- Margetts, Anna. 2011. Filming with native speaker commentary. *Documenting endangered languages: achievements and perspectives* (Trends in Linguistics: Studies and Monographs 240), 321–338. De Gruyter Mouton.
- Mayer, Mercer. 1969. *Frog, where are you?* New York: Dial Press.
- Mean. 2017. *Avidemux: Multi-platform video editor*.
<http://avidemux.sourceforge.net/>.
- Meyerhoff, Miriam. 1999. Sorry in the Pacific: Defining communities, defining practices. *Language in Society* 28(2). 225–238.
- Meyerhoff, Miriam. 2000. *Constraints on null subjects in Bislama (Vanuatu): social and linguistic factors* (Pacific Linguistics 506). Australian National University.
- Meyerhoff, Miriam. 2002. Formal and cultural constraints on optional objects in Bislama. *Language Variation and Change* 14(3). 323–346.
- Meyerhoff, Miriam. 2003. Claiming a Place: Gender, Knowledge, and Authority as Emergent Properties. In Janet Holmes & Miriam Meyerhoff (eds.), *The handbook of language and gender* (Blackwell Handbooks in Linguistics), vol. 25, 302–326. Oxford: Blackwell.
- Meyerhoff, Miriam. 2009. Animacy in Bislama? Using quantitative methods to evaluate transfer of a substrate feature. In James N. Stanford & Dennis R. Preston (eds.), *Variation in indigenous minority languages* (Studies in Language and Society 25), 369–97. John Benjamins Publishing Company.
- Meyerhoff, Miriam. 2014a. Variation and gender. In Susan Ehrlich, Miriam Meyerhoff & Janet Holmes (eds.), *Handbook of Language, Gender and Sexuality* (Blackwell Handbooks in Linguistics), 87–102. John Wiley & Sons, Incorporated.
- Meyerhoff, Miriam. 2014b. Borrowing in Apparent Time: With some comments on attitudes and universals. *U. Penn Working Papers in Linguistics*, vol. 20.2, 121–8.
- Meyerhoff, Miriam. 2015. Turning variation on its head: Analysing subject prefixes in Nkep (Vanuatu) for language documentation. *Asia-Pacific Language Variation* 1(1). 78–108. doi:10.1075/aplv.1.1.04mey.
- Meyerhoff, Miriam. 2017. Writing a linguistic symphony: Analyzing variation while doing language documentation. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 62(4). 525–549. doi:10.1017/cnj.2017.28.
- Meyerhoff, Miriam & James A. Walker. 2007. The persistence of variation in individual grammars: Copula absence in ‘urban sojourners’ and their stay-at-home peers, Bequia (St Vincent and the Grenadines) 1. *Journal of Sociolinguistics* 11(3). 346–366.
- Milroy, James & Lesley Milroy. 1985. Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21(02). 339–384. doi:10.1017/S0022226700010306.

- Mithun, Marianne. 1992. Is basic word order universal? In Doris L. Payne (ed.), *Pragmatics of word order flexibility* (Typological Studies in Language 22), 15–61. John Benjamins Publishing.
- Moeller, Sarah Ruth. 2014. SayMore, a tool for language documentation productivity. *Language Documentation and Conservation* 8. 66–74.
- Moreno Sandoval, Antonio & José María Guirao. 2006. Morphosyntactic Tagging of the Spanish C-ORAL-ROM Corpus: Methodology, Tools and Evaluation. In Yuji Kawaguchi, Susumu Zaima & Toshihiro Takagaki (eds.), *Spoken language corpus and linguistic informatics*, 199–218. John Benjamins Publishing Company.
- Moyse-Faurie, Claire. 2008. Constructions expressing middle, reflexive and reciprocal situations in some Oceanic languages. *Reciprocals and reflexives: Theoretical and typological explorations*, 105–168.
- MPI Nijmegen. 2018. *ELAN*. Nijmegen, The Netherlands: Max Planck Institute for Psycholinguistics, The Language Archive.
- Mufwene, Salikoko S. 2002. Colonization, globalization, and the future of languages in the twenty-first century. *International Journal on Multicultural Societies* 4(2). 162–193.
- Mühlhäusler, Peter. 1996. *Linguistic ecology: Language change and linguistic imperialism in the Pacific region* (The Politics of Language). London: Routledge.
- Muysken, Pieter. 2013. Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition* 16(4). 709–730. doi:10.1017/S1366728912000727.
- Næss, Åshild. 2012. Cutting and breaking in Äiwoo: Event integration and the complexity of lexical expressions. *Cognitive Linguistics* 23(2). 395–420. doi:10.1515/cog-2012-0012.
- Neverland, Alan. 2018. *AlReader*. Android.
<https://play.google.com/store/apps/details?id=com.neverland.alreader>.
- Nichols, Johanna. 1986. Head-Marking and Dependent-Marking Grammar. *Language* 62(1). 56–119. doi:10.2307/415601.
- Nichols, Johanna. 1995. Diachronically stable structural features. In Henning Andersen (ed.), *Historical linguistics, 1993*, 337–337. Los Angeles: John Benjamins Publishing.
- Nichols, Johanna. 1996. The comparative method as heuristic. In Mark Durie & Malcolm D. Ross (eds.), *The comparative method reviewed: Regularity and irregularity in language change*, 39–71. Oxford University Press.
- Nikolaeva, Irina (ed.). 2007. *Finiteness: Theoretical and Empirical Foundations*. Oxford, UK: Oxford University Press.
- Oakes, Michael P. 1998. *Statistics for Corpus Linguistics* (Edinburgh Textbooks in Empirical Linguistics). Edinburgh: Edinburgh University Press.
- Oakes, Michael P. 2009. Corpus Linguistics and Language Variation. In Paul Baker (ed.), *Contemporary corpus linguistics*, 159–183. A&C Black.
- Osmond, Meredith. 2016. Body part metaphors. In Malcolm Ross, Andrew Pawley & Meredith Osmond (eds.), *The lexicon of Proto Oceanic. The culture and environment of ancestral Oceanic society* (Asia-Pacific Linguistics 28), vol. 5 People: body and mind, 519–534.
- Palmer, Frank Robert. 1986. *Mood and Modality*. Cambridge: Cambridge University Press.

- Parker, Gary J. 1968a. Southeast Ambrym phonology. *Oceanic Linguistics* 81–91.
- Parker, Gary J. 1968b. Southeast Ambrym verb inflection and morphophonemics. In A. Capell, Gary J. Parker & A. J. Schütz (eds.), *Papers in linguistics of Melanesia* (Pacific Linguistics Series A 15).
- Parker, Gary J. 1970a. *Southeast Ambrym Dictionary* (Pacific Linguistics Series C 17). Australian National University.
- Parker, Gary J. 1970b. Morphophonemics of inalienable nouns in Southeast Ambrym. *Oceanic Linguistics* 1–10.
- Paton, W. F. 1971. *Ambrym (Lonwolwol) Grammar* (Pacific Linguistics Series B 19). Australian National University.
- Paton, W. F. 1979. *Customs of Ambrym: texts, songs, games, drawings*. (Pacific Linguistics Series D 22). Australian National University.
- Pauwels, Anne. 2016. Language Maintenance and Shift by Anne Pauwels. *Cambridge Core*. doi:10.1017/CBO9781107338869.
- Paviour-Smith, Martin. 2015. We've never seen a cyclone like this: Exploring self-concept and narrator characterisation in Aulua. In Farzana Gounder (ed.), *Narrative and Identity Construction in the Pacific Islands*, 37–58. John Benjamins Publishing Company.
- Pawley, Andrew. 1973. Some problems in Proto-Oceanic grammar. *Oceanic Linguistics* 12(1/2). 103–188.
- Pawley, Andrew. 2003. Grammatical categories and grammaticisation in the Oceanic verb complex. (Ed.) A. Riehl & T. Savella. *Cornell Working Papers in Linguistics* 19. 149–172.
- Pawley, Andrew. 2008. Compact versus narrative serial verb constructions in Kalam. *Serial verb constructions in Austronesian and Papuan languages*. Pacific Linguistics.
- Pickering, Martin J. & Holly P. Branigan. 1999. Syntactic priming in language production. *Trends in Cognitive Sciences* 3(4). 136–141. doi:10.1016/S1364-6613(99)01293-0.
- Pickering, Martin J. & Victor S. Ferreira. 2008. Structural Priming: A Critical Review. *Psychological bulletin* 134(3). 427–459. doi:10.1037/0033-2909.134.3.427.
- Plank, Frans. 1991. Of abundance and scantiness in inflection: a typological prelude. In Frans Plank (ed.), *Paradigms: the economy of inflection* (Empirical Approaches to Language Typology 9), 1–39. De Gruyter.
- Prince, Kilu von. 2012. *A grammar of Daakaka*. Berlin: Humboldt University PhD Thesis.
- Prince, Kilu von. 2015. *A grammar of Daakaka* (Mouton Grammar Library). Berlin: De Gruyter Mouton.
- Prince, Kilu von. 2017a. Dozing eyes and drunken faces: nominalized psychocollocations in Daakaka (Vanuatu). *Studies in Language* 41(2). 502–520.
- Prince, Kilu von. 2017b. Paradigm-induced implicatures of TAM markers: The case of the Daakaka distal. *Proceedings of Sinn und Bedeutung* 21, 16. Edinburgh.
- R Core Team. 2018. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org>.
- Rayson, Paul, D. Berridge & B. Francis. 2004. Extending the Cochran rule for the comparison of word frequencies between corpora. In G. Purnelle, C.

- Fairon & A. Dister (eds.), *Le poids des mots*, 926–936. Presses universitaires de Louvain.
- Reichenbach, Hans. 1947. Symbolic logic. *Berkeley: University of California*.
- Ridge, Eleanor. 2013. *Cutting and breaking events in Abma: a language of Central Pentecost Island, Vanuatu*. Oxford University MSt Thesis.
- Ridge, Eleanor. 2016. VV6: A story about two birds with a song. *YouTube*.
<https://youtu.be/pSh0F46QAfs>.
- Ridge, Eleanor. 2017a. Vatlongos Videos. *YouTube*.
https://www.youtube.com/playlist?list=PLbz9bC52hkAinBk09yGznLzF_Q-tmQRjn.
- Ridge, Eleanor. 2017b. Use of loan verbs in Vatlongos (Southeast Ambrym), Vanuatu. Oral Presentation presented at the Cambridge Conference on Language Endangerment 7 (CCLE7), Cambridge University.
- Ridge, Eleanor. 2017c. Grammatical and phonological words in the Vanuatu verb complex. Oral Presentation presented at the What is a Word Workshop, University of Zurich.
- Ridge, Eleanor. 2018a. Vatlongos. *Pangloss Collection*.
http://lacito.vjf.cnrs.fr/pangloss/corpus/list_rsc_en.php?lg=Vatlongos&name=Vatlongos.
- Ridge, Eleanor. 2018b. Language contexts: Vatlongos, Southeast Ambrym (Vanuatu). *Language Documentation and Description* 15. 87–122.
- Ridge, Eleanor. 2018c. Repetition of verbs and their objects as a constituency test in Vatlongos (Vanuatu). Oral Presentation presented at the Syntax of the World's Languages 8, INALCO, Paris.
- Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, Mass.: Massachusetts Institute of Technology.
- Ross, Malcolm D. 1996. Contact-induced change and the comparative method: cases from Papua New Guinea. In Mark Durie & Malcolm D. Ross (eds.), *The comparative method reviewed: regularity and irregularity in language change*, 180–217. Oxford University Press.
- Ross, Malcolm D. 2001. Contact-induced change in Oceanic languages in North-West Melanesia. In Alexandra Y. Aikhenvald & Robert M. W. Dixon (eds.), *Areal diffusion and genetic inheritance: problems in comparative linguistics*, 134–66.
- Ross, Malcolm D. 2003. Diagnosing prehistoric language contact. In Raymond Hickey (ed.), *Motives for language change*, 174–98. Cambridge University Press.
- Ross, Malcolm D. 2004. The morphosyntactic typology of Oceanic languages. *Language and Linguistics* 5(2). 491–541.
- Ross, Malcolm D. 2007. Calquing and metatypy. *Journal of Language Contact* 1(1). 116–143.
- Ross, Malcolm D., Andrew Pawley & Meredith Osmond (eds.). 2008. *The lexicon of Proto Oceanic. The culture and environment of ancestral Oceanic society*. Vol. 3 Plants. 5 vols. Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University.
- Round, Erich. 2015. Rhizomorphemes, meromorphemes, and metamorphemes. In Matthew Baerman, Dunstan Brown & Greville G. Corbett (eds.), *Understanding and measuring morphological complexity*. Oxford University Press.

- RStudio Team. 2018. *RStudio: Integrated development for R*. Desktop. Boston, MA.: RStudio Inc. www.rstudio.com.
- Sankoff, David & William Labov. 1979. On the uses of variable rules. *Language in Society* 8(2–3). 189. doi:10.1017/S0047404500007430.
- Sankoff, David, Sali A. Tagliamonte & Eric Smith. 2005. *Goldvarb X: A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto.
- Sankoff, G. 2006. Age: Apparent Time and Real Time. In Keith Brown (ed.), *Encyclopedia of Language & Linguistics*, 110–116. Second edition. Oxford: Elsevier. doi:10.1016/B0-08-044854-2/01479-6.
- Sarkar, Deepayan. 2008. *Lattice: Multivariate data visualization with R*. New York: Springer.
- Schmid, Monika S. 2011. *Language Attrition*. Cambridge University Press.
- Schmidt, Annette. 1985. *Young people's Dyirbal: an example of language death from Australia* (Cambridge Studies in Linguistics, Supplementary Volumes.). Cambridge University Press,.
- Schmidt, Annette. 1991. Language attrition in Boumaa Fijian and Dyirbal. In Herbert W. Seliger & Robert Michael Vago (eds.), *First language attrition*, 113–124.
- Schneider, Cindy. 2010. *A Grammar of Abma: A language of Pentecost Island Vanuatu* (Pacific Linguistics 608). Australian National University.
- Schneider, Cindy. 2017. Ideologies, practices, and their effect on dialect vitality: a case study from Vanuatu. *Journal of Multilingual and Multicultural Development* 1–16. doi:10.1080/01434632.2017.1311904.
- Schneider, Cindy & Charlotte Gooskens. 2017. Approaching micro-level planning from an intelligibility perspective: a case study from Pentecost Island, Vanuatu. *Journal of Multilingual and Multicultural Development* 38(6). 530–544. doi:10.1080/01434632.2016.1212865.
- Schütz, Albert J. 1969. *Nguna Grammar* (Oceanic Linguistics Special Publications 5).
- Seifart, Frank. 2011. Competing motivations for documenting endangered languages. In Geoffrey L. J. Haig, Nicole Nau, Stefan Schnell & Claudia Wegener (eds.), *Documenting endangered languages: Achievements and perspectives*, 18–32.
- Seiss, Melanie. 2009. On the difference between auxiliaries, serial verbs and light verbs. *Proceedings of the LFG09 Conference*, 501–519.
- Seliger, Herbert W. & Robert Michael Vago (eds.). 1991. *First language attrition*. Cambridge University Press.
- Seuren, Pieter. 1991. The definition of serial verbs. In Francis Byrne & Thom Huebner (eds.), *Development and structures of creole languages: essays in honor of Derek Bickerton*, 193–205. John Benjamins Publishing.
- SIL International. 2018. *Dictionary App Builder*. SIL International. <https://software.sil.org/dictionaryappbuilder/>.
- Silva-Corvalán, Carmen. 1991. Spanish language attrition in a contact situation with English. In Herbert W. Seliger & Robert Michael Vago (eds.), *First language attrition*, 151–71.
- Silva-Corvalán, Carmen. 1994. *Language Contact and Change: Spanish in Los Angeles*. Oxford University Press.

- Silva-Corvalán, Carmen. 2008. The limits of convergence in language contact. *Journal of Language Contact* 2(1). 213–224.
- Simon, Kathie. 2013. Breakthrough for Mele Maat. *Vanuatu Daily Post*, sec. Local News. <http://www.dailypost.vu/content/breakthrough-mele-maat> (25 March, 2014).
- Simons, Gary F. & M. Paul Lewis. 2013. The world's languages in crisis: a twenty year update. In Elena Mihás, Bernard Perley, Gabriel Rei-Doval & Kathleen Wheatley (eds.), *Responses to language endangerment: In honor of Mickey Noonan* (New Directions in Language Documentation and Language Revitalization 3), 3–19. John Benjamins Publishing.
- Slowikowski, Kamil. 2018. *ggrepel: Automatically position non-overlapping text labels with ggplot2*.
- Souag, Lameen. 2010. *Grammatical contact in the Sahara: Arabic, Berber, and Songhay in Tabelbala and Siwa*. SOAS, University of London PhD Thesis.
- Spencer, Andrew & Ana Luís. 2012. The canonical clitic. In Dunstan Brown, Marina Chumakina & Greville G. Corbett (eds.), *Canonical morphology and syntax*, 123–150. Oxford University Press.
- Stamen Design. 2018. *Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL*. Stamen Design.
- Stump, Gregory T. 2001. *Inflectional morphology: A theory of paradigm structure*. Cambridge University Press.
- Tagliamonte, Sali A. & Rachel Hudson. 1999. Be like et al. beyond America: The quotative system in British and Canadian youth. *Journal of Sociolinguistics* 3(2). 147–172. doi:10.1111/1467-9481.00070.
- Thieberger, Nicholas. 2004. *Topics in the grammar and documentation of South Efate, an Oceanic language of Central Vanuatu*. PhD Thesis.
- Thomas, A. 1990. Language planning in Vanuatu. In Richard B. Baldauf & Allan Luke (eds.), *Language planning and education in Australasia and the South Pacific* (Multilingual Matters), 234–258.
- Thomason, Sarah Grey & Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Thompson, Sandra A., Robert E. Longacre & Shin Ja J. Hwang. 2007. Adverbial clauses. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 2, 237–300. 2nd. ed.
- Thornton, Anna M. 2012a. Overabundance in Italian verb morphology and its interactions with other non-canonical phenomena. *Irregularity in Morphology (and Beyond)*, 251–269. Berlin: Akademie Verlag.
- Thornton, Anna M. 2012b. Reduction and maintenance of overabundance. A case study on Italian verb paradigms. *Word Structure* 5(2). 183–207.
- Thurston, William R. 1992. Sociolinguistic typology and other factors effecting change in northwestern New Britain, Papua New Guinea. In Thomas Edward Dutton (ed.), *Culture change, language change: Case studies from Melanesia* (Pacific Linguistics Series C 120), 123–39. Australian National University.
- Timberlake, Alan. 2007. Aspect, tense, mood. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. 3, 280–333. 2nd ed. Cambridge University Press.

- Toivonen, Ida. 2001. *Non-projecting words: Evidence from verbal particles in Swedish*. Stanford PhD Thesis.
- Tonkinson, Robert. 1968. *Maat village, Efate: a relocated community in the New Hebrides*. Eugene, OR: University of Oregon.
- Tonkinson, Robert. 1979. The paradox of permanency in a resettled New Hebridean Community. *Mass Emergencies* 4. 105–116.
- Tonkinson, Robert. 1981. Sorcery and social change in Southeast Ambrym, Vanuatu. *Social Analysis* 8. 105–116.
- Tonkinson, Robert. 1982. National identity and the problem of Kastom in Vanuatu. *Mankind* 13. 306–315.
- Tonkinson, Robert. 1985. Forever Ambrymese? Identity in a Relocated Community. *Pacific Viewpoint* 26(1). 139–159.
- Tonkinson, Robert. 2011. The research context in New Hebrides-Vanuatu. In J. Taylor & Nicholas Thieberger (eds.), *Working together in Vanuatu: research histories, collaborations, projects and reflections*, 12–26. Australian National University.
- Toulmin, Matthew. 2009. *From linguistic to sociolinguistic reconstruction: the Kamta historical subgroup of Indo-Aryan* (Pacific Linguistics). Research School of Pacific and Asian Studies, the Australian National University.
- Tourism Vanuatu. 2017. Mele Cascades. *Tourism Vanuatu*.
<https://www.tourismvanuatu.com/tours/adventures-paradise-t1/cascade-waterfalls.html>.
- UNDP. 2016. *Human Development Report: Vanuatu*. United Nations Development Programme.
<http://hdr.undp.org/en/countries/profiles/VUT> (22 November, 2017).
- Urgelles-Coll, Miriam. 2010. *The syntax and semantics of discourse markers*. A&C Black.
- Van Valin, Robert D. 1990. Layered syntax in role and reference grammar. In Jan Nuyts, A. Machtelt Bolkestein & Co Vet (eds.), *Layers and levels of representation in language theory* (Pragmatics and Beyond New Series 13), 193–231. Amsterdam, Philadelphia: John Benjamins.
- Van Valin, Robert D. 2010. Role and reference grammar as a framework for linguistic analysis. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*.
- Van Valin, Robert D. & Randy J. LaPolla. 1997. *Syntax: Structure, meaning, and function* (Cambridge Textbooks in Linguistics). Cambridge University Press.
- Vanuatu Ministry of Education. 2012. Vanuatu National Language Policy.
<https://moet.gov.vu/docs/policies>.
- Venables, W. N. & B. D. Ripley. 2002. *Modern applied statistics with S*. Fourth. New York: Springer. <http://www.stats.ox.ac.uk/pub/MASS4>.
- VNSO. 2009. National Population and Census Data.
<https://vnso.gov.vu/index.php/census-and-surveys/censuses#census-archived-reports>.
- VNSO. 2013. Hybrid Survey Factsheet. *Vanuatu National Statistics Office*.
https://vnso.gov.vu/images/Special_Report/Expenditure_Patterns/2013HybridSurveyFactSheet.pdf (31 March, 2018).
- VNSO. 2017. Mini-Census 2016. <https://vnso.gov.vu/index.php/mini-census-2016>.

- VNSO. 2018. Statistics Pocket Books. *Vanuatu National Statistics Office*.
https://vnso.gov.vu/images/PublicDocuments/Statistical_Pocket_Books/Statistics_Pocket_Book_2014.pdf.
- Vossen, Frens & Johan van der Auwera. 2014. The Jespersen cycles seen from Austronesian. In Maj-Britt Mosegaard Hansen & Jacqueline Visconti (eds.), *The diachrony of negation*, 47–84. John Benjamins.
- Vries, Lourens de. 2005. Towards a typology of tail–head linkage in Papuan languages. *Studies in Language* 29(2). 363–384.
- Walsh, D. S. 1984. Is “English-Based” an Adequately Accurate Label for Bislama? *Fifty-Fourth ANZAAS Congress*. Canberra.
- Wertheim, Suzanne. 2003. *Linguistic purism, language shift, and contact-induced change in Tatar*. University of California, Berkeley PhD Thesis.
- Wickham, Hadley. 2007. Reshaping data with the reshape package. *Journal of Statistical Software* 21(12). 1–20.
- Wickham, Hadley. 2016. *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag. <http://ggplot2.org>.
- Wickham, Hadley. 2017. *tidyverse: Easily install and load the “Tidyverse.”* <https://CRAN.R-project.org/package=tidyverse>.
- Wickham, Hadley. 2018. *stringr: Simple, consistent wrappers for common string operations*. <https://CRAN.R-project.org/package=stringr>.
- Wickham, Hadley, Romain François, Lionel Henry & Kirill Müller. 2018. *dplyr: A grammar of data manipulation*. <https://CRAN.R-project.org/package=dplyr>.
- Wickham, Hadley, James Hester & Jeroen Ooms. 2018. *xml2: Parse XML*. <https://CRAN.R-project.org/package=xml2>.
- Willans, Fiona. 2011. Classroom code-switching in a Vanuatu secondary school: conflict between policy and practice. *International Journal of Bilingual Education and Bilingualism* 14(1). 23–38.
doi:10.1080/13670050903576038.
- Wohlgemuth, Jan. 2009. *A typology of verbal borrowings* (Trends in Linguistics: Studies and Monographs). Vol. 211. Walter de Gruyter.
- Woodbury, Anthony C. 2003. Defining documentary linguistics. *Language documentation and description* 1(1). 35–51.
- Wycliffe Bible Translators. 2015. *The New Testament in the Southeast Ambrym Language of Vanuatu: Vanuvei Eo e sepinien Vatlongos na mol-Vatimol xil*. Wycliffe.
- Zeileis, A. 2004. Econometric computing with HC and HAC covariance matrix estimators. *Journal of Statistical Software* 11(10). 1–17.
- Zwicky, Arnold M. & Geoffrey K. Pullum. 1983. Cliticization vs. Inflection: English N'T. *Language* 59(3). 502–513.

Consent form

These are the questions used in the detailed consent section of the metadata questionnaire (§1.3.2), in the original Bislama and translated into English.

1. Hemi oraet se storian blo yu i stap insaet long projek ia?
Is it okay for your story to be used in this project?
2. Hemi oraet blo rikodem storian blo yu?
Is it okay to record your story?
3. Hemi oraet blo video storian blo yu?
Is it okay to video your story?
4. Hemi oraet blo askem sam kwestun lo saed blo personal histori blo yu?
Is okay to ask some questions about your personal history?
5. Hemi oraet sipos nem blo yu stap insaet lo thesis blo Eleanor mekem se hemi savi talem tankyu tumas lo yu?
Is it okay if your name is included in Eleanor's thesis so that she can thank you?
6. Hemi oraet blo serem storian blo yu witem narafala risoja we oli gat interes lo lanwis blo Southeast Ambrym?
Is it okay to share your story with other researchers who are interested in the language of Southeast Ambrym?
7. Hemi oraet blo putum storian blo yu insaet long wan akaev blo mek so se bae hemi stap longtaem?
Is it okay to put your story in an archive to ensure that it is stored for a long time?
8. Hemi oraet blo putum storian blo yu i stap witem olgeta long Vanuatu Kaljoral Senta?
Is it okay to deposit your story with the Vanuatu Cultural Centre?
9. Hemi oraet blo usum storian blo yu insaet long ol samting blo komuniti? Eksampol: ol storibuk, video, mp3
Is it okay to use your story in community material? For example, storybooks, videos, mp3s
10. Hemi oraet blo putum storian lo yu long intanet we eni man savi lisen lo hem?
Is it okay to put your story on the internet, where anyone can listen to it?