A grammar of the Fehan dialect of Tetun
An Austronesian language of West Timor
PACIFIC LINGUISTICS

FOUNDING EDITOR: Stephen A. Wurm

EDITORIAL BOARD: Malcolm D. Ross and Darrell T. Tryon (Managing Editors), John Bowden, Thomas E. Dutton, Andrew K. Pawley

Pacific Linguistics is a publisher specialising in linguistic descriptions, dictionaries, atlases and other material on languages of the Pacific, the Philippines, Indonesia and Southeast Asia. The authors and editors of Pacific Linguistics publications are drawn from a wide range of institutions around the world.

Pacific Linguistics is associated with the Research School of Pacific and Asian Studies at The Australian National University. Pacific Linguistics was established in 1963 through an initial grant from the Hunter Douglas Fund. It is a non-profit-making body financed largely from the sales of its books to libraries and individuals throughout the world, with some assistance from the School.

The Editorial Board of Pacific Linguistics is made up of the academic staff of the School’s Department of Linguistics. The Board also appoints a body of editorial advisors drawn from the international community of linguists. Publications in Series A, B and C and textbooks in Series D are refereed by scholars with relevant expertise who are normally not members of the editorial board.

To date Pacific Linguistics has published over 400 volumes in four series:

Series A: Occasional Papers; collections of shorter papers, usually on a single topic or area.

Series B: Monographs of intermediate length.

Series C: Books; publications of greater length, especially reference books such as dictionaries and grammars, and conference proceedings.

Series D: Special Publications; including archival materials, pedagogical works, maps, audiovisual productions, and materials that do not fit into the other series.

Series C, Volume 155
A grammar of the Fehan dialect of Tetun
An Austronesian language of West Timor

Catharina Lumien van Klinken
## Contents overview

<table>
<thead>
<tr>
<th>Chapter contents in detail</th>
<th>vii</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of maps</td>
<td>xix</td>
</tr>
<tr>
<td>List of tables</td>
<td>xix</td>
</tr>
<tr>
<td>List of figures</td>
<td>xxi</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>xxiii</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>xxv</td>
</tr>
<tr>
<td>Conventions for examples and formulae</td>
<td>xxvii</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 1: Introduction</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 2: Phonology and morphophonemics</td>
<td>14</td>
</tr>
<tr>
<td>Chapter 3: Word classes</td>
<td>49</td>
</tr>
<tr>
<td>Chapter 4: Derivational morphology</td>
<td>58</td>
</tr>
<tr>
<td>Chapter 5: Numerals, classifiers and the numeral phrase</td>
<td>99</td>
</tr>
<tr>
<td>Chapter 6: Pronouns and determiners</td>
<td>110</td>
</tr>
<tr>
<td>Chapter 7: The noun phrase</td>
<td>132</td>
</tr>
<tr>
<td>Chapter 8: Prepositions and the prepositional phrase</td>
<td>159</td>
</tr>
<tr>
<td>Chapter 9: The clause</td>
<td>169</td>
</tr>
<tr>
<td>Chapter 10: Auxiliaries</td>
<td>214</td>
</tr>
<tr>
<td>Chapter 11: Adverbs and verbal modifiers</td>
<td>227</td>
</tr>
<tr>
<td>Chapter 12: Serial verb constructions and prepositional verbs</td>
<td>250</td>
</tr>
</tbody>
</table>
Contents overview

Chapter 13  Complementation  277
Chapter 14  Beyond the clause  294
Appendix A:  Texts  323
Appendix B:  List of texts  333

References  337

Index  351
# Chapter contents in detail

## 1 Introduction

1.1 The Tetun language
   1.1.1 Introduction
   1.1.2 Dialects of Tetun
   1.1.3 Literature in Tetun
   1.1.4 Literature about Tetun

1.2 The Fehan dialect
   1.2.1 Setting
   1.2.2 Languages spoken in Betun
   1.2.3 Language registers
      1.2.3.1 Introduction
      1.2.3.2 Poetry
      1.2.3.3 Noble register
      1.2.3.4 Sea taboo register

1.3 This study
   1.3.1 Introduction
   1.3.2 Data

1.4 Overview of Tetun grammar

## 2 Phonology and morphophonemics

2.1 Introduction
2.2 Syllable template
2.3 Structure of the phonological word
   2.3.1 Introduction
   2.3.2 Word templates
      2.3.2.1 Introduction
      2.3.2.2 Disyllabic word template
      2.3.2.3 Underived four-syllable word template
      2.3.2.4 Trisyllabic word template
      2.3.2.5 Restrictions on initial and final consonants
      2.3.2.6 Prefix and clitic templates
   2.3.3 Lexical stress
   2.3.4 Phonetically long vowels
      2.3.4.1 Introduction
## Chapter contents in detail

2.3.4.2 Contrasts between long and short vowels  | 20
2.3.4.3 Analysis of phonetically long vowels  | 20
2.3.4.4 Orthographic representation  | 21

2.3.5 The evidence from word games  | 22
2.3.5.1 Introduction  | 22
2.3.5.2 Game 1: consonant substitution  | 22
2.3.5.3 Game 2: second syllable modification  | 23
2.3.5.4 Game 3: vowel substitution  | 23

2.4 Vowels  | 24
2.4.1 Vowel phonemes  | 24
2.4.2 Vowel allophones  | 24

2.5 Consonants  | 26
2.5.1 Consonant phonemes  | 26
2.5.2 Loan consonants  | 27
2.5.3 The consonant /w/  | 28
2.5.3.1 Contrast with /u/  | 28
2.5.3.2 Possible merger with /b/  | 28
2.5.3.3 Co-occurrence restrictions on /w/ and back vowels  | 29

2.6 Vowel sequences  | 30
2.7 /k/-initial consonant clusters  | 31
2.7.1 Introduction  | 31
2.7.2 Positions within the word  | 31
2.7.3 Phonetic quality  | 32
2.7.4 Contrasts between word-initial /C/, /kC/ and /kaC/  | 36
2.7.5 Phonological and morphological properties  | 37
2.7.6 Analysis of clusters  | 38
2.7.6.1 Introduction  | 38
2.7.6.2 Extrasyllabicity  | 39
2.7.6.3 Disallowed clusters  | 39

2.8 Phoneme frequencies  | 40
2.8.1 Vowel frequencies  | 40
2.8.2 Consonant frequencies  | 41

2.9 Variation  | 41

2.10 Morphophonemics  | 43
2.10.1 Introduction  | 43
2.10.2 Suffixation  | 43
2.10.3 Prefixation  | 43
2.10.3.1 Intervocalic /h/-insertion  | 43
2.10.3.2 Cluster-reduction on prefixed base  | 44
2.10.3.3 Avoidance of invalid consonant sequences  | 44
2.10.4 Partial reduplication  | 44
2.10.5 Compounding and full reduplication  | 45
2.10.5.1 Word template  | 45
2.10.5.2 Truncation rules  | 45
2.10.5.3 Glottal stop insertion  | 47
2.10.6 Prefixation to and reduplication of bases longer than a foot  | 47
2.10.7 Contraction of underived four-syllable words  | 48
2.10.8 Alternation of ne’el e’el ‘this’  | 48
Chapter contents in detail

3 Word classes

3.1 Introduction
3.2 Multiple membership of word classes
3.3 Nominals
3.4 Verbs
  3.4.1 Introduction
  3.4.2 Distinctions between transitive verbs, intransitive verbs and adjectives
  3.4.3 Further comments on adjectives
  3.4.4 Further cross-classification of transitive and intransitive verbs
  3.4.5 Copulas
3.5 Numerals
3.6 Numeral classifiers
3.7 Auxiliaries
3.8 Adverbs
3.9 Determiners
3.10 Relators
3.11 Tags
3.12 Interjections

4 Derivational morphology

4.1 Overview
4.2 Derivation of verbs: ha-
  4.2.1 Introduction
  4.2.2 Causative verb from verbal base: ha-
    4.2.2.1 Causative verb from underived base
    4.2.2.2 Transitive verb from derived stem
    4.2.2.3 Intransitive verb
    4.2.2.4 Omission of final consonant
  4.2.3 Verb from nominal base: ha-
4.3 Derivation of verbs: hak-
  4.3.1 Introduction
  4.3.2 Intransitive verb from transitive base: hak-
  4.3.3 No change in transitivity: hak- (-k)
  4.3.4 Reciprocal action verb: hak- -k
  4.3.5 ‘Make sound’ verb: hak- (-k), ha- (-k)
4.4 Derivation of actor-describing verbs: mak- (-n), ma- (-k)
  4.4.1 Overview
  4.4.2 Distinction between mak- and ma-
  4.4.3 Conditions on use of suffix
  4.4.4 Syntactic function of actor derivations
  4.4.5 Contrast between prefix and relative clause marker mak
4.5 Derivation of nouns
  4.5.1 Nominalisation: partial reduplication
  4.5.2 Instrument, undergoer and result nouns
    4.5.2.1 Introduction
    4.5.2.2 Using partial reduplication
x  Chapter contents in detail

  4.5.2.3  Using k(a)- (-k)  81
  4.5.2.4  Using -n  82
  4.5.2.5  Using -k  82
  4.5.2.6  Zero derivation  83
  4.5.3  Time nouns: ma-, ma- -k  83
  4.5.4  Compounds  84

  4.6  Derivation of adjectives  85
  4.6.1  Overview  85
  4.6.2  Result: k- -k  86
  4.6.3  ‘Quality of actor’: k- -k  87
  4.6.4  ‘Similarity to’ adjective base: k(a)- -k  87
  4.6.5  ‘Quality of’ noun base, ‘having’: k(a)- -k  88
  4.6.6  ‘Similarity to’: -k  88
  4.6.7  Attributive adjective: -k  89
  4.6.8  Result: ma- -k  90
  4.6.9  Result: -n  91

  4.7  Derivation of adverbs: full reduplication  92

  4.8  Other functions of full reduplication  93
  4.8.1  Reduplication of verb: ‘do heedlessly’  93
  4.8.2  Reduplication of numerals: group, distributive  93
  4.8.3  Reduplication of adjectives  94
  4.8.4  Repetition of nouns: various plural  95

  4.9  Miscellaneous low-frequency derivations  95
  4.9.1  Introduction  95
  4.9.2  Iteration: partial reduplication of verbs  95
  4.9.3  Names of fingers and toes: fa-, sa-  96
  4.9.4  An alleged prefix ba-  97
  4.9.5  Nicknames: omission of k-  97
  4.9.6  Adjective: -s  97

  5  Numerals, classifiers and the numeral phrase  99

    5.1  The numeral phrase  99
    5.2  Numerals  100
    5.2.1  Cardinal numerals  100
    5.2.2  Culturally significant numbers  102
    5.2.3  Arithmetic  102
    5.2.4  Ordinal numbers and alternative expressions  103

    5.3  Classifiers  104
    5.3.1  Introduction  104
    5.3.2  Sortal classifiers  104
    5.3.2.1  Inventory  104
    5.3.2.2  Use  106

    5.3.3  Mensural classifiers  107
    5.3.3.1  Inventory  107
    5.3.3.2  Comparison with nouns  108

    5.4  Head of a numeral phrase  109
6 Pronouns and determiners

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Overview</td>
<td>110</td>
</tr>
<tr>
<td>6.2</td>
<td>Personal pronouns</td>
<td>111</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Overview</td>
<td>111</td>
</tr>
<tr>
<td>6.2.2</td>
<td>ita: inclusive and general ‘we’</td>
<td>112</td>
</tr>
<tr>
<td>6.2.3</td>
<td>ita, ó ‘you’</td>
<td>113</td>
</tr>
<tr>
<td>6.2.4</td>
<td>Other first and second person pronouns</td>
<td>113</td>
</tr>
<tr>
<td>6.2.5</td>
<td>sia ‘they’</td>
<td>114</td>
</tr>
<tr>
<td>6.2.6</td>
<td>nia ‘he, she, it, that’</td>
<td>114</td>
</tr>
<tr>
<td>6.3</td>
<td>Demonstratives</td>
<td>115</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Introduction</td>
<td>115</td>
</tr>
<tr>
<td>6.3.2</td>
<td>ne’e ‘this’</td>
<td>115</td>
</tr>
<tr>
<td>6.3.3</td>
<td>nia ‘that’</td>
<td>117</td>
</tr>
<tr>
<td>6.3.4</td>
<td>ne’et ‘that’</td>
<td>119</td>
</tr>
<tr>
<td>6.3.5</td>
<td>nemai ‘here, this’, nebá ‘there, that’</td>
<td>119</td>
</tr>
<tr>
<td>6.4</td>
<td>Quantifying pronouns/determiners</td>
<td>120</td>
</tr>
<tr>
<td>6.5</td>
<td>Other determiners</td>
<td>121</td>
</tr>
<tr>
<td>6.5.1</td>
<td>á: definite article</td>
<td>121</td>
</tr>
<tr>
<td>6.5.2</td>
<td>ida ‘one’ as indefinite article</td>
<td>123</td>
</tr>
<tr>
<td>6.5.3</td>
<td>sia: plural marker</td>
<td>124</td>
</tr>
<tr>
<td>6.5.4</td>
<td>Comparison of pronouns and determiners</td>
<td>125</td>
</tr>
<tr>
<td>6.6</td>
<td>Epistememes (interrogative/indefinite)</td>
<td>127</td>
</tr>
<tr>
<td>6.6.1</td>
<td>Overview of epistememes</td>
<td>127</td>
</tr>
<tr>
<td>6.6.2</td>
<td>Basic epistememes: interrogative</td>
<td>128</td>
</tr>
<tr>
<td>6.6.2.1</td>
<td>sé ‘who, which person’</td>
<td>128</td>
</tr>
<tr>
<td>6.6.2.2</td>
<td>nabé ‘which, where’</td>
<td>128</td>
</tr>
<tr>
<td>6.6.2.3</td>
<td>sá ‘what’</td>
<td>129</td>
</tr>
<tr>
<td>6.6.3</td>
<td>Basic epistememes: indefinite</td>
<td>129</td>
</tr>
<tr>
<td>6.6.4</td>
<td>Complex epistememes</td>
<td>129</td>
</tr>
<tr>
<td>6.6.4.1</td>
<td>Epistememe + ida ‘one’: interrogative, indefinite</td>
<td>129</td>
</tr>
<tr>
<td>6.6.4.2</td>
<td>Repeated epistememe: indefinite</td>
<td>130</td>
</tr>
<tr>
<td>6.6.5</td>
<td>Distribution of epistemic determiners</td>
<td>131</td>
</tr>
</tbody>
</table>

7 The noun phrase

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1</td>
<td>Introduction</td>
<td>132</td>
</tr>
<tr>
<td>7.2</td>
<td>The structure of simple NPs</td>
<td>132</td>
</tr>
<tr>
<td>7.2.1</td>
<td>Introduction</td>
<td>132</td>
</tr>
<tr>
<td>7.2.2</td>
<td>NP with underived common noun head</td>
<td>132</td>
</tr>
<tr>
<td>7.2.2.1</td>
<td>General structure</td>
<td>132</td>
</tr>
<tr>
<td>7.2.2.2</td>
<td>Marking plurality</td>
<td>134</td>
</tr>
<tr>
<td>7.2.2.3</td>
<td>Order of modifiers: numerals and adjectives</td>
<td>136</td>
</tr>
<tr>
<td>7.2.2.4</td>
<td>Order of modifiers: prepositional phrases, relative clauses and determiners</td>
<td>136</td>
</tr>
<tr>
<td>7.2.2.5</td>
<td>Order of modifiers: determiners</td>
<td>137</td>
</tr>
<tr>
<td>7.2.3</td>
<td>NP with deverbal noun head</td>
<td>138</td>
</tr>
<tr>
<td>7.2.4</td>
<td>NP with pronominal head</td>
<td>138</td>
</tr>
<tr>
<td>7.2.5</td>
<td>NP with proper noun as head</td>
<td>139</td>
</tr>
</tbody>
</table>
Chapter contents in detail

7.2.5.1  Sociolinguistic rules for terms of address 139
7.2.5.2  Structure of a proper noun NP 140
7.2.6  NP without head nominal 141

7.3  Possession
7.3.1  Overview 142
7.3.2  Form of the possessor: pronouns and possessive clitics 145
  7.3.2.1  Overview of possible forms 145
  7.3.2.2  Alienability 145
  7.3.2.3  Effect of alienability on the form of the possessor 146
  7.3.2.4  Effect of position of possessor 148
  7.3.2.5  Effect of style 149
7.3.3  Form of the possessum: genitive clitics 149
  7.3.3.1  Distinction between -n and -r 149
  7.3.3.2  Effect of alienability and 1S possessors 150
  7.3.3.3  -n on non-possessed nouns 150
7.3.4  Differences between preposed and postposed possessor 151
  7.3.4.1  Effect of alienability 151
  7.3.4.2  Other factors 152
7.3.5  Adjectival ‘possessors’ 152

7.4  Premodifiers
7.4.1  Introduction 153
7.4.2  ohin ‘aforementioned’ 153
7.4.3  sura ‘every’ 153
7.4.4  Premodifying relative clause 154

7.5  Postmodifiers
7.5.1  Noun phrase 155
7.5.2  Adjective phrase 156
7.5.3  Relative clause 156

7.6  Complex NPs
7.6.1  Coordinate noun phrase 156
  7.6.1.1  Regular coordination 156
  7.6.1.2  Dual pronoun construction 157
7.6.2  Apposition of noun phrases 158

8  Prepositions and the prepositional phrase
8.1  The prepositional phrase 159
8.2  Overview of prepositions 159
8.3  Prepositions of location
  8.3.1  iha ‘at (location)’ 160
  8.3.2  rē ‘at (location)’ 162
  8.3.3  hori ‘from (location)’ 162
8.4  Prepositions of time
  8.4.1  hori ‘since (time)’ 163
  8.4.2  tu’o ‘until’ 163
  8.4.3  natōn bā, bā natōn ‘at (time)’ 164
8.5  Prepositions of similarity
  8.5.1  nú, nu’u ‘like, just like’ 164
  8.5.2  nudar ‘like, be as if’ 165
8.5.3  *hanesan* 'like, resemble'  

8.6  Other prepositions  
8.6.1  *bodik* 'for (beneficiary); as for'  
8.6.2  *hetak* 'as for'  
8.7  The prepositional phrase as predicate  

9  The clause  

9.1  Introduction  

9.2  Subjects and objects  
9.2.1  Properties of subjects and objects  
9.2.2  Ellipsis of subjects and objects  

9.3  Subject marking  
9.3.1  Overview  
9.3.2  Consistency of subject marking  
9.3.3  Inflection of adjectives and non-verbs  
9.3.3.1  Introduction  
9.3.3.2  Adjectives  
9.3.3.3  Other word classes  
9.3.4  A developing pronoun *hak-*?  
9.3.5  Subject marker versus subject  

9.4  Verbal predicates  
9.4.1  Introduction  
9.4.2  Intransitive predicates  
9.4.3  Transitive predicates  
9.4.3.1  Introduction  
9.4.3.2  SVO constituent order  
9.4.3.3  OSV constituent order: topicalisation  
9.4.3.4  'SOV' constituent order: object incorporation  
9.4.4  Ditransitive predicates  
9.4.5  Copular constructions  
9.4.5.1  Introduction  
9.4.5.2  *nī* 'be'  
9.4.5.3  *tu'an* 'grow'  
9.4.5.4  *dadi* 'become'  
9.4.6  *kona* 'touch, undergo'  
9.4.7  Existential and possessive constructions  
9.4.7.1  Introduction  
9.4.7.2  Intransitive *ihā* 'be present'  
9.4.7.3  Intransitive *n-ō* '3-exist'  
9.4.7.4  Transitive *hō* 'have'  
9.4.7.5  Overlap for *nō* '3S-have/exist'  
9.4.7.6  Comparison of *ihā* and *nō*  
9.4.7.7  *lalek* 'lack'  
9.4.7.8  Presentational NP  

9.5  Non-verbal predicates  
9.5.1  Prepositional phrases  
9.5.2  Numeral phrases  
9.5.3  Noun phrases
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.5.4</td>
<td>Possessors</td>
<td>193</td>
<td></td>
</tr>
<tr>
<td>9.5.5</td>
<td>Relative clause NPs: clefts</td>
<td>194</td>
<td></td>
</tr>
<tr>
<td>9.5.5.1</td>
<td>Structure</td>
<td>194</td>
<td></td>
</tr>
<tr>
<td>9.5.5.2</td>
<td>Function</td>
<td>195</td>
<td></td>
</tr>
<tr>
<td>9.6</td>
<td>'body-good' predicates</td>
<td>196</td>
<td></td>
</tr>
<tr>
<td>9.6.1</td>
<td>Basic description</td>
<td>196</td>
<td></td>
</tr>
<tr>
<td>9.6.2</td>
<td>Lexicalisation and productivity</td>
<td>198</td>
<td></td>
</tr>
<tr>
<td>9.6.3</td>
<td>Distribution of subject properties</td>
<td>199</td>
<td></td>
</tr>
<tr>
<td>9.6.4</td>
<td>Analysis</td>
<td>201</td>
<td></td>
</tr>
<tr>
<td>9.6.4.1</td>
<td>Analysis as clausal predicate</td>
<td>201</td>
<td></td>
</tr>
<tr>
<td>9.6.4.2</td>
<td>Arguments against analysis as subject incorporation</td>
<td>202</td>
<td></td>
</tr>
<tr>
<td>9.6.4.3</td>
<td>Arguments against analysis as topic-comment</td>
<td>203</td>
<td></td>
</tr>
<tr>
<td>9.7</td>
<td>Dependent clauses</td>
<td>204</td>
<td></td>
</tr>
<tr>
<td>9.7.1</td>
<td>Introduction</td>
<td>204</td>
<td></td>
</tr>
<tr>
<td>9.7.2</td>
<td>Reduced complement clauses</td>
<td>204</td>
<td></td>
</tr>
<tr>
<td>9.7.3</td>
<td>Small clauses</td>
<td>205</td>
<td></td>
</tr>
<tr>
<td>9.8</td>
<td>Peripheral constituents</td>
<td>206</td>
<td></td>
</tr>
<tr>
<td>9.8.1</td>
<td>Overview</td>
<td>206</td>
<td></td>
</tr>
<tr>
<td>9.8.2</td>
<td>Time</td>
<td>206</td>
<td></td>
</tr>
<tr>
<td>9.8.3</td>
<td>Location</td>
<td>207</td>
<td></td>
</tr>
<tr>
<td>9.8.3.1</td>
<td>Difference between core and peripheral location</td>
<td>207</td>
<td></td>
</tr>
<tr>
<td>9.8.3.2</td>
<td>Peripheral location</td>
<td>207</td>
<td></td>
</tr>
<tr>
<td>9.9</td>
<td>Non-declarative clauses</td>
<td>208</td>
<td></td>
</tr>
<tr>
<td>9.9.1</td>
<td>Introduction</td>
<td>208</td>
<td></td>
</tr>
<tr>
<td>9.9.2</td>
<td>Imperatives</td>
<td>208</td>
<td></td>
</tr>
<tr>
<td>9.9.3</td>
<td>Information interrogatives</td>
<td>209</td>
<td></td>
</tr>
<tr>
<td>9.9.3.1</td>
<td>Description of information interrogatives</td>
<td>209</td>
<td></td>
</tr>
<tr>
<td>9.9.3.2</td>
<td>Answers to information questions</td>
<td>210</td>
<td></td>
</tr>
<tr>
<td>9.9.4</td>
<td>Disjunctive and polar interrogatives</td>
<td>211</td>
<td></td>
</tr>
<tr>
<td>9.9.4.1</td>
<td>Description of interrogatives</td>
<td>211</td>
<td></td>
</tr>
<tr>
<td>9.9.4.2</td>
<td>Answers to polar questions</td>
<td>211</td>
<td></td>
</tr>
<tr>
<td>9.9.5</td>
<td>Speech acts using polar interrogatives</td>
<td>212</td>
<td></td>
</tr>
<tr>
<td>9.9.5.1</td>
<td>Question bias</td>
<td>212</td>
<td></td>
</tr>
<tr>
<td>9.9.5.2</td>
<td>Rhetorical questions</td>
<td>212</td>
<td></td>
</tr>
<tr>
<td>9.9.5.3</td>
<td>Other indirect speech acts</td>
<td>213</td>
<td></td>
</tr>
<tr>
<td>9.9.6</td>
<td>Greetings</td>
<td>213</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Auxiliaries</td>
<td>214</td>
<td></td>
</tr>
<tr>
<td>10.1</td>
<td>Introduction</td>
<td>214</td>
<td></td>
</tr>
<tr>
<td>10.2</td>
<td>Characteristics of auxiliaries</td>
<td>214</td>
<td></td>
</tr>
<tr>
<td>10.2.1</td>
<td>Similarities to complement-taking verbs</td>
<td>214</td>
<td></td>
</tr>
<tr>
<td>10.2.2</td>
<td>Differences from complement-taking verbs</td>
<td>215</td>
<td></td>
</tr>
<tr>
<td>10.2.3</td>
<td>Distinctions amongst auxiliaries</td>
<td>217</td>
<td></td>
</tr>
<tr>
<td>10.3</td>
<td>Aspectual auxiliaries</td>
<td>218</td>
<td></td>
</tr>
<tr>
<td>10.3.1</td>
<td>Overview</td>
<td>218</td>
<td></td>
</tr>
<tr>
<td>10.3.2</td>
<td>ho'i: progressive</td>
<td>218</td>
<td></td>
</tr>
<tr>
<td>10.3.3</td>
<td>foin 'only just'</td>
<td>219</td>
<td></td>
</tr>
<tr>
<td>10.3.4</td>
<td>sei with positive polarity: 'still'</td>
<td>219</td>
<td></td>
</tr>
</tbody>
</table>
Chapter contents in detail

10.3.5 *sei* with negative polarity: ‘not yet’ 219
10.3.6 *atu*: irrealis 220
   10.3.6.1 Introduction 220
   10.3.6.2 *atu* in preverbal position: ‘about to, want to, intend to, in order to’ 221
   10.3.6.3 *atu*: irrealis mood 222
10.4 Deontic modal auxiliaries (permission and obligation) 223
   10.4.1 Overview 223
   10.4.2 *bele* ‘can’ 224
   10.4.3 *lalika(n)* ‘need not’ 225
   10.4.4 *keta* ‘do not’ 226
   10.4.5 *musti* ‘must’ 226

11 Adverbs and verbal modifiers 227
   11.1 Overview 227
   11.2 Negators 228
      11.2.1 Overview of negation 228
      11.2.2 Clausal negators: *la*, *ha*i ‘not’ 228
         11.2.2.1 Shared features 228
         11.2.2.2 Differences between *la* and *ha*i 229
      11.2.3 Contrastive constituent negator: *lahós* ‘indeed not’ 230
   11.3 Degree modifiers 231
      11.3.1 Downtoner: *ladún* ‘not very’ 231
      11.3.2 Intensifiers 231
   11.4 Participant coreference markers 232
      11.4.1 Introduction 232
      11.4.2 Reflexive *án* 232
      11.4.3 Reciprocal *malu* 233
   11.5 Aspectual and temporal markers 235
      11.5.1 Introduction 235
      11.5.2 *ti’an*, *ti’a*, *ta* ‘already’ 235
      11.5.3 *kedan* ‘immediately’ 237
      11.5.4 *onan* ‘imminent’ 238
      11.5.5 -*n* ‘imminent’ 239
      11.5.6 *lai* ‘first, beforehand’ 240
   11.6 Adverbs of repetition and frequency 240
   11.7 Manner modifiers 242
   11.8 Deictic particles 243
   11.9 Imperative *bá* 244
   11.10 *bá* ‘so be it’ 245
   11.11 Connective adverbs 245
   11.12 Sentence adverbs of possibility 247

12 Serial verb constructions and prepositional verbs 250
   12.1 Overview of serial verbs 250
      12.1.1 Introduction 250
      12.1.2 Characteristics of serial verbs 250
Chapter contents in detail

12.1.3 Comparison with paratactic clauses 252
12.1.4 Scope of negation 254
12.1.5 Overview of construction types 254

12.2 Nuclear serialisation 256
12.2.1 Overview 256
12.2.2 Nuclear causative (type 1) 257
   12.2.2.1 Description 257
   12.2.2.2 Comparison with morphological causatives 258
12.2.3 Motion – intransitive direction (type 2) 258
   12.2.3.1 Description 258
   12.2.3.2 Comparison with verb-modifier constructions 260

12.3 Transitive verbs sharing subject and object (type 3) 260

12.4 Motion-action (type 4) 262
12.4.1 Description 262
12.4.2 Comparison with auxiliaries 263

12.5 Introducing oblique arguments 263
12.5.1 Introduction: serial verbs, prepositional verbs and prepositions 263
12.5.2 Motion – transitive direction (type 5) 264
   12.5.2.1 Introduction 264
   12.5.2.2 Same-subject motion-direction 265
   12.5.2.3 Switch-subject motion-direction 266
12.5.3 Other functions of ba 267
   12.5.3.1 Introduction 267
   12.5.3.2 ba ‘to (addressee)’ – prepositional verb 268
   12.5.3.3 ba ‘to (recipient)’ – prepositional verb 268
   12.5.3.4 ba ‘at (location)’ – prepositional verb 269
   12.5.3.5 ba ‘to (factitive result)’ – preposition? 270
   12.5.3.6 ba ‘for (beneficiary, purpose)’ – preposition 270
   12.5.3.7 ba ‘at (time)’ – preposition 271
   12.5.3.8 Summary of functions of ba 271

12.5.4 ho ‘accompany’ – verb, prepositional verb (type 6) 272
12.5.5 hodi ‘bring, use’ (type 6) 273
   12.5.5.1 Overview 273
   12.5.5.2 ‘use (instrument)’ – verb 273
   12.5.5.3 ‘use (instrument)’ – prepositional verb 274
   12.5.5.4 ‘with (comitative)’ – verb 275
   12.5.5.5 ‘by (time)’ – prepositional verb 275

12.6 Participant cooperation: ‘together’/’alone’ (type 7) 275

13 Complementation 277

13.1 Classes of complement-taking verbs 277
13.2 Overview of complement types 278
13.3 Sentential complements 279
   13.3.1 Introduction 279
   13.3.2 Clause-final complements: non-causatives 280
   13.3.3 Clause-final complements: analytic causatives 281
      13.3.3.1 Syntax 281
13.3.3.2 Comparison with morphological and nuclear serial verb causatives 283

13.3.4 Complementisers 283
13.3.4.1 Introduction 283
13.3.4.2 hosí ‘about’ 283
13.3.4.3 Purpose conjunction 285

13.3.5 Clause-initial complements 286
13.3.5.1 Introduction 286
13.3.5.2 Question + ‘know’ 286
13.3.5.3 Statement + ‘want’ 287

13.3.6 Sentential subjects 288

13.4 Reduced complements 288
13.4.1 Description 288
13.4.2 Object control: reported orders 289
13.4.3 Subject control 290
13.4.3.1 Overview 290
13.4.3.2 Clause-final complements 291
13.4.3.3 Preverbal small clause complements 292

14 Beyond the clause 294

14.1 Introduction 294

14.2 Extra-clausal constituents 294
14.2.1 Overview 294
14.2.2 Pragmatic topics 295
14.2.2.1 Introduction 295
14.2.2.2 Left-dislocation: topic 295
14.2.2.3 Left-dislocation: ‘premature subject’ 296
14.2.2.4 Topics with no clause-internal role 297
14.2.3 Right-dislocation: elucidation of a referent 298
14.2.4 Vocatives 299
14.2.5 Tags 299

14.3 Phrasal and clausal coordination 301
14.3.1 Introduction 301
14.3.2 nó ‘and’ 301
14.3.3 ká ‘or’ 302
14.3.4 Paratactic coordination 303

14.4 Clause parataxis 304

14.5 Coordinate-dependent constructions 305
14.5.1 Introduction 305
14.5.2 hodi: coordinate dependent 306
14.5.3 ha’ak ‘say’ 308

14.6 Nominalisation with ne’e 310

14.7 Adverbial clauses introduced by conjunctions 311
14.7.1 Overview 311
14.7.2 Time 311
14.7.3 Condition 312
14.7.3.1 kalo ‘if’ 312
14.7.3.2 só ‘unless’ 313
### Chapter contents in detail

14.7.3.3 surak 'on condition that'  
14.7.4 Concession: masik 'although'  
14.7.5 Reason  
  14.7.5.1 Introduction  
  14.7.5.2 tán, tán bá 'because'  
  14.7.5.3 té 'because'  
  14.7.5.4 wain 'because'  
14.7.6 Purpose  
14.8 Relative clauses  
  14.8.1 Introduction  
  14.8.2 Relative clause with *mak*  
    14.8.2.1 Clause structure  
    14.8.2.2 Non-restrictive relative clause  
    14.8.2.3 Headless relative clause  
  14.8.3 Relative clause without *mak*  
    14.8.3.1 Relativisation on subject and object  
    14.8.3.2 Time  
    14.8.3.3 Associated entity  
    14.8.3.4 Comparison with attributive adjective phrases
Maps, tables and figures

Maps

<table>
<thead>
<tr>
<th>Map</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Map 1</td>
<td>Languages of Timor</td>
<td>xxx</td>
</tr>
<tr>
<td>Map 2</td>
<td>Location of Timor</td>
<td>1</td>
</tr>
</tbody>
</table>

Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Sample of noble register vocabulary</td>
<td>8</td>
</tr>
<tr>
<td>Table 2</td>
<td>Sample of sea taboo vocabulary</td>
<td>9</td>
</tr>
<tr>
<td>Table 2.1</td>
<td>Examples of syllable templates</td>
<td>14</td>
</tr>
<tr>
<td>Table 2.2</td>
<td>Examples of word templates: two syllables</td>
<td>16</td>
</tr>
<tr>
<td>Table 2.3</td>
<td>Examples of word templates: four syllables</td>
<td>16</td>
</tr>
<tr>
<td>Table 2.4</td>
<td>Examples of word templates: three syllables</td>
<td>17</td>
</tr>
<tr>
<td>Table 2.5</td>
<td>Examples of lexical stress patterns</td>
<td>19</td>
</tr>
<tr>
<td>Table 2.6</td>
<td>Vowel phonemes</td>
<td>24</td>
</tr>
<tr>
<td>Table 2.7</td>
<td>Vowel contrasts</td>
<td>24</td>
</tr>
<tr>
<td>Table 2.8</td>
<td>Examples of high and low allophones of mid vowels</td>
<td>26</td>
</tr>
<tr>
<td>Table 2.9</td>
<td>Consonant phonemes</td>
<td>26</td>
</tr>
<tr>
<td>Table 2.10</td>
<td>Consonant contrasts word-initially, after /k/ and word-medially</td>
<td>27</td>
</tr>
<tr>
<td>Table 2.11</td>
<td>Consonant contrasts word-finally</td>
<td>27</td>
</tr>
<tr>
<td>Table 2.12</td>
<td>Assimilation of loan consonants</td>
<td>28</td>
</tr>
<tr>
<td>Table 2.13</td>
<td>Examples of vowel-/w/ combinations</td>
<td>30</td>
</tr>
<tr>
<td>Table 2.14</td>
<td>Examples of vowel sequences</td>
<td>30</td>
</tr>
<tr>
<td>Table 2.15</td>
<td>Contrasts between word-initial /C/, /kC/ and /kaC/</td>
<td>37</td>
</tr>
<tr>
<td>Table 2.16</td>
<td>Vowel frequencies (as percentages)</td>
<td>40</td>
</tr>
<tr>
<td>Table 2.17</td>
<td>Consonant frequencies (as percentages)</td>
<td>41</td>
</tr>
<tr>
<td>Table 2.18</td>
<td>Examples of variation in non-final consonants</td>
<td>42</td>
</tr>
<tr>
<td>Table 2.19</td>
<td>Examples of dropping of final vowels</td>
<td>42</td>
</tr>
<tr>
<td>Table 2.20</td>
<td>Examples of variant vowels</td>
<td>43</td>
</tr>
<tr>
<td>Table 3.1</td>
<td>Comparison of transitive verbs, intransitive verbs and adjectives</td>
<td>52</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>Derivation of nouns and adjectives</td>
<td>58</td>
</tr>
<tr>
<td>Table 5.1</td>
<td>Numerals</td>
<td>100</td>
</tr>
<tr>
<td>Table 5.2</td>
<td>Sortal classifiers</td>
<td>105</td>
</tr>
<tr>
<td>Table 5.3</td>
<td>Mensural classifiers</td>
<td>108</td>
</tr>
<tr>
<td>Table 6.1</td>
<td>Pronouns and determiners</td>
<td>110</td>
</tr>
<tr>
<td>Table 6.2</td>
<td>Personal pronouns</td>
<td>111</td>
</tr>
<tr>
<td>Table 6.3</td>
<td>Comparison of pronouns and determiners: ne’e, nia, sia and á</td>
<td>126</td>
</tr>
<tr>
<td>Table 7.1</td>
<td>Noun phrase constituents</td>
<td>133</td>
</tr>
<tr>
<td>Table 7.2</td>
<td>Summary of major factors influencing choice of possessive construction</td>
<td>144</td>
</tr>
<tr>
<td>Table 7.3</td>
<td>Syntactic correlates of alienability classes</td>
<td>146</td>
</tr>
<tr>
<td>Table 7.4</td>
<td>Effect of alienability on possessive marking of preposed possessors</td>
<td>148</td>
</tr>
<tr>
<td>Table 8.1</td>
<td>Prepositions</td>
<td>160</td>
</tr>
<tr>
<td>Table 8.2</td>
<td>Location nouns</td>
<td>161</td>
</tr>
<tr>
<td>Table 9.1</td>
<td>Subject markers</td>
<td>173</td>
</tr>
<tr>
<td>Table 9.2</td>
<td>Examples of body-good expressions</td>
<td>197</td>
</tr>
<tr>
<td>Table 9.3</td>
<td>Examples of literal and metaphorical interpretations of body-good expressions</td>
<td>198</td>
</tr>
<tr>
<td>Table 9.4</td>
<td>Examples of near synonyms: body-good expressions and adjectives</td>
<td>198</td>
</tr>
<tr>
<td>Table 10.1</td>
<td>Auxiliaries</td>
<td>217</td>
</tr>
<tr>
<td>Table 11.1</td>
<td>Adverbs of repetition and frequency</td>
<td>241</td>
</tr>
<tr>
<td>Table 11.2</td>
<td>Connective adverbs which are phonologically clause-initial</td>
<td>245</td>
</tr>
<tr>
<td>Table 11.3</td>
<td>Connective adverbs which need not be phonologically clause-initial</td>
<td>246</td>
</tr>
<tr>
<td>Table 12.1</td>
<td>Serial verb construction types</td>
<td>256</td>
</tr>
<tr>
<td>Table 12.2</td>
<td>Direction verbs in nuclear serialisation</td>
<td>259</td>
</tr>
<tr>
<td>Table 12.3</td>
<td>Contrast between transitive verbs and prepositions</td>
<td>264</td>
</tr>
<tr>
<td>Table 12.4</td>
<td>Direction verbs in core layer serialisation</td>
<td>266</td>
</tr>
<tr>
<td>Table 12.5</td>
<td>Verbal characteristics of ba in various functions</td>
<td>272</td>
</tr>
<tr>
<td>Table 13.1</td>
<td>Complement types found for each class of complement-taking predicate</td>
<td>279</td>
</tr>
<tr>
<td>Table 13.2</td>
<td>Semantic contrast between sentential and reduced complements</td>
<td>290</td>
</tr>
<tr>
<td>Table 14.1</td>
<td>Comparison of hodi, atu and the purpose conjunctions batu and ne’ebé</td>
<td>316</td>
</tr>
</tbody>
</table>
Figures

Figure 2.1  Sound waveform for *Ami talik* ‘We intertwine (arms)’  
Figure 2.2  Sound waveform for *ai ktalik* ‘vine’ (lit. ‘plant vine’)  
Figure 2.3  Sound waveform for *Nia dök* ‘He (is a) shaman’  
Figure 2.4  Sound waveform for *Nia kdök* ‘He (is) far away’  
Figure 2.5  Sound waveform for *Ha’u kadök* ‘I distance (it)’  
Figure 2.6  Sound waveform for *Nia feur* ‘He spins (a top); *Ne’e kfeu* ‘This (is a) kfeu tree’  
Figure 2.7  Sound waveform for *Ne’e kafeur* ‘This (is a) top’
Acknowledgments

The present study was made possible, and furthermore largely enjoyable, by many people. Within The Australian National University my primary thanks go to my supervisors in the Department of Linguistics in the Research School of Pacific and Asian Studies, Dr Malcolm Ross and Professor Andrew Pawley. They saw me through the long process, and stimulated many improvements in analysis and expression. Dr Darrell Tryon took on the unenviable task of commenting on the pre-final draft as a single composition. Dr Nikolaus Himmelmann kindly volunteered to read some chapters. The receipt of an Australian Postgraduate Research Award was greatly appreciated.

Dr Chuck Grimes, then a visiting fellow, together with fellow students Robert Early and Mark Donohue, helped prepare me for fieldwork and guided me to useful computer tools. Other staff and students who contributed in diverse ways included Dianna Kovacs, Joan Birnie, Professor James Fox, Cindi Farr, Phil Quick, Lea Brown, Jeff Marck, Chikao Yoshimura and Ian Heyward. People further afield who provided assistance include Dr Juliette Blevins, of the University of Western Australia, who gave useful feedback on a draft of the phonology chapter. Kevin Sherlock, in Darwin, was helpful in locating material on Tetun, while Antonio Sarmento, in Brisbane, gave me some initial lessons in Tetun.

Fieldwork in Timor was made possible through the auspices of the Artha Wacana Christian University in Kupang. Particular thanks go to the rector Dr A.A. Yewangoe, and to my immediate supervisor Drs Frits FanggidaE. They provided a base in Kupang, developed a formal proposal for fieldwork, and together with Rev. (now Rev. Dr) Tom Therik saw me through the maze of paperwork and reporting. Tom Therik's assistance was invaluable, giving introductions to people in Kupang, Atambua and Betun, finding me accommodation, and orientating me to Tetun culture and language. His friendship is much appreciated.

In Betun Mama Funan and Bapa Fuah and family gave me a secure and comfortable home base, where there was laughter, good parental advice, and generous helpings of Mama's excellent baking. Ama Bo'uk and Ina Bete of Uma Katahan offered much hospitality and taught me a lot about Tetun language and culture; Ama Bo'uk additionally introduced me to many interesting people, and accompanied me on longer trips. Special thanks are also due to their family (especially Bi Iku), and to many other members of Uma Katuas (including Bi Mea, Mau Manek, Bi Lotu and Ama Klau Ki'ik). Others who gave formal assistance in language learning, as well as providing friendship and hospitality together with their families, included Pak Hendrikus Dini, Ibu Juliana Se'uk Seran, the late Ibu Serlince Manafe, Pak Alfonsius Klau, Pak Yohanes Tahu (Anis), Bi Lan, Ina Bete Ki'ik.
Acknowledgments

and Pak Herminus. Esther Hawthaba and the family of Pak Mias Pandi assisted in transcribing texts. Numerous others welcomed me into their homes and shared with me in many ways; some deserving special mention are Na‘i Mea, Na‘i Ni‘is and Na‘i Nona in Laran, Liurai L.S. Tei Seran and family in Betun, and Pak Ferdy and Ibu Brigitta in Lafoun. Such hospitality was not restricted to Tetun people, with others of diverse ethnic backgrounds generously accepting me into their homes, particularly on visits to Atambua and Kupang.

A final expression of thanks goes to my family. It was my parents Arie and Anneke van Klinken who first awakened my interest in Indonesia, and sister-in-law Helene van Klinken who opened my eyes to the possibility of pursuing linguistic research in Timor. The last word, however, goes to my husband Rob Williams, particularly for his patience and encouragement, and for allowing me to ‘follow a dream’ at such cost to himself.
The term ‘Malay’ is used rather than ‘Indonesian’ because Malay already had some currency as a lingua franca in Timor (at least in the ports on the north coast) by the early nineteenth century (Thomaz 1981:54), long before the development of Bahasa Indonesia as a national language. In addition, more recent loans are likely to be from Kupang Malay rather than from Indonesian. All loans which are ultimately from Dutch are presumed at this stage to have come via Malay. Very commonly used loans are not acknowledged as loans in the examples. These include Malay ibu ‘mother’ (the title by which I was addressed) and kalo ‘if’, and oras ‘time’ (from Portuguese horas ‘hours’).

In §7.3.3.3 it is noted that -n is often used on location, body part and kin nouns even when these are not possessed. To simplify glosses, -n is glossed as a separate morpheme only when the noun heads a possessive NP.

---

**Abbreviations**

<table>
<thead>
<tr>
<th>Code</th>
<th>Meaning</th>
<th>Code</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>word boundary</td>
<td>Cnj</td>
<td>conjunction</td>
</tr>
<tr>
<td>[Mly]</td>
<td>loan from or via Malay1</td>
<td>Comp</td>
<td>complement</td>
</tr>
<tr>
<td>[Port]</td>
<td>loan from Portuguese</td>
<td>COORD</td>
<td>coordinator (hodi)</td>
</tr>
<tr>
<td>1PE</td>
<td>1st person plural exclusive (ami)</td>
<td>DEF</td>
<td>definite (á)</td>
</tr>
<tr>
<td>1PI</td>
<td>1st person plural inclusive (ita)</td>
<td>Det</td>
<td>determiner</td>
</tr>
<tr>
<td>1S</td>
<td>1st person singular (ha’u, k-)</td>
<td>esp.</td>
<td>especially</td>
</tr>
<tr>
<td>2P</td>
<td>2nd person plural (emi)</td>
<td>EXCL/Excl</td>
<td>exclamation</td>
</tr>
<tr>
<td>2S</td>
<td>2nd person singular (ô, m-)</td>
<td>fn./fnn.</td>
<td>footnote(s)</td>
</tr>
<tr>
<td>2S.HON</td>
<td>2nd person singular respect (ita)</td>
<td>GEN</td>
<td>default genitive marker (-n)2</td>
</tr>
<tr>
<td>3</td>
<td>3rd person, no number (Suai n-)</td>
<td>GEN.P</td>
<td>genitive for plural possessor (-r)</td>
</tr>
<tr>
<td>3P</td>
<td>3rd person plural (sia, r-)</td>
<td>HES</td>
<td>hesitation marker</td>
</tr>
<tr>
<td>3S</td>
<td>3rd person singular (nia, n-)</td>
<td>IMM</td>
<td>imminent (onan, -n)</td>
</tr>
<tr>
<td>ADJ</td>
<td>adjectivaliser (e.g. -k)</td>
<td>IMP</td>
<td>imperative (bá)</td>
</tr>
<tr>
<td>Adj</td>
<td>adjective</td>
<td>Intj</td>
<td>interjection</td>
</tr>
<tr>
<td>AdjP</td>
<td>adjective phrase</td>
<td>INTR</td>
<td>intransitiviser (hak-)</td>
</tr>
<tr>
<td>Adv</td>
<td>adverb</td>
<td>IO</td>
<td>indirect object</td>
</tr>
<tr>
<td>approx.</td>
<td>approximately</td>
<td>IRR</td>
<td>irrealis (atu)</td>
</tr>
<tr>
<td>Attrib</td>
<td>attributive adjective</td>
<td>k.o.</td>
<td>kind of</td>
</tr>
<tr>
<td>Aux</td>
<td>auxiliariy</td>
<td>lit.</td>
<td>literally</td>
</tr>
<tr>
<td>C</td>
<td>consonant</td>
<td>LOC</td>
<td>general locative (iha)</td>
</tr>
<tr>
<td>CLS:</td>
<td>classifier for following category</td>
<td>N</td>
<td>noun; no</td>
</tr>
<tr>
<td>CM:</td>
<td>data from Morris (1984b)</td>
<td>NOM</td>
<td>nominaliser (e.g. Ca-, -n)</td>
</tr>
</tbody>
</table>

---

1 The term ‘Malay’ is used rather than ‘Indonesian’ because Malay already had some currency as a lingua franca in Timor (at least in the ports on the north coast) by the early nineteenth century (Thomaz 1981:54), long before the development of Bahasa Indonesia as a national language. In addition, more recent loans are likely to be from Kupang Malay rather than from Indonesian. All loans which are ultimately from Dutch are presumed at this stage to have come via Malay. Very commonly used loans are not acknowledged as loans in the examples. These include Malay ibu ‘mother’ (the title by which I was addressed) and kalo ‘if’, and oras ‘time’ (from Portuguese horas ‘hours’).

2 In §7.3.3.3 it is noted that -n is often used on location, body part and kin nouns even when these are not possessed. To simplify glosses, -n is glossed as a separate morpheme only when the noun heads a possessive NP.
## Abbreviations

<table>
<thead>
<tr>
<th>Code</th>
<th>Meaning</th>
<th>Code</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>noun phrase</td>
<td>REFL</td>
<td>reflexive (<em>ān</em>)</td>
</tr>
<tr>
<td>Num</td>
<td>numeral</td>
<td>REL</td>
<td>relative clause marker (<em>mak</em>)</td>
</tr>
<tr>
<td>NumP</td>
<td>numeral phrase</td>
<td>RelCl</td>
<td>relative clause</td>
</tr>
<tr>
<td>O</td>
<td>direct object</td>
<td>S</td>
<td>subject; strong syllable</td>
</tr>
<tr>
<td>pers. comm.</td>
<td>personal communication</td>
<td>s.o.</td>
<td>someone</td>
</tr>
<tr>
<td>PL</td>
<td>plural (<em>sia</em>)</td>
<td>s.th.</td>
<td>something</td>
</tr>
<tr>
<td>PN</td>
<td>proper noun</td>
<td>SUF</td>
<td>suffix (<em>-n, -k</em>)</td>
</tr>
<tr>
<td>POS</td>
<td>possessive (<em>-kan,-k</em>)</td>
<td>usu.</td>
<td>usually</td>
</tr>
<tr>
<td>PP</td>
<td>prepositional phrase</td>
<td>V</td>
<td>verb; vowel</td>
</tr>
<tr>
<td>Prep</td>
<td>preposition</td>
<td>Vi</td>
<td>intransitive verb</td>
</tr>
<tr>
<td>Pro</td>
<td>pronoun</td>
<td>Vt</td>
<td>transitive verb</td>
</tr>
<tr>
<td>Psr</td>
<td>possessor</td>
<td>W</td>
<td>weak syllable</td>
</tr>
<tr>
<td>RDP</td>
<td>reduplication</td>
<td>Y</td>
<td>yes</td>
</tr>
<tr>
<td>RECIp</td>
<td>reciprocal (<em>hak- -k</em>)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Conventions for examples and formulae

Examples, as well as tables and figures, are numbered consecutively throughout each chapter.

Conventions for presenting Tetun examples are presented below.

**Tetun convention**  **Meaning**

\(=\) separates two paraphrases
\(/\) clause boundary, where relevant to the discussion
\(*\) the following example is ungrammatical
\(\ldots\) ellipsis
\(()\) surrounds asides and hesitations
\([\ldots]\) surrounds a subordinate clause, where these are under discussion
\(<ma>[ha]nanu<k>\) [] surrounds part of the base that is omitted from the derivation
\( ha-fou\) hyphen marks a morpheme boundary
\( ha-h.abut\) ’h’ is inserted by phonemic rule, and not part of either morpheme
\( hori.hirak\) ‘.’ separates two morphemes which are glossed as a single word
\(<ma>h\dot{o}<k>\) <> surrounds the two parts of a circumfix, and the single English gloss

\(N\ldots\) substitutes for names to protect anonymity

\(\underline{\text{underlining}}\) highlights the part of the example relevant to the discussion
\(\acute{a}\) acute accent marks primary stress in the relevant parts of the phonology chapter; otherwise it marks phonemically double vowels
\(\grave{a}\) grave accent marks secondary stress where stress is under discussion

Spelling follows that of Tetun writers in West Timor (both in published works and in private correspondence), except for two additions. Glottal stops, which are omitted from the standard spelling, are represented by apostrophes. And phonemically double vowels, which in

---

1 This convention is used in §4.4, on morphology. When /h/-initial verbs take subject marking, the initial /h/ is dropped; this /h/ is however not shown as ‘[h]’ in the examples, because the phenomenon is so common and phonologically regular. In any case, the rarity of vowel-initial verbs in Tetun means that no ambiguity results if /h/ is omitted from the representation of the verb.
the standard orthography are not distinguished from single vowels, are indicated by a single vowel with an acute accent (e.g. bá ‘go’ is phonemically /baː/, but written by West Timorese writers as ‘ba’).

Morpheme breaks are usually shown unless doing so would unnecessarily obscure the meaning. Except where the morphology is under discussion, this is judged to be the case if the base of a derivation is unclear to me, or if the semantic relationship between the base and the derivation is not straightforward. In addition, a final /n/ on nouns is only glossed as genitive if the noun is in a possessive construction and the /n/ is not (synchronously) inherent in the word. For further discussion of this issue see §7.3.3.3.

In interlinear glosses, full stops separate words which gloss a single Tetun word (e.g. samodo ‘green.tree.snake’), and square brackets surround the source of borrowings (e.g. [Malay]) or information on register (e.g. [sea taboo]).

In translations, gender, tense and aspect (which are often not inferable from the Tetun) are given as in the original context. Conventions for translations are presented below.

<table>
<thead>
<tr>
<th>Translation convention</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>[]</td>
<td>surrounds the translation (or paraphrase) of parts of the example omitted by ellipsis</td>
</tr>
<tr>
<td>()</td>
<td>surrounds parts of the translation which have no correlate in the glosses (e.g. explanations, or rendering of omitted arguments)</td>
</tr>
</tbody>
</table>

Following the free translation is a reference number, which identifies the text and sentence number for running texts (e.g. R5.23), or the book and page number or date for loose examples (e.g. T0.54, C5/3).

In formulae, the following conventions apply.

<table>
<thead>
<tr>
<th>Convention</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>()</td>
<td>surrounds optional elements</td>
</tr>
<tr>
<td>{}</td>
<td>surrounds alternatives</td>
</tr>
<tr>
<td>X*</td>
<td>one or more instances of ‘X’ are allowed</td>
</tr>
<tr>
<td>[X Y Z]*</td>
<td>one or more sequences of ‘X Y Z’ are allowed</td>
</tr>
<tr>
<td>underlining</td>
<td>highlights the construction head, unless only one element is obligatory</td>
</tr>
</tbody>
</table>
Map 1: Languages of Timor (after Wurm and Hattori 1981: map 40)
1 Introduction

1.1 The Tetun language

1.1.1 Introduction

The present work presents a grammar of the Fehan dialect of Tetun. This language is spoken on the island of Timor in eastern Indonesia (Map 2). The western half of the island is part of the Indonesian province of Nusa Tenggara Timur. The eastern half, which was formerly a colony of Portugal, has since 1976 constituted the Indonesian province of Timor Timur (East Timor). As this grammar goes to press, East Timor is preparing for independence.

Map 2: Location of Timor (from Therik 1995)

Map 1 shows the distribution of languages on Timor. There are about 18 languages indigenous to Timor (Thomaz 1974:293ff.; Wurm & Hattori 1981: Map 40), although counts vary depending on where one draws the boundary between ‘language’ and ‘dialect’. Most of the linguistic diversity is found in the eastern half of the island. Three languages are
non-Austronesian (loosely labelled ‘Papuan’), namely Bunak, Makasai and Fataluku. The remaining languages belong to the large Austronesian language family.

Tetun is assigned by Blust (1993) to the Central Malayo-Polynesian subgroup of Austronesian. According to Wurm and Hattori (1981: Map 40), Tetun is, at a lower level of subgrouping, part of a ‘Timor and islands’ subgroup, which also includes Rotinese, Dawan, Mambai, Kemak, Galoli and Tukudedede, amongst others.

Tetun is spoken as a first language in three discontinuous parts of central and eastern Timor. The anthropologist Fox (1984:vii) estimates the total number of native speakers at 300,000 to 400,000. In addition, Tetun is used as a lingua franca throughout much of East Timor, where over two thirds of the population are said to be conversant with it at some level (Hull 1996b:xiv). Only in the extreme east of East Timor, in the Fataluku-speaking region, is Tetun replaced by Portuguese as a second language. Fox (1984:vii) concludes that in all there are at least 650,000 speakers of the language.

One of the three Tetun-speaking areas is around Soibada, in the south of East Timor. The Tetun from this region was very influential early in the twentieth century because the only secondary college in East Timor was located there. Thus it was the centre for the training of schoolteachers for the whole colony, as well as the location where missionaries wrote down traditional oral literature (Thomaz 1981:67).

The second Tetun-speaking region is a strip crossing central Timor from the north to the south coast. It is situated mostly on the western side of the East Timor–West Timor border, in the regency (kabupaten) of Belu, but covers some neighbouring portions of East Timor also. Troeboes et al. (1987:10) estimate the number of native speakers on the West Timor side of the border to total some 150,000, while Therik’s (1995:18) estimate is 190,000 (based on 90% of the total population of Belu in the 1991 census).

The third area in which Tetun is spoken is in and around Dili, the capital of East Timor, which is situated on the north coast.

Tetun is also known in the literature by the variants ‘Tetum’ or ‘Tettum’ (derived from the Portuguese use of ‘m’ for final nasalisation, and not from any use of bilabial nasals in its pronunciation (Francillon 1967:xxii)), ‘Teto’ (das Dores 1907) and (according to B.P. Grimes (1992:618)) ‘Tetung’. Another label for this language is ‘Belu’ (sometimes ‘Belo’), which is the name of the regency (kabupaten) in West Timor in which Tetun is spoken. The name ‘Tetun’ is used in this grammar because it is widely known, and because it reflects the pronunciation of the name better than the other commonly accepted spelling ‘Tetum’.

1 It is recognised in the literature that the Central Malayo-Polynesian grouping is problematic, and hypothesised that problems in finding shared innovations may reflect history as an extensive dialect network (Blust 1993:263ff.; Pawley & Ross 1993:437). According to Tryon (1995:12), eastern Indonesia “remains perhaps the least known area in the Austronesian world today”, with extremely diverse languages whose interrelationships are not well established.

2 Dawan is also referred to in the literature by the Dutch term ‘Timorese’ (Middelkoop 1950) and as ‘Atoni’ (a name more correctly used for the people rather than for the language). Some Atoni people along the border with the Tetun-speaking area consider ‘Dawan’ an offensive Tetun-centric term, and prefer ‘Timorese’ or variants of ‘Uab Meto’. The former does not acknowledge the many other languages spoken on Timor. Since the latter has not as yet caught on in the literature (with the exception of Fox (1997)), and since ‘Dawan’ is used in linguistic and anthropological works written by Timorese authors (e.g. various mostly unpublished references in Steinhauer (1996), Therik (1995), Troeboes et al. (1987:12)), I will retain the name ‘Dawan’ in this grammar.

3 Various alternative groupings below the level of Central Malayo-Polynesian have been proposed. These include ‘Banda Sea’ and ‘Timor’ (Blust, cited in Tryon (1995:25)), the hierarchy of ‘Timor-Flores’, ‘Timor’ and ‘Central Timor’ (Grimes 1992:618), and the more detailed hierarchy of ‘Timor’, ‘Nuclear Timor’ and ‘East’ (Grimes et al. 1995).
1.1.2 Dialects of Tetun

Tetun is generally recognised as having four dialects, as shown on Map 1.

Within East Timor there are two dialects. East Tetun is spoken on the south coast, while Dili Tetun is spoken in and around the capital of Dili, on the north coast. Dili Tetun (also known as ‘Tetun Prasa’, ‘Tetum Praça’ or ‘Tetun Dili’) has had centuries of exposure to Portuguese, being the language of communication between Timorese people and the Portuguese colonisers. This is reflected in heavy Portuguese influence, particularly in areas of vocabulary and phonology. Dili Tetun lacks some of the syntactic options available in other varieties of Tetun (such as subject marking). This variety has special status, being the basis for most printed material in Tetun, and also for the lingua franca Tetun used throughout East Timor. In this grammar I occasionally use the term ‘East Timorese Tetun’ as a cover term for the dialects of East Timor, since it is not always clear from the literature whether features being written about are relevant to only one of the dialects or to East Timorese Tetun in general.

The other two dialects of Tetun are spoken in the north-south strip which coincides approximately with the regency of Belu. The northern dialect is known to Fehan dialect speakers as Foho ‘mountain’. The major government, church and education centres are situated in the north, and its dialect is the basis for the written Catholic liturgies used in churches in both the north and the south of Belu.

South Tetun, spoken in the north of the north-south strip, in turn has two sub-dialects. The one on which this grammar is based is known to its inhabitants as Fehan ‘plain’, on account of being spoken in the southern agricultural plain. It in turn has two locally-recognised variants. That of ‘Wehalí’ (Therik 1995:37) or ‘Manlea’ (Troeboes et al. 1987:11) is spoken around the town of Betun, towards the northern end of the Fehan dialect area. The ‘Besikama’ (Troeboes et al. 1987:11) or ‘Wewiku’ (Therik 1995:37) variant is spoken in the south around the town of Besikama. The main distinguishing feature of the Besikama variant is the merging of the /d/ and /t/ phonemes.

The second sub-dialect in the south is known as ‘Suai’ or ‘Kamanasa’ (Therik 1995:37; Wortelboer 1955:172). It is spoken to the east of Fehan in the region of Suai, which straddles the West Timor–East Timor border. In addition, there are within the Fehan-speaking area two towns (Kamanasa and Kletek) with sizeable populations of Suai immigrants who moved to the area during unrest at the border in about 1911 (Wortelboer 1955:172). Suai speech is distinguished from that of Fehan mainly by differences in vocabulary, and by distinctive rhythm and intonation.

The verb for ‘speak’ is recognised by Fehan people as being diagnostic of dialect differences. In Fehan it is hakés, in Suai dale, in Foho terik, and in Dili Tetun ko’a lia (lit. ‘cut word’) or kolia.

---

4 This northern dialect is known in the literature as ‘Foho’ (Troeboes et al. 1987:10), ‘Tetun Terik’ (from terik ‘speak’; Morris (1984b:x)) and ‘northern’ or ‘Tasi Feto’ (lit. ‘female sea’) Tetun (Wurm & Hattori 1981: Map 40).

South Tetun is alternatively known in the literature as ‘Tetun Belu’ (Morris 1984b:x) and ‘southern’ or ‘Tasi Mane’ (lit. ‘male sea’) Tetun (Wurm & Hattori 1981: Map 40).

5 Comments about the ‘Suai’ sub-dialect in this grammar are based on the speech of these immigrants. When I spent time in the town of Suai in East Timor just as this book was being prepared for press, it became obvious that the Tetun spoken there is grammatically closer to Dili Tetun than to the Fehan dialect. A good dialect survey is still needed.
1.1.3 Literature in Tetun

The earliest printed texts in Tetun appeared at the end of the nineteenth century (Domingues 1947; Thomaz 1981:60). Since then a sizeable amount of literature has been published, using spelling systems based on Portuguese, Dutch and Indonesian.

In East Timor, written literature in Tetun is largely religious in scope (Hull 1996b:xiv), with a range of translated liturgies and Bible portions being available. In 1981, the Vatican decided to allow the use of Tetun in the mass in East Timor, replacing Portuguese (Carey 1995:2); this has further elevated the Dili dialect of Tetun.

In West Timor, too, various religious texts are available in Tetun; these are all in the northern Foho dialect, but some are used in the south as well. While most publications are small, a large publication is a translation of Bible portions from early this century (Mathijsen 1967).

Apart from religious texts, there are bilingual publications which make Tetun texts available to non-Tetun speakers. Folktales told in the Foho dialect have been published with translations in English (Bartkowiak 1979) and Dutch (Mathijsen 1915), while stories told in East Timorese and Dili Tetun are available with translations in English (Langker 1996; Morris 1984a). Traditional Tetun poems from the region of Belu have been published with translations in Indonesian (Seran 1986) and Dutch (Riedel 1889:7f.), while modern poetry in East Timorese Tetun has been produced with English translation (da Costa 1976). Thomaz (1981:68–79) presents a selection of short traditional and modern East Timorese texts with English translation. Other East Timorese texts which I have not sighted include ones with translations in Portuguese (de Sá 1961) and both English and Portuguese (Sylvan 1988). The Sydney-based Mary McKillop Institute of East Timorese Studies is currently further promoting the publication of texts, particularly in Dili Tetun, both with and without English translation.

1.1.4 Literature about Tetun

A lot of material has been published on Tetun, although none of it is about the Fehan dialect which forms the basis for the present study.

The major grammatical work on Tetun is the 106-page general grammar of the Foho dialect by Troeboes et al. (1987). Kartawinata and Nitiasih (1990) present a brief analysis of the verb phrase.

Tilman (1996), writing in Tetun, discusses aspects of Tetun grammar. He proposes various extensions to Tetun to modernise it, including ordinal numbers (based on a prefixed da-), a passive (making use of ha’ak ‘say’), and extensive use of nominalisation.

Saliwangi et al. (1991) focus on morphology, listing meanings of various affixes, types of reduplication, and compounds.

De Sá (1952) discusses the sound systems of five Timorese languages, including Tetun, and proposes a single transcription system to cover them all. De Araújo e Corte-Real (1990) contrasts the consonants of Dili Tetun with those of English. Another publication comparing Tetun and English, in terms of both phonology and grammar, is the anonymous Asian language notes: some likely areas of difficulty for Asian learners of English (1977).

Some other publications include grammatical sketches, but have as their dominant feature lists of words in various classes, such as pronouns and time words (Martinho 1943:275–303; Soedjianto et al. 1992).

Most of the works on Tetun are, however, dictionaries, with or without short grammatical sketches. These include dictionaries in English (Handbook of Malay and Tetun 1945; Hull
Various pedagogical books are available. Some are designed to teach Dili Tetun, in Portuguese (Fernandes 1937), Indonesian (Mari belajar bahasa Timor Timor 1976; Serantes & Doko 1976c) and English (Hull 1996b). Other booklets are designed to teach Indonesian to Tetun speakers (Presto n.d.; Serantes & Doko 1976b).

Finally, there are a number of works dealing with diverse sociolinguistic questions. The anthropologist David Hicks has published various articles on Tetun oral literature, based on fieldwork in East Timor (1973, 1974, 1988).

Hull (1994) looks at the linguistic situation in East Timor, surveys the literature on and in East Timorese languages, and discusses the character of Tetun, with particular emphasis on the sources of lexical loans in Dili Tetun. Thomaz (1981) discusses the formation of Dili Tetun, with particular attention being given to the sources of new words. Hull (1996a) considers language policy for an independent East Timor, proposing that Tetun be the national language, with both Tetun and Portuguese as official languages. He commends current efforts in spelling reform (in which he is taking a lead) and in developing new vocabulary for expressing present-day realities.

The proceedings of the 1997 Tetun Language Conference in Darwin (Tetun Language Conference 1997) include a paper by Fox (1997) on the historical position of Tetun among the languages of Timor and some papers presenting ideas on how to teach Tetun within Australia. Wait (1994) looks at language maintenance amongst the Timorese community in Darwin, Australia, focusing on Tetun and Portuguese.

1.2 The Fehan dialect

1.2.1 Setting

The Fehan dialect of Tetun, on which this study is based, is spoken on a 300,000 hectare plain near the south coast of central Timor. Fieldwork was conducted in villages around the town of Betun (latitude 9.31°S, longitude 124.59°E). Betun is the capital of the administrative district (kecamatan) of Malaka Tengah, in the south of the regency (kabupaten) of Belu in west Timor. The district of Malaka Tengah has a population of some 23,000 people (Belu dalam angka 1991), of whom the vast majority are Tetun speakers (Troebes et al. 1987:9).

Most of the Tetun-speaking people in Malaka Tengah are subsistence farmers, with traditional crops including maize (corn), mung beans, bananas, sorghum, cassava, coconuts and upland rice (the latter being used mainly for ritual purposes). Sago palms and maek tubers provide alternative food sources, particularly during periods of famine. In addition, those with sufficient resources raise pigs, chickens and/or buffalo. Since World War II some

Works which I have not seen, but which are referred to in the literature, include a Tetun-Dutch dictionary (Tetum-Maleisch woordenlijstje 1894), a trilingual Tetun-Indonesian-Portuguese dictionary (Soares 1985), a grammar book for primary schools (Laranjeira 1916), an anonymous grammar book for high schools (Regras elementares de Térum 1975), and an overview of the sociolinguistic situation in East Timor (Masinambow 1980). Note that Sherlock's (1980) excellent bibliography of Timor includes many references to works on and in Timorese languages.
paddy rice has been grown under irrigation. People live together in villages along the road, travelling out to their gardens, which are usually within easy walking distance.

Tetun people with post-school education who still live in the district are mostly schoolteachers. Many maintain a garden in addition to their teaching.

The culture of the Fehan region has been the focus of two PhD theses in anthropology from The Australian National University (Francillon 1967; Therik 1995). Part of the anthropological interest lies in the fact that the social structure in the Fehan district is matrilineal, while most of the surrounding region is patrilineal. Another source of interest is that this district is the centre of the former kingdom of Wehali, which is believed to have had influence over a large part of Timor prior to the destruction of its power by a Portuguese-authorised military force in 1642 (Nordholt 1971:159, 165f.).\(^7\) All Tetun people still have hereditary rank as either noble or commoner, a distinction which has bearing in traditional contexts but not in modern ones.

The town of Betun has a single bitumen road lined by shops, almost all of which are owned by non-Tetun people. Many shop owners are Chinese residents of long standing, while government posts (such as those in the army, police, post office and hospital) tend to be filled by more recent settlers from other parts of Indonesia.

Until recently Betun was isolated from the rest of Timor, a fact commented on by Tetun friends as well as by the anthropologist Francillon (1967:18). This isolation was fostered (amongst other things) by the mountains which run from east to west along the length of the island. However, in the last decade electricity supply was introduced to Betun and is spreading from there to those who can afford it in the surrounding villages. In its wake came television in 1993. Transport too has become much easier, and by 1995 there was a bus service several times per day to the regional capital of Atambua to the north as well as a daily service to the provincial capital of Kupang.

### 1.2.2 Languages spoken in Betun

In Betun and surrounds all indigenous people speak Tetun, as do the Chinese shopkeepers, some of whom have lived there for several generations. Most younger people and many older people also speak Indonesian and/or the Malay dialect used on Timor, at least to some level.

Indonesian, as the national language, is the language of schools, radio and television. It is also increasingly the language used in church (both Catholic and Protestant), although some church services are still conducted in Tetun, using material written in the northern Foho dialect. Indonesian is the dominant language of books and newspapers; nevertheless, in practice few people read and very little written material is available unless one travels to larger towns to acquire it. Finally, Indonesian is often heard in the town of Betun, where there are recent settlers from other parts of Indonesia who do not speak Tetun.

In the villages around Betun, however, one rarely hears Indonesian, with Tetun being used for all home and community functions other than those mentioned above.

A minority of Tetun speakers also know Dawan (also known as ‘Atoni’ or ‘Timorese’), the Austronesian language spoken to the west. Most of these are of Dawan origin (such as the spouses of several senior nobles) or originate from bilingual villages on the border between the two languages.

\(^7\) Other anthropological publications about Tetun-speaking people include Fox (1982a) on the kingdom of Wehali, several on the regency of Belu (Grijzen 1904; Manehat & Neonbasu 1990; Vroklage 1952; Wortelboer 1955), and works by Hicks (1976, 1984) on southern East Timor.
Within a few kilometres of Betun are several hamlets of native speakers of Bunak, a non-Austronesian language. They came from East Timor after 1910 (Wortelboer 1955:172).

Within Betun itself there is a significant amount of intermingling of people from the two sub-dialects of south Fehan, since people from Suai have come to live in or near the town. Nevertheless, the amount of intermingling in the villages around Betun does not seem to be great. On my first field trip I ended up recording texts by speakers from both sub-dialects, resulting in considerable variation in the data. This was reduced when on the next field trip I had less exposure to Suai speakers.

1.2.3 Language registers

1.2.3.1 Introduction

The kind of language used in everyday speech and in writing, with its varying levels of formality, forms the basis of this grammar. Nevertheless, there are three other language registers to which I would like to briefly draw attention. These are poetry, the noble register and the sea taboo register.

1.2.3.2 Poetry

For traditional speakers of Tetun, poetry (kananuk) is one of the delights of life. Poetry used to be the medium of courtship, which brings light to the eyes of those reminiscing about it. Many older speakers can (and love to) recite poems associated with a large range of ordinary items and events; good practitioners are able to recombine old motifs into new patterns to suit the occasion. Poetry is also used for the public welcoming of important guests, the recitation of origin myths by the mako'an (the men who have the right to formally tell these myths), and prayers to the spirits. Many folktales incorporate some poetry, either spoken or sung.

The essence of kananuk poetry is pairs of parallel lines, in which the second line is an echo of the first, except for the replacement of one word or expression (or sometimes more) by a related one. The pairs come from a largely standard repertoire of hundreds of parallelisms. Such poetry is common throughout eastern Indonesia, and is described for Rotinese by Fox (1971, 1974, 1975, 1982b), and for a range of eastern Indonesian languages in a volume edited by Fox (1988). A large selection of short Tetun poems from the Belu regency (including many in the Fehan dialect) has been collected by Seran (1986), with translations in Indonesian.

1.2.3.3 Noble register

A noble language register (lia na'in lit. 'word/language noble' = 'noble language') is used on formal and ritual occasions, including the formal telling of origin myths, preaching of sermons and discourses in formal meetings. It is also appropriately used in daily conversations with people of noble rank (Therik 1995:40). It appears that only nobles, and those commoners who regularly interact with nobles (including all ritual specialists), are fully conversant with this register.

The essence of the noble register is special vocabulary, which shows respect to the other while humbling oneself. Much of this vocabulary involves reinterpretation of terms in use in daily speech. This is demonstrated by the examples below, which are presented with their meanings in the noble register as well as their meanings in the common register. Parallelism is
found in this register too, providing motivation for the many synonymous pairs (e.g. *akar*, *maek* ‘rice (of commoner)').

<table>
<thead>
<tr>
<th>Noble register term</th>
<th>Noble register meaning</th>
<th>Common (non-noble) meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>akar</em></td>
<td>rice (of commoner)</td>
<td>sago</td>
</tr>
<tr>
<td><em>maek</em></td>
<td>rice (of commoner)</td>
<td>k.o. edible root of low value</td>
</tr>
<tr>
<td><em>baria tahan</em></td>
<td>message (from commoner)</td>
<td>k.o. bitter plant leaf</td>
</tr>
<tr>
<td><em>ha' u ata</em></td>
<td>I (of commoner)</td>
<td>I (of servant/slave)</td>
</tr>
<tr>
<td><em>hanatar</em></td>
<td>rest, sit (of commoner)</td>
<td>rest (of animals, esp. cattle)</td>
</tr>
<tr>
<td><em>hoku</em></td>
<td>sit, lie down (of commoner)</td>
<td>lie on stomach (of animals)</td>
</tr>
<tr>
<td><em>neras</em></td>
<td>clothes (for commoner)</td>
<td>swaddling cloth for baby</td>
</tr>
<tr>
<td><em>simu seka</em></td>
<td>eat or drink (by commoner without noble)</td>
<td>receive leftovers</td>
</tr>
<tr>
<td><em>tane lamak bá X</em></td>
<td>eat or drink with noble X</td>
<td>hold up festival meat for X</td>
</tr>
<tr>
<td><em>fafudi</em></td>
<td>converse with (noble)</td>
<td>alternate (head to foot)</td>
</tr>
<tr>
<td><em>haksalak</em></td>
<td>chew betel (of noble)</td>
<td>stretch out waist</td>
</tr>
<tr>
<td><em>halamak</em></td>
<td>eat (of noble)</td>
<td>lighten (a load)</td>
</tr>
<tr>
<td><em>hasalak</em></td>
<td>eat (of noble)</td>
<td>make ascend oneself</td>
</tr>
<tr>
<td><em>halolo knotak</em></td>
<td>lie down (of noble)</td>
<td>sea (is) dry</td>
</tr>
<tr>
<td><em>hamán</em></td>
<td>accompany (noble)</td>
<td>make heavy oneself</td>
</tr>
<tr>
<td><em>hamó</em></td>
<td>offer betel (to noble)</td>
<td>body friend-GEN</td>
</tr>
<tr>
<td><em>ha-sa'e án</em></td>
<td>die (of noble)</td>
<td>thigh</td>
</tr>
<tr>
<td><em>tasi mara, meti mara</em></td>
<td>die (of noble)</td>
<td>betel container</td>
</tr>
<tr>
<td><em>ha-todan án</em></td>
<td>sit down (of noble)</td>
<td>betel container</td>
</tr>
<tr>
<td><em>isi belu-n</em></td>
<td>clothes (of noble)</td>
<td></td>
</tr>
<tr>
<td><em>kaka'ut</em></td>
<td>betel container (of noble)</td>
<td></td>
</tr>
<tr>
<td><em>kelen</em></td>
<td>cattle and horses (of noble)</td>
<td></td>
</tr>
<tr>
<td><em>lawalu</em></td>
<td>(go to the) toilet (of noble)</td>
<td></td>
</tr>
<tr>
<td><em>mam-fatin</em></td>
<td>order/words (from noble)</td>
<td></td>
</tr>
</tbody>
</table>

### 1.2.3.4 Sea taboo register

A special language register (*lia tasi* ‘word sea’ = ‘sea language’) is required to be used on fishing trips, which are mainly conducted in the mangrove swamps by groups of people. Fehan people gave various reasons for needing to use this sea taboo register, which all amounted to fear that something bad would happen if they didn’t. These included fears of getting lost in the mangrove swamps, being speared by mangrove spikes, being taken by a crocodile or by the sea, and not catching any fish. The sea taboo register is accompanied by a general preference for speaking as little as possible, and a prohibition on idle chatter. Men, women and children are all required to respect the taboo.
Table 1.2: Sample of sea taboo vocabulary

<table>
<thead>
<tr>
<th>Sea taboo term</th>
<th>Sea taboo meaning</th>
<th>Common (non-taboo) meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ai maran</td>
<td>firewood</td>
<td>dry wood</td>
</tr>
<tr>
<td>ai maran abut</td>
<td>cassava</td>
<td>dry wood root</td>
</tr>
<tr>
<td>asu</td>
<td>fish trap, fish net</td>
<td>dog</td>
</tr>
<tr>
<td>asu bukae</td>
<td>food brought from home</td>
<td>dog food for a trip</td>
</tr>
<tr>
<td>badut</td>
<td>sun, moon</td>
<td>wick lamp</td>
</tr>
<tr>
<td>biku tahan</td>
<td>prawn, shrimp</td>
<td>k.o. tree leaf</td>
</tr>
<tr>
<td>hadi’a kbonan sia</td>
<td>prepare to go home (pack up...)</td>
<td>prepare front of sarong PL</td>
</tr>
<tr>
<td>haknán</td>
<td>chew betel, pound betel</td>
<td>chew</td>
</tr>
<tr>
<td>hakraik oi</td>
<td>go down into the seawater</td>
<td>lower (one’s) face</td>
</tr>
<tr>
<td>haraik án</td>
<td>return home</td>
<td>lower oneself, humble oneself</td>
</tr>
<tr>
<td>hibak</td>
<td>eat, drink</td>
<td>move (s.th.) out of the way</td>
</tr>
<tr>
<td>hoku</td>
<td>sit, lie down</td>
<td>lie on stomach (of animals)</td>
</tr>
<tr>
<td>kabár</td>
<td>betel nut</td>
<td>astringent</td>
</tr>
<tr>
<td>kabelak</td>
<td>buffalo</td>
<td>flat</td>
</tr>
<tr>
<td>kabuar</td>
<td>horse</td>
<td>circle, round</td>
</tr>
<tr>
<td>katar</td>
<td>tobacco</td>
<td>itchy</td>
</tr>
<tr>
<td>kbonan</td>
<td>woven container to put fish in</td>
<td>front of sarong tied at the waist</td>
</tr>
<tr>
<td>kbonan talín sit/kotu</td>
<td>many fish were caught</td>
<td>to make a ‘pocket’</td>
</tr>
<tr>
<td>kesak na’in</td>
<td>crocodile</td>
<td>fish-trap lord/owner</td>
</tr>
<tr>
<td>knaban oan</td>
<td>small children</td>
<td>basket for liquids child/small</td>
</tr>
<tr>
<td>knase rahun</td>
<td>fish, eels, sharks</td>
<td>mugil. fish dust</td>
</tr>
<tr>
<td>kuku warak</td>
<td>crab</td>
<td>many (East Timor) pincers</td>
</tr>
<tr>
<td>mahokuk</td>
<td>fish trap</td>
<td>who lies on stomach (like animal)</td>
</tr>
<tr>
<td>manas</td>
<td>drinking water</td>
<td>hot</td>
</tr>
<tr>
<td>mata meak</td>
<td>fire</td>
<td>red eye</td>
</tr>
<tr>
<td>metan</td>
<td>woman</td>
<td>black, dark (colour)</td>
</tr>
<tr>
<td>metan</td>
<td>there are no fish</td>
<td>black, dark (colour)</td>
</tr>
<tr>
<td>neras</td>
<td>cloth</td>
<td>swaddling cloth for baby</td>
</tr>
<tr>
<td>nia kmeik</td>
<td>mosquito</td>
<td>it (is) pointed</td>
</tr>
<tr>
<td>silu ai maran</td>
<td>go to the toilet</td>
<td>snap dry wood</td>
</tr>
<tr>
<td>suan</td>
<td>leg</td>
<td>wooden digging stick</td>
</tr>
<tr>
<td>tilun</td>
<td>sea</td>
<td>ear</td>
</tr>
<tr>
<td>usi mane</td>
<td>man; crocodile</td>
<td>noble man</td>
</tr>
</tbody>
</table>

The sea taboo register is distinguished solely by the use of special vocabulary (of which I elicited about 60 terms), with phonology, morphology and syntax unaffected. Such vocabulary includes nominal terms for basic foods and drinks consumed on a fishing trip, betel nut and its accessories, fishing implements, certain animals (including insects), people, and terms associated with home, sea and time of day. In addition, there are a few verbs for activities which are carried out on fishing trips, such as eating, chewing betel, cooking and going fishing. There are no special function words.

All terms appear to be based on words in daily use, with a semantic relationship between the taboo and the everyday meanings often being transparent. Taboo words tend to be more generic than everyday ones.

These points are demonstrated by the examples above. There is some overlap between sea taboo and noble registers, in terms such as hoku ‘sit’; this is not so surprising since a shared
motivation is humbling of oneself. Note that some terms are adjectives or intransitive verbs in everyday use, but nouns in the sea taboo register (e.g. kabâr, kabuar, metan).

For more information on Fehan sea taboos, including additional examples of taboo vocabulary, see Therik (1995:40–43).

1.3 This study

1.3.1 Introduction

This grammar of a Timorese language supplements earlier extensive grammars of the neighbouring Timorese language Dawan (e.g. Middelkoop 1950) and of the closely related language of the nearby island of Roti (e.g. Jonker 1915), and shorter grammars of Galoli and Helong (da Silva 1900) and Maryanto (1977), both referred to by Hull (1998). The remaining Austronesian languages of the island of Timor as yet have no published descriptions.

The study ranges from phonology through morphology to the syntax of phrases, clauses and sentences. In a wide-ranging descriptive grammar of this type, it is inevitable that some statements will be based on more data, and on a deeper understanding, than others. In order to give some indication of the amount of data supporting a claim, the number of examples for the construction or word under discussion is often given. These numbers should not, however, be interpreted as giving an accurate guide to relative frequency in texts (particularly where the number is small), because the counts include elicited examples as well as examples which were written down 'on the run' because they seemed interesting for some reason.

As a reference grammar, this work follows the general principles of modern linguistics, but avoids commitment to any particular formal model. Instead it is eclectic, drawing on the work of many different theoreticians when their work is useful for the analysis and appropriate for the presentation.

1.3.2 Data

The primary target of analysis is spoken language, although some texts written by language consultants were used as data also.

Fieldwork for this study was conducted during two four-month visits to Timor (July–October 1993; May–August 1995). Before the initial trip I received a few private lessons in East Tetun from Antonio Sarmento in Brisbane, learned vocabulary from Morris's (1992) dictionary of East Timorese Tetun, began initial analysis based on the bilingual folktales published in Morris (1984a), and followed lessons in Dili Tetun from a prepuplication version of Hull (1993). Although there proved to be significant differences between these dialects of Tetun and the Fehan dialect, the pre-fieldwork exposure to Tetun was extremely useful.

---

8 Simons (1982) surveyed taboo registers in Austronesian languages. He found 'hunting' taboos (in which he includes hunting, fishing, harvesting and mining) in 50% of the 30 Western Austronesian languages surveyed, as well as in 20% of the 45 Oceanic languages. Van Engelenhoven (1995:20) mentions a sea taboo register on the island of Leti, near Timor. A sea taboo register for Sangir, spoken on the islands between Sulawesi and Mindanao, is described at some length by Grimes and Maryott (1994:286–300).

9 Levels beyond that of the sentence are not dealt with in this study, since these are beyond the scope of syntax (at least as it is traditionally defined), and are in any case worthy of separate study.
During the initial weeks of fieldwork I accompanied Tom Therik, who at the time was completing fieldwork for his PhD in anthropology (Therik 1995). He provided valuable introductions to people as well as to the language and culture, and in addition acted as interpreter. An interesting side effect of this arrangement was that the first text I was exposed to was an origin myth which proved difficult to understand even after translation. After his departure, almost all fieldwork was conducted monolingually in Tetun. This gave admirable incentive for language learning, while meaning that more detailed questioning about the language had to wait until greater proficiency was achieved.

Much of my language learning was by the ‘absorption’ method, with many people eager to talk with me about a wide range of topics, and willing to put up with the strains of initially poor communication. Some were willing to record, and these recordings form the basis of the observations in this grammar. Recordings totalled approximately 23 hours. Those texts which were fully transcribed and entered into the computer totalled over 64,000 words and are listed in Appendix B. Other recordings which were deemed of lesser value were used as backup material only. Reasons for not using these as primary data included poor quality of recording, interference from Malay or from other Tetun dialects, and content which was boring or unreliable (making it hard for consultants to persevere with working on the text). Recorded texts were supplemented by notes made whenever I overheard interesting constructions or uses of words (a habit of which people were amazingly tolerant), as well as by direct questioning concerning the language.

Inherent limitations of textual data include errors and false starts, the impossibility of proving what is not allowed by the grammar, and the scarcity of data on infrequent constructions (see Newmeyer 1983:60ff.). Transcribed texts thus needed to be checked with native speakers. In addition, direct elicitation was used to supplement data on infrequent constructions, and to provide information as to what was not possible.

Direct elicitation proved rather difficult, particularly during the first field trip. Everyone was keen to chat, and many were willing to explain words and traditions, with some proving excellent at making up example sentences to illustrate how a word could be used. Relatively few, however, were willing to answer formal questions about syntactic possibilities. Those that proved best at such discussions tended to be educated men (schoolteachers and a man who had trained for the priesthood) who were in the 30–45 age bracket; some women in this age bracket also contributed significantly.

Problems with formal elicitation were exacerbated by the fact that in general both consultants and I soon grew bored of questions such as “Can you say this?” and “What happens if you swap these words around?”. One frustrated response to an attempt to check which elements in a sentence could be questioned was “If you want questions, let me give you questions! I’ll make up a story about a court case”. Inevitably the story didn’t answer my questions, but brought out other useful data instead. As is to be expected, some attempts at elicitation brought out contradictory results (shown particularly in subject marking for adjectives), or results which seem dubious in the light of what occurs in texts (for instance, in rejecting a construction which was later found to occur quite frequently). It was not always clear whether contradictions reflected genuine inter-speaker differences, marginal grammaticality, or distinctions between grammaticality and extra-grammatical acceptability (a problem even for trained linguists working on their native languages; see Newmeyer (1983:51ff.)). Elicited data are thus treated with some caution in this grammar, unless confirmed by data from texts; in fact any examples in this grammar which were elicited specifically for the issue at hand are explicitly labelled as ‘elicited’.
Although data collection for a dictionary was never part of my aim, the dictionary file resulting from this research contains over 3,400 headwords. The amount of data on each word varies enormously, from a fairly thorough understanding of its range of meanings and syntactic possibilities to a single partially-understood example from a poem.

Two computer programs from the Summer Institute of Linguistics proved particularly useful in collating the data. Shoebox provided tools for filing dictionary material, as well as for interlinearising texts, while FIESTA allowed concordance searches on the texts. These searches enabled all examples of a word or phrase to be retrieved and copied into the word processor for analysis and sorting. Searches could similarly be done by construction (e.g. left-dislocation, topicalisation) so long as these were manually tagged during the interlinearising process.

1.4 Overview of Tetun grammar

This grammar is presented basically in order from smallest to largest units within the grammatical hierarchy. It thus begins with phonology and morphology, and then proceeds to phrase, clause and sentence level syntax. To orientate the reader, the current section offers a brief overview of Tetun grammar.

Tetun has five vowels, /i, e, a, o, u/, with /a/ accounting for over 40% of all vowel occurrences. There are thirteen consonants: /b, t, d, k, ’, f, s, h, l, r, m, n, w/. Phonological words consist of from two to four syllables, with two syllables being the most common. Lexical stress falls on the penultimate syllable. Although there are a few minimal pairs of lexemes distinguished by stress and length (e.g. hare ‘rice’ and haré ‘see’), these are analysed as differing also in their number of syllables, resulting in an exceptionless penultimate stress rule. The basic syllable template is (C)V(C). Except for /kl/, which can function as coda in the initial syllable of trisyllabic words (e.g. haklati ‘topple’), codas occur only root-finally, and are restricted to /t, k, s, r, n/.

An unusual feature of Tetun phonology is the fact that (except in compounds and contractions) the only consonant clusters are those beginning with /kl/. The second element in the cluster can be any consonant other than the glottals (/ʔ, hi/) and /k/. These clusters are found only word-initially (where the /kl/ is analysed as extrasyllabic) and at the interface between a prefix and a root.

There are four reasonably productive derivational prefixes or prefix-suffix combinations, and no productive suffixes. The prefix har- derives causative verbs from verbal bases as well as verbs from nominal bases; hak- derives intransitive verbs, predominantly from transitive verbs, while mak- -n (or ma- -k) derives actor-describing verbs from other verbs. Result adjectives can be derived from transitive verbs by k- -k. In addition, partial reduplication (of the onset followed by /a/) derives nouns from verbs. Full reduplication has a range of functions.

Tetun has little morphological marking of constructions, with the result that there is often little or no morphological evidence for distinguishing between alternative analyses of a construction. Within possessive NPs there is (under certain circumstances) both head-marking (by genitive -n or -r on the possessed noun) and dependent-marking (by possessive -k(an) on a possessor pronoun).

The only other inflection is subject marking on verbs, which takes the form of an initial consonantal prefix. For /h/-initial verbs, this prefix replaces the initial /h/. The paradigm is 1S k-, 2S m-, 3S n-, and 3P r-, with 1P and 2P retaining initial /h/. All other consonant-initial verbs take only the 1S prefix k-. Most speakers apply subject marking quite consistently to
/h/-initial verbs; however, other consonant-initial verbs are inflected by only some speakers in some circumstances, and even then inconsistently.

Most phrase-level syntax follows the order of head-modifier. Thus, within the NP, modifying nouns, adjectives, numerals, prepositional phrases and determiners follow the NP head (e.g. *uma rua ne’e* ‘house two this’ = ‘these two houses’). Most relative clauses follow the head, but a syntactically highly restricted type of relative clause can precede it. In contrast most possessors precede the head (e.g. *ha’u-kan ina-n* ‘1S-POS mother-GEN’ = ‘my mother’); however, for possessive relations other than part-whole the possessor can optionally be presented as a postmodifier (e.g. *ina ha’u-k* ‘mother 1S-POS’ = ‘my mother’). In addition to premodifying relative clauses and possessors, there are a few single-lexeme modifiers which can precede the NP head.

Tetun has a small set of sortal classifiers, used to count whole objects; these precede the numeral (e.g. *ema na’in rua* ‘person CLS:human two’ = ‘two people’).

Clauses follow strict SV constituent order, with the unmarked order being SVO. There is no system of voice.\(^{10}\) Zero anaphora is common for both subjects and objects; its interpretation is primarily pragmatic, with very few syntactic constraints. As expected in SVO languages, prepositions precede their complements and auxiliaries precede the verb.

Verb serialisation is productive in Tetun. In nuclear serialisation, two consecutive verbs make up a single clause nucleus (e.g. *buti be’o* ‘squeeze shatter’ = ‘squeeze to pieces’). Three other types of verb serialisation are particularly common. One involves two consecutive transitive verbs sharing subject and object (e.g. *há hola ha’u* ‘eat take IS’ = ‘eat me up’). The second consists of an intransitive motion verb followed by an action verb, with the two verbs specifying consecutive sub-events (e.g. *bá haré* ‘go see’). In the third, a transitive verb introduces an argument which has oblique status for the other verb, such as source, goal, instrument or co-actor. Transitive verbs fulfilling this function are crosslinguistically subject to reanalysis as prepositions; in Tetun some have mixed verbal and prepositional characteristics and are classified as prepositional verbs.

Many clauses are linked by means of intonation, repetition of words or ideas and/or connective adverbs, while some are joined by conjunctions. Complement clauses are of two types, namely sentential and reduced clauses. Relative clauses are usually introduced by an invariant subordinator *mak*. They make use of the ‘gap’ strategy for relativisation on core arguments, and use a resumptive pronoun when relativisation is on the object of a preposition.

---

\(^{10}\) The closest approximation to passivisation is the derivation of intransitive undergoer-subject verbs from some transitive verb bases (§4.3.2).
Phonology and morphophonemics

2.1 Introduction

This chapter deals with phonology up to the level of the word, including the phonological rules associated with morphological derivations. Phonology above the level of the word, and particularly intonation and its relationship with grammar, is left for another study.¹

The chapter begins with an overview of syllable and word structure. This is followed by discussion of phonemes and allophones, vowel sequences, consonant clusters, phoneme frequencies and variant pronunciations. The final section deals with morphophonemic rules.

The basic conclusions are as follows: Tetun has five vowel phonemes (/i, e, a, o, u/), and thirteen consonant phonemes (/b, t, d, k, ʃ, s, h, l, r, m, n, w/). Lexical words have from two to four syllables. Primary stress falls on the penultimate syllable, and secondary stress on the fourth-last syllable. Phonetically long vowels are analysed as geminates. Consonant clusters are found only in restricted environments, and have peculiarities which are discussed at some length. The metrical foot is found to be the domain for a number of rules, namely stress, vowel allophones, co-occurrence restrictions, language games, prefixation and reduplication.

2.2 Syllable template

The syllable template is (C)V(C). Each vowel is the head of its own syllable, and long vowels are interpreted as geminates (§2.3.4.3). Examples of each syllable type are given below, where ‘.’ represents a syllable boundary.

<table>
<thead>
<tr>
<th>Template</th>
<th>Example</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>o.e</td>
<td>N cane</td>
</tr>
<tr>
<td>CV</td>
<td>to.lu</td>
<td>Num three</td>
</tr>
<tr>
<td>VC</td>
<td>o.in</td>
<td>N face</td>
</tr>
<tr>
<td>CVC</td>
<td>ka.tak</td>
<td>V tell, speak</td>
</tr>
</tbody>
</table>

¹ The relationship between higher-level phonology and grammar is one of cooperation rather than interdependence, as has been pointed out for English both for intonation (Bolinger 1982; Pike & Pike 1982; Woodall 1984) and sentence stress (Bolinger 1972). Higher-level phonology is thus a field worthy of study in its own right.
2.3 Structure of the phonological word

2.3.1 Introduction

The following sections focus on the structure of the phonological word. In particular, they deal with the CV and foot patterns within the word, restrictions on initial and final consonants, lexical stress, and the distribution and analysis of phonomically long vowels. Finally, word games are considered, since they cast light on the phonological analysis of words, and particularly on the question as to how phonomically long vowels should be analysed.

Note that the phonological word corresponds to the grammatical word form in Tetun, except in the case of phonological words which consist of a grammatical word plus surrounding clitics.

2.3.2 Word templates

2.3.2.1 Introduction

Phonological words in Tetun have from two to four syllables. In a corpus-derived list of 3,012 unique lexemes (excluding homonyms, full reduplications, compounds and exclamations), 55% of entries had two syllables, 43% three syllables and 2% four syllables. 96% were consonant-initial and 4% vowel-initial.

2.3.2.2 Disyllabic word template

The template for the minimal prosodic word is shown below, where ‘S’ represents a strong (stressed) syllable, and ‘W’ a weak (unstressed) one.\(^2\)

\[
\text{Minimal prosodic word}
\]

\[
\begin{align*}
\text{Foot} & \quad S & \quad W \\
((k) & \quad C)V & \quad (C)V(C)
\end{align*}
\]

The minimal prosodic word in Tetun thus consists of a single disyllabic metrical foot, with an optional /k/ before the initial consonant. Word-initial consonant clusters resulting from this initial /k/ are discussed in §2.7, where it is argued that the /k/, although part of the word, is not underlyingly part of the adjoining syllable. The first syllable is necessarily open, while the second can be closed.

\(^2\) Note that word templates for Dili Tetun are quite different, with many consonant clusters having been introduced through Portuguese (de Araújo e Corte-Real 1990:65).
Table 2.2: Examples of word templates: two syllables

<table>
<thead>
<tr>
<th>Template</th>
<th>Vowel-final example</th>
<th>Consonant-final example</th>
</tr>
</thead>
<tbody>
<tr>
<td>V V(C)</td>
<td>ai N plant, wood</td>
<td>ain N leg, foot</td>
</tr>
<tr>
<td>V CV(C)</td>
<td>iita Pro 1P</td>
<td>inur N nose</td>
</tr>
<tr>
<td>CV V(C)</td>
<td>sta Pro 3P</td>
<td>hean Vt pull</td>
</tr>
<tr>
<td>CV CV(C)</td>
<td>simu Vt receive</td>
<td>metan Adj black</td>
</tr>
<tr>
<td>kCV V(C)</td>
<td>kfuji N flute, whistle</td>
<td>ktaek N respect</td>
</tr>
<tr>
<td>kCV CV(C)</td>
<td>kneni N large mat</td>
<td>knosen N rib</td>
</tr>
</tbody>
</table>

Note that the metrical foot is the domain for a range of phonological rules, dealing with lexical stress (§2.3.3), vowel allophones (§2.4.2), co-occurrence restrictions (§2.5.3.3, §2.6), and language games (§2.3.5). In addition, the base for prefixation and reduplication always consists of a single foot (§2.10.6).

2.3.2.3 Underived four-syllable word template

The maximal prosodic word consists of two metrical feet. Underived words of four syllables fit the template below.

Maximal prosodic word (underived)

```
  Weak foot                     Strong foot
   S   W                       S   W
(VC)(VC)                      (VC)(VC)
```

Note that vowel sequences (including geminate vowels) occur only within a foot, a coda is allowed only word-finally, and there is no word-initial consonant cluster.

The various possibilities for words of four syllables are exemplified below. Omissions are assumed to be accidents of the data, resulting from the low frequencies of vowel-initial words and of words with four syllables. Note that the symbol ‘ö’ in kaibök represents the sequence /ool/ (§2.3.4).

Table 2.3: Examples of word templates: four syllables

<table>
<thead>
<tr>
<th>Template</th>
<th>Vowel-final example</th>
<th>Consonant-final example</th>
</tr>
</thead>
<tbody>
<tr>
<td>V V CV V(C)</td>
<td>akitou N dove</td>
<td>kaibök N leaf vegetable</td>
</tr>
<tr>
<td>V V CV CV(C)</td>
<td>liurai N executive noble</td>
<td>tualekik N wake songs</td>
</tr>
<tr>
<td>V CV CV V(C)</td>
<td>maufinu N danger</td>
<td>labadain N spider</td>
</tr>
<tr>
<td>V CV CV CV(C)</td>
<td>banokae N k.o. sea shell</td>
<td>sibalebok N parsley</td>
</tr>
<tr>
<td>CV CV CV V(C)</td>
<td>bibiliku N drum</td>
<td></td>
</tr>
</tbody>
</table>
Compounds and reduplications also consist of two metrical feet. However, they are slightly different, in that they have some characteristics of a single word, and others of two words, and in that they allow for a range of word-medial consonant clusters. Their template is given in §2.10.5.1.

2.3.2.4 Trisyllabic word template

The template for trisyllabic words is presented below. Trisyllabic words resulting from contraction of words of four syllables have a different template, and are discussed in §2.10.5.2 and §2.10.7.

\[ \text{Trisyllabic prosodic word} \]

\[
\begin{array}{c}
W \\
(\text{CV}) \\
(\text{Cak}) \\
(\text{CV})
\end{array} \quad \begin{array}{c}
\text{Foot} \\
S \\
W
\end{array}
\]

The examples below illustrate the templatic possibilities. Omissions are assumed to be accidents of the data, due to the rarity of trisyllabic words with initial consonant clusters.

<table>
<thead>
<tr>
<th>Template</th>
<th>Vowel-final example</th>
<th>Consonant-final example</th>
</tr>
</thead>
<tbody>
<tr>
<td>V CV V(C)</td>
<td>inou</td>
<td>EXCL</td>
</tr>
<tr>
<td>V CV CV(C)</td>
<td>aruma</td>
<td>Adv</td>
</tr>
<tr>
<td>CV CV V(C)</td>
<td>bukae</td>
<td>N</td>
</tr>
<tr>
<td>CV CV CV(C)</td>
<td>batane</td>
<td>N</td>
</tr>
<tr>
<td>Cak CV CV(C)</td>
<td>hakmeo</td>
<td>V</td>
</tr>
<tr>
<td>Cak CV CV(C)</td>
<td>hakfalu</td>
<td>Vt</td>
</tr>
<tr>
<td>kCV CV V(C)</td>
<td>klalear</td>
<td>N</td>
</tr>
<tr>
<td>kCV CV CV(C)</td>
<td>klalear</td>
<td>N</td>
</tr>
</tbody>
</table>

The initial syllable in trisyllabic words is unique in a number of ways. It can have a coda, which is restricted to /kl/. In fact the only closed initial syllables are /hak/ and /mak/, which both can (but synchronically need not) come about as the result of prefixation (by hak- (§4.3), mak- (§4.4) and ha- (§4.2)). The only other closed syllables in Tetun are found word-finally, or word-initially within words of two metrical feet when the first foot is truncated (e.g. am-bei ‘old man’ from ama ‘father’ and bei ‘ancestor’, §2.10.5.2).

Alternatively the initial syllable can have an initial consonant cluster, which comes about either as the result of partial reduplication of a cluster-initial base (§2.10.4), or of prefixing the IS subject marker k- to a trisyllabic verb.

The vowel is always /a/ in prefixes and proclitics, and nearly always /a/ in the initial syllable of underived trisyllabic words. Some such words have variants with other vowels, but all verified non-compounded trisyllabic words in the corpus have at least a variant with /a/ (e.g. bukae/bakae ‘food taken on a journey’, kowά/kawά ‘crow’). This strong preference for /a/ is reflected in the assimilation of loans, with variants such as tobako/tabako ‘tobacco’,
and *polisimalisi* ‘police’. Only in this pre-tonic position does the vowel tend to be centralised to schwa (§2.4.2).³

In poetry and word games the initial syllable is, in a significant minority of occasions, interpreted as a metrical foot. In poetry this means that phonetically it receives both stress and length, while in games it means that it is interpreted as if it had two consecutive identical vowels (§2.3.5). Apart from in these contexts, there is no evidence that this weak syllable should be interpreted as belonging to a foot, because it carries no stress.⁴ For this reason I have tentatively connected the syllable directly to the word level, without the intermediary of a foot.⁵

2.3.2.5 Restrictions on initial and final consonants

In words of any size all consonants can occur foot-initially except the glottal stop (/ʔ/), which occurs only foot-medially.⁶ Coda consonants are restricted to /kl/ word-medially and to /t, k, s, r, n/ word-finally.⁷ In loans from Malay there is a tendency to omit the final glottal consonants /h/ (e.g. from *sekolah* ‘school’) and /l/ (e.g. from *pá’i* ‘mister’), to retain final /l/ and /ml/ (e.g. in *kól* ‘cabbage’, *baskom* ‘large bowl’), and to replace final /ŋ/ with /n/ (e.g. for *senán* ‘(be) content’).

2.3.2.6 Prefix and clitic templates

Prefixes have the form Ca- (e.g. causative *ha*), Cak- (e.g. *mak* ‘who’) or k (adjectival *k*). Proclitics have the form (C)a(C) (e.g. a ‘2S’, la ‘not’, at ‘IRR’, mak ‘REL’). Enclitics and suffixes consist of a single consonant (e.g. enclitic genitive -n, suffix nominaliser -n), with the exception of a single CVC enclitic (possessive -kan).

2.3.3 Lexical stress

Primary stress falls on the penultimate syllable of the word. Secondary stress falls on the initial syllable of the word, if the word consists of two metrical feet.⁸}

---

³ Note that the antepenultimate syllable similarly has special characteristics in the Central Malayo-Polynesian language Buru, where this syllable is restricted to a VC template, and where the vowel for underived roots is necessarily /e/ (Grimes 1991a:59).

⁴ Since the vowel is restricted to /a/, this syllable is not subject to other rules whose domain is the foot, namely the determination of mid-vowel allophones (§2.4.2), and co-occurrence restrictions (§2.5.3.3, §2.6).

⁵ The fact that the initial unstressed syllable is connected directly to the word level in this diagram is a matter of convenience only, and is not intended to make a theoretical point. An alternative analysis consistent with Goldsmith (1990:172, 175f.) is that the syllable initially belongs to a ‘degenerate foot’ of only one syllable. The foot is subsequently deleted, and the stray syllable "assigned to the nearest foot at word level".

⁶ Blevins (1995:242) notes that constraints on word-initial glottals are not unique, although in general single-member syllable onsets are unrestricted crosslinguistically.

⁷ That is, the coda is restricted to all the coronal and velar consonants except /d/ and /l/. The lack of final /d/ reflects a language universal tendency against word-final voiced stops. /l/ is found word-finally in East Timorese Tetun (de Araújo e Corte-Real 1990; Jonker 1906:282; Morris 1984b). /l/-final entries in Morris’s dictionary tend to have /l/-final variants listed there as well, with the latter being used in the Fehan dialect. According to Hull (1996b:285) the /l/-final variant is also used in Dili.

⁸ This stress pattern is widely recognised in the literature on Tetun (das Dores 1907:29; Fernandes 1937:9; Hull 1996b:xxviii).
four syllables, as well as in trisyllabic compounds and full reduplications where the underlying unstressed second syllable has been elided (§2.10.5.2).

Prefixes, partial reduplicants and clitics do not attract stress, and are not counted in stress assignment rules. (That is, they are extrametrical.) However, there are no syllabic suffixes and the only syllabic enclitic is possessive -kan, with the result that the only phonological words for which stress is antepenultimate are the possessive-marked pronouns. Examples of all stress patterns are presented below, where the acute accent marks primary stress (rather than geminate vowels, as elsewhere).

Table 2.5: Examples of lexical stress patterns

<table>
<thead>
<tr>
<th>Number of syllables</th>
<th>Example</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>feò</td>
<td>N woman</td>
</tr>
<tr>
<td>3: underived root</td>
<td>baliki</td>
<td>N winnowing basket</td>
</tr>
<tr>
<td>3: prefixed root</td>
<td>ha-badak</td>
<td>Vt make-short</td>
</tr>
<tr>
<td>3: compound</td>
<td>man-tolun</td>
<td>N bird-egg</td>
</tr>
<tr>
<td>3: with enclitic -kan</td>
<td>ha’u-kan</td>
<td>Pro 1S-POS</td>
</tr>
<tr>
<td>4</td>
<td>matabian</td>
<td>N spirit</td>
</tr>
</tbody>
</table>

Word stress seems to be realised mainly by loudness and pitch. The difference in strength between stressed and unstressed syllables varies considerably from one speaker and context to another.

Utterance stress nearly always falls on a syllable which bears lexical stress. However, strong stress occasionally falls on the utterance-final syllable (2.1). This is particularly noticeable in official announcements at traditional occasions.

2.1 Ne’e ha’u k-la’ò-n.
this 1S 1S-go-IMM
Now I’ll go. (Strongly stressed: said in a folktale by someone who was very upset.)(O7.49)

2.3.4 Phonetically long vowels

2.3.4.1 Introduction

There is a phonemic contrast between long and short vowels. However, this contrast always correlates with a contrast in stress, with long vowels necessarily being stressed. There are thus no pure length contrasts such as [ba]/[ba:], where both words share the same stress pattern.

In the following sections I will first demonstrate the phonemic contrast involved, then discuss how to analyse this contrast, and finally justify the orthographic representation chosen for long vowels.

A few words with variable stress have been noted. Maromak ‘God’ (both traditional and Christian) can have either penultimate or initial stress, with the latter allegedly having been introduced by Dutch priests. The negator lāhōs ‘(contrastive) indeed not’ (§11.2) is alternatively pronounced with initial stress, as lāhos.
2.3.4.2 Contrasts between long and short vowels

The contrast between long and short vowels is found in only three contexts. In each, the long, stressed, vowel is the only vowel within the metrical foot. Such long vowels are represented orthographically using an acute accent (e.g. ‘á’).

The first contrast is between lexical words consisting of a single metrical foot and unstressable function words. All lexical words are also phonological words, and so bear at least one stress. In contrast, some function words are unstressable, being phonologically clitics rather than words. (This vowel is restricted to /a/, as noted in §2.3.2.4.)

<table>
<thead>
<tr>
<th>Phonological word (one stress)</th>
<th>Clitic (no stress)</th>
</tr>
</thead>
<tbody>
<tr>
<td>át   Adj bad</td>
<td>at   Aux IRR (short form of atu)</td>
</tr>
<tr>
<td>há   Vt eat</td>
<td>ha   Pro 1S (short form of ha’u)</td>
</tr>
<tr>
<td>sá   Pro what</td>
<td>sa   Pro 3P (short form of sia)</td>
</tr>
<tr>
<td>tá   Vt chop</td>
<td>ta   Adv already</td>
</tr>
</tbody>
</table>

The second context in which a length contrast is found is in a number of lexeme pairs which differ in whether stress is placed on the second or on the initial vowel. Stress on the second syllable here is frequently, but not necessarily, the result of applying an unstressable prefix to a base with a single, stressed vowel. The next two sets of examples include some near-minimal pairs to show that the phenomenon is found for all five vowels.

<table>
<thead>
<tr>
<th>Stress on second syllable</th>
<th>Initial stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>katí Vi fly up</td>
<td>kati Vt call dogs</td>
</tr>
<tr>
<td>haré Vt see</td>
<td>hare N rice</td>
</tr>
<tr>
<td>hakát Vi fight</td>
<td>hakat Vt pace</td>
</tr>
<tr>
<td>halón Vi plead</td>
<td>halo Vt do, make</td>
</tr>
<tr>
<td>kakún N hard inner coconut shell, skull</td>
<td>kakun N outer shell (e.g. coconut/rice husk, peel, snakeskin)</td>
</tr>
</tbody>
</table>

Finally, reduplications and compounds carry two stresses, while other words have secondary stress only if they consist of four syllables. This too results in minimal pairs distinguished by stress.

<table>
<thead>
<tr>
<th>Reduplication (two stresses)</th>
<th>Non-reduplication &lt; 4 syllables (one stress)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sí-sín Adj sour (various plural)</td>
<td>sisi Vt comb</td>
</tr>
<tr>
<td>ké-kés Vi speak heedlessly</td>
<td>keke Vt spread out evenly</td>
</tr>
<tr>
<td>bá-bá-n Vi go for no reason</td>
<td>baban Vt patch up</td>
</tr>
<tr>
<td>há-hát Num four at a time</td>
<td>hahát Vt damage</td>
</tr>
<tr>
<td>ló-lós Adv truly</td>
<td>lolo N earthen water-pot</td>
</tr>
<tr>
<td>tú-túr Vi sit for no reason</td>
<td>tutur Vt carry on head</td>
</tr>
</tbody>
</table>

2.3.4.3 Analysis of phonetically long vowels

Since stress and vowel length are directly correlated, it should be possible to consider one as basic and derive the other from it. Two basic alternatives are available.

Firstly, one may lexically mark stress, at least for those words for which it is not predictable.9 Vowel lengthening (or doubling) is then a phonetic (postlexical) rule which

---

9 This is the approach reflected in the Tetun orthography of das Dores (1907:29) and Fernandes (1937:9). It is defended at length for the Oceanic language Rotuman by Blevins (1994).
Phonology and morphophonemics

applies to a stressed vowel if it is the only vowel in a metrical foot. Under this analysis the minimum prosodic word is at underlying level a single stressed syllable.

\[
\begin{array}{c|c}
\text{Foot} & \text{Foot} \\
\hline
S & S \\
(C)V(C) & (C)V(C)
\end{array}
\]

Alternatively, one can treat vowel length as basic. Within this analysis, there are three possibilities. One is to posit a set of long vowel phonemes corresponding to each of the short vowels. A preferable solution, which fits the facts mentioned below, is to analyse long vowels as a sequence of two vowels, each heading its own syllable. The minimal prosodic word is then a disyllabic foot, and stress falls predictably on the initial syllable of each foot. A third possibility is that long vowels constitute two morae but only a single syllable. Coda consonants do not count as morae, and so do not affect stress rules. Such an analysis is feasible but not necessary, as there is no evidence requiring that the mora be recognised as a metrical unit distinct from the syllable in Tetun.

There are two sources of evidence favouring an analysis of these long stressed vowels as double vowels. Firstly, within the prosodic word long vowels share the same distribution as vowel sequences, in that neither can cross a foot boundary (§2.3.2, §2.6). This is most easily accounted for if long vowels are analysed as geminate sequences. Secondly, word games treat lexical words with a single stressed vowel as if they have a sequence of two identical vowels (§2.3.5). 10

2.3.4.4 Orthographic representation

A related question is how such long vowels should be represented in the orthography. Long and short vowels are not distinguished in any way in the orthography currently in use in the Fehan district, such as in religious publications, consultants’ transcriptions of oral texts, and private correspondence. Given the low functional load of the length distinction, this lack of representation of the distinction does not raise significant practical problems.

In contrast to the popular Fehan orthography, Morris (1984b) uses double vowels to indicate length, Hull (1996b) uses a mixture of double vowels and diacritics, and das Dores (1907) and Fernandes (1937) use diacritics.

Since this is a technical rather than a popular work, length distinctions should be reflected in the orthography used. Nevertheless, in order to keep the appearance of written words as close as possible to the common Fehan orthography, I have chosen to use an acute accent to mark long vowels (e.g. ‘á’), rather than directly reflecting the phonemic analysis by using double vowels (e.g. ‘aa’).

10 Treating stress as predictable from the presence of a geminate vowel, rather than vice versa, accords well with current phonological theory, in that stress rules are generally held to follow syllabification (Goldsmith 1990:169ff.; Kenstowicz 1994:252).
Chapter 2

2.3.5 The evidence from word games

2.3.5.1 Introduction

Young people have several word games which they use to communicate with each other with the intention that older people and small children will not understand. While at least some of these are new games, and the results do not necessarily accord with Tetun phonological patterns (in allowing /ŋ/ and sequences of three vowels) they do throw some light on phonology. In particular they show that phonetically long vowels are treated as sequences of two identical vowels (e.g. compare kdök and sei in example 2.7). The games also show that initial syllables of trisyllabic words are treated ambivalently as either not belonging to a foot, or as constituting a foot on their own.

Each of the three games discussed below are templatic (Bagemihl 1995:700; McCarthy 1981), relying on placing certain constants in either vowel or consonant slots within a metrical foot. Roots of four syllables, as well as compounds and reduplications, are treated as two metrical feet, as shown by Katarina in example 2.3, man-tolun in 2.8, and lai-lais in 2.9. The pretonic syllable of trisyllabic words (whether this be part of the root, a prefix or a proclitic such as la 'not') is usually left unchanged, although it is occasionally treated as if it were a separate foot containing a sequence of two identical vowels. For instance, the irrealis proclitic at may be either left unchanged or encoded identically to the adjective at ‘bad’ (which has a stressed vowel). This inconsistency can be found within the one speaker on the one occasion of speaking, as illustrated by the initial syllable ra- in example 2.9, and the variant encoding of the proclitic la in 2.4 and 2.5.

In the examples below a period is used to indicate foot boundaries within words, to facilitate comparison of the coded Tetun on the top line with the normal Tetun on the second line.

2.3.5.2 Game 1: consonant substitution

In the first game, the foot-medial consonant slot is filled by /ŋ/ (or for some speakers /ŋ/), and the final coda is filled by /t/.

Source: (k)C1 V1 C2 V2 (C3)
Code: (k)C1 ŋ V2 t

The examples below illustrate this rule, including the variation in the treatment of trisyllabic words and of clitics. For ease of reading /ŋ/ is represented as ‘ŋ’ in the examples.

2.2 Ongkat ingkat ingkat bangkat Ungkat Kata. t h a t .
Oras ida ita bá Uma Kata. h a n .
Soon we will go to Uma Katahan.

2.3 Ingkat Kangkat.ringkat nongkit to.bangkot.
Ibu Kata. r i n a n-odi to. b a k o .
Mother Catharina 3S-bring tobacco
Mrs Catharina brings tobacco.
2.4  Hanggut kanggat.kengget langgat kanggat.tengget.  
Ha’u k-a.kés la k-a.tene.  
1S 1S-talk not 1S-know  
I don’t know how to talk.  
(Q0.170)

2.5  Hanggut la bengget ka.kengget.  
Ha’u la bele k-a.kés.  
1S not can 1S-talk  
I cannot talk.  
(Q0.170)

2.6  Langgat bengget danggit.  
La bele dadi.  
not can happen  
(It) can’t happen.  
(Q0.171)

2.7  Hanggit! Senggit kdonggot.  
Hai! Sei kdók.  
EXCL still far  
Hey! (It) is still far.  
(Q0.170)

2.3.5.3 Game 2: second syllable modification

The second game reinterprets the (C)V of the second syllable of the foot as two syllables, repeating the onset (if any), and using the vowel /o/ in the first copy and /e/ in the second. This can result in sequences of three vowels, which are not elsewhere permissible in Tetun.

Source:  
((k)C₁) V₁ (C₂) V₂ (C₃)

Code:  
((k)C₁) V₁ (C₂) o (C₃) e (C₃)

2.8  Ha’o’e ka.karore maoen.tololën ne’o’e.  
Ha’u k-a.kara man.-tolun ne’e.  
1S 1S-like bird-egg this  
I like this bird-egg.  
(X0.72)

2.9  Sioe ra.oae laoe.laoes. = Sioe ra.oelaoe laoe.laoes.  
Sia r-a.lai lai.-lais Sia r-a.lai lai.-lais.  
3P 3P-run RDP-quick 3P 3P-run RDP-quick  
They run fast.  
(X0.73)

2.3.5.4 Game 3: vowel substitution

In the third game the two vowels of the foot are fixed as /a/ and /i/ respectively. The consonants remain unchanged.

Source:  
((k)C₁) V₁ (C₂) V₂ (C₃)

Code:  
((k)C₁) a (C₂) i (C₃)

2.10  Ha’i ati kbai Ami Ka.tahin.  
Ha’u atu k-bá Uma Ka.tahin.  
1S IRR 1S-go House Katahan  
I will go to Uma Katahan.  
(N0.162)
2.4 Vowels

2.4.1 Vowel phonemes

Tetun has five vowel phonemes, /i, e, a, o, u/.

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>i</td>
<td></td>
<td>u</td>
</tr>
<tr>
<td>Mid</td>
<td>e</td>
<td></td>
<td>o</td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

Vowel contrasts are exemplified in the following minimal and near-minimal pairs, for the initial (stressed) and final (unstressed) syllables of disyllabic words.

<table>
<thead>
<tr>
<th>Initial (stressed) syllable</th>
<th>Final (unstressed) syllable</th>
</tr>
</thead>
<tbody>
<tr>
<td>silu Vt snap</td>
<td>kloki N hanging larder</td>
</tr>
<tr>
<td>belu N companion</td>
<td>loke Vt open</td>
</tr>
<tr>
<td>balu N some</td>
<td>loka N single girls' room</td>
</tr>
<tr>
<td>bolu Vt call</td>
<td>loko V boast</td>
</tr>
<tr>
<td>kulu N k.o. fruit</td>
<td>loku N men's bracelet</td>
</tr>
</tbody>
</table>

2.4.2 Vowel allophones

The central low unrounded vowel /a/ is generally phonetically close to [a], and has no allophones. However, it often shortens and centralises towards schwa in an open pretonic syllable of a trisyllabic phonological word (e.g. sakunar [səkunar] ‘scorpion’). This includes the unstressable proclitic ta (usually [tə], probably a cliticised form of aspectual ti’a ‘already’; §11.5.2).

In some words the pretonic /a/ can be omitted altogether, resulting in an initial consonant cluster. Resulting clusters in the corpus include /bl/ (balår, blår ‘astonished’), /br/ (baria, bria

---

11 These five vowels are recognised by all who have written on Tetun phonology. Further vowels have been claimed for Dili Tetun, for which I have no evidence in the Fehan dialect. Nasal vowels (presumably allophones) have been recognised by das Dores (1907:9), Fernandes (1937:11) and Hull (1996b:xxi). Examples given suggest that the nasal allophone occurs before an underlying nasal consonant in both native and Portuguese loan vocabulary. De Araújo e Corte-Real (1990:54) adds the schwa as a phoneme. If these extra vowels are correct for Dili Tetun, they are likely to have come about through centuries of contact with the Portuguese language in Dili.

12 Hull (1996b:xix) says that unstressed /a/ in final /ar/ and /an/ is often weakened to schwa. This does not hold for the Fehan dialect. Das Dores (1907) orthographically distinguishes two varieties of ‘a’ but does not specify rules for them.

13 When pressed, consultants were unsure as to which of the cardinal vowels /a/ or /e/ should occur in the citation form of this proclitic, just as sometimes they were undecided about the unstressed initial syllable of trisyllabic words. Since the clitic is not used in written Tetun, there is no spelling convention for it.
'a kind of creeper with bitter fruit'), and /sk/ (sakili, skili 'tickle'). In addition, /fl/ (falilo, fliilo 'sound of a wooden flute') and /tf/ (tafou, tfou 'spit downwards, making the sound t(afou') are found in the corpus in onomatopoeic words. Although the full range of possible consonant clusters is not known, it does not extend as far for the Fehan dialect as it does for the Tetun of Dili, for which de Araújo e Corte-Real (1990:64) gives a wide range of examples which I have not encountered in Fehan, such as /hn/ (hananu, hnau 'sing') and /hl/ (halai, hlai 'run, escape').

The two mid vowels, front unrounded /e/ and back rounded /o/, each have a high and a low allophone, conditioned by the following environment. The low allophones [e, 0] occur in the final syllable of a metrical foot, or in the non-final syllable if the vowel in the following syllable is also a mid vowel. The high allophones [e, o] occur elsewhere. 14

The only allophonic rules presented in the literature are by Troeboes et al. (1987:16ff.). They say that a low allophone of the mid vowels is found in a syllable preceding or following a syllable with a vowel /e, o, a/, while a high allophone is found before /i, u/. These rules leave final syllables unaccounted for if they follow a high vowel, and falsely predict a low allophone if the following vowel is /a/ (e.g. hela 'leave', hola 'fetch').

14 The common truncation of the pronoun emi [emi] '2p' to em [em] indicates that the allophone rule applies before (optional) truncation of the final vowel.

The existence of these allophones for /e/ and /o/ is widely recognised in the literature on Tetun (das Dores 1907:9; Fernandes 1937:10; Francillon 1967; Mathijsen 1906:xiv), but no satisfactory rule has yet been presented. Morris (1984b) in his dictionary entries reliably distinguishes 'o' for low [ɔ] and 'ou' for high [o], but makes no such orthographic distinction for /e/.
Table 2.8: Examples of high and low allophones of mid vowels

<table>
<thead>
<tr>
<th>High allophone</th>
<th>Low allophone</th>
</tr>
</thead>
<tbody>
<tr>
<td>/e/</td>
<td></td>
</tr>
<tr>
<td>[keta] V</td>
<td>[wee] N</td>
</tr>
<tr>
<td>[ferik] N</td>
<td>[bele] Aux</td>
</tr>
<tr>
<td>[etu] N</td>
<td>[fe.te] N</td>
</tr>
<tr>
<td>[lekirauk] N</td>
<td>[wee-hali] PN</td>
</tr>
<tr>
<td></td>
<td>water</td>
</tr>
<tr>
<td></td>
<td>can</td>
</tr>
<tr>
<td></td>
<td>woman</td>
</tr>
<tr>
<td></td>
<td>placename (lit. ‘water-banyan’)</td>
</tr>
<tr>
<td>/o/</td>
<td></td>
</tr>
<tr>
<td>[hola] Vt</td>
<td>[moos] Adj</td>
</tr>
<tr>
<td>[hosi] Vt</td>
<td>[kole] Adj</td>
</tr>
<tr>
<td>[hosu] Vi</td>
<td>[foho] N</td>
</tr>
<tr>
<td>[ora-'.oras] Adv</td>
<td>[kokeek] Vi</td>
</tr>
<tr>
<td></td>
<td>clean</td>
</tr>
<tr>
<td></td>
<td>tired</td>
</tr>
<tr>
<td></td>
<td>mountain</td>
</tr>
<tr>
<td></td>
<td>crow (cock)</td>
</tr>
</tbody>
</table>

The high front unrounded vowel /i/ varies in quality from [i] to [I]. There is a tendency for [I] to occur in closed word-final (unstressed) syllables (e.g. [talm] ‘string’, [fenk] ‘old woman’), and [i] to occur elsewhere. However, this does not appear to be strictly allophonic. 15

The high back rounded vowel /u/ has no allophonic variation, being phonetically [u]. 16

2.5 Consonants

2.5.1 Consonant phonemes

Tetun has thirteen consonant phonemes, as follows. 17

Table 2.9: Consonant phonemes

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Coronal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stops</td>
<td>voiceless</td>
<td>t</td>
<td>k</td>
<td></td>
</tr>
<tr>
<td></td>
<td>voiced</td>
<td>b</td>
<td>d</td>
<td></td>
</tr>
<tr>
<td>Fricatives</td>
<td>voiceless</td>
<td>f</td>
<td>s</td>
<td>h</td>
</tr>
<tr>
<td>Lateral</td>
<td></td>
<td>l</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill</td>
<td></td>
<td>r</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The consonants are exemplified in the following minimal and near-minimal pairs for all positions within disyllabic words. Omissions reflect disallowed positions. These result from

15 Mathijsen (1906:xiv), writing of the Foho dialect of northern Belu, orthographically recognises a high and low variant of /i/, but his entries are not consistent with Fehan pronunciation.

16 Troeboes et al. (1987:16ff.), who base their description on the Foho and Fehan dialects, say that both /i/ and /u/ have a high allophone in open syllables and a low allophone in closed syllables. I have found no evidence for this in the Fehan dialect.

17 Note that a voiced-voiceless pair for stops exists only for the coronal pair /t, d/, which are at the universally least marked place of articulation (Clements 1990:313).

Around Besikama, in the south of the Fehan dialect area, /t/ and /d/ have merged into a single phoneme, which phonetically tends towards an alveolar flap.
the following facts: (i) that the glottal stop does not occur word-initially (§2.3.2.5), (ii) that glottal and velar consonants do not follow /k/ (§2.7.1), and (iii) that there are restrictions on permitted coda consonants (§2.3.2.5).

Table 2.10: Consonant contrasts word-initially, after /k/ and word-medially

<table>
<thead>
<tr>
<th>Initial</th>
<th>After k-</th>
<th>Medial</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>baka</td>
<td>Vt</td>
</tr>
<tr>
<td></td>
<td>split in two</td>
<td>kbadak</td>
</tr>
<tr>
<td>t</td>
<td>taka</td>
<td>Vt</td>
</tr>
<tr>
<td></td>
<td>close</td>
<td>ktakak</td>
</tr>
<tr>
<td>d</td>
<td>dakar</td>
<td>Vt</td>
</tr>
<tr>
<td></td>
<td>guard</td>
<td>kdahur</td>
</tr>
<tr>
<td>k</td>
<td>kaka</td>
<td>Vt</td>
</tr>
<tr>
<td></td>
<td>point out</td>
<td>--</td>
</tr>
<tr>
<td>f</td>
<td>fakar</td>
<td>Vt</td>
</tr>
<tr>
<td></td>
<td>tip out</td>
<td>kfatan</td>
</tr>
<tr>
<td>s</td>
<td>sakat</td>
<td>Vt</td>
</tr>
<tr>
<td></td>
<td>cut strips</td>
<td>ksadan</td>
</tr>
<tr>
<td>h</td>
<td>hakat</td>
<td>Vt</td>
</tr>
<tr>
<td></td>
<td>pace</td>
<td>--</td>
</tr>
<tr>
<td>l</td>
<td>laka</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td>shine</td>
<td>klaran</td>
</tr>
<tr>
<td>r</td>
<td>raka</td>
<td>Vt</td>
</tr>
<tr>
<td></td>
<td>sell (meat)</td>
<td>krakat</td>
</tr>
<tr>
<td>m</td>
<td>m-akat</td>
<td>Vt</td>
</tr>
<tr>
<td></td>
<td>2S-pace</td>
<td>kmakur</td>
</tr>
<tr>
<td>n</td>
<td>n-akat</td>
<td>Vt</td>
</tr>
<tr>
<td></td>
<td>3S-pace</td>
<td>knaban</td>
</tr>
<tr>
<td>w</td>
<td>wani</td>
<td>N</td>
</tr>
<tr>
<td></td>
<td>bee</td>
<td>kwakat</td>
</tr>
<tr>
<td></td>
<td>sago palm</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td>--</td>
</tr>
</tbody>
</table>

Table 2.11: Consonant contrasts word-finally

<table>
<thead>
<tr>
<th>Final</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>m-ōt</td>
</tr>
<tr>
<td>k</td>
<td>mōk</td>
</tr>
<tr>
<td>s</td>
<td>mōs</td>
</tr>
<tr>
<td>r</td>
<td>mōr</td>
</tr>
<tr>
<td>n</td>
<td>mōn</td>
</tr>
<tr>
<td></td>
<td>m-ō</td>
</tr>
</tbody>
</table>

All stops are unaspirated, except when there is utterance final devoicing, which is not uncommon for some speakers. No allophonic variation is evident.

2.5.2 Loan consonants

The vast majority of loan words used in the Fehan dialect of Tetun are from Malay, especially from the non-standard dialect spoken in Timor, rather than from Bahasa Indonesia. Some of these loans are originally from Dutch, but are incorporated into Tetun via Malay. The number of loans used, and the degree to which they are assimilated to Tetun phonology, differs widely from speaker to speaker. For speakers who use unassimilated loans

---

18 In Dili Tetun, in contrast, there is extensive incorporation of Portuguese vocabulary, resulting in a different set of non-native phonemes. De Araújo e Corte-Real (1990:55) lists /p, v, g, j, z, š/; while Hull (1996b:34–37) lists /p, v, g, j, z, x, rr, ll, ñ, y/.
there are, in addition to the thirteen Tetun consonants, some loan consonants from Malay. These consonants, along with the consonants to which they tend to be assimilated, are listed below.

<table>
<thead>
<tr>
<th>Phoneme</th>
<th>Malay Example</th>
<th>Tetun Example</th>
<th>Malay Assimilated</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>polisi</td>
<td>balisi</td>
<td>N</td>
<td>police</td>
</tr>
<tr>
<td>g</td>
<td>guru</td>
<td>kuru</td>
<td>N</td>
<td>teacher</td>
</tr>
<tr>
<td>j</td>
<td>juta</td>
<td>duta</td>
<td>Num</td>
<td>million</td>
</tr>
<tr>
<td>c</td>
<td>meja</td>
<td>meda</td>
<td>N</td>
<td>table</td>
</tr>
<tr>
<td>η/ _ #</td>
<td>sanán (‘senang’)</td>
<td>Sina</td>
<td>N</td>
<td>China, Chinese</td>
</tr>
<tr>
<td>η/ _ g</td>
<td>mingu</td>
<td>mingu</td>
<td>N</td>
<td>week, Sunday</td>
</tr>
<tr>
<td></td>
<td>tangal</td>
<td>tangal</td>
<td>N</td>
<td>date</td>
</tr>
</tbody>
</table>

The data are not sufficient to judge how other Malay phonemes (/y, ñ/) would be assimilated, or for identifying conditioning factors for /η/ and /c/, which appear not to be assimilated in some phonological contexts while, potentially, they are assimilated in others.

### 2.5.3 The consonant /w/ 

#### 2.5.3.1 Contrast with /u/

The phoneme /w/ must be classed as a consonant since it occurs in all consonant positions except the highly restricted word-final position. Like other consonants, and unlike vowels, it can occur intervocically and cannot carry stress.

The primary contrast between the consonant /w/ and the phonetically similar vowel /u/ lies in the syllabicity, and hence stressability, of the vowel. I have found only the following minimal and near-minimal pairs distinguishing (unstressed) /w/ and (stressed) /u/.

<table>
<thead>
<tr>
<th>/w/</th>
<th>/u/</th>
</tr>
</thead>
<tbody>
<tr>
<td>kwák</td>
<td>Adj wide-open</td>
</tr>
<tr>
<td>wár</td>
<td>N neck of coconut tree</td>
</tr>
<tr>
<td>wé</td>
<td>N water</td>
</tr>
</tbody>
</table>

Morris (1984b) represents both phonemes as ‘u’, a spelling which confused some consultants.

#### 2.5.3.2 Possible merger with /bh/

Of all the consonants, /w/ carries the least functional load, representing only 1% of consonants (in a word list of 3,012 entries).
There is some evidence, mainly from other dialects, that /w/ is merging into /b/.

At present these consonants do contrast, as exemplified in the following minimal pairs:

<table>
<thead>
<tr>
<th>/w/</th>
<th>/b/</th>
</tr>
</thead>
<tbody>
<tr>
<td>wé</td>
<td>bér</td>
</tr>
<tr>
<td>wa'ik</td>
<td>ba’ik</td>
</tr>
<tr>
<td>kwá</td>
<td>k-bá</td>
</tr>
</tbody>
</table>

 água  | Adv also

water  | older, mature

However, there are also a number of words for which both /b/ and /w/ are heard within the Fehan dialect.

<table>
<thead>
<tr>
<th>/b/</th>
<th>/w/</th>
</tr>
</thead>
<tbody>
<tr>
<td>labadain</td>
<td>lawadain</td>
</tr>
<tr>
<td>bei</td>
<td>wei</td>
</tr>
<tr>
<td>serbisu</td>
<td>serwisu</td>
</tr>
<tr>
<td>Babiku</td>
<td>Wewiku</td>
</tr>
</tbody>
</table>

Both Morris (1984b:xi) and Hull (1996b:xviii) note interdialectal differences in the use of /b/ and /w/ within East Timor. Many of Morris’s dictionary entries using /b/ have a /w/ in the corresponding word in the Fehan dialect (e.g. East Timor be, Fehan wé ‘water’). According to Hull (1996b:xxvii), /w/ has merged completely into /b/ in Dili Tetun.

### 2.5.3.3 Co-occurrence restrictions on /w/ and back vowels

With the exception of the exclamations woul and wul, /w/ can neither be preceded nor followed by a back vowel within a metrical foot (*/wo, wu, ow, uw/). Sequences such as /oa/ can, as far as their phonetics is concerned, be interpreted as either /oa/ or /owa/. However, given the failure of /w/ to co-occur with back vowels in other sequences where it could not be interpreted as a glide vowel (e.g. *iwu, owo/), it is simplest to interpret the phonetic [w] in these sequences as a non-phonemic glide only. This interpretation is in accordance with the commonly used orthography.

The following table gives examples of all vowel-/w/ combinations found in the corpus. Missing sequences marked by a dash are disallowed (except in exclamations), while blank entries are assumed to be accidents of the data.

---

19 A similar merger appears to have occurred in the neighbouring language Dawan (also known as ‘Timorese’). Middelkoop (1950:387) notes that /w/ occurs only in archaic forms or as an intervocalic glide (‘verbings-medeklinker’, lit. ‘joining consonant’), and that words written with /w/ by the traveller Müller in 1837 were in 1950 pronounced with /b/. In Rotinese, too, /w/ is restricted, occurring only in a few interjections and loan words (Fox & Grimes 1995).

20 Troebes et al. (1987:28) say that /w/ does not occur morpheme-medially. However, this is invalidated by the inclusion of awan ‘tomorrow’ and sawan ‘morning’ in their examples on page 84. The sequence /wu/ is missing in many languages (Hayes 1989:300).

21 Nevertheless, I came across one young woman who consistently wrote oa ‘child’ as ‘owa’.
The only sequences of /w/ with a back vowel across a foot boundary in the corpus are kowa and turuwaku, which both denote varieties of wild birds. These are assumed to include a /w/ (rather than a non-phonemic glide) because the corpus has no other unambiguous cases of onsetless syllables introducing the second foot of a word.

### 2.6 Vowel sequences

All two-vowel sequences are permissible except a high vowel followed by a mid vowel (i.e. */ie, ue, io, uo/), and the sequence */ao/. Examples of all sequences are given below. (Note that a vowel spelled with an acute accent, such as ‘á’, represents a geminate vowel.)

<table>
<thead>
<tr>
<th>-i</th>
<th>-e</th>
<th>-a</th>
<th>-o</th>
<th>-u</th>
</tr>
</thead>
<tbody>
<tr>
<td>i-</td>
<td>sin</td>
<td>lia</td>
<td>N</td>
<td>kiu</td>
</tr>
<tr>
<td>Adj</td>
<td>sour</td>
<td>N word</td>
<td></td>
<td>N chick</td>
</tr>
<tr>
<td>e-</td>
<td>sei</td>
<td>sé</td>
<td>mean</td>
<td>kreon</td>
</tr>
<tr>
<td>Aux</td>
<td>still</td>
<td>Pro who</td>
<td>Adj red</td>
<td>Adj thin</td>
</tr>
<tr>
<td>a-</td>
<td>kair</td>
<td>kaer</td>
<td>há</td>
<td></td>
</tr>
<tr>
<td>N fishing</td>
<td>Vt grasp</td>
<td>Vt eat</td>
<td>Vt call loudly</td>
<td></td>
</tr>
<tr>
<td>o-</td>
<td>soi</td>
<td>soe</td>
<td>oa</td>
<td>fó</td>
</tr>
<tr>
<td>Adj rich</td>
<td>Vt throw out</td>
<td>N child</td>
<td>Vt give</td>
<td>Adj new</td>
</tr>
<tr>
<td>u-</td>
<td>sui</td>
<td>bua</td>
<td>N</td>
<td>hú</td>
</tr>
<tr>
<td>Vt impale</td>
<td>N betel nut</td>
<td>Vt blow on</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The restriction prohibiting a sequence of a high and a mid vowel still holds when a consonant intervenes. The domain of this rule is the metrical foot. The only exception in the corpus is the onomatopoeic falilo ‘sound of a wooden flute’. A sequence of high and mid vowel is permitted because it forms part of the onomatopoeia.

---

22 Troebes et al. (1987:19) say that /ao/ does occur, but cite only sao ‘tea’, a Portuguese borrowing which is not used in the Fehan dialect.
vowel with intervening consonant is, however, allowed across a foot boundary (e.g. akitou 'dove', bakuro 'prickly acacia' and reduplications such as beibeik 'constantly').

In contrast a sequence of /a/ and /o/ is permitted when a consonant intervenes, even within a foot (e.g. la’o ‘walk’, halo ‘do, make’).

Perceptually, vowels in a sequence are more distinct when they have different values for backness, specifically in the sequences front plus non-front (/ea, eo, eu, ia, iu/) and back plus non-back (/oa, oe, ua, ui/). The exception is the rising diphthong /oi/, in which the vowels are somewhat less distinct. The sequence /ae/ is also less distinct but not a glide. A rising vowel sequence is heard as a glide when it consists of /a/ followed by a high vowel (/ai/, au/) or of a mid vowel followed by a high vowel of the same backness (/ei/, /ou/).

2.7 /k/-initial consonant clusters

2.7.1 Introduction

Tetun has two types of consonant clusters. One is found word-initially in contractions (§2.4.2) and word-medially in compounds (§2.10.5.2) and contractions (§2.10.7). In these the cluster comes about as a result of omitting an antepenultimate vowel in a word which underlyingly has three or four syllables.

The second type of cluster has wider distribution, but is restricted in that it always has /k/ as the initial consonant. The second element can be any other consonant except the glottals (/’/ and /h/) and the phoneme /k/ itself. It is these clusters that form the topic of discussion in the following sections. I first survey the positions in which these clusters are found, then their phonetic quality, their phonemic contrast with non-clusters, their unique phonological and morphological properties, and finally the issue of how such clusters should be analysed.

2.7.2 Positions within the word

/k/-initial clusters are found in only two positions within the word, namely word-initially and at the interface between a prefix and a base.23

Word-initial clusters come from three sources. Firstly, 16% of the 1,664 disyllabic lexemes in the dictionary corpus (i.e. 269 entries) begin with a /k/-initial consonant cluster. Of these words, a quarter are adjectives denoting qualities, including size, shape and temperature, while the remainder are nouns, mostly denoting physical entities. Noticeably absent are verbs denoting actions or processes. This uneven distribution of word class and semantics suggests that the /k/ is, or historically was, a derivational prefix, attached to consonant-initial roots.

A circumfix /-k/-k which forms nouns and adjectives is synchronically recognisable in many of the lexemes (e.g. k-feto-k 'effeminate' from feto 'woman'; §4.5.2.3, §4.6). In the Fehan dialect this prefix is restricted to disyllabic roots. Nevertheless, one occasionally hears non-Fehan derivational sequences of a /k/ followed by a reduplicated root, resulting in cluster-initial trisyllabic words (e.g. k-la-leno-k 'mirror' from leno 'read, look at'; §4.5.2.3).

Many other cluster-initial lexemes do not result from a synchronically recognisable derivation. Some, in fact, have a corresponding lexeme without the initial /k/ with which

23 The only confirmed words in the corpus with consonant clusters that do not fit these patterns are lakna ‘it riddle’ and aksisi ‘type of hard-wood tree’. I am uncertain of the morphological make-up of these two words. It is possible, however, that the latter is a compound in which the initial /a/ is a truncation of ai ‘plant’.
there is no semantic relationship, and from which they could not be derived. This is illustrated by the examples below.

<table>
<thead>
<tr>
<th>C</th>
<th>kC</th>
</tr>
</thead>
<tbody>
<tr>
<td>nü</td>
<td>N coconut</td>
</tr>
<tr>
<td>fui</td>
<td>Vt pour</td>
</tr>
<tr>
<td>mí</td>
<td>Vi urinate</td>
</tr>
</tbody>
</table>

A second source of word-initial /k/ is the IS subject marker on verbs that begin with a consonant other than /h/ (e.g. k-mai ‘1S-come’; §9.3.1). While the vast majority of such verbs are disyllabic, this k- can be added to trisyllabic verbs also (e.g. k-sakili ‘1S-tickle’, k-taksér ‘1S-estimate’ [a Dutch loan]), thus introducing the antepenultimate syllable.

Thirdly, the cluster may occur word-initially in a reduplication of a disyllabic /k/-initial base (e.g. kra-rakat ‘cruelty’ from krakat ‘mad, angry’, kmeti-metis ‘tightly’ from kmetis ‘tight’).

Word-medial clusters arise at the interface of a monosyllabic prefix and a base, where either the prefix is /k/-final and the base starts with a consonant (e.g. mak-tuir ‘who follows’), or the base begins with a consonant cluster (e.g. ha-kbokur ‘make-fat’).

### 2.7.3 Phonetic quality

The charts below, produced using the Summer Institute of Linguistics’ WinCECIL program (Hunt 1995), show the speech sound waveforms for words with and without initial consonant clusters. Specifically, they demonstrate the contrasts /t/-/kt/, /d/-/kd/-/kad/ and /f/-/kf/-/kaf/, based on carefully enunciated elicited speech by a woman approximately 30 years of age. In the charts each phoneme symbol is placed above the section of the chart at which that phoneme begins. Judgement as to where phonemes begin was based on comparison of the charts with very slow replays of the recordings, as well as with corresponding spectrograms and magnitude charts.

The charts (and others not included in this publication) show that the closure of the cluster-initial /k/ is represented by a period of total silence (the remaining waves in the chart at this point being ‘noise’). This silence is followed in most cases by a brief plosion of very low amplitude. The following consonant starts shortly after the plosion (usually 15–40 milliseconds after it). The consonant following a /k/ within a cluster is significantly shorter in duration than that same consonant when it occurs as sole onset.

The only cases in which no plosion was noted for /k/ were those in which it precede another stop, with the clearest instance being that shown in Figure 2.2 for ktalik. In this case the presence of /k/ is still shown by the fact that the closure before the plosive release of /kt/ is far longer than the closure for /t/ in talik.

Where /k/ occurs before a stop (/t, b, d/) and to a lesser extent the fricative /f/, I (and other non-native speakers) frequently have difficulty in recognising it as such. This was especially so during the early months of fieldwork. The clusters I have the most difficulty discerning are thus those that have a low sonority contrast between the successive consonants.

Where the cluster is preceded by a vowel, the /k/ is more discernible, and functions as the coda to the preceding syllable (§2.7.6.2).

---

24 Writing about the Fehan dialect, Francillon (1967:xix) says the /k/ is “only very slightly sounded, almost unnoticeable when it precedes an initial consonant”.
Figure 2.1: Sound waveform for *Ami talik* ‘We intertwine (arms)’

Figure 2.2: Sound waveform for *ai ktalik* ‘vine’ (lit. ‘plant vine’)
Figure 2.3: Sound waveform for *Nia dók* ‘He (is a) shaman’

Figure 2.4: Sound waveform for *Nia kdók* ‘He (is) far away’

Figure 2.5: Sound waveform for *Ha’u kadók* ‘I distance (it)’
Figure 2.6: Sound waveform for *Niafeur* 'He spins (a top)'; *Ne’e kfeu* 'This (is a) kfeu tree'

Figure 2.7: Sound waveform for *Ne’e kafeur* 'This (is a) top'
2.7.4 Contrasts between Word-initial /C/, /kC/ and /kaC/

The status of the /k/ in word-initial consonant clusters tends to be unstable. Many words occur in the corpus both with and without an initial consonant cluster (e.g. *ktodan, todan 'heavy'), or with an initial cluster and with an unstressed /a/ (often phonetically [ə]) between the two consonants (e.g. *kle'uk, kale'uk 'crooked'). For a handful of words all three are found.25

There are several reasons for this variation. Speakers are in wide agreement as to which words have an initial cluster, as opposed to a single consonant, in the Fehan citation form. There is dialect variation in this regard, with the omission of initial /k/ being widely recognised as a characteristic of Suai speech (e.g. Fehan *kwalan, Suai *walan 'naked'). For a few words there is a difference according to village, even within the Fehan area (e.g. Laran *kmaek, Betun *maek 'kind of edible tuber mainly eaten during famines').

In addition to dialect differences and resultant dialect mixing, /k/ is sometimes omitted in casual speech. Many consultants agree that it is permissible to omit the /k/ from many words when speaking, but not when writing (for which the citation form should be used). In addition to this variation in speaking, there is, as pointed out earlier, often a difficulty in hearing /k/ in certain contexts even when it is produced.

There is less agreement amongst native speakers as to whether a word begins with a consonant cluster or whether there is an intervening vowel (e.g. consultants disagreed over *klete, kalete 'bridge'). This confusion is a result of two conflicting tendencies. On the one hand, pretonic vowels are often centralised and shortened, or even omitted, in casual speech (§2.4.2). This results in words with pretonic vowels in the citation form sometimes being heard as having a consonant cluster (e.g. karawa [karawa, krawa] 'k.o. monkey').26 On the other hand, some speakers insert an epenthetic /a/ within words which in the citation form start with a consonant cluster. This is perceived to be a non-Fehan characteristic (e.g. of Foho speakers, or of those people who speak Malay most of the time) (e.g. Fehan *kd6k, uncommon *kad6k 'far').27

Despite the overlap between a single consonant and a cluster, and between a consonant cluster and two consonants with an intervening vowel, there are many minimal and near-minimal pairs and even triads proving that these three are distinct at the lexical level. Some of these are presented below.

---

25 Morris (1984b:104) recognises the problem, stating for East Timor: “In some regions there exists the practice of adding k to words beginning with a consonant, without changing their significance...Also some words beginning with ka are shortened to k, and vice versa”. Similarly, in the neighbouring language of Dawan (or ‘Timorese’), Middelkoop (1950:385) has noticed an optional /kl/ before some words.

26 The pretonic vowel in this sequence is also omitted in some words when Fehan speakers write. For instance, kawd ‘crow’ is spelled ‘kwa’ (although it contrasts phonologically with *kwd ‘medicine’), and karawa ‘k.o. small monkey’ as ‘krawa’; however, kaliuk ‘above all’ is written ‘kaliuk’.

27 Hull (1996b:xxviii), writing of Dili Tetun, says it “dislikes the indigenous consonant clusters of Tetun Terik...and will often insert a glide vowel to facilitate pronunciation, a speech habit not regularly noted in the standard spelling”: Dili Tetun does not use subject marking on verbs, so has no consonant clusters arising from that source.

Within the Fehan dialect, poetry distinguishes clearly between consonant clusters and /kaC/, since it allows antepenultimate syllables to be stressed. Thus, for instance, trisyllabic *kati [kati] ‘fly up’ can be poetically stressed as *[kâtti:], while disyllabic *krade ‘wild duck’ can take only a single penultimate stress.
Table 2.15: Contrasts between word-initial /C/, /kC/ and /kaC/

<table>
<thead>
<tr>
<th>C</th>
<th>kC</th>
<th>kaC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belak</td>
<td>PN man's name</td>
<td>kbelak N metal disk</td>
</tr>
<tr>
<td>bón</td>
<td>N burrow</td>
<td>kbón N smoke, steam</td>
</tr>
<tr>
<td>talik</td>
<td>Vt entwine</td>
<td>ktalik Adj which is a vine</td>
</tr>
<tr>
<td>dök</td>
<td>N shaman</td>
<td>kdök Adj far</td>
</tr>
<tr>
<td>feur</td>
<td>Vt rotate</td>
<td>kfeur N k.o. tree</td>
</tr>
<tr>
<td>sala</td>
<td>Vi err</td>
<td>ksalak Adj immature (beast)</td>
</tr>
<tr>
<td>hû</td>
<td>Vt blow</td>
<td>—</td>
</tr>
<tr>
<td>latik</td>
<td>N earthworm</td>
<td>klatik Adj toppled</td>
</tr>
<tr>
<td>rai</td>
<td>N earth</td>
<td>kraik N, Adj below, low</td>
</tr>
<tr>
<td>maus</td>
<td>Adj tame</td>
<td>kmaus Adj easy to get</td>
</tr>
<tr>
<td>nú</td>
<td>N coconut</td>
<td>knú N bundle (fish)</td>
</tr>
<tr>
<td>wa’ik</td>
<td>Adj older28</td>
<td>kwa’ik Adj many</td>
</tr>
</tbody>
</table>

2.7.5 Phonological and morphological properties

So far we have seen these phonological properties of /k/-initial consonant clusters:

1. They are restricted to certain positions in the word, these being defined partially in morphological terms (a result of subject inflection, reduplication, prefixation) and partially in terms of phonology (introducing the stressed syllable in disyllabic lexemes).

2. There is an unusual amount of interdialectal variation with respect to consonant clusters, with clusters being avoided by omission of the /k/ in Suai speech, and by vowel epenthesis following the /k/ in the dialect of Dili. This suggests that consonant clusters have an uneasy status within Tetun.

In addition, consonant clusters have the following unique characteristics within Tetun morphology:

3. Most speakers inflect verbs reasonably consistently for /h/-initial verbs, for which the inflection does not result in a consonant cluster (e.g. kein ‘1S-wait’ from hein ‘wait’). However, where it would result in a cluster (by attaching 1S k- to a non-/h/-initial verb, e.g. k-mai ‘1S-come’) such inflection is, textually, much reduced in frequency. This again suggests an uneasy status for consonant clusters.

4. When verbs with a word-medial cluster are inflected with 1S k-, the word-medial /k/ is optionally deleted (e.g. kaklati or kalati ‘1S-topple’ from haklati). It is not deleted for any other subject inflection (e.g. 2S maklati, *malati). For further information about subject marking see §9.3.

28 Twins are given the same ancestral name, to which is appended wa’ik ‘older’ for the younger twin, and ki’ik ‘little’ for the elder (e.g. Bui Wa’ik and Bui Ki’ik). The reversal is for reasons of taboo.
5. In reduplication the /k/ of a cluster-initial base is retained only on the (initial) reduplication, not on the (final) base (§2.10.4, §2.10.5.2). Cluster-initial /k/ is similarly deleted from the second member in compounds (§2.10.5.2).

<table>
<thead>
<tr>
<th>Reduplicant/Root</th>
<th>Root</th>
<th>Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RDP</td>
<td>kleur Adj long (time)</td>
<td>kla-leur N length (time)</td>
</tr>
<tr>
<td>RDP</td>
<td>kraik Adj short, low</td>
<td>krai-raik Adj low</td>
</tr>
<tr>
<td>batar N</td>
<td>ktasak Adj ripe</td>
<td>ba-tasak N young maize</td>
</tr>
</tbody>
</table>

6. If the causative prefix ha- is added to a base with an initial consonant cluster, there is considerable variation between words, and to a lesser extent amongst speakers, as to whether the /k/ is retained. This correlates somewhat with the part of speech of the base. The /k/ is generally retained if the base is a noun (14 examples). The one exception in the corpus is hatuik 'plant in rows', for which haktuik was accepted as a less common variant form.

<table>
<thead>
<tr>
<th>Base: N</th>
<th>ha- derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>kbaluk N</td>
<td>hakbaluk Vt split, slice</td>
</tr>
<tr>
<td>knés N</td>
<td>haknés Vt split, slice finely</td>
</tr>
<tr>
<td>ktuik N</td>
<td>ha(k)tuik Vt plant in rows</td>
</tr>
</tbody>
</table>

The /k/ is usually lost if the base is an adjective (14 examples omit /k/, 5 retain it).

<table>
<thead>
<tr>
<th>Base: Adj</th>
<th>ha- derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>kبوكur Adj fat</td>
<td>hakبوكur Vt fatten</td>
</tr>
<tr>
<td>kmukit Adj poor</td>
<td>hamukit Vt impoverish</td>
</tr>
<tr>
<td>ktodan Adj heavy</td>
<td>hatodan Vt seat [noble] lit. ‘make-heavy (self)’</td>
</tr>
<tr>
<td>kwér Adj smooth</td>
<td>hawér Vt smooth</td>
</tr>
</tbody>
</table>

This discrepancy in the retention of cluster-initial /k/ upon prefixation presumably has a diachronic source, as there appears to be no correlation with the phonology of the base, nor with whether the base includes the synchronic prefix k- (which none of those in the above examples do). In addition, there is some inter-speaker variation with regard to retention of /k/ on prefixing. This may, however, result from dialect mixing, since in Suai speech the initial /k/ of clusters is omitted in both the base and the derivation.

2.7.6 Analysis of clusters

2.7.6.1 Introduction

The question arises as to how to analyse these consonant clusters in such a way that their unique characteristics become more than a jumble of unrelated observations.\(^{30}\) The following sections deal with how the initial /k/ is to be slotted into syllable and word templates, and why certain consonant clusters are disallowed.

---

\(^{29}\) The rule truncating the initial member of some compounds, including this one, is presented in §2.10.5.2.

\(^{30}\) Troebes et al. (1987:23) consider /k/-initial consonant clusters to be unitary phonemes, an analysis which is not illuminating.
2.7.6.2 Extrasyllabicity

It has been widely recognised in the literature that in some languages the principles of syllabification which apply word-internally (in Tetun, a strict (C)V(C) pattern) do not apply at the edges of phonological words. Extra segments permitted at the edges of the word are called ‘extrasyllabic’ (Clements & Keyser 1983:121), and are invisible to some phonological processes. The initial /kj/ of consonant clusters fits this description.

In particular, extrasyllabic segments are often invisible to reduplication rules (McCarthy & Prince 1986:44ff.), just as cluster-initial /kj/ is in Tetun. Tetun reduplication rules copy either a foot (§2.10.5) or a syllable onset followed by /a/ (§2.10.4). At the stage at which these rules apply, the initial /kj/ is, as an extrasyllabic segment, not attached to the foot or syllable to be copied. It thus remains unattached in its word-initial position, and fails to take part in the reduplication.31

If a prefix is added to a cluster-initial base, the base-initial /kj/ (if retained) will not be extrasyllabic, but will instead syllabify with that prefix (e.g. ha-ki-liran ‘send off sparks’, where '.' represents a syllable boundary). This is in accordance with the general principle that extrasyllabicity does not apply word-internally (Kenstowicz 1994:274).

By the postlexical level of a derivation (i.e. after all morphological processes have applied), all extrasyllabic segments must be attached to a syllable (Itô 1988:53) or, in some versions of the theory, to the word level (Goldsmith 1990:108). At this stage the underlying constraints on syllable structure do not apply. At this postlexical stage cluster-initial /kj/ is, where it follows an open syllable, syllabified in running speech as a coda to the preceding syllable. Where it does not follow an open syllable, the cluster functions as a complex syllable onset.32

In addition to incorporating segments as extrasyllabic consonants, there are several other methods by which segments which do not fit the syllable structure of a language can be dealt with (Blevins 1995:223f.). Two which were mentioned earlier as being used by various speakers of Tetun are epenthetic vowel insertion (to create a new syllable) and deletion of the stray segment. In fact it is likely that for some speakers, and especially for some non-Fehan varieties of Tetun (such as Suai and Dili), there are no extrasyllabic consonants, with all consonants able to be syllabified within a (C)V(C) template.

2.7.6.3 Disallowed clusters

A second issue to be addressed is the reason for /kj/ being able to precede any labial or coronal consonant, but not the velar stop /kJ/ or the glottal consonants /h/ and /H/.

The lack of geminate /kk/ reflects not only the lack of other geminate consonants in Tetun, but also the relative rarity of word-initial geminates crosslinguistically. /kj/ must

31 Van der Hulst and Klamer (1996) propose a similar analysis for reduplication in the Central Malayo-Polynesian language Leti.
32 This syllabification of cluster-initial /kj/ with a preceding open syllable is mirrored in the transcriptions of some younger assistants, especially when transcribing words with which they were unfamiliar. One non-native speaker with a good ear for phonetic detail regularly wrote the 15 k- inflection attached to the preceding pronoun (e.g. ‘ha’uk ba’ instead of ‘ha’u kba’ ‘I go’), much to the amazement of a more mature consultant.
therefore be deleted when it follows /k/, as illustrated in the derivations below. For a formal specification of this rule see §2.10.3.3.

Affix Base Derivation

\( \text{hak-} -k \) RECIP \( \text{kohi} \) Vt catch \( \text{hakohik} \) Vt wrestle

\( k- \) 1S \( \text{kuru} \) Vt scoop \( \text{kuru} \) Vt 1S-scoop

The absence of the sequence /k'/ reflects the absence of glottal stops root-initially in Tetun.

Although /kh/ does not occur at surface level in Tetun, there is evidence that it occurs underlingly, and that /h/ is deleted whenever it occurs after a consonant. This simplifies the statement of subject marking rules, since a subject inflection effectively replaces the initial /h/ in all /h/-initial verbs, but precedes non-/h/-initial verbs (§9.3.1). Except for subject marking by IS \( k- \), there are no indisputable instances of consonant-final prefixes attaching to /h/-initial bases; however, three possible derivations are listed below for habit 'squeeze' and hihi 'call to horses'. The rule which deletes /h/ is formally specified in §2.10.3.3.

Affix Base Derivation

\( m- \) 2S \( \text{habit} \) Vt squeeze \( \text{mabit} \) Vt 2S-squeeze

\( k- \) 1S \( \text{habit} \) Vt squeeze \( \text{kabit} \) Vt 1S-squeeze

\( k- \) Adj \( \text{habit} \) Vt squeeze \( \text{kabit} \) Adj at an acute angle

\( \text{hak-} \) ? \( \text{habit} \) Vt squeeze \( \text{hakabit} \) Vt carry under arm (squeezing)

\( \text{hak-} -k \) make sound \( \text{hihi} \) Intj call to horses \( \text{hakihi} \) Vt call horses

2.8 Phoneme frequencies

2.8.1 Vowel frequencies

Vowel frequencies are presented below, based on a dictionary corpus of 3,012 lexemes (excluding homonyms, full reduplications, compounds and exclamations). The frequencies indicate the percentage of vowels represented by each vowel phoneme in the overall dictionary corpus, as well as in the initial and final syllables of disyllabic words. Since frequencies differ for closed and open final syllables, these are presented separately.

<table>
<thead>
<tr>
<th>All Words</th>
<th>Disyllabic Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>All syllables (3,012 words)</td>
<td>Initial syllable (1,664 words)</td>
</tr>
<tr>
<td>( i )</td>
<td>15</td>
</tr>
<tr>
<td>( e )</td>
<td>12</td>
</tr>
<tr>
<td>( a )</td>
<td>44</td>
</tr>
<tr>
<td>( o )</td>
<td>13</td>
</tr>
<tr>
<td>( u )</td>
<td>17</td>
</tr>
</tbody>
</table>

Note that /a/ has the highest functional load, representing over 40% of all vowels. This is partly because it accounts for virtually all vowels in the initial syllables of trisyllabic words (§2.3.2.4). /a/ is about 50% more common in closed than in open final syllables.
2.8.2 Consonant frequencies

The table below presents frequencies of occurrence for all consonants. The first column lists overall frequencies, based on the dictionary corpus of 3,012 lexemes. The remaining columns show frequency by position within the word, based on the 1,162 lexemes that satisfy the CVCV(C) template. A dash indicates that the consonant cannot occur in that position.

<table>
<thead>
<tr>
<th>All Words</th>
<th>Words with CVCV(C) Template</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All consonants (3,012 words)</td>
</tr>
<tr>
<td>b</td>
<td>6</td>
</tr>
<tr>
<td>t</td>
<td>10</td>
</tr>
<tr>
<td>d</td>
<td>4</td>
</tr>
<tr>
<td>k</td>
<td>21</td>
</tr>
<tr>
<td>s</td>
<td>9</td>
</tr>
<tr>
<td>h</td>
<td>11</td>
</tr>
<tr>
<td>l</td>
<td>8</td>
</tr>
<tr>
<td>r</td>
<td>8</td>
</tr>
<tr>
<td>m</td>
<td>5</td>
</tr>
<tr>
<td>n</td>
<td>11</td>
</tr>
<tr>
<td>w</td>
<td>1</td>
</tr>
</tbody>
</table>

2.9 Variation

A significant number of words have variant forms, mostly differing by only one consonant or vowel. Much of this variation comes about through dialect mixing, with speakers recognising the source dialect. Other variation involves the use of short forms, including cliticisation.

The most variation is found for word-final consonants. Amongst the most common are a choice between final /n/ and final /k/ (e.g. kedan, kedak ‘immediately’, fatin, fatik ‘place’), final /n/ and no coda (e.g. ulu(n) ‘head’, lalika(n) ‘need not’), or final /k/ and no coda (e.g. hotu(k) ‘all, finished’, kane(k) ‘wound’). Since final /k/ and /n/ are, or at least were, suffixes in some of these words, the present phonological variation could have had its source in different morphemes. An apparently purely phonological variation is between the presence and absence of final /rt/ in some lexemes (e.g. ema(r) ‘person’, tu(r) ‘follow’).

In addition to variation in the word-final consonant, there is also variation in other consonants. Examples are given below. Variation resulting from the merging of the /d/ and /t/ phonemes in the speech of Besikama (fn. 17) is obviously widespread throughout the lexicon.

33 Jonker (1906:286) similarly notes frequent variation between the Tetun of Dili and that of Belu (i.e. the Fehan and Foho dialects) with regard to final consonants.
There is some variation in vowels. One source is the dropping of final vowels, especially when cliticising a word. Such variation is illustrated below. Note that the tendency for the vowel of a clitic to be /a/ (§2.3.2.4) accounts for the variant clitic forms shown for nia and sia. In addition, the relative clause marker ma’ak is usually cliticised as monosyllabic mak, as in the table above.

**Table 2.19: Examples of dropping of final vowels**

<table>
<thead>
<tr>
<th>Final vowel</th>
<th>Citation form</th>
<th>Variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>a/-</td>
<td>ita</td>
<td>it</td>
</tr>
<tr>
<td></td>
<td>keta</td>
<td>ket</td>
</tr>
<tr>
<td></td>
<td>nia</td>
<td>ni, na</td>
</tr>
<tr>
<td></td>
<td>sia</td>
<td>si, sa</td>
</tr>
<tr>
<td>i/-</td>
<td>ami</td>
<td>am</td>
</tr>
<tr>
<td></td>
<td>emi</td>
<td>em</td>
</tr>
<tr>
<td>u/-</td>
<td>atu</td>
<td>at</td>
</tr>
<tr>
<td></td>
<td>batu</td>
<td>bat</td>
</tr>
</tbody>
</table>

Finally, there is for some words variation in the vowel phonemes, as shown in the examples below.
Table 2.20: Examples of variant vowels

<table>
<thead>
<tr>
<th>Variant</th>
<th>Common form</th>
<th>Variant</th>
<th>Comment on variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>a/e</td>
<td>fa’en</td>
<td>fa’an</td>
<td></td>
</tr>
<tr>
<td>a/o</td>
<td>án</td>
<td>cón, ó</td>
<td></td>
</tr>
<tr>
<td>loke</td>
<td>Vt open</td>
<td>lake</td>
<td></td>
</tr>
<tr>
<td>ó</td>
<td>Pro 2S a</td>
<td>a</td>
<td>clitic</td>
</tr>
<tr>
<td>u/a</td>
<td>atu</td>
<td>ata</td>
<td></td>
</tr>
<tr>
<td>w/e</td>
<td>batu</td>
<td>bata</td>
<td></td>
</tr>
<tr>
<td>e/i</td>
<td>ne’an</td>
<td>ni’an</td>
<td></td>
</tr>
</tbody>
</table>

2.10 Morphophonemics

2.10.1 Introduction

In this section the phonological rules that apply at morpheme boundaries are discussed and exemplified.

2.10.2 Suffixation

Consonantal suffixes and enclitics cannot be added to consonant-final bases. This ensures that codas never consist of more than one consonant.

2.10.3 Prefixation

2.10.3.1 Intervocalic /h/-insertion

The consonant /h/ is always added between a vowel-final prefix and a vowel-initial base (which necessarily consists of a foot). This ensures that vowel sequences do not occur across a foot boundary. The rule is formally specified below, followed by examples of its application.

Prefix | Base
--- | ---
Ca- | h V...

34 The /h/-insertion rule is also presented by Saliwangi et al. (1991:43) and Troeboes et al. (1987:65).

In the corpus causative ha- and da- are the only prefixes found to precede vowel-initial roots (of which there are in any case not very many). There are no examples of partial reduplication for vowel-initial roots. Note that da- is found in the corpus in only the two listed derivations; however, the fact that no other Tetun morpheme ends in /h/ is strong evidence that the form of this prefix is da- and not dah-.
2.10.3.2 Cluster-reduction on prefixed base

When a prefix is added to a base with an initial consonant cluster, the cluster-initial /k/ is in some instances deleted. For a discussion of the conditions and for examples see §2.7.5.

2.10.3.3 Avoidance of invalid consonant sequences

There are two circumstances under which the initial consonant of a base is deleted following a consonant-final prefix. Firstly, the consonant /h/ is deleted following any consonant-final prefix; the most common application of this rule in the corpus is for subject marking prefixes. Secondly, the consonant /k/ is deleted following consonant-final derivational prefixes (all of which end in /k/) or the 1S subject marker k-. Although all consonant-final prefixes which can precede a /k/-initial base themselves end in /k/, I tentatively generalise these two rules as follows to apply following any consonant-final prefix. For discussion and examples of the application of this rule see §2.7.6.3.

2.10.4 Partial reduplication

In partial ("Ca- ") reduplication, the onset is copied followed by a constant vowel /a/. Any extrasyllabic /k/ does not participate in the reduplication, and remains in word-initial position.

---

35 There are no confirmed examples in the corpus with vowel-initial roots.
2.10.5 Compounding and full reduplication

2.10.5.1 Word template

Both compounds and full reduplications have some features of a phonological word, and some of two words.\(^{36}\)

The two members each make up a metrical foot. As noted in §2.10.6, reduplication of bases of more than two syllables uses only the final foot. The corpus contains no phonological compounds composed of members of more than two syllables.

Like underived phonological words, primary stress is on the penultimate syllable. Unlike other words of four syllables, there can be a word-initial consonant cluster, and the initial consonant of the second foot can be a glottal stop (inserted by the rule in §2.10.5.3). The word template for non-truncated compounds and full reduplications is presented below, followed by examples.

![Word template](image)

<table>
<thead>
<tr>
<th>RDP/Base</th>
<th>Base</th>
<th>Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RDP</td>
<td>adverbial</td>
<td>kmetis</td>
</tr>
<tr>
<td>RDP</td>
<td>little</td>
<td>oan</td>
</tr>
<tr>
<td>á</td>
<td>?</td>
<td>isin</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RDP/Base</th>
<th>Base</th>
<th>Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>RDP</td>
<td>adverbial</td>
<td>kmeti-metis</td>
</tr>
<tr>
<td>RDP</td>
<td>little</td>
<td>o(a)-'.oan</td>
</tr>
<tr>
<td>á</td>
<td>?</td>
<td>á-.isin</td>
</tr>
</tbody>
</table>

2.10.5.2 Truncation rules

Truncation rules apply at the interface of two bases. The rules are summarised here and discussed below.

---

\(^{36}\) This is not surprising, since in other ways too compounds have characteristics of both syntactic and lexical expressions, with both the components and the result existing as complete words (Bybee 1985b:106).
The following examples illustrate the application of these rules. The first column indicates which rules are applied, while the underlining highlights which parts of the bases are omitted in the derivation.

<table>
<thead>
<tr>
<th>Rules</th>
<th>Base</th>
<th>Base</th>
<th>Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>i, ii</td>
<td>knês</td>
<td>RDP plural</td>
<td>knês</td>
</tr>
<tr>
<td>i, ii</td>
<td>kmesak</td>
<td>RDP very</td>
<td>kmesak</td>
</tr>
<tr>
<td>i, ii</td>
<td>lahaæ</td>
<td>N strand</td>
<td>kwér</td>
</tr>
<tr>
<td>i, ii, iii</td>
<td>batar</td>
<td>N maize</td>
<td>ktomak</td>
</tr>
<tr>
<td>i, ii, iii</td>
<td>ktomak</td>
<td>RDP adverbial</td>
<td>ktomak</td>
</tr>
<tr>
<td>ii</td>
<td>lais</td>
<td>RDP adverbial</td>
<td>lais</td>
</tr>
<tr>
<td>ii</td>
<td>loron</td>
<td>RDP adverbial</td>
<td>loron</td>
</tr>
<tr>
<td>ii, iii</td>
<td>di'ak</td>
<td>RDP adverbial</td>
<td>di'ak</td>
</tr>
<tr>
<td>ii, iii</td>
<td>tebes</td>
<td>RDP adverbial</td>
<td>tebes</td>
</tr>
<tr>
<td>ii, iii</td>
<td>de'ian</td>
<td>RDP heedlessly</td>
<td>de'ian</td>
</tr>
<tr>
<td>iii</td>
<td>ina</td>
<td>N mother</td>
<td>bei</td>
</tr>
<tr>
<td>iii</td>
<td>mama</td>
<td>N betel</td>
<td>fatin</td>
</tr>
<tr>
<td>iii</td>
<td>ai</td>
<td>N wood</td>
<td>rahun</td>
</tr>
</tbody>
</table>

The truncation rules have two sometimes conflicting phonological results. On the one hand, consonant clusters are avoided by the removal of any cluster-initial (extrasyllabic) /kl/ on the second base (rule i), and by the obligatory removal of the coda from the initial one
Clusters are also avoided by the normally optional deletion of the whole of the second (unstressed) syllable of the initial base (rule iii).

On the other hand, the initial base may be truncated only to (C)VC by partial application of rule iii, deleting the V but not the onset of the second syllable. In this case the final C is reinterpreted as the coda of the preceding syllable. This can result in a coda consonant which would be invalid elsewhere, and in a consonant sequence at the meeting of the two bases. In some cases the degree of truncation of the initial base is optional, while in others it appears to be fixed.

2.10.5.3 Glottal stop insertion

Where the initial member in a compound or reduplication is vowel-final and the second is vowel-initial a glottal stop is inserted intervocally, as in the examples below. Note that this glottal stop insertion rule applies only after truncation of the initial base. For instance, uit-uit ‘RDP-little’ = ‘gradually’ is truncated to ui-uit resulting in a vowel sequence, before glottal stop insertion can apply to give the final form ui-'uit'. (A full stop is used to separate glottal stop from base.)

<table>
<thead>
<tr>
<th>Base</th>
<th>Base</th>
<th>Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>...V</td>
<td>?V...</td>
<td>(obligatory within a phonological word, optional at word boundaries)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Base</th>
<th>Base</th>
<th>Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>át</td>
<td>RDP</td>
<td>various</td>
</tr>
<tr>
<td>ida</td>
<td>Num</td>
<td>one</td>
</tr>
<tr>
<td>oj</td>
<td>RDP</td>
<td>various</td>
</tr>
<tr>
<td>ai</td>
<td>N</td>
<td>wood</td>
</tr>
</tbody>
</table>

This rule also applies between vowel-final proclitics and following vowel-initial words (e.g. na ina-n [na'inan] ‘3S mother-GEN’ = ‘his/her mother’). Glottal stops are often inserted intervocally at word boundaries also (e.g. Ne’e utu [ne'e'utu] ‘This (is a) louse’). Glottal stops are only represented in the orthography if they occur within an orthographic word.

2.10.6 Prefixation to and reduplication of bases longer than a foot

Both prefixation and reduplication apply to a metrical foot only. Where a stem has more than two syllables, it is the final foot that is retained. In all such cases in the corpus, the stem is trisyllabic and the initial syllable looks phonologically like a prefix, even where what follows is not synchronically recognisable as a root. For instance, in the examples below, all the bases look like they begin with the prefix ha- or hak-, but what one would expect to be the root is in some cases not a Fehan word (e.g. *nasa).

Note that in some derivations (using ma- -k or reduplication) the /k/ of a word-medial consonant cluster is retained, becoming extrasyllabic, and not omitted along with the rest of

---

37 When words are repeated to indicate iteration, the final consonant of the non-final word(s) is sometimes deleted also (e.g. haké hakés ‘talk talk’ = ‘keep talking’; lora-lora loran -lororon ‘RDP-day RDP-day’ = ‘every day for a long time’).

38 Middelkoop (1950:399) similarly notes that in the neighbouring language of Dawan compounds consisting of two disyllabic members are often reduced to three syllables by means of deletion, with accompanying weakening of the stress on the initial member.
the syllable of which it is a coda (as shown by *maklatik* and *klelelek*). In prefixation by *ha-* , however, the /k/ is omitted (as shown by *hatesi*). To facilitate comparison, the part of the base that is omitted is underlined in the following examples.

```
Prefix/RDP       Affix/RDP       Derivation
   ma- -k   who   haklati  Vi tople    maklatik  Adj which is toppled
   k- -k   result  habelak  Vt flatten   kbelak  Adj flattened
   RDP    heedlessly  hakerek  Vt write   ke-kerek  V write heedlessly
   RDP -n    heedlessly  hanasa  V laugh    nasa-nasan  V laugh for no reason
   RDP    heedlessly  haklelek  Vt abuse   kle-lelek  Vt abuse for no reason
   RDP    NOM  hakés  Vi talk      ka-kés  N talking
   ha- make  haktesi  Vi broken    ha-tesi  Vt break
```

### 2.10.7 Contraction of underived four-syllable words

Deletion of the antepenultimate vowel is optional for underived words of four syllables. As in truncation of derived words, it can result in otherwise unacceptable codas and in consonant clusters.

**4 syllables**

**Contracted form**

- *koroné*      *korné*  N government head [Dutch]
- *bitilili*    *bitilili*  N dragonfly

### 2.10.8 Alternation of ne'e/e'e ‘this’

The determiner *ne'e* ‘this’ optionally takes the form *e'e* following a consonant-final word (e.g. *oras* (n)e'e ‘this time’, *batar* (n)e'e ‘this maize').

```
Word
-----
    ... C ne'e (optional)
    ↓
e'e
```
3 Word classes

3.1 Introduction

This chapter presents an overview of word classes (or 'parts of speech'). Many of the classes will be discussed in more detail in subsequent chapters.

Word classes are determined by a number of criteria. These are the range of syntactic functions they can fulfil, their distribution within syntactic units, the categories which are specified for members of that class (e.g. subject marking for transitive and intransitive verbs), and derivational possibilities. As is to be expected, the various criteria sometimes conflict. In this case I give priority to the syntactic criteria, and note the discrepancies. For some classes, such as numerals, the single criterion of distribution is sufficient to fully define the class, with all other criteria being superfluous.

Open word classes tend to have an associated semantic range (e.g. with nouns including the names of persons, places and things) as well as an associated discourse function (Hopper & Thompson 1984). These tendencies are, however, not sufficiently strong to be useful as diagnostic criteria.

Two large superordinate classes are distinguished. The superordinate class of nominals includes common nouns, proper nouns, time nouns and pronouns. That of verbs includes transitive and intransitive verbs as well as adjectives, copulas and actor-describing verbs. The remaining word classes are numeral, numeral classifier, auxiliary, adverb, determiner, relator (including subordinating conjunctions, prepositions, prepositional verbs, coordinating conjunctions and the relative clause marker), tag and interjection.

Common nouns, proper nouns and three classes of verbs (transitive verbs, intransitive verbs and adjectives) form large open classes. The classes of time nouns and adverbs are much smaller, and it is not clear to what extent they are open. The remaining classes are small and virtually 'closed', although new members have been added to some through borrowing (e.g. the conjunction kalo 'if', and interjection ya 'yes').

3.2 Multiple membership of word classes

Tetun does not have much overlap in membership amongst the various word classes. There are some homonyms with unrelated meanings (e.g. bei N 'ancestor'; Adv 'however'), and some 'everyday' words which have a modified word class and meaning within sea taboo terminology (§1.2.3.4). Apart from these, overlap in class membership falls into two categories.

---

1 This is in contrast to some other eastern Indonesian languages, such as Buru (C. Grimes, pers. comm.) and Tukang Besi (Donohue 1995:82ff.), for which there is said to be a large noun-verb overlap.
In the first, the difference in word class correlates with some difference in meaning. The following combinations of word classes have been found.

1. Verbs and nouns. Most fall into two classes, namely verbs and their associated instrument, undergoer or result nouns (e.g. *luku* Vt, N ‘plough’; §4.5.2.6), or verbs and their associated nominalisation within a possessive NP (e.g. *nia-kan moris* ‘3S-POS live = ’his/her life’; §4.5.1). Other pairs of verbs and nouns with related meanings are:

\[
\begin{array}{lll}
\text{badaen} & \text{Vt} & \text{do work of a craftsman} \\
\text{karian} & \text{Vi} & \text{work} \\
\text{kateri} & \text{Vt} & \text{cut with scissors} \\
\text{kawá} & \text{Vi} & \text{caw (saying ‘kawá’)} \\
\text{kiu} & \text{Vi} & \text{squeak (make sound of chick)} \\
\text{tuku} & \text{Vt} & \text{strike} \\
\text{udan} & \text{Vi} & \text{rain} \\
\text{ukun} & \text{Vt} & \text{rule, govern} \\
\text{uluk} & \text{Vi} & \text{go first}
\end{array}
\]

N craftman
N work
N scissors (from *teri* Vt ‘trim’)
N crow
N chick
N o’clock, hour
N rain
N rule
N former times

2. Numeral classifiers and common nouns. See §3.6 for comment.

3. Verbs, prepositional verbs and prepositions. These are mostly related by grammaticisation (e.g. *bá* ‘go, to, for, at (time)’. Overlap between prepositions and verbs is noted in §8.2, while prepositional verbs, and their relationship to verbs and prepositions, are discussed in §12.5.

4. Other sets of words with the same form and related meanings but different word classes are listed below.

\[
\begin{array}{ll}
\text{ká} & \text{Cnj or (§14.3.3)} \\
\text{lale} & \text{Intj no} \\
\text{hotu} & \text{Pro all (§6.4)} \\
\text{keta} & \text{Aux do not (§10.4.4)} \\
\text{join} & \text{Aux only just (§10.3.3)}
\end{array}
\]

Tag maybe, or not (§14.2.5)
Adv else (§11.11)
Adv after that (§11.11); Vi finished
Adv perhaps (§11.12)
Adv then (§11.11)

In the second category of dual class membership, the only reason for analysing a word as belonging to the two word classes is that it has the syntactic distribution of both. There is (in most cases) no semantic difference between the two distributions. In these instances I have chosen to be a ‘splitter’ rather than a ‘lumper’, analysing these as having dual class membership, rather than as constituting special classes of words which have dual distribution. The alternative approach is of course equally possible. Classes for which such dual membership occur are:

1. Pronouns and determiners (e.g. *ne’e* ‘this’; §6). Note that for *nia* ‘3S, that’ and *sia* ‘3P, PL’, there is a semantic difference between the pronoun and the determiner.

2. Prepositions and subordinating conjunctions (e.g. *hori* ‘since’, *to’o* ‘until’; §8.2)

### 3.3 Nominals

Noun phrases are the constituents that typically fill the roles of subject and object within a clause, and are the only type of constituent which can function as complement of the locative
preposition *ih*. Nominal predicates contrast with other predicates in that they cannot readily be negated by premodifying *la* ‘not’.

All classes of nominals can function as head of an NP, and only nominals can head complete NPs. That is, an NP can only lack a nominal head if the context makes the reference or denotation sufficiently clear (e.g. *naruk sia* ‘long PL’ = ‘long (one)s’). Only nominals can function as possessor within an NP, and only nominals can be modified by numerals, or be marked as definite by the definite article *á*.

With the exception of pronouns, nominals can function as modifier within an NP. In this case the nominal is closer to the head than any other modifiers (e.g. *ema Isra’el uluk ne’e sia* ‘person Israel former.times this PL’ = ‘these former people of Israel’).

Nominals fall into four classes.

1. The class of **common nouns** (e.g. *manu* ‘bird’) is an open class, “in which occurs the names of most persons, places, and things” (Schachter 1985:7). Only common nouns and some time nouns may be possessed.

   **Titles** (e.g. *bi* ‘elder sister’, *liurai* ‘ruler’, *na’i* ‘noble’) are a subclass of common nouns, which can occur before a proper noun in polite terms for people (§7.2.5.2).

   **Deverbal nouns** (e.g. *babót* ‘size’) are common nouns derived from verbs by partial reduplication. They are found almost exclusively as possessed nouns (§7.2.3).

2. **Proper nouns** (e.g. *Bete, Wehali, Hendrikus*) constitute a large open class in which are included the names of specific people and places. Since many names, especially of people, have origins outside the Tetun-speaking area, they frequently do not follow Tetun phonological patterns. Tetun naming conventions are discussed in §7.2.5.1.

3. **Time nouns** (e.g. *loron* ‘day’, *awan* ‘tomorrow’, *rafoni* ‘day after tomorrow’) are unique amongst nouns in that they can function as peripheral constituents without being introduced by a preposition (§9.8.2). In this respect they have the same distribution as both time NPs containing modifiers (e.g. *loron ida* ‘day one’ = ‘one day’) and time adverbs (e.g. *loro-loron* ‘daily’). Nevertheless, time nouns share with other nouns the ability to function as complement of prepositions (e.g. *hort uluk* ‘since former.times’), and to be followed by what at least seems to be the definite article *á* (e.g. *mú ohin á* ‘like just.now DEF’ = ‘like just before’). As such, analysis as nouns is preferable to the more traditional classification of such words as being both nouns and time adverbs (as pointed out by Huddleston (1984:335) for English).

4. **Pronouns** constitute a closed class of nominals which function only as heads of NPs, and not as modifiers. Many pronouns are, however, also classed as determiners, in which capacity they do modify NPs. The four subclasses of pronouns are discussed in Chapter 6.

### 3.4 Verbs

#### 3.4.1 Introduction

Verbs usually function as predicate heads, but can also function as NP modifiers.

Verbs have valency. That is, each verbal lexeme licenses a restricted number and type of arguments (including oblique arguments), over which it in general places selectional
restrictions. All verbs allow a subject. Other arguments are restricted to certain subclasses of verbs.

The major classes of verbs (except copulas) all have members which can be nominalised by partial reduplication (§4.5.1), and members which can be intensified by modifiers of degree (§11.3).

Only predicates headed by verbs are linked by hodi 'COORD' (§14.5.2). Some auxiliaries (e.g. irrealis atu when it means 'want to, about to'; §10.3.6.2) can only be followed by verbs.

There are three open classes of verbs, namely transitive verbs, intransitive verbs and adjectives. Distinctions between them are discussed in §3.4.2. A fourth verb class, of copulas, is small and closed, and is discussed in §3.4.5. In addition, there is a class of derived stative verbs which describe actors; these have a unique syntactic distribution and morphological composition, and are discussed in §4.4.

Note that auxiliaries (§3.7) and prepositional verbs (§3.10) both share characteristics with some verbs. They are not analysed as verbal classes because they cannot occur in a complete clause without an accompanying verb.

### 3.4.2 Distinctions between transitive verbs, intransitive verbs and adjectives

The characteristics of the three main classes of verbs are summarised in the table below. Note that only morphological processes which are at least moderately productive are included. Characteristics marked ‘*’ are shared between adjectives and nouns.

<table>
<thead>
<tr>
<th></th>
<th>Transitive verb</th>
<th>Intransitive verb</th>
<th>Adjective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Used predicatively</td>
<td>Nearly always</td>
<td>Nearly always</td>
<td>Approx. 60%</td>
</tr>
<tr>
<td>Takes subject marking</td>
<td>Yes</td>
<td>Yes</td>
<td>No (see below)</td>
</tr>
<tr>
<td>NP modifier uses relative clause marker</td>
<td>Usual</td>
<td>Usual</td>
<td>Seldom</td>
</tr>
<tr>
<td>Base of ma(k)-actor-describing verb</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Full reduplication means 'plural, varied'</td>
<td>No</td>
<td>No</td>
<td>Yes *</td>
</tr>
<tr>
<td>Result of k-k derivation</td>
<td>No</td>
<td>No</td>
<td>Yes *</td>
</tr>
<tr>
<td>Can begin with kC consonant cluster</td>
<td>No</td>
<td>No</td>
<td>Yes *</td>
</tr>
<tr>
<td>Transitive</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Base of ‘Instrument’ noun</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Base of ha-transitive verb</td>
<td>Seldom</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>2nd verb in nuclear causative series</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

From the table it is evident that intransitive verbs share some properties with transitive verbs, and others with adjectives; they do not have unique characteristics of their own.

There are two main criteria distinguishing the three classes of verbs, from which many of the other properties follow: (i) the tendency to function predicatively, and (ii) transitivity.

---

2 There appear to be no verbs with a valency of zero. Weather terms, which have zero valency in some languages, in Tetun can be expressed either by a noun (e.g. Udan mai 'Rain comes'), or by a verb taking rai 'earth' as subject (e.g. ...rai ã sei la udan 'earth DEF still not rain' = '...it isn't raining yet/the rains haven't started yet').
Transitive and intransitive verbs are unique amongst word classes in that they nearly always function predicatively. This follows from their semantics, in that words denoting actions or processes (transitive and intransitive verbs) are more likely to be comment-worthy, while those denoting more time-stable properties (adjectives) are more useful for restricting reference (Dixon 1994:31). Although adjectives also often function predicatively, they are almost as frequently found as modifiers within NPs.

Since subject marking is restricted to predicates, it is not surprising that the ability to take subject marking correlates with the frequency with which a word class is used predicatively. Transitive and intransitive verbs take subject marking for certain person-number combinations under certain phonological conditions (§9.3.1). Adjectives do not take subject marking in the texts on which this study is based; however, during elicitation consultants differed as to whether such marking was possible (§9.3.3.2). The few words which are synchronically not verbs but which nevertheless take subject marking are discussed in §9.3.3.3.

Another concomitant of the strong preference for predicative function is that when transitive and intransitive verbs are used to modify an NP, they have a strong tendency to be introduced by the relative clause marker mak. In contrast, adjectives (as well as certain other word classes such as numerals and common nouns) tend not to be placed in a relative clause when used attributively.

The second major feature cross-classifying verbs is transitivity. This ability to take NP complements sets transitive verbs apart from other verbs, and unites them with prepositional verbs and prepositions (§3.10). Only transitive verbs can be the base for partial reduplication deriving instrument or undergoer nouns (§4.5.2.2). On the other hand, only intransitive verbs and adjectives are commonly used as the base in transitive causative derivations (§4.2.2), and function as the second element in causative nuclear verb serialisations (§12.2.2).

There are only a few verbs which are classified as both transitive and intransitive. Almost all of these begin with ha-, suggesting that the dual transitivity is a function of the ha- prefix. For a list of these verbs see §4.2.2.3.

### 3.4.3 Further comments on adjectives

An additional distinction between adjectives and the other classes of verbs is that adjectives (like nouns) can begin with a consonant cluster (e.g. kdók ‘far’, ktodan ‘heavy’). In some cases this is the result of affixation by the circumfix k- -k (§4.6). The various parameters identifying adjectives do not exactly match. Exceptions which begin with consonant clusters and which are semantically non-active but which nevertheless are transitive are kre’is ‘close to’ (e.g. Sia kre’is Maromak ‘3P close.to God’ = ‘They are close to God’), któs ‘miserly (with respect to)’ (e.g. Nia któs hemu ‘3S miserly drink’ = ‘He is miserly with respect to giving drink’), and krakat ‘angry (at)’.

Note that most of Dixon’s (1982) semantic groupings of ‘adjectival’ notions can be expressed by adjectives in Tetun. These include value (e.g. di’ak ‘good’, át ‘bad’), age (at least of things; e.g. foun ‘new’), colour (e.g. kmodok ‘yellowish’), dimension (e.g. bót ‘big’, luan ‘wide’), speed (e.g. lais ‘fast’), physical property (e.g. kmeik ‘sharp-pointed’, monas ‘hard’), and human propensity (e.g. matenek ‘clever’, klokok ‘dumb (ignorant, or unable to speak)’).

Some quantity expressions too are adjectives (e.g. wa’in ‘many, much’).

---

3 Many human propensity concepts (e.g. ‘angry’, ‘happy’) are, however, expressed by clausal predicates, such as laran moras ‘interior sick’ = ‘sad’ (§9.6). In addition, some expressions are intransitive verbs, as shown by the fact that they take subject marking (e.g. moe ‘ashamed’, moras ‘sick’).
3.4.4 Further cross-classification of transitive and intransitive verbs

Both transitive and intransitive verbs can be further subclassified according to whether the verb takes oblique arguments such as recipient (e.g. fô Vt 'give') or locative goal (e.g. tau Vt 'put', mai Vi 'come'). In addition, verbs can be classified according to what type of complement clause they accept, if any.

3.4.5 Copulas

Copulas (§9.4.5) are two-place verbs. Unlike transitive verbs, they take a following predicative complement rather than an NP object. The predicative complement may, depending on the copula, be an NP, adjective phrase, or prepositional phrase. Unlike objects of transitive verbs, the complement may not be fronted to before the subject, and selectional restrictions apply between the subject and the predicative complement, rather than between the copula and complement.

3.5 Numerals

Numerals (§5) are unique in that only numerals can follow the classifier in a numeral phrase, fill slots in the formula for a complex numeral, or be terms in arithmetic formulae. In addition, full reduplications of numerals (§4.8.2) have different meanings to reduplication of other word classes.

Like verbs and nominals, numerals can function either attributively or predicatively. In attributive use they are not preceded by the relative clause marker mak, and (except for ida 'one') follow adjectives (§7.2.2.3). In predicative use, which accounts for only 2% of examples in the corpus, numerals do not take subject marking.4

3.6 Numeral classifiers

Numeral classifiers stand immediately before a numeral within a numeral phrase. All classifiers are also members of the class of common nouns. However, as classifiers they do not head NPs, and so have none of the defining properties of common nouns.

There are two classes of numeral classifiers, both of which are discussed in more detail in §5.3.

1. **Sortal classifiers** (e.g. na'in for persons) constitute a small closed class of words used to count whole objects, whose collocations with the words they count are not fully predictable. There is always some semantic relationship between a sortal classifier and the common noun with the same phonological form. So, for instance, tahan 'leaf' is the classifier for things which are, like leaves, thin and flat.

2. **Mensural classifiers** (e.g. daun 'hand (bananas)', tu 'threaded bundle') form a larger class of semantically transparent common nouns which refer to the form of the entities being counted, rather than to the entities themselves.

---

4 Numerals could alternatively be analysed as either a class of adjectives, or a class within some superordinate class which also includes adjectives. However, this would give the classes of the superordinate class very few properties in common. Apart from the above-mentioned unique properties of numerals, numerals differ from verbs in that they don't function as the root of derivations, cannot be intensified, and are not attested in serial verb constructions.
3.7 Auxiliaries

Auxiliaries specify aspect (e.g. irrealis atu) or deontic modality (e.g. bele ‘can’). They are usually found immediately preceding a verb, although some can also occur before the subject. Some, like verbs, allow subject marking. Auxiliaries constitute a diverse class, with members varying widely as to which properties they share with verbs. An alternative to considering auxiliaries as a primary part of speech is therefore to analyse them as a non-prototypical subclass of verbs.

The unique properties of auxiliaries are discussed at length in §10.2.

3.8 Adverbs

Adverbs function as modifiers of constituents other than nouns (Schachter 1985:20). They cannot themselves head predicates, have no valency, and cannot be modified (e.g. negated or intensified). There are very few instances of morphological derivation from adverbial bases. Adverbs do not in general take subject marking; there are, however, a few exceptions, discussed in §9.3.3.3. Some adverbs are derived from bases of other word classes by full reduplication (§4.7).

Subclasses of adverbs are discussed in §11.

3.9 Determiners

Determiners indicate definiteness and/or number, and follow all other NP modifiers with the exception of prepositional phrases and relative clauses. Except for combinations of two determiners from the group ne’e ‘this’, sia ‘PL’ and hotu-hotu ‘all’, an NP can contain only one determiner (§7.2.2.5). Unlike other determiners, which only modify NPs, ne’e ‘this’ can also nominalise a clause (§14.6). The subclasses of determiners are discussed in §6.

3.10 Relators

Relators link two constituents. The first three classes of relators typically introduce peripheral constituents or oblique arguments within the clause.

1. **Subordinating conjunctions** (e.g. surak ‘on condition that’, ne’ebé ‘so that’; §14.7) introduce adverbial clauses, that is, full clauses which have an adverbial function within the superordinate clause. Depending largely on the semantic relation, the clause they introduce may be initial, final, or in some cases internal to the host.

2. **Prepositions** (e.g. iha ‘in, at’, hori ‘since, from’) require an NP complement. The complement may neither be fronted nor omitted, with the result that when the complement of a preposition is relativised, the slot following the preposition is filled by a resumptive pronoun (in contrast to verbs, for which a relativised object is omitted from the relative clause; §14.8.2). Semantically, prepositions have no ‘action’ component.

All prepositions can introduce peripheral constituents within the clause. Some can, in addition, introduce oblique arguments or head predicates. Unlike verbs, they cannot take

---

3 Some prepositions also allow prepositional phrase and/or clausal complements. In this case they are analysed as having dual membership of the class of prepositions and the class of subordinating conjunctions.
subject marking. Prepositional phrases can function as modifiers within NPs, without use of the relative clause marker mak.

3. **Prepositional verbs** (e.g. bō 'to', hodī 'use') introduce oblique arguments and peripheral constituents. They have some characteristics of transitive verbs, and some of prepositions, with the exact mix of characteristics varying considerably. For details see §12.5.1.

4. **Coordinating conjunctions** (e.g. nō 'and', kā 'or'; §14.3) coordinate a range of constituent types, including NPs and clauses. They are placed between the constituents they coordinate.

5. The **relative clause marker** mak introduces relative clauses, linking them into the NP which they postmodify.

### 3.11 Tags

Tags follow the final clause in a sentence, linked to the preceding constituents intonationally, but falling outside the structure of the clause. Each is associated with a characteristic intonation pattern. Tags are listed in §14.2.5.

### 3.12 Interjections

Interjections take no part in clausal syntax, and typically (except in the case of hesitation markers) constitute utterances (or at least intonation groups) on their own. They need not conform to normal Tetun phonological patterns. Semantically there are five subclasses of interjections.

1. **Pro-clauses** typically form single-word responses to a preceding utterance. They include hou (or ou, hō; often with following rising tag é) 'yes' to signal agreement, lale 'no', so’in ‘That’s OK’, and he’e to respond to greetings. In addition ia (or yā) ‘yes’ is sometimes borrowed from Malay ya.

2. **Exclamations** express attitudinal rather than propositional meanings. They are nearly always utterance-initial, but can occur elsewhere as well, usually between (rather than within) clauses. Common exclamations, with very approximate characteristic conditions of use, include adī, inō and wā (all with rising intonation) to express astonishment, aikē (with rise-fall intonation) to express annoyance, and hai and hei to draw attention.

3. **Fillers** too tend to be utterance-initial; however, they appear to function as some kind of discourse marker rather than expressing attitudinal meanings. They are not used in writing, and their meanings are unclear. They include ah (with falling intonation), and ó and ou, which are similar to English oh and often precede answers to questions.

4. **Hesitation markers** (e.g. é, a, nú-é) occur fairly freely within utterances, and appear to fill gaps in speech while the speaker determines what to say. They are not used in writing.

---

6 The exception is the complementiser hōstī ‘about’, which is syntactically a preposition/conjunction but can nevertheless take subject marking, albeit with disputed rules (§13.3.4.2). This is presumed to reflect its diachronic origin as a verb.

7 Note that in addition to these conjunctions, the verb hodī functions as a relator of clauses (§14.5.2), without losing its classification as a verb.
5. **Calls to animals** give instructions to animals. Some are roots of verbal derivation with *ha(k)*- (-k) (§4.3.5). Calls for animals to come include *krr* or *krú* for chickens, *mé* for goats, *hi hi hi hi* for horses, *œ* for buffalo, *é* for dogs, and *ó* (in conjunction with a name) for pigs. *huít* is a call to instruct hunting dogs to start chasing prey.
4 Derivational morphology

4.1 Overview


There are three affixes which derive verbs. One is ha-, which productively derives causative verbs from verbs and nouns. Another is hak-, which has as its most productive function the derivation of intransitive verbs from transitive verb bases. Finally, there is mak- (-n), with a variant ma- (-k), which derives a unique class of actor-describing verbs from other verbs.

There are seven affixes which derive nouns and adjectives. These affixes overlap considerably in function, as shown in the table below. Somewhat productive uses of affixes are indicated by ‘Y’, while apparently unproductive ones are listed as ‘(Y)’.

<table>
<thead>
<tr>
<th>Base</th>
<th>Derivation</th>
<th>Ca- -k/n</th>
<th>k(a)- -k</th>
<th>k(a)- (-k)</th>
<th>ma- (-k)</th>
<th>-k</th>
<th>-n</th>
<th>-s</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>N Property</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>N Instrument/object</td>
<td>Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Num</td>
<td>N Time</td>
<td></td>
<td></td>
<td></td>
<td>(Y)</td>
<td>(Y)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>Adj Result</td>
<td></td>
<td></td>
<td></td>
<td>(Y)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>Adj Quality of actor</td>
<td>(Y)</td>
<td></td>
<td></td>
<td>(Y)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adj/N</td>
<td>Adj Similarity to base</td>
<td>(Y)</td>
<td></td>
<td></td>
<td>(Y)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>Adj Having base</td>
<td>(Y)</td>
<td></td>
<td></td>
<td>(Y)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V/Num</td>
<td>Adj Attributive</td>
<td></td>
<td></td>
<td></td>
<td>(Y)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As indicated in the table, somewhat productive affixes deriving nouns are k(a)- (-k) and partial reduplication (‘Ca-’) to derive instrument or object nouns from verbs, partial reduplication to derive abstract property nouns from verbs, and ma- and ma- -k to derive time nouns from numerals. Nouns can also be derived by compounding. The only productive derivation of adjectives is via the circumfix k- -k.

Derivation of adverbs is by means of full reduplication of bases. Other functions of full reduplication are discussed in this chapter also, although they are on the border between derivational morphology and syntax.
In this chapter affixes are primarily sorted by the part of speech of the resulting derivation, with the more productive affixes discussed first within each section. Full reduplication which does not change word class is discussed towards the end of the chapter, followed by highly fossilised affixes which do not semantically fit into any of the earlier groupings.

Certain layout conventions are relevant only to this chapter. In the lists of examples, the base is presented first followed by the derived word. Suffixes are separated by hyphens only if they are part of circumfixes (and hence not applicable in all cases), so that the reader can readily ascertain which of the examples in the list have such a suffix. If no mention is made of the number of examples, this indicates that the listed examples represent all clear examples in the corpus.

In interlinearised examples the two parts of a circumfix, as well as their single English gloss, are surrounded by angular brackets (<>). Square brackets in interlinearised examples surround parts of the base which do not surface in the derivation. So, for instance, <ma> [ha]nanu<k> ‘<who>sings’ is derived from ma- -k and hananu ‘sing’ but does not phonologically contain the bracketed /ha/.

4.2 Derivation of verbs: ha-

4.2.1 Introduction

The causative prefix ha- is highly productive, and applies to verbal and nominal bases.¹

4.2.2 Causative verb from verbal base: ha-

4.2.2.1 Causative verb from underived base

The prefix ha- derives transitive causative verbs from adjectives (over 70 examples) and intransitive verbs (40 examples). The object of the derived verb corresponds to the subject of the root (i.e. S → O), while the derivation acquires a new actor subject.

The spontaneous recorded derivation haruki, from the reasonably frequent Malay borrowing rugi (assimilated as ruki) ‘suffer loss’, illustrates the productivity of this prefix.

<table>
<thead>
<tr>
<th>Adj/Vi</th>
<th>Vt</th>
</tr>
</thead>
<tbody>
<tr>
<td>át</td>
<td>hahát</td>
</tr>
<tr>
<td>be'o</td>
<td>habe'o</td>
</tr>
<tr>
<td>butan</td>
<td>habutan</td>
</tr>
<tr>
<td>kbadak</td>
<td>habadak</td>
</tr>
<tr>
<td>kdór</td>
<td>hadór</td>
</tr>
<tr>
<td>kmukit</td>
<td>hamukit</td>
</tr>
<tr>
<td>lakon</td>
<td>halakon</td>
</tr>
<tr>
<td>lôs</td>
<td>halôs</td>
</tr>
<tr>
<td>mate</td>
<td>hamate</td>
</tr>
</tbody>
</table>

¹ Monteiro (1985:xiii), Saliwangi et al. (1991) and Troeboes et al. (1987) all discuss this prefix in some detail, and reach much the same conclusions as are presented here. Both Saliwangi et al. and Troeboes et al. list the various inflections of ha- (e.g. 1S ka-, 2S ma-) as if they are separate prefixes, and consequently intermingle causative meanings with the meanings of the non-causative prefixes ka- and ma-.
Chapter 4

<table>
<thead>
<tr>
<th>Adj/Vi</th>
<th>Vt</th>
</tr>
</thead>
<tbody>
<tr>
<td>metan</td>
<td>hametan</td>
</tr>
<tr>
<td>moe</td>
<td>hamoe</td>
</tr>
<tr>
<td>mós</td>
<td>hamós</td>
</tr>
<tr>
<td>rani</td>
<td>harani</td>
</tr>
<tr>
<td>sa’e</td>
<td>hasa’e</td>
</tr>
<tr>
<td>sai</td>
<td>hasai</td>
</tr>
<tr>
<td>tasa</td>
<td>hatasa</td>
</tr>
<tr>
<td>to’o</td>
<td>hato’o</td>
</tr>
<tr>
<td>toba</td>
<td>hatoba</td>
</tr>
<tr>
<td>tubu</td>
<td>hatubu</td>
</tr>
</tbody>
</table>

4.1 Niá tun n-ika bá rai kłaran.
3S descend 3S-back go earth middle
She descended back to earth (from heaven). (P2.122)

4.2 Ita ha-tún kair bá wé.
1PI make-descend fish.line go water
We lower the fishing line into the water. (H0.36)

Derivation from transitive verb roots is far less common. The undergoer of the root verb is either omitted (as in example 4.3) or introduced by a serial verb (4.5), with the result that the causative verb is transitive rather than ditransitive.

Vt: root

koma  craving
kre’is close to
memi  say (a name)
tama  entering

Vt: derivation

hakoma  make (s.o.) crave (s.th.)
hare’is  bring (s.th.) closer; come closer
hamemi  name (a child)
hatama  insert

4.3 Nia n-ák “lale” bat n-oí n-a-koma Malae.
3S 3S-say no so.that 3S-COORD 3S-make-crave non.native
He said “no” in order to make the Chinese strongly desire (to buy his dog). (T0.122)

4.4 Karau tama to’os...
buffalo enter garden
Buffalo entered the garden. [So I went to guard it.] (O2.61)

4.5 ...ha-tama Bita Nahak bá iha tafatik bá...
make-enter Bita Nahak go LOC noble.house go
[They] took Bita Nahak over into the noble house... (AA1.107)

4.2.2.2 Transitive verb from derived stem

There are a number of cases in which the stem for the causative derivation is itself derived. In this case the initial syllable of the stem, which constitutes a prefix or reduplication, is omitted when ha- is added (§2.10.6). The presumed root of the stem need not (synchronically at least) be a word.

The following examples illustrate derivation from stems with the adjectival prefix ka-, and from a reduplicated intransitive verb nó-nók whose unreduplicated form is not synchronically recognised as a word in the Fehan dialect.
Derivational morphology

Root: Adj/Vi: stem | Vt: derivation
--- | ---
le'u | Vt coil | kabelak | flat | habelak | flatten
*nök | Vt be quiet | nonök | be quiet | hanök | quieten

A prefix more frequently found in stems of ha- derivations is hak- (15 examples; §4.3).2 Since the roots of hak- derivations are themselves transitive verbs, the causative verb is often close in meaning to the root, where one exists. The three columns in the following list present the transitive root of the hak- stem (where one exists), the intransitive hak- stem, and the derived transitive ha- verb. Note that where the root is not recognised as a word by Fehan consultants (indicated by an asterisk), it is in several cases listed in Morris’s (1984b) dictionary of East Timorese Tetun, his glosses being prefaced by his initials ‘CM:’.

Vt: root | Vi: stem (hak-) | Vt: derivation
--- | --- | ---
doko | rock (e.g. baby) | hakdoko | be rocked, tremble | hadoko | rock (e.g. baby)
*laik | daklaik | partly dry | halai | partly dry
*lati | haklati | topple | halati | lay down horizontally
*leka | CM: overturn | hakleka | face-upward | haleka | put face-upward
*nana | CM: open | haknana | open | hanana | open
silu | snap | haksilu | snap | hasilu | snap
sira | tear | haksira | tear | hasira | tear
tesi | break | haktesi | broken | hatesi | break
tutan | join (pieces) | hatutun | connected | hatutan | pass on (words)

4.6 Ita hawai haklaik na’an lai. Na’an n-aklaik.
1PI dry partly.dry meat now meat 3S-partly.dry
We first dry partly-dry the meat. (So) the meat is partly dry. (Q0.58 elicited)

There is a residue of transitive verbs which appear to have very similar meanings to their transitive roots, rather than being causative.

Vt: root | Vt: derivation
--- | ---
fa’e | divide (any number) | hafa’e | split in two (of midnight)
fe’u | kick | hafetu | kick
ho’uk | agree to (short form of haho’uk) | haho’uk | agree to
lolo | stretch out, pass | halolo | straighten, correct
talik | entwine, cross (limbs) | hatalik | tie/twist (esp. ropes) together
tuda | throw spear at | hatuda | fight war
tudu | nominate, select, point at | hatudu | show, demonstrate, point out

It is possible that some of these words are derived via intransitive hak- stems of which I am unaware; however, those hak- derivations which are found in the corpus (hakfa’ek ‘split in two’ and haklolo ‘stretch forward (legs)’) are transitive, and so do not qualify as mediating stems. In the case of haho’uk ‘agree to’, the apparent root ho’uk is often held by consultants to be a colloquial short form of the verb.

---

2 Mathijsen (1906:xii), in contrast, says that hak- words are derived from ha- stems.
4.2.2.3 Intransitive verb

There are some verbs beginning with *ha-* which are intransitive, and have no causative semantic component. The following examples have a very similar meaning to the apparent base.

**Base**  
**Vi**

<table>
<thead>
<tr>
<th>Base</th>
<th>Adj</th>
<th>last</th>
<th><strong>Vi</strong></th>
<th>Adj</th>
<th>be quiet (reduplication)</th>
<th><strong>Vi</strong></th>
<th>be quiet (not 'quieten')</th>
</tr>
</thead>
<tbody>
<tr>
<td>ikus</td>
<td>Adj</td>
<td>last</td>
<td>hahikus</td>
<td>go last</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nanotok</td>
<td>Vi</td>
<td>be quiet</td>
<td>hanotok</td>
<td>be quiet</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Some of the intransitive *ha-*initial verbs in the corpus begin with a *ham-* prefix in Cliff Morris's (1984b) dictionary of East Timorese Tetun. This suggests that at least some intransitive derivations with *ha-* have come about as a result of a diachronic merging of *ha-* and *ham-* in the Fehan dialect, a merging which would be consistent with the lack of sonorant-initial consonant clusters in this dialect. The following examples are selected because the Fehan dialect uses *ha-* for both a causative and an intransitive (undergoer subject) verb, while Morris lists *ha-* for the causative and *ham-* for the intransitive.

**Vi: Fehan**  
**Vi: CM**  
**Vt**

<table>
<thead>
<tr>
<th>Vi: Fehan</th>
<th>Vi: CM</th>
<th>Vt</th>
</tr>
</thead>
</table>
| hanok     | hamnok | hanok 
| harik     | hamrik | harik |
| haro      | hamro  | haro |
| hata'uk   | hamta'uk | hata'uk |

4.7  
...ambei á n-arik iha ai hún bót ida.  
man DEF 3S-stand LOC plant origin big one  
...the man is standing at a big tree trunk.  

4.8  
...foin n-arik Uma Kfau Lulik.  
then 3S-erect house k.o.tree taboo  
[He brought our noble ancestor back, then] established the House of Kfau Lulik.  

(K6.9)  
(E0.9)

There are a number of other verbs (not listed under *ham-* in Morris's dictionary) for which the same form has both a causative and a non-causative (usually intransitive) meaning. In fact almost all verbs showing this transitivity pattern begin with /ha/, indicating that it is a function of the *ha-* prefix. The examples below list the intransitive and causative verbs, preceded by the ultimate root, where that is known. For *hamutu*, the apparent root *mutu* is often held by consultants to be a short form of the longer verb; however, there are syntactic differences between the two, with *mutu* being an adverb and *hamutu* a verb.

**Root**  
**Vi**  
**Vt**

<table>
<thead>
<tr>
<th>Root</th>
<th>N</th>
<th>room</th>
<th>Vi</th>
<th>Adj</th>
<th>light</th>
<th>Vt</th>
<th>black</th>
<th>Adv together</th>
<th>Vt</th>
<th>together</th>
</tr>
</thead>
<tbody>
<tr>
<td>ke'an</td>
<td>N</td>
<td>room</td>
<td>hak' an</td>
<td>be separate</td>
<td>hak' an</td>
<td>separate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kotu</td>
<td>Vi</td>
<td>finished</td>
<td>hakotu</td>
<td>finished</td>
<td>hakotu</td>
<td>conclude</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kroman</td>
<td>Adj</td>
<td>light</td>
<td>haroma/</td>
<td>become light</td>
<td>haroma/</td>
<td>make light</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>hakroman</td>
<td></td>
<td>hakroman</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>hakusak</td>
<td>crumpled</td>
<td>hakusak</td>
<td>crumple</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>hametan</td>
<td>bruised</td>
<td>hametan</td>
<td>blacken</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mutu</td>
<td>Adj</td>
<td>black</td>
<td>hamutu</td>
<td>together</td>
<td>hamutu</td>
<td>cause to be together</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>haris</td>
<td>bathe</td>
<td>haris</td>
<td>bathe (s.o.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
While the non-causative verbs in the above examples are intransitive, the corpus also contains several pairs in which the non-causative verb is itself transitive.

While the non-causative verbs in the above examples are intransitive, the corpus also contains several pairs in which the non-causative verb is itself transitive.

### 4.2.2.4 Omission of final consonant

There are several derived verbs in which the final consonant of the base is omitted. These are presented below along with the consonant-final derivation (if any) with which there is a semantic contrast. For several there is evidence that the alleged base itself has, or at least diachronically had, a suffix which is not applicable to the causative verb. Thus *nesan* 'same' is related to both *hanesa* 'put in order' and *knesak* 'same', both of which phonologically suggest a root *nesa*. Similarly, the alleged base *kmamuk* 'empty' itself looks as if it contains the adjectival circumfix *k- -k*.  

<table>
<thead>
<tr>
<th>Adj</th>
<th>Vt</th>
<th>Vt</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>di’ak</em></td>
<td>good</td>
<td><em>hadi’a</em> improve, fix</td>
</tr>
<tr>
<td><em>foun</em></td>
<td>new</td>
<td><em>hadi’ak</em> reconcile, improve (health)</td>
</tr>
<tr>
<td><em>kmamuk</em></td>
<td>empty</td>
<td><em>hafou</em> renew</td>
</tr>
<tr>
<td><em>nesan</em></td>
<td>same</td>
<td><em>hamamu</em> empty in one go</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>hamamuk</em> empty bit by bit</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>hanesa(n)</em> put in order, make the same</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>hanesan</em> Prep resemble, be like</td>
</tr>
</tbody>
</table>

### 4.2.3 Verb from nominal base: ha-

When applied to nominal bases, the prefix *ha-* derives both intransitive and transitive verbs. All such bases phonologically comprise a single metrical foot (i.e. they are disyllabic).

As may be observed in the examples below, the final /n/ of nominal bases is in many cases omitted in the causative derivation. The omission applies for all body parts, location terms and kinship terms, for which the /n/ is a (sometimes fossilised) marker of inalienable possession (§7.3.3.3). For bases of other semantic classes it is not clear what conditions the omission or retention of the /n/. However, it is likely that those nouns whose final /n/ is omitted in the causative derivation diachronically had a genitive -n or a nominal -n suffix.  

3 According to Saliwangi et al. (1991:45) base-final /k/ and /n/ are both potentially omitted in causative derivations. Omission of /k/ has not been observed in Fehan nominal bases, but limited evidence of it has been found for adjectival bases of causative derivations (§4.2.2.4).
Intransitive verbs derived by *ha-* refer to producing that which is denoted by the base (20 examples).

<table>
<thead>
<tr>
<th>N</th>
<th>Vi</th>
</tr>
</thead>
<tbody>
<tr>
<td>abut</td>
<td>hahabut</td>
</tr>
<tr>
<td>alas</td>
<td>hahalas</td>
</tr>
<tr>
<td>fulin</td>
<td>hafuli</td>
</tr>
<tr>
<td>kbón</td>
<td>hakkó</td>
</tr>
<tr>
<td>kdahur</td>
<td>hakkdahur</td>
</tr>
<tr>
<td>kliran</td>
<td>haktiran</td>
</tr>
<tr>
<td>oan</td>
<td>hasoru</td>
</tr>
<tr>
<td>sorun</td>
<td>hatolu</td>
</tr>
</tbody>
</table>

4.9 *Manu inan koe bat nia at* n-a-tolu.

bird female cackle so.that 3S IRR 3S-make-egg

The hen cackles to (announce that) it will lay an egg. (N0.98)

Transitive causative derivations from nominal bases relate the meaning of the base in some way to that of the undergoer. The most common relations (over 30 examples) are making the undergoer become (like) the base, or causing the undergoer to have the base.

<table>
<thead>
<tr>
<th>N</th>
<th>Vt: make (like)</th>
</tr>
</thead>
<tbody>
<tr>
<td>kdan</td>
<td>hakkán stack, put in a pile</td>
</tr>
<tr>
<td>kiak</td>
<td>hakkiaik make an orphan</td>
</tr>
<tr>
<td>klaut</td>
<td>haklaut cut (betel/cassava) into slices</td>
</tr>
<tr>
<td>liman</td>
<td>haliman tie corn cobs into one 'hand'</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N</th>
<th>Vt: cause to have</th>
</tr>
</thead>
<tbody>
<tr>
<td>folin</td>
<td>hafoli respect</td>
</tr>
<tr>
<td>futar</td>
<td>hafutar adorn</td>
</tr>
<tr>
<td>hán</td>
<td>hahán feed, bribe</td>
</tr>
<tr>
<td>kneter</td>
<td>hakneter honour</td>
</tr>
<tr>
<td>lét</td>
<td>halét make far apart</td>
</tr>
<tr>
<td>naran</td>
<td>hanaran name</td>
</tr>
</tbody>
</table>

4.10 *Laho n-a-bón rai ne’e kle’an basuk.*

rat 3S-make-hole earth this deep very

The rat burrows very deeply into the ground. (J0.13)

Other derivations refer to using that which is denoted by the base as an instrument.

<table>
<thead>
<tr>
<th>N</th>
<th>Vt</th>
</tr>
</thead>
<tbody>
<tr>
<td>di’a</td>
<td>hadi’a trap (wild animals, birds)</td>
</tr>
<tr>
<td>kair</td>
<td>hakair fish with fishing line</td>
</tr>
<tr>
<td>kelun</td>
<td>hakelu carry against lower arm</td>
</tr>
<tr>
<td>kliiun</td>
<td>haklii carry by strap over shoulder</td>
</tr>
<tr>
<td>kuhus</td>
<td>hakuhus cook in a steamer</td>
</tr>
<tr>
<td>ta’u</td>
<td>hatu dye (cotton) in mud</td>
</tr>
<tr>
<td>tali(n)</td>
<td>hatali tie together</td>
</tr>
</tbody>
</table>
Some derived verbs refer to using the denotatum of the base as a container, into which the undergoer is put.

N

<table>
<thead>
<tr>
<th>N</th>
<th>Vt</th>
</tr>
</thead>
<tbody>
<tr>
<td>kbonan</td>
<td>front of sarong/shirt</td>
</tr>
<tr>
<td>klu'ak</td>
<td>hole</td>
</tr>
<tr>
<td>knuan</td>
<td>sheath</td>
</tr>
<tr>
<td>ktuik</td>
<td>row of plants</td>
</tr>
</tbody>
</table>

4.12  

Ami  bái ha-tuik  fore.  
1PE  go make-row(plants) legume  
We're going to plant beans (in rows).  

Yet others refer to treating someone or something as one would treat the denotata of the base (over 10 examples). The undergoer is either presented as a direct object (example 4.13) or, more frequently, introduced by a serial verb (4.14).  

Finally, there is a range of other relationships, showing that the meaning of these derivations is not quite predictable.

---

4  Data are insufficient to determine whether all these derivations can be used transitively.

The only example in the corpus of any derivation from a compound base is ha-feto-malun 'treat as sister'. The stem feto-malun is an inherently reciprocal kinship term referring to the relationship between two sisters (or classificatory sisters, such as wives of two brothers, or daughters of two sisters). It derives from feto 'female, woman' and malun. The latter is not a free lexeme, but does occur in the compound mane-malun, referring to the same relationship between brothers.

5  Note that hakiak can also mean 'cause to be an orphan', showing that the prefix ha- can be used with two distinct meanings on the one base.
4.3 Derivation of verbs: hak-

4.3.1 Introduction

The verbal prefix hak- has four separate uses. Without an accompanying suffix it derives intransitive undergoer-subject verbs from transitive verbs. A second use is to derive verbs from other verbs with no change in transitivity and little change in meaning; in this case hak- is sometimes accompanied by the suffix -k. When it derives verbs indicating reciprocal action from transitive roots, hak- is always accompanied by -k. Finally, hak- (or ha-), with or without a suffix -k, derives verbs referring to the making of sounds from roots which indicate the sound that is made.

4.3.2 Intransitive verb from transitive base: hak-

The most productive function of the prefix hak- is to derive intransitive verbs with undergoer subjects from transitive verb bases (over 30 examples). The subject of the derivation thus corresponds with the object of the base (i.e. O → S). This function of hak- is widely recognised in the literature on Tetun (Mathijsen 1906:xii; Monteiro 1985:xiv; Saliwangi et al. 1991:25; Troeboes et al. 1987:43).

Vt

<table>
<thead>
<tr>
<th>dalan</th>
<th>road</th>
</tr>
</thead>
<tbody>
<tr>
<td>karin</td>
<td>small change (money)</td>
</tr>
<tr>
<td>knau</td>
<td>second planting season</td>
</tr>
<tr>
<td>lamak</td>
<td>festival meat given to noble</td>
</tr>
</tbody>
</table>

Vt

<table>
<thead>
<tr>
<th>hadalan</th>
<th>explain; mark a route</th>
</tr>
</thead>
<tbody>
<tr>
<td>hakarin</td>
<td>give small change (for money)</td>
</tr>
<tr>
<td>haknau</td>
<td>plant before 2nd planting season</td>
</tr>
<tr>
<td>halamak</td>
<td>eat [noble term]</td>
</tr>
</tbody>
</table>

Vt

<table>
<thead>
<tr>
<th>bois</th>
<th>squeeze out with hand</th>
</tr>
</thead>
<tbody>
<tr>
<td>budi</td>
<td>let slip</td>
</tr>
<tr>
<td>fakar</td>
<td>tip out</td>
</tr>
<tr>
<td>fahu</td>
<td>overturn, turn</td>
</tr>
<tr>
<td>fera</td>
<td>split</td>
</tr>
<tr>
<td>feur</td>
<td>spin, rotate</td>
</tr>
<tr>
<td>fokit</td>
<td>jerk up</td>
</tr>
<tr>
<td>fota</td>
<td>hit</td>
</tr>
<tr>
<td>loke</td>
<td>open</td>
</tr>
<tr>
<td>losu</td>
<td>extract</td>
</tr>
</tbody>
</table>

Vt

<table>
<thead>
<tr>
<th>hakbois</th>
<th>escape from one's grasp</th>
</tr>
</thead>
<tbody>
<tr>
<td>hkbudi</td>
<td>fly away, slip away</td>
</tr>
<tr>
<td>hakfakar</td>
<td>spill out</td>
</tr>
<tr>
<td>hakfalu</td>
<td>capsize, overturn</td>
</tr>
<tr>
<td>hakfera</td>
<td>partially split</td>
</tr>
<tr>
<td>hakfeur</td>
<td>spin, rotate</td>
</tr>
<tr>
<td>hakfokit</td>
<td>freed (e.g. from being bogged)</td>
</tr>
<tr>
<td>hakfota</td>
<td>show chop marks, be cracked</td>
</tr>
<tr>
<td>hakloke</td>
<td>open</td>
</tr>
<tr>
<td>haklosu</td>
<td>suddenly extracted</td>
</tr>
</tbody>
</table>

---

6 This is what Comrie (1985:325) calls an 'anticausative' derivation.

In the literature on Tetun, a number of analyses list hak- as having causative meanings (Monteiro 1985:xiv–xv; Morris 1984:xxviii; Saliwangi et al. 1991; Troeboes et al. 1987:42ff.). There is no evidence for this in the Fehan dialect. Some of the causative hak- examples given in the literature have no /k/ in Fehan (e.g. Monteiro lists hakmetan 'blacken' where Fehan has hametan, and Troeboes et al. list hakas 'heighten' where Fehan has hakas). In others the /k/ comes from the base rather than from the prefix (e.g. Monteiro's hakbaluk is morphologically ha-kbaluk 'make into strips').

Several reports also say that hak- can derive verbs from nouns, with a meaning such as 'become like base' (Monteiro 1985:xiv–xv; Saliwangi et al. 1991; Troeboes et al. 1987:42ff.). Again the meaning, and several of the examples, correspond to the Fehan prefix ha-. For instance, Saliwangi et al.'s haksusu 'suck' is hasusu in Fehan, while Monteiro's hakisi 'freeze, cause to coagulate' in Fehan is hahisi. The only example in the corpus which possibly supports a 'become like base' analysis for hak- is the lexeme pair hakfunan 'mouldly' and funan 'flower'.

I have no evidence for the prefix variant haka- which occurs in examples in de Araújo e Corte-Real's (1990:71) description of Dili Tetun.
Derivational morphology

<table>
<thead>
<tr>
<th>Vt</th>
<th>Vi</th>
</tr>
</thead>
<tbody>
<tr>
<td>saki</td>
<td>haksaki</td>
</tr>
<tr>
<td>silu</td>
<td>haksilu</td>
</tr>
<tr>
<td>sira</td>
<td>haksira</td>
</tr>
<tr>
<td>sobu</td>
<td>haksobu</td>
</tr>
<tr>
<td>tesi</td>
<td>haktesi</td>
</tr>
<tr>
<td>tutan</td>
<td>haktutan</td>
</tr>
<tr>
<td>cut lengthways</td>
<td>split in two</td>
</tr>
<tr>
<td>break off</td>
<td>break (e.g. bone)</td>
</tr>
<tr>
<td>tear</td>
<td>tear</td>
</tr>
<tr>
<td>demolish</td>
<td>collapse</td>
</tr>
<tr>
<td>break, chop</td>
<td>break</td>
</tr>
<tr>
<td>connect</td>
<td>connected</td>
</tr>
</tbody>
</table>

4.15  Ita feur ai ka-feur. = Ita halo ai ka-feur  
1PI spin wood NOM-spin 1PI make wood NOM-spin

n-ak-feur.
3S-INTR-spin
We spin the top (a spinning toy). = We make the top spin.  (N0.29)

4.16  Nia monu, kidan n-ak-tesi.
3S fall backside.bones 3S-INTR-break
(When) he fell, the bone(s) in his backside broke.  (I0.75)

4.17  Ita habit liman n-ak-tesi. hodi ai rua.
1PI squeeze arm 3S-INTR-break use wood two
We splint (lit. 'squeeze between two things') a broken arm, using two pieces of wood.

4.18  Ai ne' e n-ak-so'i. foin ita ha'ak "ai n-awia".  
wood this 3S-INTR-break.off then 1PI say wood 3S-dangle
Only when the branch is partly cut through (e.g. by a knife or by the wind) do we say (the) "branch dangles".  (T0.41)

4.19  Baliun isin n-ak-losu. kati kona o.
axe body 3S-INTR-extract fly.up touch 2S
(If) the axe head comes off, (it will) fly over and hit you.  (Q0.86)

There is no implication that an actor is involved. Thus, for instance, haksira 'tear' can be used regardless of whether the undergoer tore of its own accord, or was torn on purpose by someone. The warning in example 4.19 is that the axe head may fly off while the speaker is chopping wood; here there is no implication that there is an agent removing the axe head.

There is not a consistent aspectual interpretation. As illustrated in the examples above, the verb can refer to a continuing process (4.15), a punctual event (4.16, 4.19), or a resulting state (4.17, 4.18). In addition, several of the derivations from cutting verbs were explained as showing evidence of having been cut, without the cutting having been complete (e.g. hakso'i 'partly severed', hakfota 'show chop marks, cracked', hakfera 'partially split').

7 Where hak- derives intransitive verbs, the literature proposes a range of implied meanings, somewhat similar to those often proposed for Indonesian ter-. These include happening of its own accord (Monteiro 1985:xiv–xv; Saliwangi et al. 1991), happening in an unexpected or unintentional manner (Saliwangi et al. 1991), and happening unintentionally, being done for someone else, or being done for oneself (Troboes et al. 1987:45f.). I do not believe these are inherent to the meaning of the derivations, but accept that some of these implications can come about in context as a result of no actor being specified.
4.3.3 No change in transitivity: hak- (-k)

In addition, hak- can derive verbs from other verbs without a change in transitivity. The derivation has either a very similar meaning to the base or a semantically more restricted one. It is not clear whether this derivation is synchronically productive. Some derivations with vowel-final bases have a -k suffix, which is separated below by a hyphen for easier recognition. It appears that the -k is applied for transitive verbs only (the only exception being hakbaku); however, this could well be an accident of the limited data.¹

<table>
<thead>
<tr>
<th>Vi: base</th>
<th>Vi</th>
<th>Vt: base</th>
<th>Vt</th>
</tr>
</thead>
<tbody>
<tr>
<td>lén</td>
<td>haklén</td>
<td>baken</td>
<td>drape (over anything)</td>
</tr>
<tr>
<td>ro’an</td>
<td>hakro’an</td>
<td>baku</td>
<td>strike [Suai]; overturn</td>
</tr>
<tr>
<td>sala</td>
<td>haksala</td>
<td>basak</td>
<td>throw (at); collide with</td>
</tr>
<tr>
<td>sés</td>
<td>haksés</td>
<td>fa’e</td>
<td>divide (any number)</td>
</tr>
<tr>
<td>siku</td>
<td>haksiku</td>
<td>habit</td>
<td>squeeze between two things</td>
</tr>
<tr>
<td></td>
<td></td>
<td>kahi</td>
<td>restrain verbally; massage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>kasu</td>
<td>perform healing ceremony</td>
</tr>
<tr>
<td></td>
<td></td>
<td>lás</td>
<td>wave (arms) sideways</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sera</td>
<td>lay down (s.th.) in front of</td>
</tr>
</tbody>
</table>

Vt: base

<table>
<thead>
<tr>
<th>Vt</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>hakbaken</td>
<td>strike over shoulder</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hakbaku</td>
<td>strike [Fehan]; overturn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hakbasak</td>
<td>throw at; collide with; flap wings at</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hakfa’e-k</td>
<td>separate into two</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hakabit</td>
<td>carry under one arm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hakahí-k</td>
<td>restrain verbally</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hakasú-k</td>
<td>perform healing ceremony</td>
<td></td>
<td></td>
</tr>
<tr>
<td>haklás</td>
<td>shoo out (e.g. fan, chase away)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>haksera-k</td>
<td>present (an offering)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.20 Na'i feto tasi sa’e ti’a, n-a’ak atu baku rahu

Na’i feto tasi sa’e ti’a, n-a’ak atu baku rahu

The noble woman having got angry (lit. ‘sea ascend’), said she would strike this flute to pieces. (G2.40 from Suai)

4.21 Nia ti’a, r-ak-baku: r-ak-baku ferik e’e.

Then, they hit: hit the old lady. (R4.51)

4.3.4 Reciprocal action verb: hak- -k

A third use of hak- is to derive intransitive verbs which are inherently reciprocal from transitive bases. For this the suffix -k is always used if the base ends in a vowel. Since the

¹ Some intransitive verbs appear on the surface to derive by hak- (or ha-) from adjectives beginning with /k/ (e.g. hakbakar ‘obese’, kbakar ‘obese’). However, it seems likely that both the hak- verb and the k- adjective derive from a common transitive verb root for which there is no record in the corpus, or which in some cases no longer exists as a semantically related word. An example for which the common transitive root is known is haksira Vi ‘tear’ and kisirak Adj ‘torn’, which both derive from sira Vi ‘tear’.
verbs denote reciprocal action, the subject must be plural. There is some variation between the verbs as to whether reciprocal malu can or should follow the verb.

<table>
<thead>
<tr>
<th>Vt</th>
<th>Vreciprocal</th>
</tr>
</thead>
<tbody>
<tr>
<td>át</td>
<td>hakát</td>
</tr>
<tr>
<td>dudu</td>
<td>hakdudu-k</td>
</tr>
<tr>
<td>fota</td>
<td>hakfota-k</td>
</tr>
<tr>
<td>kohi</td>
<td>hakohi-k</td>
</tr>
<tr>
<td>libur</td>
<td>haklibur</td>
</tr>
<tr>
<td>sala</td>
<td>haksala-k</td>
</tr>
<tr>
<td>sona</td>
<td>haksona-k</td>
</tr>
<tr>
<td>sudi</td>
<td>haksudi-k</td>
</tr>
<tr>
<td>te'ur</td>
<td>hakte'ur</td>
</tr>
<tr>
<td>tei</td>
<td>haktei-k</td>
</tr>
<tr>
<td>tolo</td>
<td>haktolo-k</td>
</tr>
<tr>
<td>tuda</td>
<td>hatchuda-k</td>
</tr>
<tr>
<td>tuku</td>
<td>haktuku-k</td>
</tr>
</tbody>
</table>

Many of the above derivations refer to methods of fighting which are or were standard within the culture, and in some cases were the object of gambling. Not surprisingly it is not entirely productive. Thus, for instance, the semantic relationship between hakte'ur ‘fight’ and te'ur ‘chase’ is not clear, and the breadth of meaning of hakdudu-k ‘push, accuse, urge each other’ is not predictable from the meaning of dudu ‘push’. An attempt to derive a reciprocal verb *hak-fetik from fetik ‘flick’ was not accepted by consultants.

4.22  To'o wé bá Bei Lilin nó Bei Rai Kfu’ak sia
reach water go Mr candle and Mr earth clod 3P

r<ak>dudu<k> malu.
3P<RECIP>push each other
(When) they reached the water, Mr Candle and Mr Dirt Clod—they urged each other to go (into the water) first. (F5.9)

4.23 Sia na’in rua n<ak>tuda<k>.
3P CLS:human two 3<RECIP>throw(spear)
The two of them threw spears at each other. (N0.22 elicited)

4.24 ...hotu-hotu <hak>sudi<k> iha ama tamuku nia-kan
RDP-all <RECIP>gather LOC father village.head 3S-POS
uma.
house
[This evening] everyone (must) gather in the village head’s house. (T0.21)

4.3.5 ‘Make sound’ verb: hak- (-k), ha- (-k)

A number of verbs referring to the making of sounds are derived by applying a prefix ha­ or hak- to the sound made. In some cases a suffix -k is added. Some such derived verbs are intransitive, while others are transitive, referring to the action of making that sound to call an addressee (usually to call an animal). There is no apparent logic to the choice of prefix, or to the presence or absence of the suffix.
Sound | V: make sound
---|---
hae | laughing sound | hahae-k | Vi | laugh (Fehan laugh for women)
he'e | reply to greeting | hahe'e | V | reply he'e (to greeting)
hi hi hi hi | call to horses | hakihis-k | Vt | call horses
huit | call to dogs | hahuit | Vt | call huit to set hunting dogs running
krú, krr | call to chickens | hakhrú | Vi, Vt | rustle (e.g. leaves); call chickens
kudu kú kú kú | dove’s sound | hakudu-k | Vi | make dove sound
lale | no | halale | Vt | say no to, deny falsely
mé | bleat; | hakmé | Vi | bleat;
call to goats | | | Vt | call goats
mú | sound of pigeon | hamú | Vi | make pigeon sound, roar (of sea)
oe | bellowing sound | hahoe(k) | Vi, Vt | bellow; call buffalo

4.25 Kabau n-ahoe. Ema n-ahoe bá kabau.
buffalo 3-bellow person 3-bellow go buffalo
Buffalo bellow ("ngoe"). People call to buffalo ("ngoe"). (N0.66)

There are in addition a number of verbs beginning with ha- or hak- which refer to the making of sounds, but for which I am unaware of a root. It is possible that at least some of these follow a similar pattern to those listed above, although for some others the apparent root is not a synchronic word (e.g. *farat, *rake)
hakedek | croak (of frog)
hakeo | splatter (e.g. of river flowing over rocks, rain on tin roof)
hakforat | snort (of horse, through nose)
hakilar | yelp (of dog in pain)
hakmeo | miaow (of cat)
hakmoi | call pigs (‘o’ or pig name)
haknukun | call chicks (of mother hen)
hakoron | snore
hakrake | yell out (usually in fear or excitement, e.g. ‘Adó!’)
hakrí | tinkle (e.g. of metal wings flapping, of telephone)
hakroto | rumble (e.g. purr; sound of drum or gun)
hakwé | squeal (of pig)
hasisu | hiss (e.g. of tape recorder)
hatoto | pop (e.g. of water or rice or oil boiling, gun, drum)

4.4 Derivation of actor-describing verbs: mak- (-n), ma- (-k)

4.4.1 Overview

The circumfix mak- (-n), and its variant ma- (-k), derives words describing typical subject referents, from both transitive and intransitive verbs (e.g. mak-fa’en ‘who sells’ from fa’en ‘sell’).¹⁰ There is a strong tendency for the stem verbs to have actor (as opposed to undergoer)

---

⁹ Phonemically irregular patterns, such as the vowel-less krr, are found in animal calls in many languages (e.g. van den Berg 1989:193).

¹⁰ These affixes are widely recognised as deriving actors, including habitual actors (Monteiro 1985:xv; Morris 1984b:xviii; Troeboes et al. 1987:38, 49).
subjects, for which reason the derivations will be loosely referred to as ‘actor-describing verbs’ or ‘actor derivations’.

The following sections first present examples of such derivations, and then comment on the distinction between mak- and ma-, and on the conditions under which the associated suffixes (-n and -k) are used. This is followed by a discussion of the unique syntactic distribution of these words, and the problematic issue of their word class. In particular, most actor derivations have the verbal characteristics of transitivity and of occurrence in serial verb constructions, yet are not able to function predicatively. Finally, similarities and differences between the prefix and the relative clause marker mak are noted.

An actor derivation may describe someone who has performed the action of the verb on a particular occasion (e.g. makawen ‘(one) who married’ referred in the corpus to one who had married that day), or one who can and habitually does perform that action (e.g. makleon ‘guardian (e.g. guardian saint/angel)’), or one who intends to perform that action (e.g. makfalen ‘seller’, a term which, as in English, is applicable even if no sale is made). Such aspectual variation is presumed to be pragmatically determined.

In the examples that follow, ‘X’ indicates that an object NP is required, while ‘(X)’ indicates that such an object is optional.

<table>
<thead>
<tr>
<th>V: non-/h/-initial</th>
<th>Actor-describing V</th>
</tr>
</thead>
<tbody>
<tr>
<td>daka₁¹</td>
<td>Vt guard</td>
</tr>
<tr>
<td>dakar</td>
<td>Vt guard</td>
</tr>
<tr>
<td>fa’en</td>
<td>Vt sell</td>
</tr>
<tr>
<td>kaer</td>
<td>Vt hold</td>
</tr>
<tr>
<td>kahi</td>
<td>Vt massage</td>
</tr>
<tr>
<td>kawen</td>
<td>Vt marry</td>
</tr>
<tr>
<td>ko’a</td>
<td>Vt cut</td>
</tr>
<tr>
<td>koto</td>
<td>Vt work for a noble</td>
</tr>
<tr>
<td>koto</td>
<td>Vt work for a noble</td>
</tr>
<tr>
<td>koto</td>
<td>Vt sever</td>
</tr>
<tr>
<td>lakon</td>
<td>Vi shine</td>
</tr>
<tr>
<td>lakon</td>
<td>Vi disappear</td>
</tr>
<tr>
<td>le’at</td>
<td>Vt visit, inspect</td>
</tr>
<tr>
<td>leba</td>
<td>Vt carry on pole</td>
</tr>
<tr>
<td>leo</td>
<td>Vt protect</td>
</tr>
<tr>
<td>mate</td>
<td>Vi die</td>
</tr>
<tr>
<td>rona</td>
<td>Vt hear</td>
</tr>
<tr>
<td>sai</td>
<td>Vt exit</td>
</tr>
<tr>
<td>sala</td>
<td>Vi err</td>
</tr>
<tr>
<td>semo</td>
<td>Vi fly</td>
</tr>
<tr>
<td>soi</td>
<td>Adj rich</td>
</tr>
<tr>
<td>sorì</td>
<td>Vt separate fighters</td>
</tr>
<tr>
<td>sosà</td>
<td>Vt buy</td>
</tr>
<tr>
<td>tabar</td>
<td>Vt join</td>
</tr>
</tbody>
</table>

₁¹ Daka and dakar (in the following example) are common variants of the one verb.

₁² Evidence that mako’an may be derived from ko’a is that speech is in other contexts also compared to cutting. In particular, one of the noble houses with responsibility for judging disputes is called Mako’a Rai (lit. ‘who-cuts earth’). The act of reaching judgement in a dispute is known as hakotu lia (lit. ‘chop through word’). Finally, in East Timor, ko’a lia (lit. ‘cut words’) means ‘speak’ (Morris 1984b).
### Chapter 4

<table>
<thead>
<tr>
<th>V: non-/h/-initial</th>
<th>Actor-describing V</th>
</tr>
</thead>
<tbody>
<tr>
<td>taman</td>
<td>Vt</td>
</tr>
<tr>
<td>té</td>
<td>Vi</td>
</tr>
<tr>
<td>tesi</td>
<td>Vt</td>
</tr>
<tr>
<td>titu</td>
<td>Vt</td>
</tr>
<tr>
<td>tuir</td>
<td>Vt</td>
</tr>
<tr>
<td>tür</td>
<td>Vi</td>
</tr>
</tbody>
</table>

4.26  
*Ema mak-fa'en n-a'ak "...". Ema <mak>sosa<n> nia*  
person who-sell 3S-say person <who>buy 3S  
n-a'ak "...".  
3S-say  
The seller says "[My coconut oil is 900 rupiahs per bottle]". The buyer—he says "[If possible, 600 rupiahs]".  

(K0.180)

### V: /h/-initial

<table>
<thead>
<tr>
<th>V: /h/-initial</th>
<th>Actor-describing V</th>
</tr>
</thead>
<tbody>
<tr>
<td>há</td>
<td>Vt</td>
</tr>
<tr>
<td>halo</td>
<td>Vt</td>
</tr>
<tr>
<td>he'uk</td>
<td>Vi</td>
</tr>
<tr>
<td>hein</td>
<td>Vt</td>
</tr>
<tr>
<td>hemu</td>
<td>Vt</td>
</tr>
<tr>
<td>hili</td>
<td>Vt</td>
</tr>
<tr>
<td>hó</td>
<td>Vt</td>
</tr>
<tr>
<td>hö</td>
<td>Vt</td>
</tr>
<tr>
<td>hodi</td>
<td>Vt</td>
</tr>
<tr>
<td>hoku</td>
<td>Vi</td>
</tr>
<tr>
<td>hola</td>
<td>Vt</td>
</tr>
<tr>
<td>hotus</td>
<td>Vt</td>
</tr>
<tr>
<td>husu</td>
<td>Vt</td>
</tr>
</tbody>
</table>

### V: /ha/-initial

<table>
<thead>
<tr>
<th>V: /ha/-initial</th>
<th>Actor-describing V</th>
</tr>
</thead>
<tbody>
<tr>
<td>hadidu</td>
<td>Vt</td>
</tr>
<tr>
<td>hadomi</td>
<td>Vt</td>
</tr>
<tr>
<td>hafuli</td>
<td>Vi</td>
</tr>
<tr>
<td>hahoa</td>
<td>Vt</td>
</tr>
<tr>
<td>hahoris</td>
<td>Vt</td>
</tr>
<tr>
<td>hahoris</td>
<td>Vt</td>
</tr>
<tr>
<td>haka'as</td>
<td>Vi</td>
</tr>
<tr>
<td>hakát</td>
<td>Vi</td>
</tr>
<tr>
<td>hakés</td>
<td>V</td>
</tr>
<tr>
<td>haklati</td>
<td>Vi</td>
</tr>
</tbody>
</table>

---

13 Note that, unless they have a -k suffix, the derivations in this column have the same phonological form as the 2S inflection of the base verb (e.g. mahoris ‘who gives birth’ or ‘2S-give birth’).
Derivational morphology

V: /hi/-initial Actor-describing V

| hakroman | vi   | shine       | makroman | who shines       |
| hakwe    | vi   | squeal      | makwe-k  | pig [sea taboo]  |
| halai    | vt   | run away (from) | malai-k | refugee         |
| hamu     | vi   | roar (of the sea) | mamu-k  | waves [sea taboo] |
| hana'o   | vt   | steal       | mana'o-k | thieving, thief; habitual thief\(^{14}\) |
| hananu   | v    | sing        | mananu-k | singers in choir |
| harik    | vi   | stand       | marik    | who stands [ritual language] |
| haroma   | v    | make light; become light | maroma-k | God/god\(^{15}\) |
| hatene   | vt   | know        | matene-k | clever          |
| hatoto   | vi   | pop, bang   | matoto-k | which sounds like gunfire |
| hawia    | vi   | dangle      | mawia-k  | that hangs down  |

4.27  \textit{Ita} <\textit{ma}>\textit{hili}<\textit{k}> sia ksotir.
1PI <\textit{who}>pick PL fortunate
We who pick up (the fallen mung beans) are fortunate (if many beans fall from the plant). \((\text{V0.85})\)

4.4.2 Distinction between mak- and ma-

As illustrated in the examples, there are two differences between mak- and ma-. Firstly, the choice between them is fully determined by the phonology of the base: ma- applies to /hi/-initial roots and to stems beginning with ha- or hak-, while mak- applies to disyllabic bases beginning with any other consonant.\(^{16}\) This phonological conditioning is puzzling, since the only other area in which /hi/-initial verbs are distinguished from other verbs is in subject marking, with /hi/-initial verbs being inflectable for a range of person-number combinations, while verbs beginning in other consonants can be inflected only for 1S -k \((\text{§9.3.1})\).

A second difference is that where there is an associated suffix, mak- pairs with -n, and ma- with -k.\(^{17}\) Again there is no apparent logic to this conditioning. The suffixes -k and -n, when not paired with prefixes, overlap in function, with both unproductively deriving nouns

\(^{14}\) 
\textit{Mana'ok} can denote a character of habitual thieving, or identify a thief who stole something on a particular occasion.

\(^{15}\) 
\textit{Maromak} 'God/god' (both Christian and traditional) was considered to derive from \textit{kroman} Adj, N 'light' by one Fehan Christian well-versed in tradition, and also by the anthropologist T. Therik (pers. comm.), who (like Francillon (1967)) considers it literally to mean 'the Luminous'. By deriving it instead via a derived verb \textit{haroma}, this example presents none of the semantic, part of speech and phonological eccentricities that make derivation directly from \textit{kroman} unattractive.

\(^{16}\) 
Derivations from /k/-initial bases are of course phonologically consistent with interpretation as having either ma- or mak-, but since two of the examples (mak'\_an, mak'\_oton) have an -n suffix which is normally associated with mak-,-k/-initial bases are analysed as taking mak-.

Since the prefix derives words describing actors, it does not apply to words with initial consonant clusters, as these are all either adjectives or nouns. I have no examples with vowel-initial roots, such verbs being in any case rare.

Apart from derivations based on derived ha- and hak- stems, there appear to be no derivations from bases of more than two syllables. My attempts at such derivation (*ma-tand\_t 'who-look.up', *ma-falahok 'which-indistinct') were rejected, and replaced by relative clauses (ma'ak tand\_t, mak falahok).

\(^{17}\) 
There appears to be dialect variation in this, since for the dialects of Belu Troboes et al. (1987:51, 53) associate both -k and -n with mak-, while Morris's (1984b) examples from East Timorese Tetun involve mak- -k (e.g. mak'sa'ek 'climber' from sa'e 'ascend'). Two examples in the corpus support this variation while indicating a very strong preference for the pairing of mak- with -n (rather than with -k) in the Fehan dialect. These are mak'\_sala\_k 'who errs', which was subsequently replaced by the speaker with mak'\_sala-n, and mak'\_soi\_k 'noble' (lit. 'who-rich'), for which the speaker similarly preferred /nl/-final mak'\_soi-n.
and adjectives. Both co-occur (with no apparent conditioning) with partial reduplication, but only -k co-occurs with the relatively productive adjectival prefix k-, the verbal prefix hak-, and the time prefix ma-.

It is not clear why the choice of prefix form and the choice of suffix should be directly correlated. In particular, there appears to be no semantic difference between /h/-initial verbs and other verbs which could contribute to an explanation. The following are a short list of near-synonyms to show the semantic overlap between these two groups of verbs.

<table>
<thead>
<tr>
<th>/h/-initial</th>
<th>Not /h/-initial</th>
</tr>
</thead>
<tbody>
<tr>
<td>habu'a  Vt</td>
<td>te'ur  Vt</td>
</tr>
<tr>
<td>hakara Vt</td>
<td>bér  Vt</td>
</tr>
<tr>
<td>hean Vt</td>
<td>firi  Vt</td>
</tr>
<tr>
<td>hetan Vt</td>
<td>simu  Vt</td>
</tr>
<tr>
<td>hili Vt</td>
<td>boi  Vt</td>
</tr>
</tbody>
</table>

**4.4.3 Conditions on use of suffix**

While the form of the suffix (-k or -n) is determined by the initial consonant of the base, there are two conditions which determine whether an actor derivation has a suffix at all. Firstly, as is consistently the case in the morphology (§2.10.2), the suffixes are only applied if the base is vowel-final. Thus, for instance, one cannot add a suffix to consonant-final ma-horis 'who-give.birth' (*feto ma-hori-k, *ma-horis-k).

A second condition on the application of the suffix, which appears to be unique to this type of derivation, is that no suffix is applied if the word has a following object NP. This point is illustrated by the contrast between examples 4.28 and 4.30 which have suffixes, and 4.29 and 4.31 which have object NPs but no suffixes.

4.28  
<mak>/kloto<n> sia  
<who>work.for.noble PL  
(person) who work for nobles  
(I26/6)

4.29  
mak-/kloto lamak sia  
who-work.for.noble meat.for.noble PL  
(person) who work at cooking for nobles  
(I26/6)

4.30  
Ema <ma>hò<k> sia foin r-ó.  
person <who>have PL then 3P-have  
Only rich people (lit. 'people who possess') have (old Dutch coins).  
(QO.15)

4.31  
Ema ma-hó kabau malaé ... Ma-hó ha'í sia...  
person who-have buffalo non-native ... who-have not PL  
People who had horses [carried the sago trunks on them]; those who didn't have any [hailed the sago themselves].  
(QO.105)

The object NP is always short, being either a single nominal (4.29) or a noun-modifier pair which is conventionally used as a name (4.31; masi manas 'salt hot' = 'mixture of chilli and salt'). It is usually non-referring. However, a clear example of a referential object NP is the pronoun nia '3S' in the following example, which was given to explain to me that ma-hó-k could mean 'who live with (him/her)' as well as meaning 'rich' (lit. 'who have', as in 4.31).
4.32  Bä husu bá ema <ma>hó<k> sia.
go request to person <who>accompany PL
Go and ask the people who are with (her—i.e. her adopted parents)

= ema ma-hó nia sia
person who-accompany 3S PL
= the people who are with her

= ema ma-hein hó nia sia.
person who-wait accompany 3S PL
= the people who live with her. (V0.120)

The final line of the above example (4.32) shows mahein in serialisation with the following verb hó. Serialisation is also found in the following standard poetic idiom from ritual language, which describes the noble maromak oan (4.33). These examples suggest that a suffix is not added if the verb is the initial element in a serialisation. However, such a proposal remains tentative for lack of further evidence.

4.33  do'ú oan kawak oan; ma-há toba ma-hemu
cwild.pigeon small k.o.bird small who-eat lie.down who-drink
toba
lie.down
small wild pigeon, small kawak bird; who eats lying down, who drinks lying down (i.e. who does no work) (GO.52)

The second NP in example 4.31 above (the only one of its kind in the corpus) suggests that the suffix is not applied when the verb is negated (at least by postmodifying ha'i).

4.4.4 Syntactic function of actor derivations

Actor derivations are found in four syntactic positions. From their distribution it is apparent that most actor derivations are members of a unique word class, which has the verbal characteristics of being either transitive or intransitive (depending on the transitivity of the base) and of occurrence in serial verb constructions, but which is restricted to functioning attributively within NPs. This class is labelled ‘actor-describing verbs’ for want of a better description. In addition to these stative verbs, some actor derivations have been lexicalised as nouns, and some as adjectives which can function predicatively as well as attributively.

The most common position for actor derivations is as post modifier of a noun (4.34; 71 examples). The head noun is frequently a semantically general one such as ema ‘person’ (4.30, 4.32).

4.34  ró <mak>semo<n>
boat <who>fly
boat that flies (i.e. aeroplane) (I22/7)

The second position is immediately following the relative clause marker mak (4.35, 4.36, 4.40, 4.42; 33 examples).19

18 Note that some adjectives (which are, like actor derivations, a subclass of verbs) are similarly unable to function predicatively (§4.6.7).

19 Klamer (1994:255) mentions that Kambera, spoken on the nearby island of Sumba, similarly allows the relative clause marker ma- to be followed by a homophonous and semantically overlapping prefix ma-.
So they look for a shaman who massages swollen spleens [and bring the shaman back].

Such apparent relative clauses headed by actor derivations are unique in two ways, both of which indicate that the actor derivation is not in fact predicative, and so does not head a relative clause. In the first place, most actor derivations (unlike heads of relative clauses) cannot function predicatively. For instance, the attempted clauses *Ema ne’e mahoris ‘person this who-give.birth’ and *Ema ne’e mananuk ‘person this who-sing’ were both rejected by consultants. A second unique feature is that, even when following the relative clause marker mak, such derivations cannot take aspectual or peripheral modifiers such as predicates can. The actor derivation in this construction is thus analysed as an attributive modifier within the NP, while mak functions as an attributive marker.

The third syntactic function for actor derivations is as head of an NP. There are only a few which characteristically function as head, either as proper nouns (e.g. Makleban ‘Water Carrier’, which is the name of a constellation), or as terms of address (maksoin, lit. ‘who-rich’, for nobles). Remaining examples fall into two categories. Some are religious, ritual, noble or sea taboo terms (e.g. maklituk ‘guardian’, nia-kan mak-tuir ‘3S-POS who-follow’ = ‘his followers/disciples’, sea taboo mahokuk ‘fish trap’). Such registers appear to more generally reinterpret adjectives as nouns (§1.2.3.4). The remaining examples are best analysed as headless NPs. This is because many of the actor derivations which occur initially in NPs (and so appear to be in the head slot) themselves have verbal characteristics such as object NPs or negation (4.29, 4.31, 4.37). Such characteristics are not found in nominal NP heads.

The fourth slot in which actor derivations occur is as head of a predicate. The possibility of functioning as predicate head is restricted to a very few intransitive derivations, all of which have trisyllabic stems beginning with /hal (e.g. makilik ‘ticklish’; 4.38).

20 It is not clear whether this is an accident of the data. However, it is clear that most actor derivations prefer the modifier slot to the head slot in an NP. One other apparent derivation which is characteristically used as a noun is mako’an ‘oral historian’ (fn. 12).
4.38 Ema ne’e <ma>[hajidu]<k> to’o.
    person this <who>beg very
This person is very beggarly (i.e. keeps asking for things).

4.4.5 **Contrast between prefix and relative clause marker mak**

The actor prefix and the relative clause marker *mak* have much in common. Phonologically they have the same shape and are both unstressed. Semantically they are translatable as ‘who, which’. And grammatically both actor derivations and relative clauses characteristically function as noun modifiers. The similarity is illustrated by the following three examples.

4.39 bei <mak>mate<n> sia-kan kakaluk
    ancestor <who>die 3P-POS betel.bag
the betel nut bags of the dead ancestors

4.40 ha’u-kan ama-n mak <mak>mate<n>
    1S-POS father-GEN REL <who>die
my father who died

4.41 oan mak mate sei ki’ik oan
    child REL die still small small
a child who dies while still very young

There are, however, a significant number of differences between them, which all support the analysis of the prefix as a constituent within a grammatical word, and the relative clause marker as a syntactic constituent of a clause.

Phonologically, there are three differences. Firstly, *ma-* and *mak-* apply only to bases of a certain phonological shape. In contrast relative clause *mak* is not phonologically restricted by the following word, as illustrated by the following examples.

<table>
<thead>
<tr>
<th>Phonology of word</th>
<th>Relative clause</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vowel-initial</td>
<td><em>mak át</em></td>
</tr>
<tr>
<td>/h/-initial</td>
<td><em>mak ha-susar</em></td>
</tr>
<tr>
<td>Other-consonant-initial</td>
<td><em>mak n-odi</em></td>
</tr>
<tr>
<td>Cluster-initial</td>
<td><em>mak kmán</em></td>
</tr>
<tr>
<td>Trisyllable</td>
<td><em>mak tanát</em></td>
</tr>
</tbody>
</table>

Secondly, the relative clause marker but not the prefix has a variant *ma’ak* (4.42).  

---

21 According to Kähler (1974:260f.) some languages in eastern Indonesia mark a relative clause by adding a prefix to its verb or adjective (e.g. Kambera *ma-*; Roti *ma-*). While I suspect that these languages too draw a distinction between the prefix and the relative clause marker (as noted for Kambera by Klamer (1994:255)), such a statement shows that the similarity between derivational prefixes and the relative clause marker is not restricted to Tetun. P. Quick (pers. comm.) notes that in Pendau (Sulawesi) also the clitic relative clause marker and the agentive prefix are homophonous, both being phonologically to.

22 I have no evidence for the prefix variant *maka-* which occurs in examples in de Araújo e Corte-Real’s (1990:71) discussion of Dili Tetun. This variant, being disyllabic, would not fit the pattern of Fehan’s known prefixes, which are all either monosyllabic or single consonants.
4.42 Kabau ma’ak mak-lakon ne’e, (*kabau mak ma’ak-lakon)
buffalo REL who-disappear this

se nia-k?
who 3S-POS
This buffalo which disappeared, whose was (it)? (V0.121)

And finally, as is to be expected, the prefix is more tightly bound phonologically to the base than relative clause mak is to a following verb. One correlate of this is that in non-fluent speech pauses and hesitation sounds can occur after the relative clause marker, but are not found after the prefix.

In addition to the phonological differences, there are a number of grammatical ones.

Under certain conditions actor derivations have a suffix -k or -n. There is no comparable suffix (or clitic) for relative clauses.

The prefix is restricted to deriving words describing the subject of the base verb. The relative clause marker can substitute for subjects (4.44) as well as objects (4.43), amongst others.

4.43 manu ma’ak ita atu ho’o ne’e
bird REL 1PI IRR kill this
the chicken which we are about to kill (T0.60)

The base in a derivation can take an object NP, negation, or a following serial verb, but I have no example for any other modification. In contrast, relative clauses can contain a wide range of modifiers and arguments (4.44).

4.44 Haré ema mak loron kalan hakát la hanawa, keta tuir.
see person REL day night fight not stop do. not follow
(If you) see people who fight night and day without stopping, don’t go along with (them). (T0.131)

Actor derivations are restricted to verbal bases, while relative clauses can have a variety of predicate heads, including transitive and intransitive verbs, adjectives (e.g. ema mak át ‘person who (is) bad’) and nouns (e.g. mak rán ‘which (originated from) blood’).

Finally, the base in a derivation does not take subject marking, while the verb in a relative clause does (e.g. ai mak n-arik mesan ‘plant which 3S-stand alone’).

4.5 Derivation of nouns

4.5.1 Nominalisation: partial reduplication

Adjectives and intransitive and transitive verbs are nominalised by partial (‘Ca-’) reduplication, in which the onset of the penultimate syllable is copied, followed by the vowel /a/. Note that, as is consistently the case in Tetun phonology, the initial syllable of a trisyllabic base is omitted from the derivation (as per hakés and hato’o in the examples below; §2.10.6), and the initial /k/ of a consonant cluster is not reduplicated (as per krakat and ktodan below; §2.10.4).

Where the base ends in a vowel, a suffix -k or -n is used, the choice showing no apparent synchronic conditioning.

Such nominal derivations nearly always occur in possessive constructions (4.45), where the reduplication appears to be optional (4.46).
### Adj

<table>
<thead>
<tr>
<th>N</th>
<th>Adj</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>beik</td>
<td>stupid</td>
<td>1</td>
</tr>
<tr>
<td>bót</td>
<td>big</td>
<td>2</td>
</tr>
<tr>
<td>di’ak</td>
<td>good, beautiful</td>
<td>3</td>
</tr>
<tr>
<td>katar</td>
<td>itchy</td>
<td>4</td>
</tr>
<tr>
<td>kleur</td>
<td>long (time)</td>
<td>5</td>
</tr>
</tbody>
</table>

### Adj

<table>
<thead>
<tr>
<th>N</th>
<th>Adj</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>krakat</td>
<td>angry, wild</td>
<td>1</td>
</tr>
<tr>
<td>ktodan</td>
<td>heavy</td>
<td>2</td>
</tr>
<tr>
<td>naruk</td>
<td>long</td>
<td>3</td>
</tr>
<tr>
<td>soi</td>
<td>rich</td>
<td>4</td>
</tr>
<tr>
<td>susar</td>
<td>be in difficulty</td>
<td>5</td>
</tr>
<tr>
<td>wa’in</td>
<td>many, much</td>
<td>6</td>
</tr>
</tbody>
</table>

### Vi

<table>
<thead>
<tr>
<th>N</th>
<th>Vi</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>hakés</td>
<td>talk</td>
<td>1</td>
</tr>
<tr>
<td>kotu</td>
<td>be concluded</td>
<td>2</td>
</tr>
<tr>
<td>moris</td>
<td>live</td>
<td>3</td>
</tr>
<tr>
<td>tanis</td>
<td>cry</td>
<td>4</td>
</tr>
<tr>
<td>tür</td>
<td>sit</td>
<td>5</td>
</tr>
</tbody>
</table>

### Vt

<table>
<thead>
<tr>
<th>N</th>
<th>Vt</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>bada</td>
<td>select time</td>
<td>1</td>
</tr>
<tr>
<td>fiar</td>
<td>believe, trust</td>
<td>2</td>
</tr>
<tr>
<td>halo</td>
<td>do, make</td>
<td>3</td>
</tr>
<tr>
<td>hato’o</td>
<td>supply (goods)</td>
<td>(to ‘o suffice)</td>
</tr>
</tbody>
</table>

### Derivational morphology

**4.45**

\[ Nia-kan \text{ na-naruk} \text{ naruk basuk.} \]

3S-POS NOM-length long very

Its (a klutis lizard’s) length is very long.

(N0.80)

**4.46**

\[ To’os ne’e, nia-kan naruk á hira? \]

garden this 3S-POS long DEF how much

This garden, what is its length?

(H0.57)

### 4.5.2 Instrument, undergoer and result nouns

#### 4.5.2.1 Introduction

There are four affixes which derive nouns from verbs, where the noun denotes an object which is the characteristic instrument, undergoer or result of the activity denoted by the verb. The productive derivations are via partial reduplication (with suffix -k or -n), and k(a)- (with

---

23 A schoolteacher consultant suggested that this should be written ‘ktatodan’, but that the initial /k/ is not pronounced. The preferred omission of /kl/ presumably relates to the relative uncommonness of initial consonant clusters in trisyllabic words, and to the low sonority contrast between /kl/ and /tl/, resulting in reduced perception of the initial /kl/.
or without the suffix \(-k\). Some fossilised derivations exist for the suffixes \(-k\) and \(-n\), while other nouns involve zero derivation.

In a few instances alternative derivations, with different meanings, exist for the same verbal root.

<table>
<thead>
<tr>
<th>V</th>
<th>N</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>há</td>
<td>eat</td>
<td>ha-há-k</td>
</tr>
<tr>
<td>mí</td>
<td>urinate</td>
<td>ka-mí-k</td>
</tr>
<tr>
<td>té</td>
<td>defecate</td>
<td>té-k</td>
</tr>
</tbody>
</table>

4.5.2.2 Using partial reduplication

Partial reduplication derives typical instruments or undergoers from transitive verbs, as is noted by Monteiro (1985:xxii) and Troeboes et al. (1987:55–57). This reduplication co-occurs with either of the suffixes \(-k\) or \(-n\), with no apparent synchronic basis for the choice between them. As elsewhere, the suffix applies only to vowel-final bases. Despite the synchronic uncertainty over the suffix, this derivation appears to be productive, as evidenced by its use to derive words for modern implements (da-dada-n ‘zip’ from dada ‘pull’), and by its application to an Indonesian root tawar ‘neutralise’ to derive ta-tawar ‘antidote’, instead of borrowing the Indonesian noun penawar.

The following examples list typical instruments (about 30 examples) and undergoers.

**Vt**

| bobi | pound (e.g. betel) |
| f'o'at | catch in noose |
| fofe | paddle |
| hak'a'u | carry in two hands |
| ko'us | carry in front |
| kodo | filter (sago) |
| korus | grate |
| kusa | button |
| sukat | measure |
| suru | shovel, ladle |
| dada | pull |
| há | eat |
| hi'it | guess |
| hó | have |
| keta | wean |
| sakat | cut into strips |
| stmu | receive |
| solok | send |

<table>
<thead>
<tr>
<th>N: instrument</th>
<th>N: undergoer</th>
</tr>
</thead>
<tbody>
<tr>
<td>babobi-n</td>
<td>dadada-n</td>
</tr>
<tr>
<td>fafo'at</td>
<td>hahá-k</td>
</tr>
<tr>
<td>fafofe-n</td>
<td>ai hahi'it</td>
</tr>
<tr>
<td>kak'a'un</td>
<td>hahó-k</td>
</tr>
<tr>
<td>kako'us</td>
<td>oa kaketa-n</td>
</tr>
<tr>
<td>kakusak</td>
<td>sasakat</td>
</tr>
<tr>
<td>sasukat</td>
<td>sasimu-n</td>
</tr>
<tr>
<td>sasuru-n</td>
<td>sasolok</td>
</tr>
</tbody>
</table>

According to Jonker (1906:318) vowel final bases in reduplication are often, but not always, followed by \(-n\), or by its variants \(-k\) or \(-t\) (though I have no evidence for the \(-t\) in Fehan). He notes that such variation in final consonants also occurs in neighbouring languages, and, like myself, recognises no synchronic explanation for it. Monteiro (1985:xxii) and Troeboes et al. (1987:55–57) associate only \(-n\) with this reduplication in their discussions of the Tetun dialects of East Timor and Belu. Thomaz (1981:81, fn. 11) concludes that the final added consonant is unproductive, and not, apparently, applicable to loan words.
4.47 Nia mai ha'i so'in é. Surak nia fó nia-kan
3S come not OK TAG so.long.as 3S give 3S-POS

sa-solok á.
RDP-send DEF
(If) he doesn’t come that’s OK. So long as he sends his gift. (X0.51)

4.5.2.3 Using k(a)- (-k)

The prefix k(a)-, with or without an accompanying suffix -k, similarly derives typical instruments or undergoers from verbs (15 examples). There is no apparent logic to the presence or absence of the suffix, or to the choice between initial k- and ka-.

<table>
<thead>
<tr>
<th>Vt</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>feur</td>
<td>rotate, spin</td>
</tr>
<tr>
<td>fufu</td>
<td>card cotton</td>
</tr>
<tr>
<td>lele</td>
<td>float</td>
</tr>
<tr>
<td>luni</td>
<td>lay head on</td>
</tr>
<tr>
<td>sela</td>
<td>put saddlecloth on (horse)</td>
</tr>
</tbody>
</table>

4.48 Ita le'u tali bá ka-feur kakorok á, foin le'ak
1PI encircle rope go NOM-spin neck DEF then release

bá rai.
go earth
We wind a string around the neck of the top, then release (it) on the ground (to spin). (I0.121)

Some such derived nouns with the prefix ka-, as well as many other nouns with initial ka- for which I am not aware of a derivation (e.g. kanauk ‘wares’), have alternative forms involving partial reduplication (e.g. na-nauk), or an initial k- followed by partial reduplication (e.g. k-na-nauk).26 In both types of reduplication, there is an accompanying suffix -k only if that suffix also co-occurs with the ka- prefix for that base. The reduplicated forms are almost always considered non-Fehan in origin, although they are on occasion used by Fehan speakers. It thus appears that while Fehan has many partial reduplications deriving nouns (§4.5.2.2), where there is an alternate form with a ka- prefix the latter derivation is preferred.

The following examples list some alternative instrument/undergoer derivations. Most of the reduplicated (‘Ca-‘ and ‘kCa-’) forms are taken from Morris’s (1984b) dictionary of East Timorese Tetun. Missing variants could well be accidents of the data, except for the omission of *khahúk, which would contravene Tetun phonology in having a /kh/ sequence (§2.7.1).

25 Morris (1984b) lists a variant karluni, which has unusual phonology for Tetun, and if correct suggests this is a borrowing.

26 Jonker (1906:282) similarly notes that reduplication derives instruments, and that in East Timor a k- is often added word-initially. Troeboes et al. (1987:41–42) list ka- for this function.
Some vowel-final (transitive and intransitive) verbs and their corresponding instrument, undergoer, or result nouns are distinguished only by a final -n suffix on the noun (Monteiro 1985:xviii; Troeboes et al. 1987:47-48).²⁷ It is possible that many of these final In! were historically instances of the genitive marker (§7.3.3.3). This is particularly so for body parts such as sikun ‘elbow’ and hitin ‘lap’, since most body parts synchronically end in /n/, and are in many languages obligatorily possessed.

Vt

<table>
<thead>
<tr>
<th>Vt</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>futu</td>
<td>futun</td>
</tr>
<tr>
<td>hā</td>
<td>hān</td>
</tr>
<tr>
<td>heiti</td>
<td>ulu hitin</td>
</tr>
<tr>
<td>hiti</td>
<td>hitin</td>
</tr>
<tr>
<td>kaba</td>
<td>kaban</td>
</tr>
<tr>
<td>kakai</td>
<td>kakain</td>
</tr>
<tr>
<td>rahu</td>
<td>rahun</td>
</tr>
<tr>
<td>siku</td>
<td>sikun</td>
</tr>
<tr>
<td>susu</td>
<td>susun</td>
</tr>
</tbody>
</table>

4.49 *Ita halolo mate-n á.
1PI stretch.out die-NOM DEF
We lay out the corpse (in the coffin). (F22/9)

4.5.2.5 Using -k

In a very few instances nouns are derived from verbs by the suffix -k, without any prefix or reduplication.

Vt

<table>
<thead>
<tr>
<th>Vt</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>hada</td>
<td>hadak</td>
</tr>
<tr>
<td>hiri</td>
<td>hirik</td>
</tr>
<tr>
<td>tē</td>
<td>tēk</td>
</tr>
</tbody>
</table>

²⁷ For the related language Rotinese, Jonker (1906:265) analyses such verbs as deriving from the noun by deletion of the final consonant.
4.5.2.6 Zero derivation

Instrument, undergoer or result nouns may have the same phonological form as the verb (over 30 examples). This is the case even for vowel-final words, to which suffixes could in principle be applied.28

<table>
<thead>
<tr>
<th>N</th>
<th>Vt</th>
</tr>
</thead>
<tbody>
<tr>
<td>baki pile</td>
<td>baki pile up</td>
</tr>
<tr>
<td>butuk pile</td>
<td>butuk put in a pile</td>
</tr>
<tr>
<td>dadain thin strips</td>
<td>dadain cut into thin strips</td>
</tr>
<tr>
<td>falun parcel</td>
<td>falun wrap up</td>
</tr>
<tr>
<td>hakat hand span, pace</td>
<td>hakat measure hand span, pace</td>
</tr>
<tr>
<td>kadi whetstone</td>
<td>kadi sharpen using whetstone</td>
</tr>
<tr>
<td>lolok flock</td>
<td>lolok flock together</td>
</tr>
<tr>
<td>luku plough</td>
<td>luku plough</td>
</tr>
<tr>
<td>mama betel</td>
<td>mama chew (particularly betel)</td>
</tr>
<tr>
<td>sulan cork, stopper</td>
<td>sulan stop up hole using cork/stopper</td>
</tr>
</tbody>
</table>

4.5.3 Time nouns: ma-, ma- -k

The prefix ma- attached to numeral roots derives words indicating the number of days between the present time and the specified time.29 Unless the derivation is preceded by the preposition hori 'since', the time indicated is in the future.

<table>
<thead>
<tr>
<th>Root</th>
<th>Future</th>
<th>Past</th>
</tr>
</thead>
<tbody>
<tr>
<td>tolu three</td>
<td>matolu three days hence</td>
<td>hori matolu-k three days ago</td>
</tr>
<tr>
<td>hât four</td>
<td>mahât four days hence</td>
<td>hori mahât four days ago</td>
</tr>
<tr>
<td>lima five</td>
<td>malima five days hence</td>
<td>hori malima-(k) five days ago</td>
</tr>
<tr>
<td>nén six</td>
<td>manén six days hence</td>
<td>hori manén six days ago</td>
</tr>
<tr>
<td>hitu seven</td>
<td>mahitu seven days hence</td>
<td>hori mahitu seven days ago</td>
</tr>
</tbody>
</table>

4.50 Ha’u ma-tolu foin mai.
I’ll come in three days time. (N0.181 elicited)

The suffix -k is appended for times in the recent past if the root has no final consonant. The suffix is obligatory for the root tolu ‘three’, optional for lima ‘five’, and disallowed for hitu ‘seven’ (*hori ma-hitu-k). A -k suffix appears to be fossilised on a few other past time expressions also. These are listed below.

<table>
<thead>
<tr>
<th>Base</th>
<th>Past time</th>
</tr>
</thead>
<tbody>
<tr>
<td>*sehi</td>
<td>hori.sehi-k yesterday</td>
</tr>
<tr>
<td>rafoni two days hence</td>
<td>hori (rafoni)-k two days ago</td>
</tr>
<tr>
<td>hira several, how many</td>
<td>hori hira-k recently, when</td>
</tr>
</tbody>
</table>

28 Conceptually it appears to be the noun which is basic, and the verb which is derived (a conclusion also reached by Klamer (1994:106) for the related Central Malayo-Polynesian language Kambera).

29 Sometimes one hears wain (or wai) instead of ma- in such time expressions, this generally being regarded as non-Fehan. This is based on the free root wain ‘day’, which is used in Fehan only by older people and those versed in traditional oral literature. Evidence that wain does not function as a prefix is that there is no associated -k suffix for past time. Hence one hears hori wain tolu ‘three days ago’, rather than *hori wain tolu-k. Ma-, in contrast, does not occur as a free-standing lexeme.
In derivations using ma-, the root numeral can range from three upwards, with the upper limit being disputed by consultants. In practice, numerals from seven up appear not to be used, with periods of seven days or longer being specified using the introduced terms for ‘week’, namely Indonesian minggu or Portuguese misa (lit. ‘Catholic mass’).

### 4.5.4 Compounds

In Tetun compounds, two members of (potentially) open classes combine to form a single word. Phonologically, compounds are distinguished from multi-word phrases in four ways (§2.10.5):

1. The coda of the final syllable of the initial member is deleted. Further truncation of the final syllable is normally optional.
2. Any cluster-initial /k/ on the second member is deleted.
3. Like other single words, compounds consist of at most four syllables, and carry only one primary stress, which falls on the penultimate syllable. Secondary stress falls on the initial syllable of the first member.
4. Occasionally the vowel in a reduced initial member loses its secondary stress and is replaced by the default vowel /a/. This is illustrated by ta-nasu ‘palm sugar’ below, and in §4.9.3.

As words, compounds are conventional names for entities. They can shift in meaning after they are coined, with the result that the meaning is not necessarily predictable from that of the members. Some members in compounds are not recognised as synchronic words in the language.

Nearly all compounds follow the normal syntax of noun phrases, as illustrated in the examples below.

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Members</th>
<th>Compound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Possessor + N-n</td>
<td>manu tolun</td>
<td>man-tolun</td>
</tr>
<tr>
<td></td>
<td>lima fuan</td>
<td>lima-fuan</td>
</tr>
<tr>
<td></td>
<td>ulu lahan</td>
<td>u-lahan</td>
</tr>
<tr>
<td>N + modifier</td>
<td>batar ktasak</td>
<td>ba-tasak</td>
</tr>
<tr>
<td></td>
<td>batar kтомak</td>
<td>ba-tomak</td>
</tr>
<tr>
<td></td>
<td>asu fанu</td>
<td>asu-fanu</td>
</tr>
<tr>
<td></td>
<td>tua nasu</td>
<td>ta-nasu</td>
</tr>
<tr>
<td>N + N</td>
<td>ina ama</td>
<td>ina-'.ama</td>
</tr>
</tbody>
</table>

Various expressions meaning ‘the east’ and ‘the west’ fit a structure very similar to that of possessor followed by possessed noun, except that the second member is a verb rather than a noun. The verb is not nominalised by reduplication, but is followed by a genitive -n, just as vowel-final possessed nouns are in this context. Note that the existence of a range of synonymous expressions satisfies the requirements of poetic parallelism, illustrated in example 4.51.
Syntax Members
Possessor + N-n lora lakan sun flame lora-lakan the east
Possessor + V-n lora sa'e sun ascend lora-sa'e-n the east
loro lén sun glow lora-lén the east
loro monu sun fall lora-monu-n the west
loro toba sun lie down lora-toba-n the west [Suai]

4.51 To'os Lalawar e' iha rai ulun bá. Iha lora-lakan.
garden house garden this LOC earth head go LOC sun-flame
loro-lén.
sun-glow
This Garden of Lalawar is over at Earth Head. In the sun's flame, the sun's
glow (i.e. in the east).

Many multi-word NPs are compound-like in that they represent conventional names for
things, but are semantically regular and do not show phonological evidence of constituting
single words. These include many possessor-noun sequences (e.g. lima kanutak 'hand/arm
nail' = 'fingernail'). In addition they include NPs with premodifying relative clauses (e.g.
haris fatik 'bathe place '; §7.4.4) and NPs with postmodifying clauses in which there is no
syntactic slot for the head noun (e.g. besi tā rai 'tool chop earth' = 'broad hoe'; §14.8.3.3).

There is a small class of exocentric nominal compounds in which the compound consists
of a transitive verb followed by a noun which is interpreted as its generic object. The resultant
term denotes a class of things which are used for the purpose specified by the compound.

Syntax Members
V + N beni ai cover protect foot beni-'ai thong, sandal
taka ulu cover head taka-'ulu hat
heti knotak bind waist heti-knotak30 cloth tied around waist
heti ré bind forehead heti-ré headband
heti ulu bind head heti-'ulu cloth tied around head

There are two time expressions which show phonological evidence of compounding. They
form a terminology set with expressions which clearly consist of two words, such as lora
malirin 'sun cool' = 'late afternoon'.

Syntax Members
N + Adj lora kraik sun low lo-raik late afternoon
Possessor + N-n lora tén sun excrement lo-tén early evening

4.6 Derivation of adjectives

4.6.1 Overview

Of the five affixes used to derive adjectives, only the use of the circumfix k- -k to derive
result adjectives from transitive verbs appears to be moderately productive. Unproductive
affixes which derive adjectives are -k, ma- -k, -n and -s.

30 It is unclear why cluster-initial /kl/ is retained here when it is omitted in other compounds; however, it is
possible that it has something to do with the initial root being vowel-final and non-truncated.
There is considerable overlap in the meanings of these affixes. Examples of alternative derivations from the same base are presented here (where ‘Attrib’ indicates an adjective which is used only attributively). In some cases alternative derivations appear to have different collocational restrictions rather than different meanings. This is illustrated by ruak - ruas ‘two’ and hituk - hitus ‘seven’, whose collocations are discussed in footnotes 36 and 42.

<table>
<thead>
<tr>
<th>Base</th>
<th>Adj</th>
<th>Base</th>
<th>Adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>ama</td>
<td>N father</td>
<td>ama-k</td>
<td>Adj male-like</td>
</tr>
<tr>
<td>hana’o</td>
<td>Vt steal</td>
<td>k-na’o-k</td>
<td>Adj, N thieving, thief</td>
</tr>
<tr>
<td>hawia</td>
<td>Vt dangle</td>
<td>k-wia-k</td>
<td>Adj dangling</td>
</tr>
<tr>
<td>hitu</td>
<td>Num seven</td>
<td>hitu-k</td>
<td>Attrib seven</td>
</tr>
<tr>
<td>mesia</td>
<td>Adv solely</td>
<td>k-mesa-k</td>
<td>Adj sole</td>
</tr>
<tr>
<td>monu</td>
<td>Vi fall</td>
<td>monu-k</td>
<td>Adj toothless</td>
</tr>
<tr>
<td>moru</td>
<td>Adj bitter</td>
<td>moru-k</td>
<td>Attrib bitter</td>
</tr>
<tr>
<td>rua</td>
<td>Num two</td>
<td>rua-k</td>
<td>Attrib two</td>
</tr>
</tbody>
</table>

4.6.2 Result: k- -k

A major function of the circumfix k- -k is to derive result adjectives from transitive verbs. The adjective describes a state that would come about in the undergoer of the base verb. There is, however, no implication as to how this condition came about, and certainly no implication that there is an agent involved. For instance, k(a)le’uk ‘crooked’ (from le’u ‘coil’) can describe such things as a snake, a winding road, a buffalo horn, a bent finger or a broken leg.

<table>
<thead>
<tr>
<th>Vt</th>
<th>Adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>feur</td>
<td>rotate</td>
</tr>
<tr>
<td>fifi</td>
<td>force open [impolite]</td>
</tr>
<tr>
<td>habit</td>
<td>squeeze</td>
</tr>
<tr>
<td>le’u</td>
<td>coil, wind</td>
</tr>
<tr>
<td>rahu</td>
<td>to pieces [serial verb]</td>
</tr>
<tr>
<td>silu</td>
<td>snap</td>
</tr>
<tr>
<td>sira</td>
<td>tear</td>
</tr>
<tr>
<td>sobu</td>
<td>demolish</td>
</tr>
<tr>
<td>su’u</td>
<td>fold</td>
</tr>
<tr>
<td>taka</td>
<td>close, cover</td>
</tr>
<tr>
<td>tesi</td>
<td>break</td>
</tr>
<tr>
<td>kfeur</td>
<td>whirl-(pool, wind) [rare]</td>
</tr>
<tr>
<td>kfifi-k</td>
<td>showing (teeth)</td>
</tr>
<tr>
<td>kabit</td>
<td>at an acute angle</td>
</tr>
<tr>
<td>k(a)le’u-k</td>
<td>crooked, winding</td>
</tr>
<tr>
<td>krahu-k</td>
<td>finely broken up</td>
</tr>
<tr>
<td>ksilu-k</td>
<td>snapped</td>
</tr>
<tr>
<td>ksira-k</td>
<td>torn</td>
</tr>
<tr>
<td>ksobu-k</td>
<td>in very poor condition (house)</td>
</tr>
<tr>
<td>ks’u’u-k</td>
<td>stooped</td>
</tr>
<tr>
<td>ktesi-k</td>
<td>curved downwards</td>
</tr>
<tr>
<td>ktesi-k</td>
<td>broken</td>
</tr>
</tbody>
</table>

Where the verb begins with the causative prefix ha-, the k- replaces the ha- (§2.10.6).

Base of ha- | Vt: ha- | Adj
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>kabelak flat</td>
<td>habelak</td>
<td>flatten</td>
</tr>
<tr>
<td>*ləti</td>
<td>halati</td>
<td>topple</td>
</tr>
<tr>
<td>*leka</td>
<td>haleka</td>
<td>put open-side up</td>
</tr>
<tr>
<td>hawá</td>
<td>open wide</td>
<td></td>
</tr>
<tr>
<td>hawia</td>
<td>dangle</td>
<td></td>
</tr>
<tr>
<td>kbelak</td>
<td>flattened</td>
<td></td>
</tr>
<tr>
<td>klati-k</td>
<td>toppled</td>
<td></td>
</tr>
<tr>
<td>kleka-k</td>
<td>curved upwards</td>
<td></td>
</tr>
<tr>
<td>kwá-k</td>
<td>wide open</td>
<td></td>
</tr>
<tr>
<td>kwia-k</td>
<td>dangling [rare]</td>
<td></td>
</tr>
</tbody>
</table>

This derivation from transitive verbs is reasonably productive, with a somewhat constant meaning, an almost exceptionless form of the prefix (as k-, rather than ka-), and a consistent co-occurrence with the suffix -k (for vowel-final bases).
4.52  Ita hatán ai kale’uk hodi fatuk ne’ebé dadi
1PI put.on.top.of wood crooked use rock so.that become
lós.
straight
We put a rock on a crooked stick so that (it) becomes straight. (V0.63)

4.6.3 ‘Quality of actor’: k- -k

The circumfix k--k also derives adjectives from verbs (whether transitive or intransitive), where the adjective refers to either a resulting condition of the actor, or to some quality associated with a characteristic actor. There are only a few examples for each of these.\(^{31}\) It seems that some of these adjectives are restricted to functioning attributively; however, more data are required to confirm this.

<table>
<thead>
<tr>
<th>Vi</th>
<th>Attrib: condition of actor</th>
</tr>
</thead>
<tbody>
<tr>
<td>hakseke</td>
<td>hop</td>
</tr>
<tr>
<td>mosu</td>
<td>appear</td>
</tr>
<tr>
<td>sa’e</td>
<td>ascend</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Adj/Attrib: quality of actor</th>
</tr>
</thead>
<tbody>
<tr>
<td>hana’o</td>
</tr>
<tr>
<td>hano’u</td>
</tr>
<tr>
<td>labu</td>
</tr>
<tr>
<td>sui</td>
</tr>
<tr>
<td>talik</td>
</tr>
</tbody>
</table>

4.53  Ita hateke ema kna’ok.
1PI peer.at person thieving
We peer out at a thief. (E29/8)

4.6.4 ‘Similarity to’ adjective base: k(a)- -k

The circumfix k(a)- -k relates a number of adjectives to base adjectives (or adverbs or degree verbs) with a similar meaning.\(^{32}\) In some of these, the derivation indicates some sort of imitation, temporary or non-inherent character which resembles the character described by the base. In particular, the colour terms metan ‘black’ and mean ‘red’ are neutral as to the reason for the colour, while the derived terms are mainly applicable if the colour came about as a result of stains or dirt. In some instances there is a final /n/ on the base which is omitted in the derivation; this is assumed to have been a diachronic suffix on the base (possibly akin to the nominalising suffix used to form attributive adjectives in ancestral North New Guinea Cluster (Ross 1988:184)).

---

\(^{31}\) Troeboes et al. (1987:41–42) list ka- as deriving actor nouns from verbs, but apart from kna’ok ‘thief’, all examples in my list are used attributively (e.g. ai ktalik ‘plant vine’ = ‘vine’).

\(^{32}\) Both Monteiro (1985:xvii, xix) and Troeboes et al. (1987:41–42) list ka- as deriving adjectives from adjectives or nouns, with a meaning of resemblance. Monteiro also notes its co-occurrence with the suffix -k.
<table>
<thead>
<tr>
<th>Base</th>
<th>Adj</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>bubu</strong></td>
<td>Adj</td>
<td>swollen (e.g. leg)</td>
</tr>
<tr>
<td><strong>lanu</strong></td>
<td>Adj</td>
<td>drunk, dizzy</td>
</tr>
<tr>
<td><strong>liu</strong></td>
<td>Vdegree</td>
<td>very; pass, surpass</td>
</tr>
<tr>
<td><strong>mean</strong></td>
<td>Adj</td>
<td>red</td>
</tr>
<tr>
<td><strong>mesa</strong></td>
<td>Adv</td>
<td>solely</td>
</tr>
<tr>
<td><strong>metan</strong></td>
<td>Adj</td>
<td>dark-coloured</td>
</tr>
<tr>
<td><strong>moru</strong></td>
<td>Adj</td>
<td>bitter</td>
</tr>
<tr>
<td><strong>naruk</strong></td>
<td>Adj</td>
<td>long, tall</td>
</tr>
<tr>
<td><strong>roman</strong></td>
<td>Adv</td>
<td>in the light</td>
</tr>
<tr>
<td><strong>seti</strong></td>
<td>Adj</td>
<td>tight (fit), crowded</td>
</tr>
<tr>
<td><strong>tasa</strong></td>
<td>Adj</td>
<td>cooked</td>
</tr>
<tr>
<td><strong>wa in</strong></td>
<td>Adj</td>
<td>many</td>
</tr>
</tbody>
</table>

4.54  

<table>
<thead>
<tr>
<th>Nia</th>
<th>kabun</th>
<th>kabubuk.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3S</td>
<td>stomach swollen</td>
<td>His stomach is swollen (e.g. due to illness).</td>
</tr>
</tbody>
</table>

4.6.5 ‘Quality of noun base, ‘having’: k(a) -k

The circumfix k(a)- -k also derives some adjectives from nouns. The adjective describes a quality in some way similar to a quality of the denotata of the base noun.

<table>
<thead>
<tr>
<th>N</th>
<th>Adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>feto</td>
<td>woman who behaves like a woman</td>
</tr>
<tr>
<td>foho</td>
<td>mountain</td>
</tr>
<tr>
<td>leten</td>
<td>top, above</td>
</tr>
<tr>
<td>lolon</td>
<td>trunk; long piece</td>
</tr>
<tr>
<td>masin</td>
<td>salt</td>
</tr>
<tr>
<td>rai</td>
<td>earth</td>
</tr>
</tbody>
</table>

Alternatively the adjective may indicate the possession of that which is denoted by the base noun. Here the final /n/ on the bases is most likely a genitive marker, which has been fossilised on these inherently possessed nouns (§7.3.3.3).

<table>
<thead>
<tr>
<th>N</th>
<th>Adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>sikun</td>
<td>elbow</td>
</tr>
<tr>
<td>lidun</td>
<td>corner</td>
</tr>
<tr>
<td>taran</td>
<td>thorn</td>
</tr>
</tbody>
</table>

4.6.6 ‘Similarity to’: -k

The suffix -k can derive an adjective from a noun or from another adjective. The meaning often has an element of resemblance to that denoted by the base.

The suffix applies only to bases which end in a vowel or in /n/. In most cases such final /n/ can be recognised as synchronically not part of the root (e.g. a genitive marker on inalienably

33 The relatedness of meaning is evident from the fact that both ktrasak ‘ripe’ and tasa ‘cooked’ have the same antonym, matak ‘unripe, uncooked, unrelated (by birth or marriage)’. 
possessed body parts), while in others (e.g. in colour terms) it is assumed to be a fossil. The meaning of these derivations is not predictable.

<table>
<thead>
<tr>
<th>N</th>
<th>Adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>ama father</td>
<td>amak male (roof strut, crab)</td>
</tr>
<tr>
<td>aman male (animal)</td>
<td></td>
</tr>
<tr>
<td>asu dog</td>
<td>asuk canine (tooth)</td>
</tr>
<tr>
<td>bula crazy</td>
<td>bulak (act) as if crazy</td>
</tr>
<tr>
<td>ina mother</td>
<td>inak female (roof strut, crab)</td>
</tr>
<tr>
<td>inan female (animal)</td>
<td></td>
</tr>
<tr>
<td>kabu-n stomach-GEN</td>
<td>kabuk pregnant</td>
</tr>
<tr>
<td>manu bird</td>
<td>manuk bird-like</td>
</tr>
<tr>
<td>molin outside (bare ground)</td>
<td>molik bald, bare</td>
</tr>
<tr>
<td>mú sound made by a pigeon</td>
<td>do'ú múk pigeon (do'ú) that says 'mú'</td>
</tr>
<tr>
<td>ulu-n head-GEN</td>
<td>uluk former; V go first</td>
</tr>
<tr>
<td>wē water, liquid</td>
<td>wēk watery</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Adj</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>mate Vi die, dead</td>
<td>matek numb, paralysed</td>
</tr>
<tr>
<td>mean red</td>
<td>meak reddish</td>
</tr>
<tr>
<td>mutin white, whitish (general)</td>
<td>mutik whitish (esp. of body parts)</td>
</tr>
</tbody>
</table>

4.55 ...kabun á nú ema kabuk.

[After she swallowed everything,] (her) stomach was like a pregnant person. (R5.138)

One system by which a man can be named is with his own name followed by his father’s name, which in certain instances is followed by -k. Examples are Klau Bere-k (Klau son of Bere), Bere Bau-k and Naha Tae-k. This use of -k could be related to the ‘resemblance’ meaning of -k, or alternatively to the use of -k (as a variant of -kan) to mark possessive pronouns (§7.3). Some names cannot take final -k (e.g. *Klau-k). The presence or absence of -k appears to be lexicalised (something which would be possible since there is only a small pool of names); data are, unfortunately, insufficient to search for possible conditioning factors.

4.6.7 Attributive adjective: -k

Another use of the suffix -k is to derive adjectives which are only used attributively. Bases come from a range of parts of speech, with the exception of nouns.

---

34 This agrees with Morris (1984b:xv), who says that adjectives “may be formed from other words that end in a vowel by adding k. Nouns and other words ending in n should be changed to k, but are often left as they are”.

Monteiro (1985:xviii) and Troebes et al. (1987:48), writing about the Foho dialect, say that -k derives stative adjectives from active intransitive verbs; however, their examples were not recognised in the Fehan dialect.

35 As one consultant pointed out, roof struts are not truly male or female, but have qualities resembling them. Crabs cannot be directly recognised as male or female, but their sex is inferred from the shape of the lower shell.
**4.6.8 Result: ma- -k**

Some adjectives are derived by the circumfix ma- -k. Unlike most actor derivations using ma- -k (§4.4), these freely occur predicatively.

Most of the derived adjectives have a corresponding verb beginning with ha- (the causative prefix), to which they tend to be related semantically as describing the result of the action indicated by the verb. It is alternatively possible that the root of the ha- verb is itself the root for the ma- -k adjective. If this is the case, then the root can come from a range of word classes, and the derived adjective indicates a condition in some way characterised by the root.

Regardless of the analysis accepted, it seems that this use of ma- -k is not productive, with the semantic relationship between base and derivation being unpredictable. Two of the derivations listed below are in fact nouns (maliras ‘bat, bird’ and makkukar ‘knot calendar’), but are listed here because sea taboo and ritual registers commonly reanalyse adjectives as nouns.

---

36 The only collocations for ruak ‘two’ and toluk ‘three’ in the corpus are those in example 4.56, while hituk ‘seven’ was noted only in the expression samea ulu hituk ‘snake head seven’ = ‘seven-headed snake’.

37 A derivational relationship between basu and basuk was disputed by a linguistically astute consultant, but is supported by the fact that there are two other verbs of motion which can similarly be used as intensifiers, these being to’o ‘arrive, reach’ and liu ‘go past’.
The table shows examples of the suffix -n being used with verbs to form adjectives that describe the condition of the undergoer or the actor, depending on whether the base is transitive or intransitive.

4.6.9 Result: -n

Like -k, the suffix -n can be attached to some verbs (both transitive and intransitive) when they act as modifiers within an NP. It too is lexicalised. Thus, for instance, *we hemu-n 'water drink-n' was rejected by consultants. The adjective usually describes the resulting condition of the undergoer, whether this be from a transitive base (e.g. tunu-n 'roasted') or an intransitive one (e.g. monu-n 'fallen'). The one counterexample is luku-n 'plough', which describes the actor (the plough buffalo). Examples given below are of whole NPs.

---

38 Morris (1984b) lists kili as a verb meaning 'tickle', but it was not recognised in the Fehan dialect.
4.7 Derivation of adverbs: full reduplication

Full reduplication, for which the phonological rules are described in §2.10.5, derives adverbs from bases of various classes. The most common base class, and the only one for which the derivation is perhaps productive, is the class of adjectives.

**Adj**

<table>
<thead>
<tr>
<th>N + Vt-n</th>
<th>Noun gloss</th>
<th>Vt gloss</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>batar tunu-n</td>
<td>maize (corn)</td>
<td>bake, roast</td>
<td>roasted maize</td>
</tr>
<tr>
<td>kabau luku-n</td>
<td>buffalo</td>
<td>plough</td>
<td>plough buffalo</td>
</tr>
<tr>
<td>ktubi sona-n</td>
<td>cake</td>
<td>fry</td>
<td>fried cakes</td>
</tr>
<tr>
<td>wé suma-n</td>
<td>water</td>
<td>dig (well)</td>
<td>man-made well</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N + Vi-n</th>
<th>Noun gloss</th>
<th>Vi gloss</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>batar botu-n</td>
<td>maize (corn)</td>
<td>pop</td>
<td>popped maize</td>
</tr>
<tr>
<td>fore monu-n</td>
<td>mung bean</td>
<td>fall</td>
<td>fallen bean seeds</td>
</tr>
<tr>
<td>fore tubu-n</td>
<td>mung bean</td>
<td>sprout</td>
<td>bean sprouts</td>
</tr>
<tr>
<td>wé tasa-n</td>
<td>water</td>
<td>fully cooked</td>
<td>boiled water</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>V</th>
<th>Adv</th>
</tr>
</thead>
<tbody>
<tr>
<td>bosok</td>
<td>boso-bosok</td>
</tr>
<tr>
<td>sala</td>
<td>sa-sála</td>
</tr>
</tbody>
</table>

Several derivations from nominal bases denoting times have the meaning ‘every such unit’. A remaining one (ora-'.oras) in contrast indicates ‘very little time’.

The base may also be a verb.

**N: time**

<table>
<thead>
<tr>
<th>V</th>
<th>Adv</th>
</tr>
</thead>
<tbody>
<tr>
<td>kalan</td>
<td>kala-kalan</td>
</tr>
<tr>
<td>loron</td>
<td>loro-loron</td>
</tr>
<tr>
<td>wain</td>
<td>wai-wain</td>
</tr>
<tr>
<td>oras</td>
<td>ora-'.oras</td>
</tr>
</tbody>
</table>

A number of common adverbs have the phonological form of a reduplication, although the root is unknown. These are nimanimak ‘continually, eternally’, nainaik ‘slowly’, and beibeik ‘constantly’.
4.8 Other functions of full reduplication

4.8.1 Reduplication of verb: ‘do heedlessly’

Full reduplication of verbs occurs in construction with a preceding na'i (sometimes na'īn, although this is less acceptable).39 This construction indicates that the action is being done aimlessly, without reason, or heedless of the prescribed rules.

Phonologically such reduplication follows the normal rules of Tetun reduplication, in that only the last two syllables of the verb are repeated (4.60; *nananu-nananu ‘RDP-3S.sing’), the final consonant of the initial copy is omitted (e.g. bó-bók ‘disturb for no reason’ from bók ‘disturb’), and the initial copy is optionally shortened further (e.g. na-nanu ‘sing for no reason’ from hananu ‘sing’). Occasionally a trisyllabic word is repeated in full when dictating, as the full repetition is the form that consultants consider should be used in writing. If the verb is inflected, the subject marking appears on both copies (4.62; one example).

A final -n is often added, apparently optionally, if the verb ends in a vowel (e.g. bá-bá(n) ‘go for no reason’ from bá ‘go’).

4.60 Nia na'i nasa-nasa-n.
3S just RDP-laugh(=hanasa)-just
She just laughs (even though there is nothing to laugh at). (V0.83)

4.61 Nia la bele na'i fasi-fasi.
3S not can just RDP-wash
It can’t just be washed heedlessly. (It must be done when prescribed by tradition.) (Q0.106)

4.62 Keta na'i m-aló-m-aló-n Ibu nia-kan buku.
do not just RDP-2S-make-just mother 3S-POS book[Mly]
Don’t just use Ibu’s book. (e.g. Don’t browse through it; you have no right or need to use it.) (N0.65)

4.63 Sira na'i bó-bók ha'i ita.
3P just RDP-disturb not 1P I
They don’t just disturb us for no reason. (V0.89)

4.64 Rua-rua simu malu join ne'e ha'ak 'hakāt'.
RDP-two reply each other then this say fight
Only when two (participants) respond to each other is it called ‘fighting’.

39 Note that na'i is a ‘construction marker’ (Matthews 1981:60) of this one construction, occurring nowhere else with comparable meaning. Despite its restricted distribution it shares several characteristics with verbs. In particular, it takes subject marking for 1S subjects, and is itself reduplicated if the following verb is elided (la na'i-na'īn ‘doesn’t just happen’). Auxiliaries and the preverbal negator la ‘not’ precede na'i (4.61), while the postverbal negator ha'i follows the reduplicated verb (4.63).
4.65  *Ita tau musan* há-hát, *lima-lima*.
1PI put seed RDP-four RDP-five
We plant four or five seeds at a time.  (T0.43)

There can be a distributive aspect to the meaning.

4.66  *Nén fa’e bá rua, ema ida tolu-tolu,*
six divide go two person one RDP-three
(When) six is divided amongst two, each person (receives) three.  (N0.151)

For powers of ten, reduplication can signify plurality, with *ida-’ida* meaning ‘a few’ (4.67), *atus-atus* ‘hundreds’, and *rihun-rihun* ‘thousands’ (4.68).

4.67  *Isin ida-’ida foin Ibu bosok sia tó?*
occasion RDP-one then mother deceive 3P
You (‘Ibu’) only tease them occasionally, don’t you?  (V0.105)

4.68  *Ema rihun-rihun né mate hotu r-alo mós.*
person RDP-thousand this die all 3P-make finished
These thousands of people all died.  (U3.13)

Ambiguity is possible, with *ida-’ida* able to mean ‘one at a time’, ‘one each’ or ‘several’.

4.69  *Hola ida-’ida dei.*
take RDP-one only
Take only one at a time/one each/a few.  (N0.174)

4.8.3 Reduplication of adjectives

Reduplication of adjectives has a variety of functions. It can indicate plurality (10 examples). This does not necessarily imply a large number of referents.

4.70  *...la bá iha uma se-seluk.*
not go LOC house RDP-other
[Ibu came only to this one house;] (she) didn’t go to other houses.  (Q0.126)

4.71  *Kabau aman bó-bót sia fa’en ti’an.*
buffalo male RDP-big PL sell already
The large male buffalo (whether few or many) have been sold.  (X0.89)

Frequently the reduplication in addition signifies variety amongst the referents being described. As one consultant said when I implied that *si-sín* ‘RDP-sour’ could apply to a single type of fruit:

4.72  *N-ó ai fuan oi-’oi-k mak sin foin ha’ak ‘si-sín’.*
3-exist plant fruit RDP-type-SUF REL sour then say RDP-sour
(If) there are various types of fruit which are sour, only then does (one) say ‘sour-sour’.  (N0.139)

There is limited evidence that reduplication of adjectives can signify intensification, a function which is listed for Tetun by several authors (Monteiro 1985:xxii; Morris 1984b:xiv; Troeboes et al. 1987:56). This may accompany a change in word class from adjective to adverb (§4.7).
Alternatively, two examples (including an adverb) suggest that reduplication can signify reduction (Monteiro 1985:xxii).

\[
\begin{align*}
\text{Adj} & \quad \text{Adj/Adv} \\
\text{kтомак} & \quad \text{完全} \\
\text{mesan} & \quad \text{单独} \\
\text{wa’ин} & \quad \text{大量, 多数}
\end{align*}
\]

As ы-ы ‘RDP-sour’ illustrates, reduplication of the one base can result in both a ‘plural and varied’ meaning, and in a ‘little bit’ meaning.

### 4.8.4 Repetition of nouns: various plural

Repetition of nouns is one means of indicating plurality (4.73; 28 examples). As for reduplication of adjectives, there is frequently an implication that the referents are varied (4.74).

Phonologically this meaning is usually expressed by full repetition of the noun. Unlike in morphological reduplication, words of three syllables are repeated in full, and the final consonant of an initial base is usually retained (e.g. ы’ан ы’ан ‘stable stable’ = ‘stables (of horses)’). However, occasionally the pattern is phonologically one of reduplication (e.g. ы-ы ы ‘piece-piece’ = ‘pieces’).

4.73 \[ \text{Bolu ы ы ы, alex baku.} \]
\[ \text{call woman woman come make food for trip} \]
\[ \text{They called women to come, (and) make provisions.} \] (G1.80)

4.74 \[ \text{Ума малей н-o ы’ан ы’ан.} \]
\[ \text{house non-native 3-have room room} \]
\[ \text{Non-Fehan houses have many and varied rooms (e.g. for dining, sleeping, cooking). (Note that Fehan houses have a single multipurpose room.)} \] (N0.132)

### 4.9 Miscellaneous low-frequency derivations

#### 4.9.1 Introduction

The following sections illustrate various low-frequency morphological devices which have functions other than those discussed above.

#### 4.9.2 Iteration: partial reduplication of verbs

Partial reduplication of verbs can add iterative meaning (10 possible examples). There is no co-occurring suffix.
V: base

<table>
<thead>
<tr>
<th>V</th>
<th>base</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>baku</td>
<td>strike</td>
<td>babaku</td>
</tr>
<tr>
<td>kose</td>
<td>rub</td>
<td>kakose</td>
</tr>
<tr>
<td>tabar</td>
<td>meet</td>
<td>tatabar</td>
</tr>
<tr>
<td>tei</td>
<td>stamp foot</td>
<td>tatei</td>
</tr>
<tr>
<td>tidin án</td>
<td>land on backside</td>
<td>tatidin án</td>
</tr>
</tbody>
</table>

4.75 Asu ikun ba-baku.

The dog's tail wags.

More commonly, iteration is expressed by fully repeating the verb. The verb may be repeated several times, and there can be pauses between them.

4.76 Sia la'o la'o la'o r-akur to'o fo ho ktaek

hitu...

They walked and walked and walked, crossing seven mountains...

4.9.3 Names of fingers and toes: fa-, sa-

The morphemes fa- and sa- derive names for fingers and toes, with the two forms being mostly interchangeable.

<table>
<thead>
<tr>
<th>Base</th>
<th>N: sa-</th>
<th>N: fa-</th>
<th>Finger/toe</th>
</tr>
</thead>
<tbody>
<tr>
<td>kawa’ik</td>
<td>sawa’ik</td>
<td>saturdu</td>
<td>thumb, big toe</td>
</tr>
<tr>
<td>tudu</td>
<td>point, select</td>
<td>faklaras</td>
<td>index</td>
</tr>
<tr>
<td>klaran</td>
<td>middle</td>
<td>sa’eki</td>
<td>middle and fourth</td>
</tr>
<tr>
<td>Eki</td>
<td>woman’s name</td>
<td>fa’eki</td>
<td>fourth</td>
</tr>
<tr>
<td>?</td>
<td></td>
<td>sakilik</td>
<td>little</td>
</tr>
</tbody>
</table>

The morpheme fa- is likely to be derived from the noun fua-n ‘finger, toe-GEN’ (also ‘fruit; heart’), as suggested by Jonker (1906:285). This would come about in the first instance by reduction of fuan to a single syllable fu as the initial element in a compound (§2.10.5.2). The existence of a glottal stop in fa’eki suggests that this word is still phonologically a compound rather than a prefixed derivation, since /h/ rather than a glottal stop is added intervocally in derivations using the prefix ha- (§2.10.3.1). Further reduction of fu to fa would come about if that syllable were unstressed (§2.3.2.4). The stress pattern suggests fa- has been reinterpreted as a derivation with an unstressed prefix, rather than as a compound with secondary stress on the initial root. Such distancing of the derivation from the original compound is supported by the variant forms with sa- (which is disallowed for the compound-like fa’eki), and by the fact that fua(n) can optionally co-occur with the prefix within the one NP (e.g. lima fua fatudu ‘hand/arm finger/toe point’ = ‘index finger’).40

---

40 A similar process appears to have occurred for ai ‘plant’, in such expressions as ai a-tuin ‘k.o. plant used to make drums’, and ai a-kiss ‘k.o. hardwood tree used for making posts’. Here too the general noun can be repeated once it has acquired unstressed status as a- within the (former) compound.
4.9.4 An alleged prefix ba-

The sequence ba- (with or without a suffix -k) is listed as a prefix by Monteiro (1985:xvii) and Troeboes et al. (1987:46). Both list it as deriving adjectives from nouns, where the adjective specifies the possession of the base. Monteiro also lists it as deriving qualitative adjectives from verbs, and Troeboes et al. as deriving causative verbs from adjectives, and nouns denoting typical objects from verbs. Their examples were not recognised in the Fehan dialect, and only the following possible derivations were found in the corpus.

<table>
<thead>
<tr>
<th>Base</th>
<th>Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>kroman</td>
<td>N light</td>
</tr>
<tr>
<td>lēt</td>
<td>N opportunity, space</td>
</tr>
<tr>
<td>liki</td>
<td>Vt winnow</td>
</tr>
<tr>
<td>rahu-n</td>
<td>N hair-GEN</td>
</tr>
<tr>
<td>baroma</td>
<td>N spectacles</td>
</tr>
<tr>
<td>bālēt</td>
<td>Adj free, available (have time)</td>
</tr>
<tr>
<td>baliki</td>
<td>N basket (e.g. for winnowing)</td>
</tr>
<tr>
<td>barahuk</td>
<td>Adj hairy, messy-haired</td>
</tr>
</tbody>
</table>

4.9.5 Nicknames: omission of k-

Where nicknames are based on adjectives beginning with a consonant cluster or an antepenultimate syllable /ka/, the initial /k/ or /ka/ tends to be omitted. The fact that the name is based on the adjective is clear to consultants.

<table>
<thead>
<tr>
<th>Adj</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>k(a)latar</td>
<td>Latar</td>
</tr>
<tr>
<td>kawa'ik</td>
<td>Wa'ik</td>
</tr>
<tr>
<td>kbo'uk</td>
<td>Bo'uk</td>
</tr>
<tr>
<td>kla'ok</td>
<td>La'ok</td>
</tr>
<tr>
<td>klāk</td>
<td>Lāk</td>
</tr>
</tbody>
</table>

4.9.6 Adjective: -s

The following examples suggest that -s is a frozen suffix primarily deriving adjectives, as pointed out by Jonker (1906:285). Many of the words listed below as ‘bases’ themselves have final -n or -k, which may well be fossilised suffixes too.

<table>
<thead>
<tr>
<th>Base</th>
<th>Adj</th>
</tr>
</thead>
<tbody>
<tr>
<td>halai V</td>
<td>V run</td>
</tr>
<tr>
<td>iku-n N</td>
<td>N tail-GEN</td>
</tr>
<tr>
<td>kfaluk Adj</td>
<td>widowed, single</td>
</tr>
<tr>
<td>klaran Adj</td>
<td>middle [usual term]</td>
</tr>
<tr>
<td>kmetin Adj</td>
<td>dense, having no holes</td>
</tr>
<tr>
<td>mean Adj</td>
<td>Adj red</td>
</tr>
<tr>
<td>mutin Adj</td>
<td>Adj white, whitish</td>
</tr>
<tr>
<td>talik Vt</td>
<td>Vt entwine</td>
</tr>
<tr>
<td>lais</td>
<td>fast</td>
</tr>
<tr>
<td>ikus</td>
<td>last, final</td>
</tr>
<tr>
<td>kfulus</td>
<td>widowed, single</td>
</tr>
<tr>
<td>klaras</td>
<td>middle [rare]</td>
</tr>
<tr>
<td>kmetis</td>
<td>tight, unmoving</td>
</tr>
<tr>
<td>meas</td>
<td>red (e.g. red mark on skin)</td>
</tr>
<tr>
<td>(k)mutis</td>
<td>pale, faded</td>
</tr>
<tr>
<td>fehuk talis</td>
<td>vine with edible tubers (fehuk)</td>
</tr>
</tbody>
</table>

41 This placename was explained in an origin myth as deriving from the word for ‘school of fish’.
The following attributive adjectives derived from numerals are all used as NP modifiers, without classifiers.\(^42\)

<table>
<thead>
<tr>
<th>Num</th>
<th>Attrib</th>
</tr>
</thead>
<tbody>
<tr>
<td>rua</td>
<td>two</td>
</tr>
<tr>
<td>tolu</td>
<td>three</td>
</tr>
<tr>
<td>hât</td>
<td>four</td>
</tr>
<tr>
<td>hitu</td>
<td>seven</td>
</tr>
</tbody>
</table>

\(^{42}\) The term *ruas* regularly occurs in the expression *ita ruas* 'we two'. All 55 instances of *ruas* in the corpus follow a plural pronoun (§5.2.1). There was disagreement amongst consultants as to whether *ita tolus* 'we three' constituted a valid expression. The word *hás* 'four' was recorded only in some ritual names, these being *Le'un Hás* (lit. 'four house-groups'), the name of a village which has four noble houses, and *Dato Hás Ain Ulun* (lit. 'four nobles leg head'), which refers to the combined kingdoms of Wehali and Wewiku, which are elsewhere described as *Ferik Hât Katuas Hât* 'four mature.women, four mature.men'. The only context for *hitus* 'seven' was *feito hitus*, lit. 'seven women', which refers to the seven sisters about whom many myths are told.
5 Numeral, classifiers and the numeral phrase

5.1 The numeral phrase

Numeral phrases consist minimally of a numeral, which may be preceded by a classifier. The classifier is syntactically optional (i.e. omissible without change of meaning) in the case of sortal classifiers; however, the classifier is the obligatory head of the phrase (indicated by underlining) in the case of mensural classifiers. The distinction between these two classes is discussed in §5.3.1, while the issue of headedness is discussed in §5.4.

Numeral Phrase, Sortal → (Sortal.Classifier) → Numerical
Numeral Phrase, Mensural → Mensural.Classifier → Numerical

Numeral phrases can function as NP modifiers (5.1, 5.2), as predicates (5.3), or in headless NPs (5.4).

5.1 ...fahi ida, tua botir hät, nó tais lolon ida. pig one wine bottle four and cloth CLS:trunk one
[He accepted the fine:] a pig, four bottles of wine and one handwoven cloth. (J4.39)

5.2 Dadi nia kolu kolu kolu faru tahan tolu... so 3S take.off take.off take.off clothes CLS:sheet three
So she slowly undressed, taking off three layers of clothes [, and there were still four layers of clothing to go]. (G4.105)

5.3 Na feton na’in hitu. 3S man’s.sister CLS:human seven
His sisters were seven. (i.e. He had seven sisters.) (O1.29)

5.4 Na’in rua di’ak, na’in ida ikun ne’e mak...
CLSHuman two good CLS:human one tail this REL
Two (of the brothers) were good (looking), it was the youngest one who [was deformed]. (U5.2)

In the remainder of this chapter, numerals and classifiers will be considered in turn.
5.2 Numerals

5.2.1 Cardinal numerals

Tetun has a decimal numeral system.¹ The basic numerals are as follows:

<table>
<thead>
<tr>
<th>Numeral</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ida</em></td>
<td>1</td>
</tr>
<tr>
<td><em>rua</em></td>
<td>2</td>
</tr>
<tr>
<td><em>tolu</em></td>
<td>3</td>
</tr>
<tr>
<td><em>hat</em></td>
<td>4</td>
</tr>
<tr>
<td><em>lima</em></td>
<td>5</td>
</tr>
<tr>
<td><em>nen</em></td>
<td>6</td>
</tr>
<tr>
<td><em>hitu</em></td>
<td>7</td>
</tr>
<tr>
<td><em>walu</em></td>
<td>8</td>
</tr>
<tr>
<td><em>siwi</em></td>
<td>9</td>
</tr>
<tr>
<td><em>sanulu</em></td>
<td>10</td>
</tr>
<tr>
<td><em>atus</em></td>
<td>100</td>
</tr>
<tr>
<td><em>rihun</em></td>
<td>1,000</td>
</tr>
<tr>
<td><em>hir</em></td>
<td>how much, how many, several, however many</td>
</tr>
</tbody>
</table>

*A taboo avoidance term for ‘3’ is *kabau* (§5.2.2).

For numerals beyond a thousand, *juta* (or *duta*) ‘million’ is borrowed from Indonesian.² Such large numerals are in any case seldom used, since few local people have access to that much money, and the bride price, which in northern Belu can exceed a million rupiahs (approximately $700 Australian in 1995), is very low in the matrilocal Fehan area.

The formula for numerals from ten to just under a million is as follows, where at least one of the first three bracketings (i.e. for thousands, hundreds or tens) must be specified. In the formula ‘Digit’ covers the numerals from 1 to 9. The examples that follow the formula are lined up in columns with it for easier comparison.

Consultants were unaware of the quinary system which Morris (1984b:xix) says exists alongside the decimal system in East Timor.

Mathijsen (1906:9) lists *be'in* (written 'beei n') as ‘10000’. Several knowledgeable consultants recognised it as a very large number, but could not agree on what exact value it had. Morris (1984b) lists the following large numbers, which I have not come across: *reben* ‘ten thousand, the number above *rihun*’; *tokon* ‘the next number above *reben* (possibly one hundred thousand (100,000) or one million (1,000,000)); *ala(n)* ‘a number larger than *tokon*, indicates a large number, an infinite number’; *nanun* ‘a billion (million million)’. Tilman (1996:48) proposes *mamuk* (lit. ‘empty’) for zero, *rutokon* for ‘million’ and *toltokon* or *tolnanun* for ‘million million’.

---

¹ Consultants were unaware of the quinary system which Morris (1984b:xix) says exists alongside the decimal system in East Timor.

² Mathijsen (1906:9) lists *be'in* (written 'beei n') as ‘10000’. Several knowledgeable consultants recognised it as a very large number, but could not agree on what exact value it had. Morris (1984b) lists the following large numbers, which I have not come across: *reben* ‘ten thousand, the number above *rihun*’; *tokon* ‘the next number above *reben* (possibly one hundred thousand (100,000) or one million (1,000,000)); *ala(n)* ‘a number larger than *tokon*, indicates a large number, an infinite number’; *nanun* ‘a billion (million million)’. Tilman (1996:48) proposes *mamuk* (lit. ‘empty’) for zero, *rutokon* for ‘million’ and *toltokon* or *tolnanun* for ‘million million’.
**5.5**

<table>
<thead>
<tr>
<th>Numeral</th>
<th>(rihun Numeral)</th>
<th>(atus Digit)</th>
<th>(Digit-nulu)</th>
<th>(resin Digit)</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>'1,000'</td>
<td>rihun sanulu resin nén</td>
<td>atus siwi</td>
<td>resin tolu</td>
<td>'16,000'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rihun atus ida rua-nulu</td>
<td>atus walu</td>
<td>resin lima</td>
<td>'120,000'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rihun ida</td>
<td>atus ida</td>
<td>resin rua</td>
<td>'1,993'</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>resin tolu</td>
<td>'800'</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>resin lima</td>
<td>'125'</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>resin rua</td>
<td>'12'</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>lima-nulu</td>
<td>'50-what?'</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>resin hira</td>
<td>'several 10s'</td>
<td></td>
</tr>
</tbody>
</table>

All numerals are regular according to the above formula, with two exceptions: 'ten' is expressed by *sanulu* rather than *ida-nulu*, and 'thirty' by the obligatory abbreviation *tolu-nulu* rather than incorporating the full digit (*tolu-nulu*). The bound root -nulu ‘tens’ is sometimes abbreviated to -nu. In colloquial speech -nulu is omissible after the prefix *sa- ‘one’* (i.e. for numerals from 11 to 19, e.g. *sa-(*-nulu)* resin lima ‘15’); it is, however, not omissible after other digits.

The word *resin* ‘extra, leftover’ introduces the final digit (representing the ‘units’ column) in a numeral. It can also follow relatively imprecise numerals to indicate ‘more than this number’ (e.g. *fulan ida resin* ‘month one extra’ = ‘over a month’; *rihun sanulu resin* ‘thousand ten extra’ = ‘more than 10,000’).

**5.6**

Sesawan á, ami ho’o karau ida, fahi sanulu resin... 
morning DEF 1PE kill buffalo one pig ten extra 
In the morning we killed a buffalo, more than ten pigs (i.e. up to 20) [, we had a party, and then it was finished]. *(R1.80)*

The numeral *hira* is used for unknown quantities, both in questions meaning 'how much, how many' (5.7) and as an indefinite expression meaning 'several, however many' (5.8). In this it semantically parallels the epistememe pronouns such as *sé ‘who, anyone, someone’* (§6.6).

**5.7**

*Nia fulan hira ti’an?*  
3S moon how many already  
How many months (old) is he? *(T0.117)*

**5.8**

*Emi na’in hira ne’e oin nú malu dei.*  
2P CLS:human how many this face like each other only  
You several look like each other. *(The speaker was actually referring to four people.)* *(X0.13)*

Several of the numbers also have adjectival forms derived from a numeral root and a suffix -s (5.9; §4.9.6) or -k (e.g. *ruak* ‘two’; §4.6.7). These occur only as attributive adjectives and cannot take a classifier (*ita na’in ruas ‘1PI CLS:human two’).

**5.9**

*Ita ruas bá harís lai.*  
1PI two go bathe first  
Let’s we two go and bathe now. *(F5.5)*
5.2.2 Culturally significant numbers

Saying tolu ‘three’ in public is considered rude and an offence to men. This is due to the fact that it sounds the same as the noun tolu ‘egg’, which could be interpreted as denoting testicles. The common avoidance term is kabau (lit. ‘buffalo’), with de’an tán (lit. ‘speak angrily because’) an alternative avoidance term sometimes incorporated from the Foho and Suaid dialects. The term kabau cannot be used in the expression for ‘thirty’ (*kabau-nulu), for which the offending term tolu is in any case abbreviated (to-nulu). Kabau can be used in higher complex numerals (e.g. atus kabau ‘three hundred’), but tolu is usually used in these contexts.

The linguistic taboo on ‘three’ is reflected in the avoidance of three in daily life. For instance it is very poor form to offer someone a betel container containing exactly three betel nut slices, or to photograph three people together.

In ritual, seven is a significant number. It is, for example, the number of corn cobs brought into the sacred house by each person during the ceremony of hatama batar mana’ik (lit. ‘bringing in the sacred maize’), and the number to which one counts before cutting the umbilical cord of a newborn baby. In ritual language, hitu ki’ik (lit. ‘little seven’) refers to the three slices of betel nut, three betel pepper leaves and 3,000 rupiahs given for healing of a minor illness, while hitu bót (lit. ‘big seven’) refers to the seven slices of betel nut, seven betel pepper leaves and 7,000 rupiahs given after healing of a major illness. The significance of seven is further reflected in poetic language, where hitu (lit. ‘seven’) is used to mean ‘many’.

5.2.3 Arithmetic

Arithmetic equations make use of a transitive verb between the two operands. For addition this is tu’an ‘add’, or tau tán ‘put on top of’ (5.10). Subtraction is done using hasai ‘remove’ (lit. ‘make-exit’; 5.11). Division is expressed by fa’e ‘divide’ (5.12). Consultants disagreed as to whether and how multiplication could be carried out in Tetun.

The result of the arithmetic operation may be introduced by a pause (5.10) or by a semantically appropriate verb such as hela ‘leave, remain’ for subtraction (5.11) or hetan ‘get, receive’ for division (5.12).

5.10 Rua tu’an hát, nén. = Rua tau tán hát, nén.  
Two, add four, (equals) six. (i.e. 2 + 4 = 6)  (H0.64)

5.11 Nén ha-sai hát hela rua.  
six make-exit four leave two  
Six, remove four, leaves two. (i.e. 6 - 4 = 2)  (H0.64)

5.12 Nén fa’e tolu hetan rua.  
six divide three get two  
(When) six is divided amongst three, (each one) receives two. (i.e. 6/3 = 2)  (H0.64)

---

3 According to the anthropologist Francillon (1967:432), kabau ‘buffalo, horse’ is short for kabau ida nó balu ‘horse one and half’ = ‘one and a half horses’. The logic of this is that a horse carries two packs (one on either side), so that one and a half horses is equivalent to three packs.
Approximation is expressed by a wide variety of means, including use of the numeral *hira* ‘several, however many’ (5.8), reduplication of a numeral signifying a power of ten (§4.8.2), appending *resin* ‘extra, more than’ to the number (5.6), preceding the number with *to’o* ‘reach, up to’ (5.14), or listing numerals in the approximate range (5.15).

But when we flick (in a game), there are up to four or five of us (playing).

(When) six is divided amongst two, each person (receives) three. (i.e. $6/2 = 3$)

But when we flick (in a game), there are up to four or five of us (playing).

(With regard to) this turning of the soil: firstly, we ram digging sticks into the ground.

5.13 Nén fa’e bá rua, ema ida tolu-tolu.
six divide go two person one RDP-three

5.14 Mais ami tilak, ami na’in to’o hát lima.
but 1PE flick 1PE CL:human reach four five

5.15 Kahi isin *rua* tolu moras kók ne’e n-akraik.
massage times two three sick swollen.spleen this 3S-deflate

5.16 Fila rai ne’e, *uluk fohon* ita kuda suan bá rai.
turn earth this firstly 1PI ram digging.stick go earth

5.2.4 Ordinal numbers and alternative expressions

During elicitation some speakers maintained that *ka-* is a prefix that derives ordinal numbers from numeral roots (e.g. *ka-tolu* ‘third’ from *tolu* ‘three’). Others argue that this is a recent borrowing from Indonesian *ke*. The latter is supported by the fact that I have not noted such derived ordinals in actual use, as well as by their absence from the dictionaries of Mathijisen (1906) and Morris (1984b). Instead, if used at all, ordinal numbers are borrowed from Indonesian (e.g. *ketiga* or *katika* ‘third’). Alternatively, they are indicated by prefacing the Tetun or Indonesian numeral with Indonesian *nomor* ‘number’ (e.g. *feto nomor tolu* ‘woman number three’ = ‘(the) third woman’).

Within families and groups of people, the specification of order in age is normally restricted to the three-way distinction *ulun* ‘head, first, eldest’, *klaran* ‘middle’, and *ikun* ‘tail, last, youngest’, or the two-way distinction *kawa’ik* ‘older’ and *ki’ik* ‘younger, small’.

Within procedural accounts, the first step in doing something is often introduced by *uluk fohon* ‘firstly’ (lit. ‘former top’). However, there are no equivalent introductions for subsequent steps, other than the ubiquitous *nia ti’a* ‘then’.
5.3 Classifiers

5.3.1 Introduction

Numeral classifiers are nominals which have the unique characteristic of being able to occur before a numeral in a numeral phrase. Since Tetun has no other types of classifiers, the shorter term ‘classifier’ will be used to refer to this class.

Tetun has two types of classifiers, sortal and mensural (Lyons 1977:463). Sortal classifiers (e.g. na’in for persons) are members of a small closed class of words which classify whole objects. Their collocations with the nouns they classify are not fully predictable. Mensural classifiers (e.g. futun ‘bundle’) form a much larger class to which new members are readily added (e.g. glás ‘glass’ via Malay); these are common nouns which refer to the form of the entities being counted, rather than to the entities themselves.

There is some overlap in the two classes. For instance, fuan ‘fruit’ is a sortal classifier for roundish objects (such as eggs, batteries and mangoes) as well as a mensural classifier for fruits (as opposed to other parts of a plant). Thus hás fuan hát ‘mango fruit four’ could in principle be analysed as having either a sortal classifier (‘four mangoes’) or a mensural classifier (‘four mango fruits (as opposed to mango trees’)).

The two types of classifiers will be discussed in turn.

5.3.2 Sortal classifiers

5.3.2.1 Inventory

The number of sortal classifiers in Tetun is limited, with Table 5.2 presenting a reasonably full list of those in common use. These words classify only concrete common nouns.

As can be seen in the table, there are only two classifiers for animate beings. One is na’in (lit. ‘noble, owner’), used for persons, including personified animals and other personified entities. The other, matan (lit. ‘eye, source’), is for pigs and buffalo, the two large domestic animals which are traditionally eaten. Other animals do not have a classifier. Not surprisingly the human classifier na’in accounts for the vast majority of sortal classifiers in the corpus.
Inanimate objects are classified by shape.\footnote{These classifiers accord with Conklin's (1981:226ff.) assertion that plant-based categories are important in the classifier systems of Western Austronesian languages (in which she includes languages of Timor). The Tetun classification also fits Croft's (1994:152ff.) claim that the primary distinction in sortal classifier systems is based on animacy, while the second is based on shape. In addition, it supports his observation that where classifier systems use shape, they make at least a three-way distinction between one-dimensional (long), two-dimensional (flat) and three-dimensional (round) shapes.} Again many objects (e.g. tanasak ‘basket’) do not have a corresponding classifier. There is no general ‘default’ classifier such as is found, for instance, in Malay (Hopper 1991b).

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Noun meaning</th>
<th>Classifier for</th>
<th>Examples</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>na’in</td>
<td>noble, owner</td>
<td>persons</td>
<td>ema</td>
<td>person</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>feto</td>
<td>woman</td>
</tr>
<tr>
<td>matan</td>
<td>source, eye</td>
<td>large domestic animals</td>
<td>kabau</td>
<td>buffalo</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>fahi</td>
<td>pig</td>
</tr>
<tr>
<td>lolon</td>
<td>trunk</td>
<td>long objects</td>
<td>lilin</td>
<td>candle</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>na’an</td>
<td>fish</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>knose</td>
<td>rib</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ruin</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>tais</td>
<td></td>
</tr>
<tr>
<td>tahan</td>
<td>leaf</td>
<td>thin flat objects</td>
<td>faru</td>
<td>clothing</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>fuik</td>
<td>betel pepper</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>buku</td>
<td>book [Malay]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>knose</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ruin</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>tais</td>
<td></td>
</tr>
<tr>
<td>fuan</td>
<td>fruit, heart</td>
<td>whole roundish objects</td>
<td>bua</td>
<td>betel nut</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>nú</td>
<td>coconut</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>kól</td>
<td>cabbage [Malay]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>knose</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ruin</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>tais</td>
<td></td>
</tr>
<tr>
<td>musan</td>
<td>seed</td>
<td>very small round objects</td>
<td>ai tahan</td>
<td>tablet</td>
</tr>
</tbody>
</table>

The following examples show sortal classifiers within NPs. Sortal classifiers used without head nouns are illustrated in 5.3, 5.4 and 5.22.

5.17 \(\text{kbú} \quad \text{fuan} \quad \text{hát} \)
\(\text{rice parcel} \quad \text{CLS:round} \quad \text{four} \)
\(\text{four rice parcels} \) \(\text{(U3.5)}\)

5.18 \(\text{tais} \quad \text{lolon} \quad \text{ida} \)
\(\text{cloth} \quad \text{CLS:trunk} \quad \text{one} \)
\(\text{one (handwoven) cloth} \) \(\text{(J4.30)}\)

5.19 \(\text{sia} \quad \text{na’in} \quad \text{rua} \)
\(\text{3P} \quad \text{CLS:human} \quad \text{two} \)
\(\text{they two (people)} \) \(\text{(K12.8)}\)

Since sortal classifiers are all homophonous with common nouns, ambiguity is possible between a classifier and a noun interpretation. This potential ambiguity is resolved by stress patterns, with the word carrying phrasal stress if it is a noun (e.g. ema na’in ida ‘person noble one’ = ‘a noble’) but not if it is a classifier (e.g. ema na’in ida ‘person CLS:human one’ = ‘one person’).
5.3.2.2 Use

Syntactically sortal classifiers are optional in that their omission does not change the descriptive meaning. However, their use is considered polite, especially when referring to humans. The preference for a classifier is particularly strong for NPs with plural pronoun heads (5.19), where the absence of a classifier (e.g. *sia rua* ‘3P two’) is considered rude. This stated preference is supported by the corpus, with 85% of the 105 enumerated pronoun examples having a classifier. When humans are referred to by nouns (rather than pronouns) the absence of a classifier appears to be more acceptable, judging by the fact that only 62% of the 185 such enumerated nouns have a classifier.

A classifier is used with *ida* ‘one’ only when one emphasises the fact that it is ‘one’, as opposed to some other numeral (9 examples for persons; 5.20), and not when it is used as an indefinite article (5.24; 147 examples for persons; §6.5.2).

5.20 Uma.kain ida mane na’in sanulu. Mane na’in ida mak karian...
household one man CLS:human ten man CLS:human one
REL work
A household (has) ten men. There is one man who works. [The others just wander around aimlessly.] (F8.33)

In elicitation classifiers were accepted for large numerals such as 150 (5.21); the largest unelicited numeral with a classifier in the corpus is 50 (*ema na’in lima-nulu* ‘person CLS:human five-tens’ = ‘fifty people’).

5.21 kabau matan atus ida lima-nulu
buffalo CLS:animal hundred one five-tens
150 head of cattle
(Q0.99 elicited)

Classifiers can be used with the taboo avoidance term *kabau* ‘three’ (5.22), the indefinite/interrogative *hira* ‘several, how many’ (5.8), and reduplicated numerals (5.23).

5.22 Nambé na’in rua ká, na’in kabau.
I.suppose CLS:human two or CLS:human three
I suppose two or three. (Answer to: “How many children does he have?”) (Q0.56)

5.23 ...n-aruka sia to’o na’in to-tolu, ká na’in
3S-order 3P reach CLS:human RDP-three or CLS:human
há-há!
RDP-four
[The priest] ordered them (to go in groups of) up to three or four. (T0.39)

While classifiers can be used anaphorically without a head noun (5.4), this is not a major function in discourse. Of 210 NPs incorporating the human classifier *na’in*, only 5% have no head noun, while 40% have a pronominal head (5.19), and 55% have a nominal one (5.20). In any case a human classifier is unlikely to assist much in disambiguating reference

---

8 There appears to be dialectal variation in this, in that a very articulate speaker from the Suai region (whose examples are not included in the counts in this section) used the human classifier far less consistently than Fehan speakers did.
Numerals, classifiers and the numeral phrase

(a problem noted for singular referents in Japanese by Downing (1986:355)), since the main alternative contenders for the reference interpretation are likely to be other human participants.

Textually, classifiers are most commonly used in NPs with definite reference. This is illustrated by the fact that many have pronominal heads, while 65% of those with nominal heads have some marker of definiteness (e.g. *ne’e ‘this’, a possessor, or a relative clause). Nevertheless, classifiers can be used with new referents. This is illustrated by the statement commonly found at the beginning of stories to the effect that a couple had a certain number of children (5.24), one or more of whom are subsequently central to the story.

5.24  *Sira té r-ola oan na’in rua: feto ida, mane ida.*
3P defecate 3P-take child CLS:human two girl one boy one
They gave birth to two children: a girl (and) a boy. (U4.72)

5.3.3 Mensural classifiers

5.3.3.1 Inventory

Mensural classifiers specify the form in which the counted item occurs, by identifying a part, container or collection. These classifiers are all semantically transparent common nouns and form an open class, with loans being readily incorporated.

Semantically they may be divided into three groups. Part classifiers specify which part of the entity is being considered. Container classifiers identify a type of container. Collection classifiers distinguish size and type of grouping. As the table below shows, a single noun can collocate with a range of classifiers. For instance, the staple food *batar* ‘maize (corn)’ can be measured as a whole plant, as a cob or in various types and sizes of bundling.

5.25  *tua kusi ida*
wine earthen.jar one
one earthen jar (full) of wine (F1.8)

5.26  *rahenek sa-sukat nén, samén sa-sukat ida*
sand NOM-measure six cement[Mly] NOM-measure one
six measures of sand, one measure of cement (AA3.5)

---

9 This is in marked contrast to Malay. In modern Indonesian, classifiers are said to occur only in indefinite NPs (Sneddon 1996:135), while in a nineteenth century Malay text studied by Hopper (1986:314), classifiers occur with definite nouns only if “there is a strong focus on enumeration”.

5.3.3.2 Comparison with nouns

Unlike sortal classifiers, mensural classifiers cannot be omitted without change of meaning, since it is the classifier which indicates what is being counted. Since mensural classifiers are semantically transparent, it is not always clear whether they should in fact be interpreted as a classifier, or instead as simply a common noun which specifies a standard measure.

Where there is no preceding noun specifying the substance to be counted, as in 5.27, the phrase containing the numeral can be interpreted either as an NP headed by a measure noun (‘one hand’) or as an NP with elliptical noun head and a classifier phrase (‘one hand of (unspecified substance)’). The former is preferred as the simpler analysis.

5.27...\textit{Liman} \textit{ida} \textit{kona} \textit{fulin} \textit{sanulu} \textit{resin} \textit{rua}.
\hfill (V0.81)

Where the measure term occurs between a noun and a modifying numeral, two structural interpretations are in principle possible. Either it forms a constituent with the following numeral (so being a classifier; 5.28i), or it relates directly to the preceding noun (as a

<table>
<thead>
<tr>
<th>Type</th>
<th>Classifier</th>
<th>Meaning</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part</td>
<td>fulin</td>
<td>head (grain), bunch (banana)</td>
<td>maize, rice, sorghum, banana</td>
</tr>
<tr>
<td></td>
<td>hún</td>
<td>trunk, origin</td>
<td>coconut tree, betel palm</td>
</tr>
<tr>
<td></td>
<td>klinun</td>
<td>section of trunk</td>
<td>sago trunk</td>
</tr>
<tr>
<td></td>
<td>knés</td>
<td>small piece</td>
<td>piece of banana, wood</td>
</tr>
<tr>
<td></td>
<td>rohan</td>
<td>short length</td>
<td>wood, bamboo</td>
</tr>
<tr>
<td>Container</td>
<td>blék</td>
<td>tin (measures 2 1/2 litres) [Mly]</td>
<td>harvested maize, rice, legumes</td>
</tr>
<tr>
<td></td>
<td>falun</td>
<td>parcel</td>
<td>tobacco, noodles, tablets</td>
</tr>
<tr>
<td></td>
<td>glás</td>
<td>glass [Mly]</td>
<td>liquid</td>
</tr>
<tr>
<td></td>
<td>kusi</td>
<td>earthen jar (equals about 12 bottles)</td>
<td>wine</td>
</tr>
<tr>
<td></td>
<td>tebok</td>
<td>bowl, plate</td>
<td>rice, cooked meat</td>
</tr>
<tr>
<td></td>
<td>to’os</td>
<td>garden (full)</td>
<td>cassava, legumes</td>
</tr>
<tr>
<td>Collection</td>
<td>bunus</td>
<td>large bundle tied together</td>
<td>maize (about 100 cobs), fish, meat, chickens (about 10)</td>
</tr>
<tr>
<td></td>
<td>liman</td>
<td>small bundle tied together</td>
<td>maize (about 12 cobs), grass</td>
</tr>
<tr>
<td></td>
<td>futun</td>
<td>bundle tied together</td>
<td>grass, firewood</td>
</tr>
<tr>
<td></td>
<td>ahuk</td>
<td>clump (of plants)</td>
<td>maize, banana, beans</td>
</tr>
<tr>
<td></td>
<td>butuk</td>
<td>pile (of loose items)</td>
<td>betel nut, mandarins, onions</td>
</tr>
<tr>
<td></td>
<td>daun</td>
<td>hand (of bananas)</td>
<td>bananas</td>
</tr>
<tr>
<td></td>
<td>doen</td>
<td>large portion of meat</td>
<td>meat</td>
</tr>
<tr>
<td></td>
<td>klubun</td>
<td>group (of people)</td>
<td>people</td>
</tr>
<tr>
<td></td>
<td>lalu’an</td>
<td>herd, animal pen</td>
<td>cattle, goats, horses</td>
</tr>
</tbody>
</table>
possessed noun; 5.28ii). The two possibilities, with an unambiguous example of each, are presented diagrammatically below.

![Diagram of noun phrases]

The noun interpretation is possible when the measure term can be interpreted as the head in a whole-part noun sequence (§7.3.2.3). This is the case in example 5.28ii above, where a classifier interpretation of *hún* ‘trunk’ is prevented by the adjective that intervenes between *hún* and the numeral. In contrast, a classifier interpretation is forced in 5.28i, because a noun sequence *tua* botir would not mean ‘bottle of wine’.

Some examples are in fact indeterminate. The NP *ai hún rua* ‘plant/tree trunk two’ can be interpreted as either ‘two trees’ (where *hún* is the classifier for trees), or ‘two tree-trunks’ (where *hún* is a possessed noun). Similarly, the NP *batar Julin hitu* ‘maize (corn) cob seven’ could equally be interpreted as ‘seven cobs of corn’ (with a classifier interpretation), or as ‘seven corn-cobs’ (with a noun interpretation). This structural indeterminacy, however, results in very little semantic difference, as the above examples show.

### 5.4 Head of a numeral phrase

As noted above, mensural classifiers have much in common semantically and syntactically with common nouns. It is only the possibility of them occurring in the same syntactic slot as classifiers (i.e. between NP head and numeral) that leads me to analyse them as having dual word class membership, of the class of common nouns and the class of classifiers. The close similarity to common nouns points to the mensural classifier being the head of the numeral phrase, just as a common noun is the head of an enumerated NP. Such an analysis is consistent with the fact that the classifier is obligatory. In contrast, sortal classifiers have nothing which points to them being head of the phrase, since they are not obligatory, and the numeral phrase has the same external distribution as a numeral.\(^\text{12}\)

---

\(^{10}\) In principle the measure term could also be interpreted as a modifier of the preceding noun (§7.5.1). However, I have found no classifiers which could be interpreted that way semantically.

\(^{11}\) This indeterminacy is a common situation for languages in the region, according to Conklin (1981:279). She observes that for ‘Western Austronesian’ languages (in which she includes languages of Timor), lexically transparent markers, morphemes which also serve as nouns, are predominant and “It is not always possible to draw a clear distinction between generic usage and classification”. Dixon (1986:106) similarly notes that “There are often some nouns which seem to lie on the threshold between classifiers and specific nouns”.

\(^{12}\) The same analytic distinction is drawn by Karhunen (1994:44) for the Austronesian language Padoe. The author describes sortal classifiers as “a kind of auxiliary word” and mensural classifiers as “measure nouns” which head a “measuring phrase”.
6 Pronouns and determiners

6.1 Overview

Pronouns are a closed class of words which can head NPs, but which, unlike common nouns, cannot function as NP modifiers. They accept a more limited range of modifiers than do common nouns (§7.2.4), and in particular cannot be modified by the definite article å. There are four subclasses of pronouns, distinguished on both syntactic and semantic grounds. These are personal pronouns, demonstrative pronouns, quantifying pronouns and interrogative/indefinite pronouns (‘epistememes’).

Determiners are a closed class of words which function as modifiers within NPs, indicating definiteness and/or number. This class includes the demonstratives, quantifying hotu-hotu ‘all’, the definite article å, the indefinite article ida, the plural marker sia and epistemic determiners.

Pronouns and determiners are discussed together in this chapter on account of the fact that there is much overlap in membership between these two classes. That is, many forms can function both as NP heads (where they are classed as pronouns) and as NP modifiers (where they are classed as determiners).¹

Table 6.1: Pronouns and determiners

<table>
<thead>
<tr>
<th>Subclass</th>
<th>Lexeme</th>
<th>Pronoun?</th>
<th>Determiner?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal pronoun</td>
<td>1st and 2nd person</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>sia ‘3P, PL’</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Personal pronoun/demonstrative</td>
<td>nia ‘3S, that’</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Demonstrative</td>
<td>ne’e ‘this’</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>ne’et ‘that’</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>nemai ‘here, this’</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>nebá ‘there, that’</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Quantifying pronoun/determiner</td>
<td>hotu-hotu ‘all’</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>hotu ‘all’</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ida-.idak ‘each’</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Other determiner</td>
<td>á ‘DEF’</td>
<td></td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>ida ‘one, a’</td>
<td></td>
<td>Y</td>
</tr>
<tr>
<td>Epistememe</td>
<td>sé ‘who, which person’</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>nabé ‘which, where’</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td></td>
<td>sá ‘what’</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

¹ Of course one could alternatively posit a word class which has distribution as both heads and modifiers of NPs, so avoiding assigning dual class membership to words like sá ‘what’.

² *Ida ‘one’ can head NPs in its capacity as a numeral (e.g. ida ne’e ‘one this’ = ‘this one’).*
In this chapter the personal pronouns are surveyed first, followed by demonstratives, quantifying terms and those determiners which are not discussed in any of the other sections. The chapter concludes with a discussion of epistememes.

6.2 Personal pronouns

6.2.1 Overview

Personal pronouns are a closed class of nominals used to refer to the speaker, the addressee, and to other persons or things whose referents are presumed to be clear from context (Schachter 1985:25). Syntactically they have four unique features which distinguish them from other word classes. Firstly, only personal pronouns and the noun *ema* ‘person’\(^3\) can be followed by the possessive enclitic *-kan* or its variant *-k* (§7.3). Secondly, only plural pronouns can be modified by the numeric adjective *ruas* ‘two’ (e.g. *sia ruas* ‘they two’), with nouns instead being modified by the numeral form *rua* ‘two’ (e.g. *uma rua* ‘house two’ = ‘two houses’). Thirdly, personal pronouns are the only class of nominals to be inherently specified for person or number. And finally, they are the only class of nominals whose members have two forms, the citation form being a full phonological word, and the phonologically reduced form tending towards clitic status.\(^4\) The personal pronouns are listed below.

<table>
<thead>
<tr>
<th>Citation form</th>
<th>Reduced form</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ha</em>’u*</td>
<td><em>ha, h</em>(^5)</td>
<td>1S</td>
</tr>
<tr>
<td>ó</td>
<td>a</td>
<td>2S</td>
</tr>
<tr>
<td><em>ita</em></td>
<td><em>ni, na</em></td>
<td>2S.HON</td>
</tr>
<tr>
<td><em>nia</em></td>
<td><em>it</em></td>
<td>3S</td>
</tr>
<tr>
<td><em>ami</em></td>
<td><em>am</em></td>
<td>1PE</td>
</tr>
<tr>
<td><em>emi</em></td>
<td><em>em</em></td>
<td>2P</td>
</tr>
<tr>
<td><em>sta, sira</em>(^6)</td>
<td><em>si, sa</em></td>
<td>3P</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Like common nouns, personal pronouns function as head of NPs functioning as subject or object of a clause, object of a preposition, possessor in a larger NP, and (under appropriate

---

\(^3\) The word *ema* ‘person’, when used without modifiers, tends to mean ‘someone, someone else’. It is unique in that it can either take a possessive enclitic directly, like a pronoun (e.g. *ema-kan uma-n* ‘someone’s house-GEN’; 6 examples), or be followed by a personal pronoun *nia* ‘3S’ or *sia* ‘3P’ to which the clitic is attached, as is the case for common nouns (e.g. *ema nia-kan ro’os* ‘someone’s garden’; 18 examples). It thus appears to be somewhat grammaticised as an indefinite pronoun, which shares the property of direct possessive marking with personal pronouns, but not the property of person-number marking. A common noun meaning ‘person’ is a common source for indefinite pronouns crosslinguistically (Haspelmath 1997:27f).

\(^4\) Note that subject markers, which in the grammars of some related languages are analysed as bound pronouns, are in this grammar analysed as verb agreement morphemes. They are discussed in §9.3.

\(^5\) The form *h* is used only as a possessive proclitic to vowel-initial nouns (e.g. *h ina* ‘my mother’). Vowel-initial verbs to which it could be a proclitic subject are rare.

\(^6\) In the corpus *sia* (which is listed by Mathijsen (1906) for the Foho dialect of northern Belu) is significantly more common than *sira* (which is the form listed by Cliff Morris (1984b) for East Timor). The difference is presumably one of source dialect.
discourse conditions) as single-word utterances. The citation form of the pronoun can be used in all these contexts. This full form must be used whenever the pronoun is emphasised, focused, or used contrastively, and so is required in contexts such as the initial NP in cleft constructions, topicalised objects, and left-dislocated NPs. It must also be used if the NP is modified (e.g. ha‘u ne‘e ‘1S this’ = ‘I here’), or if the pronoun takes possessive marking (e.g. ha‘u-kan ‘my, mine’). The full form is also the only form used in writing, and most speakers used it consistently when speaking carefully to me in the early stages of language learning.

The phonologically reduced forms, in contrast, can only be used for subjects and for preposed possessors, and even then only if the pronoun is not emphasised or modified. In such contexts consultants prefer the reduced form to the citation form.

Full pronoun forms are illustrated in 6.2 and 6.7, while reduced forms are shown in 6.1 and 6.10, amongst others.

As mentioned above, the reduced form phonologically tends towards being a proclitic. While it can retain stress and the vowel quality found in the full form (both features of full words), it can alternatively be an unstressed clitic to the following word, with the vowel /a/ that is found in all other proclitics and in prefixes (§2.3.2.6). This accounts for the variant vowels found in the third person reduced pronouns. Cliticisation is more common for possessors than for subjects but is possible for both.

The various personal pronouns will now be discussed and exemplified, beginning with first and second person pronouns and followed by the third person pronouns.

### 6.2.2 ita: inclusive and general ‘we’

The pronoun *ita* (or *it*) has several uses. One is as the first person inclusive plural ‘we’.

6.1  

<table>
<thead>
<tr>
<th>Ita</th>
<th>na’in</th>
<th>hitu ne’e,</th>
<th>it</th>
<th>ha-la’e</th>
<th>hotu</th>
<th>ó</th>
</tr>
</thead>
<tbody>
<tr>
<td>1PI</td>
<td>CL:human</td>
<td>seven this</td>
<td>1PI</td>
<td>make-husband</td>
<td>all 2S</td>
<td></td>
</tr>
</tbody>
</table>

*la‘en ne’e.*

husband this

We seven (i.e. we six speakers plus you), we will all marry your husband.

(O7.32)

Since inclusiveness is a culturally valued characteristic, *ita* may be used when strictly speaking the addressee is not included. This is illustrated by the following example (6.2), where women were by their choice of pronoun including me in the description of their way of life, even though the description clearly did not apply to me. When questioned on such uses, speakers recognised that this extension of the strict meaning of *ita* was a politeness phenomenon.

6.2  

<table>
<thead>
<tr>
<th>Ita</th>
<th>feto</th>
<th>la‘o</th>
<th>ha‘i.</th>
<th>Ita</th>
<th>iha</th>
<th>uma</th>
<th>karian.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1PI</td>
<td>woman</td>
<td>walk</td>
<td>not</td>
<td>1PI</td>
<td>LOC</td>
<td>house</td>
<td>work</td>
</tr>
</tbody>
</table>

We women don’t go out. We are at the house working.

(J2.36)

Related to this, *ita* is commonly used (rather like formal English ‘one’) in explanations of how things are normally done, or used to be done, or when describing hypothetical situations. In many such cases the addressee, the speaker or both are clearly not intended to be included in the class of people discussed. Thus in the following example from an explanation to me as to why men are fined for premarital sex if they subsequently refuse to marry the girl, the elderly male speaker used inclusive *ita* when speaking of such men. It is not always clear whether a gloss of ‘we’, ‘you’ or ‘one’ is more appropriate in such cases.
6.3 *Ita* *loke ta feto ne’e ti’an tó.*

1PI open already woman this already TAG[Mly]

‘We/you/one’ (the hypothetical man) had already taken this girl sexually (lit. ‘opened this girl’) after all. (I 0.130)

### 6.2.3 *Ita, ó ‘you’*

The pronoun *ita* is also used as a polite pronoun for second person singular ‘you’ (6.4). According to consultants, *ita* should be used for older people, including parents, and for other people in high positions such as nobles and outsiders. In practice *ita* is used rather little as a second person pronoun, with ó (or a) ‘2S’ being common for addressing older kin (6.5). Nevertheless, *ita* does appear to be used fairly consistently in politely addressing foreigners. An alternative to personal pronouns is the use of titles (e.g. *ama* ‘father’) as terms of address; these are discussed in §7.2.5.1.

### 6.4 *Ama Liu, ha’u k-mai k-amán ita …*

father Liu 1S 1S-come 1S-accompany(.noble) 2S.HON

Father Liu, I have come to accompany you… (Said by a noble girl to a young nobleman.) (O1.38)

### 6.5 *Ina, loron ha’u k-usu ó, ó m-eli.*

mother day 1S 1S-request 2S 2S 2S-conceal

Mother, daily I asked you, (and) you concealed (the truth). (G1.60)

A term of even higher respect than *ita* is *ita bót* (lit. ‘big you’), which is restricted to addressing God or the spirits in formal prayers (6.6), or to addressing the top nobles in formal contexts such as meetings.

### 6.6 *Ami hodi husu terima.kasi bá ita bót iha leten…*

1PE bring request thanks[Mly] to 2S.HON big LOC top

bá ás bá... go high go

We bring our thanks to Thee in the heights up above... (C1.9)

### 6.2.4 Other first and second person pronouns

The remaining first and second person pronouns are illustrated in the example below.

### 6.7 *Belu. Ó m-ó na’i feto sia, emi toba iha uma.*

friend 2S 2S-accompany noble woman PL 2P sleep LOC

house

Friend. You (singular) and the noblewomen, you (all) sleep at the house.

*Ha’u k-ó klosan sia, amí hodi kuda bá.*

1S 1S-accompany male.servant PL 1PE bring horse go

I and the manservants, we (exclusive) will take the horses out (and stay with them). (K9.64)

Corresponding to the highly respectful second person *ita bót* is a self-humbling expression consisting of a first person pronoun (*ha’u‘1S’ or *ami‘1PE’) followed by *ata* ‘servant, slave,
low rank of commoner’ (6.8). According to consultants these expressions are supposed to be used by commoners when speaking with nobles. However, only some commoners who regularly interact with nobles are truly conversant with this form of reference to self.

6.8 Ami ata liu lai.
1PE slave/commoner further first
We lowly commoners will go on now. (Said before departing from a noble on the road.)

6.2.5 sia ‘they’

The third person plural pronoun sia (with variant sira and short forms si and sa) ‘3P’ is used primarily to refer to persons (6.9, 6.10). For at least some speakers it is acceptable for animals as well, with the corpus including examples referring to dogs, chickens, cocks and rats (6.11). One consultant allowed that sia could be used even for ants, while others considered it unacceptable for horses and chickens. It cannot refer to plants or inanimate entities. The form sia (or sira) is also used as a plural marker within NPs, with no such animacy restrictions (§6.5.3).

6.9 Sia kawen ti’a, sira té r-ola oan
3P marry[Mly] already 3P defecate 3P-take child
ná’in rua.
CLShuman two
After they had married, they gave birth to two children. (U4.72)

6.10 Nia.má si la’o.
then 3P walk
Then they left. (O0.100)

6.11 Lale ita rai lerek sia r-á.
else 1PI lay.down forsake 3P 3P-eat
Else (if) we just lay down (food) anyhow, they (rats) eat (it). (Q0.90)

6.2.6 nia ‘be, sbe, it, that’

The third person singular pronoun nia (or ni, na) ‘3S’, in contrast, is used to refer not only to humans (6.12) but also at least to animals, plants and objects (6.13).

6.12 Nia bá n-usu ni ina-n.
3S go 3S-request 3S mother-GEN
She went and asked her mother. (G1.59)

6.13 Nia kbit liu besi. Nia-kan kbit liu besi.
3S strong further iron 3S-POS strong further iron
It (aksisi wood) is stronger than iron. Its strength surpasses that of iron. (L0.109)
This singular pronoun (and singular subject marking) can be used to refer to more than one referent where the referents are acting together. Examples include references to six brothers who act together (6.14), to a mother pig with her piglets, and to parents.7

6.14 Dadi, mané nén né nia n-alo lós to'os...
so man six this 3S 3S-make only garden
So these six men, they (lit. 'he') just worked the gardens... (G1.5)

This word is also a demonstrative pronoun, in which capacity it can additionally refer to times, places (6.15), and (in the expression nú nia ‘like that’) to events (6.16). As a demonstrative, nia ‘that’ contrasts with ne’e ‘this’. Its distal interpretation is reinforced by its compatibility with the distal deictic particle bá (iba nia bá ‘LOC 3S go’ = ‘there’) but not with proximal mai (*iba nia mai).

6.15 ...bá iha rai ulun, bá to'o nia, túr bá nia.
go LOC earth head go reach 3S sit at 3S
[So they went from here] to the ‘head’ of the earth, (when they) reached it, (they) sat at that (place).

6.16 Iha leo ida nú nia: fukun mama, kose hotu.
LOC hamlet one like 3S elder chew(usu.betel) rub all
In a hamlet it is like this: the elder chews betel, and rubs (it) on all (the children).

The borderline between personal pronoun and demonstrative nia is fuzzy, for which reason both are given the same gloss ‘3S’. Nevertheless, when the pronoun is in the phonologically reduced form (ni, na), or is immediately followed by possessive -k(an), it is clearly a personal pronoun, since these are possibilities not shared by any other pronoun type.

Discussion of nia as determiner is deferred until §6.3.3, to facilitate comparison with the other determiners.

6.3 Demonstratives

6.3.1 Introduction

There are two demonstratives which clearly belong to the Fehan dialect, namely ne’e ‘this’ and nia ‘that’.8 In addition, there is ne’et ‘that’, which appears to be marginal to the Fehan dialect, as well as the complex demonstratives nemai ‘here, this’ and nebá ‘there, that’, which are only marginally codified. All five demonstratives are classed as both pronouns and determiners.

6.3.2 ne’e ‘this’

The demonstrative ne’e ‘this’ shows considerable phonological variation. As determiner, it is often pronounced (and by some consultants written) e’e after a consonant-final word (e.g. batar e’e ‘maize this’ = ‘this maize’, manu ha’u-k e’e ‘bird 1S-POS this’ = ‘this my bird’). In

---

7 Such use of nia (or possessive nia-kan) for a group of referents is not particularly unusual (6 examples). It is, however, possible that it is only marginal, since one tertiary educated consultant considered it incorrect.
8 Both ne’e and nia are also connective adverbs with meanings approximating ‘so, then’.
both pronouns and determiners the glottal stop is readily omitted in speech (but never in writing). These two factors give rise to the phonological variants ne’e, e’e, né and é.

As pronoun, ne’e ‘this’ is only very rarely used to refer to persons, a task for which the personal pronoun nia ‘3S’ is available. NPs in which ne’e is a determiner, however, often have personal reference (sia ne’e ‘3P this’ = ‘they here’, final clause in 6.20).

Unlike other determiners, ne’e nominalises subordinate clauses as well as modifying NPs. Its use in clauses is discussed in §14.6, and illustrated in the second clause in 6.20 (where it cannot modify the preceding noun kabau ‘buffalo’ because that NP is non-specific).

Functionally, ne’e ‘this’, whether as pronoun or determiner, marks either proximal deixis or anaphoric reference (i.e. ‘given’ NPs in the sense of Chafe (1976)). When used deictically, it indicates proximity to the speaker (6.17), with the prepositional phrase iha ne’e (lit. ‘LOC this’) being a common expression meaning ‘here’ (6.18). In time expressions, ne’e indicates the time of speaking (6.19).

6.17 Ne’e sē nia-kan uma?
this who 3S-POS house
Whose house is this? (Said pointing.)  

6.18 Ibu lora-loron mai iha ne’e.
mother RDP-day come LOC this
You (‘Ibu’) come here every day.  

6.19 Iha nemai, rai áoras e’e nú nabé? Udan ti’an
LOC here earth DEF time this like which rain already
kā sei?
or not yet
What is the weather (lit. ‘earth’) like here now? Have the rains started, or not yet?  

In addition to deictic use, ne’e may be used anaphorically. In this case the referent of the NP has usually been mentioned very recently (somewhere in the previous three clauses). Anaphoric reference is found not only to persons (final clause in 6.20) and physical entities, but also to actions (second clause in 6.20) and propositions (6.21). In the examples, antecedents are indicated by double underlining, while ne’e in the anaphor has single underlining.

6.20 Sia loron r-ein kabau. Sia r-ein kabau ne’e.
3P day 3P-wait buffalo 3P 3P-wait buffalo this
Every day they looked after (lit. ‘waited for’) buffalo. While they did this looking after buffalo,
sia r-ö ferik oan ido.
3P 3P-accompany mature.woman small one
they were with an old woman,
mais ferik oan ne’e buan.
but mature.woman small this witch
but this old woman was a witch.  

(L0.36)  
(L0.57)  
(U6.1)
6.21 \textit{Molin n-aksurit mós ti'a, ne'e n-ák nia} defe cate$^9$ 3S-diarrhoea finished already this 3S-say 3S
\textit{molin ti'a moras wén ti'an.} defe cate already sick juice already
When the diarrhoea is finished, this means he (the patient) has excreted the sickness’ ‘juice’ (and will be well). (K2.12)

Anaphoric \textit{ne‘e} is also common with proper nouns (e.g. \textit{Suri Tuan ne‘e} ‘this Suri Senior’).
As in colloquial English, some speakers use the determiner \textit{ne‘e} when first introducing a participant into a story. The lack of syntactic marking of the participant as new is matched by the lack of any special grammatical structure (such as presentational clause structure). This is illustrated by the following first line of a traditional story about six brothers, as well as in 6.56. Other consultants consider this use of \textit{ne‘e} for new participants to be incorrect, substituting indefinite \textit{ida} ‘one, a’ for definite \textit{ne‘e}.

6.22 \textit{Ferik nég nó beí nég té n-ola oan} mature woman this and grandfather this defe cate 3-take child
\textit{mane nén.} man six
This woman and this man gave birth to six boys. (G1.2)

6.3.3 \textit{nia ‘that’}

As a third person singular pronoun and as a distal demonstrative pronoun, \textit{nia} ‘that’ has already been discussed in §6.2.6. This section therefore deals only with \textit{nia} as determiner.

As determiner, \textit{nia} is far less common than \textit{ne‘e} ‘this’, with only some 60 examples of \textit{nia} compared to over 1,500 for \textit{ne‘e}. It seems to be used disproportionately much by speakers who have been influenced by Indonesian, which leads me to suspect that in addition to its native Tetun use as determiner, the determiner \textit{nia} is also being affected by interference from Indonesian \textit{itu} ‘that’.

One possible area of interference is in the deictic use of \textit{nia}, to mean ‘that’ (as opposed to proximal \textit{ne‘e} ‘this’; e.g. \textit{bá oras nia} ‘at that time’). The few deictic examples in the corpus, as well as deictic explanation during elicitation, all come from a speaker who was heavily influenced by Indonesian. Unlike \textit{ne‘e} ‘this’, \textit{nia} appears not to be used for discourse deixis (i.e. for referring to previously mentioned events or propositions).

Instead, the determiner \textit{nia} is usually used to refer anaphorically to previously mentioned entities. Based on evidence from texts, it is almost always used for tracking minor participants, props or setting, in which the referent has been introduced (or at least mentioned) somewhere in the previous three clauses. In many cases there is no subsequent mention after the \textit{nia}-marked NP. In the examples below, NPs with double underlining are the antecedents for the anaphors in which \textit{nia} is underlined.

---

$^9$ The basic meaning of \textit{molín} is as a location noun meaning ‘outside’. Its use as a verb meaning ‘defecate’ presumably comes from the polite expression for this, \textit{bá molín} (lit. ‘go outside’).
Chapter 6

6.23  Bá, n-aré lós (ə) samodo né kakun dei. Na'in go 3S-see just HES green.tree.snake this shell only owner
la haré ti'an.
not see already
(She) went, and saw only the skin of this green tree snake. The owner (i.e. the
snake) was no longer to be seen.

Dadi n-ola ti samodo nia kakun á, so 3S-take already green.tree.snake that shell DEF
n-odi bá tau bá kluni nú ne’e.
3S-bring go put go pillow like this
So having taken (i.e. picked up) the green tree snake's skin, (she) brought (it)
(and) put (it) in a pillow like this.

Tau bá kluni nia, lita n-ato'os ti'a, nia toba.
put go pillow that sew 3S-tight already 3S lie.down
(Sh) put it in the pillow, (and) having sewed (it) up tightly, she lay down.

(R5.4)

6.24  Ami iha na'in feto uma-n á, mak ami há kbú
1PE LOC noble woman house-GEN DEF REL 1PE eat riceparcel
né na'an iha nia, ami há hemu iha na'in feto
and meat LOC 3S 1PE eat drink LOC noble woman
uma-n nia.
house-GEN that
We were at the noblewoman's house, where we ate rice parcels and meat; we
ate and drank at that noblewoman's house. (Z1.27)

While nia is usually used where there is a singular referent, it can also be used for plural
referents. The limited data (4 examples) are consistent with this plural use being restricted to
referents who form a group, as is the case for the pronoun nia.

6.25  Nia ti'a, Uduk Liurai Balak Liurai té bá oan feto:
3S already Uduk ruler Balak ruler defecate go child girl
Then Uduk Liurai Balak Liurai gave birth to girls
(em ha'ak) Luruk Muti Ki'ik, Luruk Muti Kawa'ik né.
2P say Luruk Whitey small Luruk Whitey older this
(which you say are) called Luruk Whitey Junior and Luruk Whitey Senior.

Feto na'in rua nia, ida fó bá iha Wawa'í, ida
woman CLShuman two that one give go LOC Wawa'í one
fó bá iha Manulea.
give go LOC Manulea.

Those two women, one (they) sent (lit. 'gave') to Wawa'í, one (they) sent to
Manulea. (F6.10)
6.3.4 ne\'et 'that'

The demonstrative ne\'et 'that' is used deictically for entities which are an intermediate distance away. It cannot, for instance, be used of objects that one is holding, nor of things which are very distant.\(^{10}\) This word is rarely used, with the six unelicited examples in the corpus all being from speakers who were born outside the Fehan dialect area, and consultants differing as to whether they accepted it as a word. As such I suspect that it is not part of the native Fehan vocabulary.\(^{11}\)

6.26 Ne\'et át ti\'an, é.
That is already bad, eh! (X0.145)

6.27 M-alai to'o rai kasa\'ek ne\'et, m-ein ha\'u bá nia.
Run to that slight rise, and wait for me at that (place). (X0.35)

6.3.5 nemai 'here, this', neba 'there, that'

The complex demonstratives nemai 'here, this' (22 examples) and neba 'there, that' (60 examples) clearly consist of the demonstrative ne'e 'this' followed by the deictic particles, proximal mai and distal bá respectively (§11.8). These morpheme sequences can constitute either two phonological words (ne'e mai, ne'e bá) or a single word. In the latter case the initial morpheme is shortened to an unstressed ne, and as a result the initial vowel is sometimes replaced by the default antepenultimate vowel /a/ (namai, nabá).\(^{12}\)

The incorporation of the deictic appears to stress the closeness or distance of the referent rather than necessarily adding finer distinctions of distance to the basic bipartite deictic system of Tetun (ne'e 'this', nia 'that').

Grammatically there is limited evidence to suggest that these morpheme sequences are marginally codified as single units (as assumed by Monteiro (1985:xxix) when he lists them as single words). This evidence is that they have slightly wider distribution than is allowed for other sequences of NP plus deictic particle. In particular, other NPs followed by deictic particles are restricted to occurring as objects of verbs, prepositional verbs or locative prepositions. Although NPs containing the demonstratives nemai 'here' and neba 'there' usually function in these positions (6.19), the corpus contains a few examples for other positions also, such as subject of a clause (6.28; 2 examples) or object of the preposition nû 'like' (nû nemai 'like this'; 2 examples).

---

\(^{10}\) As illustrated by example 6.27, ne\'et is not orientated towards the addressee, as is the case for the second term in some three-term demonstrative systems (Anderson & Keenan 1985:282).

\(^{11}\) It is listed as 'there, over there' in Morris's (1984b) dictionary of East Timorese Tetun, but is not present in Mathijsen's (1906) listing for the Foho dialect of northern Belu.

\(^{12}\) The distal demonstrative can alternatively be constructed from nia bá 'that go'. I am not aware of any difference in meaning between ne'e bá (lit. 'this go'), nia bá (lit. 'that go') and neba 'there'.
6.28  Oa ida ne’e ulun fatuk bei ladún, mais ida nebá  
child one this head rock however not very but one there  
ulun fatuk basuk.  
head rock very  
This one child is stubborn but not very; but the one there is very stubborn.  
(T0.107)

The demonstratives nemai ‘here’ and nebá ‘there’ are, like other demonstratives, classed as both pronouns (6.19) and determiners (6.28). Although locatives are normally expressed as prepositional phrases (e.g. iha nebá ‘LOC there’ = ‘there’), two examples in the corpus suggest that these demonstratives can also function adverbially as locatives without an introductory preposition (6.29).

6.29  ...Bei Kakae nó kahúk isin (oi!) toba nebá.  
Mr cockatoo and blowgun contents EXCL lie down there  
...Mr Cockatoo and the blowpipe ammunition (wow!) (came to) lie down way over there.  
(K8.67)

6.4 Quantifying pronouns/determiners

Tetun appears to have only three quantifying pronouns, namely hotu-hotu ‘all’ hotu ‘all’, and ida-‘idak ‘each’.13

The term hotu-hotu (sometimes hotu-hotuk) ‘all, everyone, everything’ can head NPs as a pronoun (6.30, 6.31; 21 examples), or modify them as a determiner (6.32, 6.33; 52 examples).

6.30  Hatene, é? Hatene ha’i bá sá? Hotu-hotu hatene!  
know TAG know not for what RDP-all know  
(They) know, eh! Why wouldn’t (they) know? Everyone knows!  
(P4.39)

6.31  Iha leo ida fukun né mama, kose hotu-hotu.  
LOC hamlet one elder this chew(usu. betel) rub RDP-all  
In a hamlet the elder chews betel nut, and rubs it on all (the children).  
(B2.22)

6.32  Ema hotu-hotu r-atais tais, tau faru...  
person RDP-all 3P-wear.cloth cloth put clothes  
All the people wore (traditional handwoven) cloths, and wore clothes...  
(Z5.87)

The unreduplicated form hotu ‘all’ has a unique distribution, in that it is a floating quantifier. It usually immediately follows the head of the predicate (6.34); in core layer serial verb constructions it follows the first verb (6.33).

---

13 Other quantitative concepts are expressed by alternative word classes. For instance, ‘many, much’ is expressible by the adjective wa’in (na’an wa’in basuk ‘fish many very’ = ‘very many fish’), ‘every’ by the premodifying adjective sura (§7.4.3), ‘some’ by the noun bulú ‘part, side’, ‘several’ by the numeral hura ‘several, how many’ (§5.2.1), and ‘no’ by combining negation with an indefinite NP (Ema ida la n-adau ha’u ‘person one not 3S-rob IS’ = ‘Nobody robbed me’; §11.2.2). The word ruma, which Morris (1984b) lists as meaning ‘some, any, several, few’, is in the Fehan dialect an adjective meaning ‘others’ (ema ruma sia ‘person others PL’ = ‘(the) other people’).
With intransitive predicates *hotu* quantifies the subject (64 examples). The example below illustrates the difference in distribution between the floating quantifier *hotu* (which follows the first verb) and the determiner *hotu-hotu* (which immediately follows the head of the subject NP), where both quantify the subject.

6.33  
*Ami bá hotu hakoi maten. = Ami hotu-hotu bá hakoi maten.*

1PE go all bury corpse 1PE RDP-all go bury corpse

We all go and bury the dead person.  

For transitive predicates *hotu* usually quantifies the object, regardless of whether the object is explicit (6.34, 6.47; 15 examples) or not (6.35; 23 examples).

6.34  
*Nia ti’a, n-usu hotu sia na’in nén ne’e, r-o’i.*

3S already 3S-request all 3P CLS:human six this

3P-do.not.want

Then, (he) asked all six of them, (and they) refused.  

6.35  
*Nia-kan la’en kna’ok basuk. Nia n-ana’o n-ola hotu.*

3S-POS husband thieving very 3S 3S-steal 3S-take all
Her husband was a terrible thief. He stole everything.  

However, it can quantify the subject, particularly when the verb does not have high transitivity (7 examples).

6.36  
*Sira bá hotu to’os.*

3P go all garden

They all went to the garden.  

The quantifying pronoun *ida-’idak* (or *i-’idak*) ‘each’ is derived from the numeral *ida* ‘one’, and takes singular subject agreement (14 examples). Data are insufficient to ascertain whether *ida-’idak* is also a determiner.

6.37  
*Ida-’idak mai n-odi nia-kan batar futin hitu.*

each come 3S-bring 3S-POS maize head(grain) seven

Each comes bringing his (own) seven maize cobs.  

6.38  
...*sia r-anorin hananu iha ida-’idak nia-kan fatik...*  

3P 3P-learn sing LOC each 3S-POS place

[Choir singers who lived far away,] they learnt the singing each in their own places [, and only came together on the day that the full choir sang.]  

6.5 Other determiners

6.5.1 á: definite article

The determiner á has the characteristics of a definite article, but has limited geographical distribution, much inter-speaker variation, and phonological restrictions, which all contribute
to making it less common in the corpus than one might expect of a definite article (530 examples).14

Its highly limited geographical distribution is reflected by its omission from the East Timorese dictionaries of Morris (1984b) and Hull (1996b), and from Mathijsen’s (1906) and Monteiro’s (1985) dictionaries of the Foho dialect of northern Belu. Literate native speakers of the Fehan dialect are aware that á is not used in these written dialects of Tetun. This fact may contribute to the mixed reactions of speakers when questioned about á. Some use it even when speaking very carefully in ‘dictation’ style, according to others it is acceptable to say but not to write it, while one went so far as to omit every instance of it when reading aloud from a transcription. The inter-speaker difference is illustrated in the texts in Appendix A, with Text 1 having ten instances of á, while Text 2, which is much longer (and does not come from the extreme anti-á speaker mentioned above), has none.

This article is unique phonologically in that it is always the final element in a phonological phrase, and tends to be stressed and uttered slowly.15 This phonological restriction affects its syntactic distribution, meaning, for instance, that á cannot mark a preposed possessor NP (which would normally fall under the same phonological phrase as the following possessed NP).

Unlike demonstratives, the definite article á is neither deictic nor necessarily anaphoric. Definiteness can come about from a variety of sources. Textually, about half of NPs with á refer to previously mentioned entities, as in the following examples (for which the antecedent, where included, is indicated by double underlining).

6.39 La’o la’o la’o la’o daudaun to’o alas laran.
walk walk walk walk continue reach forest interior
(T hey) walked and walked and walked, and kept walking until (they were in) a forest.

6.40 Tama alas laran á má manu kokorék.
enter forest interior DEF and then bird crow
(T hey) entered the middle of the forest, and then a cock crowed. (D1.24)

6.41 ...feto á n-aré ha’i mane á, mane á n-aré ha’i
woman DEF 3S-see not man DEF man DEF 3S-see not
feto á.
woman DEF
[The young woman didn’t get pregnant during courtship, as] the woman didn’t see the man, (and) the man didn’t see the woman. (O5.68)

Most of the other half refer to entities whose existence can be inferred from the existence of some other entity which has already been mentioned (i.e. Prince’s (1981b) ‘inferable’ class, or the ‘associative’ anaphora of Hawkins (1978:99)). This inference may be through such inalienable relationships as part-whole (Bei Bibi ne’an á ‘Mr Goat’s teeth DEF’, also 6.41), kin (aman á ‘father DEF’ = ‘the father’), dimensions (To’os ne’e, nia-kan naruk á

14 The fact that Himmelmann (1997:197) found no evidence of definite articles in his survey of Austronesian languages suggests that this article is unusual within the language family, in addition to being idiosyncratic within Tetun.

15 It is thus not a clitic. Several younger transcription assistants repeated the final consonant of the preceding word as the onset to á (e.g. ‘oan na’, ‘rais sa’), just as one did for ne’ele’e ‘this’ (e.g. ‘Lasak ke’e’). This spelling was not approved by educated older consultants.
Pronouns and determiners

'garden this 3S-POS long DEF' = 'this garden, it’s size'), or name (feto á naran á ‘woman DEF name DEF’ = ‘the woman, (her) name'). Alternatively, the NP may refer to entities which are in some other way conventionally associated with the topic under discussion. Thus, for instance, in an explanation about how to make coconut oil (Text 1 in Appendix A), botir ‘bottle’, wén ‘juice, oil’ and karon ‘cloth filter’ were all marked by á on their first mention.

6.41 ...sia n-akoi n-ela Suri Tuan bá ohin foho ne’e
3P 3-bury 3-leave Suri old at just.now mountain this
dikin á...
tip DEF
[Then] they buried and left Suri Senior at the summit of this aforementioned mountain...

(K10.137)

Finally, á can mark NPs with absolutely unique referents, such as the sun, moon and earth, regardless of whether they have been previously mentioned (6.19, 6.42). This is in contrast to ne’e ‘this’, which can mark such NPs only if the referent is either nearby (hence deictically) or previously mentioned (hence anaphorically).

6.42 Niamá monu bá rai á.
then fall go earth DEF
Then (it) falls to the ground.

(P0.93)

The definite article does not often co-occur with other means of specifying definiteness. It can readily co-occur with possessive marking (6.24, 6.55), although most possessive NPs do not have a definite article. Unlike ne’e ‘this’, it is rarely used for proper nouns. It cannot co-occur with other determiners.

The vast majority of NPs marked by á refer to non-human props. In fact only 20% refer to persons. Often, although the referent can be assumed to exist as a unique individual(s) the exact identity of the referent is communicatively unimportant.16 This is so, for instance, for the forest in 6.39, and for the bottles and cloth filters in the explanation of how to make coconut oil (mentioned above).

An NP marked by á can have either singular or plural referents. It is, however, rare to mark such an NP with any indication of plurality. Modification of the NP by numerals occurs in one textual example (6.43) and was accepted in elicitation; however, the need for the determiner to occur NP-finally (mentioned above) would inhibit plural marking with sia, which normally follows the determiner ne’e ‘this’ within the NP.

6.43 Feto na’in nén á. liras sira-k si iha.
woman CL:human six DEF wing 3P-POS PL be.present
The six women (sisters who flew down from heaven), their wings were there.

(Z6.14)

6.5.2 ida ‘one’ as indefinite article

The indefinite article ida indicates singular indefinite specific reference (6.44). It differs slightly in distribution from the definite determiner ne’e ‘this’, in that the indefinite article always precedes relative clauses (6.45) while definite ne’e ‘this’ usually follows relative clauses (except non-restrictive or ‘afterthought’ clauses).

16 This is what Givón (1990:423ff.) calls ‘semantically referential’ but ‘pragmatically non-referential’.
A woman and a man gave birth to seven children.

There is an earth person who has come up above here, Father.

This article is homophonous with the numeral 'one'. In many instances it is not possible to distinguish between an indefinite and a numeric interpretation of *ida*, for which reason both are glossed as 'one'. Semantically the difference is that the numeral *ida* is neutral with respect to definiteness, while the article indicates indefinite, though specific, reference. Syntactically, there are several differences. Firstly, the numeral but not the article can be preceded by a numeral classifier, although in practice it is not common to use sortal classifiers for singular NPs (§5.3.2). Secondly, the numeral precedes determiners and adjectives (e.g. *mane ida at e'e* 'man one bad this' = 'this bad man'), while the article follows adjectives (e.g. *buat di'ak ida* 'thing good one' = 'a good thing'), and does not co-occur with other determiners. Finally, wider context can force one interpretation rather than the other. For instance, where existence is being predicated of an NP that NP must have an indefinite interpretation (6.45). In contrast a list of items required for a fine (e.g. 'three cloths, a/one pig') encourages a numeric interpretation.

### 6.5.3 *sia*: plural marker

The plural marker *sia* (with variant *sira*, occasionally *si* in rapid speech; over 600 examples) has the same phonological form as the third person plural pronoun. As pronoun, *sia* can refer only to humans and, at least for some speakers, animals (§6.2.5). As plural marker, however, *sia* can be used of any referent, regardless of animacy. In the corpus, approximately 80% of examples are for persons (6.46), 3% for animals and 17% for inanimate referents (6.47). The plural marker is sometimes used of things which in English are uncountable, such as water (even if it is all in the one container; 6.48), or skin (e.g. if one blackens all of one's skin). As plural marker, *sia* is not restricted to the third person (6.49).

---

17 Grammaticisation of 'one' to an indefinite marker is common crosslinguistically, as pointed out by Givón (1981), amongst others.

18 Plural marker *sia* thus differs from its variant *sira* in East Timor, as it is reported by Hull (1993:31). He says that it is used only for definite nouns (e.g. *livru sira* 'the books'), and that in rural Tetun and colloquial Dili Tetun it is 'mostly omitted when the plural refers to animals and things'. That is, the plural marker there tends to have the same animacy and definiteness restrictions as the pronoun.
6.46 *Tán ha’u-kan ina-n sia ama-n sia katuas*
because 1S-POS mother-GEN PL father-GEN PL mature.man

*ti’an…*
already
Because my mother/aunts and father/uncles are already old… (Note: *ina*
includes mother and mother’s sisters, while *ama* includes father and father’s
brothers, amongst others.)

(F3.23)

6.47 …*subar hotu kakehe sia iha biti kidun á.*
hide all fan PL LOC mat bottom DEF
[Tonight…] hide all the fans under the (sleeping) mat.

(U1.22)

6.48 …*n-arís n-aló we ne’e sia mós ti’a...*
3S-bathe 3S-make water this PL finished already
[Kitten...] having bathed using up all the water [. followed them to the dance.]

(Z2.85)

6.49 *Nú ami ne sia. at sekola la ho’i.*
like 1PE this PL IRR school not do.not.want
Like us, (our parents) didn’t want us to go to school.

(O5.142)

As is discussed in §7.2.2.2, *sia* is rarely used with any other indication of plurality. The
only determiners with which it can co-occur are *ne’e ‘this’ and *hotu-hotu ‘all’ (§7.2.2.5).
The head of an NP marked by *sia* may describe all of the referents of the NP (e.g. all
referents of *asu sia ‘dog PL’ are dogs), or only one of the referents (6.50). In the latter case,
reference is to the specified individual plus associated ones. The individual specified need not
be the most senior member of the group, but only a salient one.

6.50 …*Feto Ikun sia r-e’uk...*
woman tail PL 3P-dance
…Youngest Sister and associated people (her sisters) danced...

(U4.46)

Clauses in which the subject NP contains the plural marker *sia* can consist of the same
sequence of words as a clause in which *sia* is used as a summary pronoun for a left-
dislocated topic NP (6.51; §14.2.2). The two constructions differ in that the plural marker is
tightly bound phonologically to the noun phrase of which it is a part, while the summary
pronoun is normally preceded by a pause and is not syntactically part of the preceding NP.

6.51 *Feto nén, sia r-oran laran moras to’o.*
woman six 3P 3P-sense interior sick very
The six sisters (lit. ‘women’), they felt very sick at heart.

(E0.142)

6.5.4 *Comparison of pronouns and determiners*
The following table presents generalisations on some of the more common pronouns and
determiners. ‘Y’ (‘yes’) answers are highlighted. Blank entries indicate lack of information,
while question marks indicate apparent inter-speaker differences.
Table 6.3: Comparison of pronouns and determiners: *ne’e*, *nia*, *sia* and *á*

<table>
<thead>
<tr>
<th></th>
<th>Pro: <em>ne’e</em> ‘this’</th>
<th>Det: <em>ne’e</em> ‘this’</th>
<th>Pro: <em>nia</em> ‘3S, that’</th>
<th>Det: <em>nia</em> ‘that’</th>
<th>Pro: <em>sia</em> ‘3P’</th>
<th>Det: <em>sia</em> ‘PL’</th>
<th>Det: <em>á</em> ‘DEF’</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Semantics:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Person</td>
<td>Rare Y Y Y Y Y</td>
<td>Y Y Y Y Y Y</td>
<td>Y Y Y Y Y Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inanimate</td>
<td>Y Y Y Y Y Y</td>
<td>Y Y Y Y Y Y</td>
<td>Y Y Y Y Y Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>Y Y Y Y Y Y</td>
<td>Y Y Y Y Y Y</td>
<td>Y Y Y Y Y Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Y Y Y Y Y Y</td>
<td>Y Y Y Y Y Y</td>
<td>Y Y Y Y Y Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Function:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Situational deixis</td>
<td>Y Y Y N? Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(‘pointing’ to an entity)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discourse deixis</td>
<td>Y Y Y Y19 Y</td>
<td>N N N N N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(to proposition or event)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tracking participants</td>
<td>Y Y Y Y Y Y</td>
<td>N Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(anaphora)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unique entity (e.g. ‘the sun’)</td>
<td>N N N N N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Associative-anaphoric</td>
<td>N N N N N Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(e.g. ‘tree’...’the trunk’)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Singular/plural:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-occurs with <em>sia</em> ‘PL’</td>
<td>Y Y S/P S/G S/G</td>
<td>N N N N N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Singular/Group/Plural</td>
<td>Y Y S/P S/G S/G P</td>
<td>N N N N P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In terms of semantics, the table shows that restrictions which apply to the third person pronouns (the rarity of reference to persons for *ne’e* ‘this’ and the restriction to persons for *sia* ‘3P’) do not apply to the corresponding determiners.

With regard to function, *ne’e* ‘this’ has a wider range of uses than does *nia* ‘that, 3S’, in that it can more readily be used for situational deixis (e.g. ‘this house here’), and can in addition be used for discourse deixis (referring to previously mentioned propositions or events). The plural marker does not have the anaphoric or deictic functions of the corresponding plural pronoun.

With regard to singular or plural reference, the table shows that *ne’e* ‘this’ can be used for either singular or plural reference (the latter often being indicated explicitly by the plural marker *sia*; e.g. *feto ne’e sia* ‘woman this PL’ = ‘these women’). Its deictic opposite *nia* ‘that’, however, is restricted to singular (or group, see §6.2.6) reference, a restriction presumably imposed by it also being the third person singular personal pronoun. The definite article *á* can be used for either singular or plural reference, but cannot be marked as plural by *sia*.

---

19 Discourse deixis in the corpus is found only in *nú nia* ‘like that’ (e.g. *karian nú nia* ‘work like that’ = ‘work as I have already told you about’).
6.6 Epistememes (interrogative/indefinite)

6.6.1 Overview of epistememes

Tetun has several words which are used for unknown quantities, entities or times, in either interrogative or indefinite contexts. For want of a more standard term to include both interrogatives and indefinites, I shall use the term ‘epistememe’, suggested by Mushin (1995). The basic Tetun epistememes are listed below.

<table>
<thead>
<tr>
<th>Pronoun</th>
<th>Type</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>hira</td>
<td>Numeral</td>
<td>how much, how many, several, however many</td>
</tr>
<tr>
<td>hori.hirak</td>
<td>Time noun</td>
<td>when (past), in recent time</td>
</tr>
<tr>
<td>wain.hira</td>
<td>Time noun</td>
<td>when (future)</td>
</tr>
<tr>
<td>sê</td>
<td>Pro, Det</td>
<td>who, which person, someone, anyone, whoever</td>
</tr>
<tr>
<td>nabé</td>
<td>Pro, Det</td>
<td>which, where, somewhere, anywhere, whichever, wherever</td>
</tr>
<tr>
<td>sá</td>
<td>Pro, Det</td>
<td>what, something, anything, whatever</td>
</tr>
</tbody>
</table>

Clearly the class of epistememes crosses word class boundaries. Hira ‘how much’ has the distribution of a numeral, and as such is discussed in §5.2.1. The two temporal epistememes, hori.hirak for past time and wain.hira (or wai.hira, wa.hira) for future time, are illustrated below.²⁰

6.52 Kabau ó-k ne’e lakon **hori.hirak**?
   buffalo 2S-POS this disappear when(.past)
   When did your buffalo disappear? (H1.3)

6.53 ...tân **hori.hirak** á nia la karian. [The boss hit him,] because recently he didn’t work. (V0.166)
   because recent.time DEF 3S not work

6.54 **Wai.hira** ó to’o bâ?
   when(.future) 2S arrive go
   When will you arrive there? (AA4.131)

The remaining three epistememes have dual classification as pronouns and determiners. They are unique amongst pronouns and determiners in that each can form a complex epistememe incorporating a following *ida* ‘one’ (e.g. sê *ida*) or a repetition of itself (e.g. sê sê). There appears to be no difference between epistemic pronouns and determiners except for their distribution within the NP.

In the following sections the interrogative use of the basic epistemic pronouns and determiners will be considered first, and then the indefinite use. This is followed by discussion of the complex epistememes, and finally by a discussion of the distributional restrictions on epistemic determiners. In most cases examples will be given of the epistememe as both pronoun and determiner, without necessarily commenting on the syntactic distinction.

---

²⁰ Both temporal epistememes incorporate the morpheme *hira* ‘how much’. In the term *hori.hirak* it is in construction with *hori* ‘since (time)’ and a suffix -k, while *wain.hira* contains the root wain ‘day’, a term which consultants recognised but which is not commonly used in the Fehan dialect.
6.6.2 Basic epistememes: interrogative

6.6.2.1 sé ‘who, which person’

The word sé ‘who, which person’ is used only for persons, as either pronoun (6.55) or determiner (6.56).

6.55 Emi haré sé né sa’e bá uma laran ami-k á?
2P see who this ascend go house interior 1PE-POS DEF
Did you see who went up into our house?

(U4.38)

6.56 “Ó kalo m-akara mane né at mai ó.” K-ák “Mane sé?”.
2S if 2S-like man this IRR come 2S 1S-say man who
“If you like, this man will come to (visit) you.” I said “Which man?”.

(F3.45)

6.6.2.2 nabé ‘which, where’

As a pronoun, nabé ‘which, where’ is most commonly used for location. In this case it follows a verb of motion (6.57) or a locative preposition. The latter results in the standard locative phrase iha nabé ‘where’ (lit. ‘LOC where’; 6.58). Another common expression incorporating nabé is nú nabé ‘how?’ (lit. ‘like which’; 6.59).

6.57 Ó bá nabé?
2S go where
Where are you going? (Both expressions are said to be equally polite.)

(H0.159)

6.58 Ai! Feto ne’e iha nabé ti’an?
EXCL woman this LOC where already
Ai! Where is this girl now?

(R5.9)

6.59 ..Nia lu an nú nabé? Kle’an nú nabé?
3S wide like which deep like which
[Now we’ll measure Maubesi.] How wide is it? How deep?

(AA4.130)

Apart from in these contexts, nabé tends to mean ‘which’, requesting a selection of one referent from a group (whether stated or implied), as in 6.60. There appear to be no semantic restrictions on such NPs; for instance, persons (6.60), things (6.61), location, and time (6.67) are all attested in the corpus.

6.60 Sia sei r-atene na’a in mak nabé ata mak nabé.
3P still 3P-know noble REL which slave REL which
They still knew which (individuals) were noble and which were slave.

(Q0.175)

21 Judging by the published dictionaries of Tetun, nabé is derived by cliticising ne’e ‘this’ to bé (with unknown meaning). This would relate it formally to nebá / ne’e bá / nabá (lit. ‘this go’) ‘there’ and nemai / ne’e mai / namai (lit. ‘this come’) ‘here’ (§6.3.5). Hull (1996b) lists this expression as ne’e bé ‘which’, while Mathijssen (1906) records ne’ebé (which he writes ‘néébé’) ‘which, he/she who, where’, and ne’e bé ida ‘somewhere’. Morris (1984b) lists ne’e bé ‘who, which, where’.

6.61 Nú wén ida nabé ne’e?!
coconut juice one which this
Which ‘coconut juice’ is this?! (The term nú wén can refer to coconut milk,
juice squeezed from coconut flesh or coconut oil.) (N0.112)

6.6.2.3 sá ‘what’
The word sá ‘what’ is the most general epistememe, being used for anything other than the identity of persons. This includes things (6.64), location (thus overlapping with nabé; 6.57), names (e.g. naran sá ‘name what’), and activities (halo sá ‘do what’; 6.62). It also includes reason, with tán sá (lit. ‘because what’) and bá sá (lit. ‘for what’) meaning ‘why’, and nú sá (lit. ‘like what’) meaning ‘what’s up, why’.

6.62 Ó m-alo sá?
2S 2S-do what
What are you doing? (U5.17)

6.63 At bá katak lia sá?
IRR go tell word what
What message are you going to tell? (G1.51)

6.6.3 Basic epistememes: indefinite
Apart from in interrogatives, these three epistememes are frequently found in irrealis contexts (meaning ‘anyone; anywhere; anything’). Such contexts include object of negative predicates (second instance in 6.64). In addition, they occur in positive contexts (as ‘whoever, someone; wherever, somewhere; whatever, something’; 6.65).

6.64 “Ó m-odi sá?” “K-odi ha’i sá!”
2S 2S-bring what 1S-bring not what
“What did you bring?” “(I) didn’t bring anything!” (I5/7)

6.65 Hari Senin é bá sosa sá sá, mama ká sá
day[Mly] Tuesday this go buy what what betel or what
ká, hodi mai.
or bring come
On Tuesday go and buy something—betel or whatever—and bring (it) here.

Although such non-interrogative indefinite meanings can be expressed by basic epistememes, as in the examples above, they are more commonly expressed by the epistememe followed by indefinite ida ‘one’, or by the epistememe being repeated. These complex epistememes will be discussed next.

6.6.4 Complex epistememes

6.6.4.1 Epistememe + ida ‘one’: interrogative, indefinite
An epistememe followed by indefinite ida ‘one’ is most often used after negative predicates (6.66, final clause in 6.70), but is also acceptable with positive predicates (6.67)
and in questions (6.68). The fact that sé ida can be used with plural denotation (6.69) despite containing ida ‘one’ suggests that such sequences are lexicalised.

6.66 *Nia ttu la n-aré sá ida.*  
3S look not 3S-see what one  
He looked (but) didn’t see anything. (L0.3 elicited)

6.67 *...loron blet nabé ida at sobu (nú é) Kwaur day suitable which one IRR demolish like HES alone Ra’in ne’e...*  
[Ra’in this  
[They agreed on] some suitable day to demolish this Kwaur Ra’in (a building)...  
(J6.3)

6.68 *Sé ida ema tür ne’e?*  
who one person sit this  
Who is this person sitting here? (K9.93)

6.69 *Ema ne’e sia sé ida?*  
person this PL who one  
Who are these people? (H0.134)

6.6.4.2 Repeated epistememe: indefinite

Repeated epistememes, however, are used only indefinitely and not in questions.

6.70 *“At bá nabé?” “Ha’u la bá sá sá!” = “Ha’u la IRR go where 1S not go what what 1S not bá sá ida!”*  
“Where are (you) going?” “I’m not going anywhere!” (L0.2)

LOC land where where 3-exist bad  
In every place (no matter where) there is evil. (L0.77)

6.72 *Ha’u sé sé dei, la k-ata’uk.*  
1S who who just not 1S-fear  
I’m not afraid of anyone. (T0.86)

The whole NP containing an epistemic determiner can be repeated with the same effect as repeating the epistememe.

6.73 *Lia sá lia sá katak ha’u.*  
word what word what tell 1S  
Whatever message (it is), tell me. (S2.130)

Formally related to the repetition sá sá ‘anything’ is the common noun sá-sá ‘belongings’ (lit. ‘RDP-what’), which can be modified just like other nouns.

6.74 *Hahulin emi-kan sá-sá ne’e lai.*  
put.away 2P-POS RDP-what this first  
Put away your belongings now. (I0.135)
6.6.5 Distribution of epistemic determiners

The distribution of epistemic determiners has something in common both with that of other determiners and with that of descriptive adjectives. Their classification as determiners is thus tentative.

Like determiners they follow descriptive adjectives (e.g. *tuan bót sé?* ‘important.man big who’ = ‘which ‘big man’?; 4 examples). Like both descriptive adjectives and the indefinite article *ida* (and unlike the demonstrative determiner *ne’e* ‘this’), they precede relative clauses (6.75; 18 examples).22

(6.75) *Ai* nabé *ida* mak sa’e bá ai seluk, *ne’e hák* plant whichever one REL ascend go plant other this say

‘*ai* ktalik’.
plant vine

Any plant which climbs up another plant, this is called a ‘vine’. (T0.26)

Like descriptive adjectives (§7.2.2.3), *nabé* ‘which, where’ can either precede or follow *ida* ‘one’. In the sequence *nabé ida*, *ida* is analysed as part of a complex epistememe meaning ‘which’ (6.67, 6.75). When the order is reversed (as *ida nabé* ‘one which’), there is more emphasis on the uniqueness of the individual being selected, such that *ida* clearly has a numeric (rather than indefinite) interpretation, and the sequence is better translated ‘which one’ (6.61, 6.76).

(6.76) *Lia* ida *nabé* mak at n-ó folin?
word one which REL IRR 3S-hav e price

Which one story is it that has value? (This was said with regard to weighing up different versions of origin myths.) (Q0.35)

Evidence is at yet inconclusive as to whether epistemic determiners can co-occur with determiners.

---

22 This construction of an epistememe followed by a relative clause is mostly found in the corpus in talk of Christian matters, or in explanations (as in 6.75).
The noun phrase

7.1 Introduction

Noun phrases (NPs) are usually headed by a noun or pronoun, although the head is omissible. NPs typically function as arguments of verbs and are the only type of phrase which can function as complement for all prepositions.

This chapter begins with an overview of the structure of simple NPs, followed by a discussion of possession, and then of various NP modifiers. Finally, complex NPs will be dealt with, including coordinated and appositional NPs.

7.2 The structure of simple NPs

7.2.1 Introduction

The internal structure of an NP is considerably affected by the type of the head. Underived common nouns allow the widest range of modifiers, and so will be discussed first, along with restrictions on combination of modifiers. Deverbal nouns are unique in that they can take arguments. Pronouns and proper nouns allow fewer modifiers than do common nouns. Finally NPs without head nominals will be discussed.

7.2.2 NP with underived common noun head

7.2.2.1 General structure

The formula for a simple NP with common noun head is as follows. The longest sequence of modifiers normally found in an NP would consist of up to two determiners and two other modifiers (7.5).

\[
\text{NP}.\text{Common}.\text{Noun} \rightarrow (\text{Premod}) \text{Psr} \text{PreCl} \text{N} \text{Gen} \text{Psr} \text{NumP1} \text{AdjP1} \text{NumP2} \text{AdjP2} \text{Det1*} \text{PP} \text{RelCl*} \text{Det2*}
\]

The order of Tetun modifiers goes against the universal proposed by Hawkins (1983:93). He says that, for prepositional languages, adjectives, demonstratives and numerals are easier to place in premodifier position than are the heavier modifiers, namely possessors and relative clauses. Similar suggestions have been made, based on empirical observations, by Greenberg (1963) (for relative clauses and adjectives only) and by Kuno (cited in Mallinson & Blake 1981:300ff.). In contrast, Tetun, though a prepositional language, requires that adjectives, demonstratives and numerals follow the head, while allowing possessors and a restricted class of relative clauses to occur in premodifier position.

\[\text{1}\] The order of Tetun modifiers goes against the universal proposed by Hawkins (1983:93). He says that, for prepositional languages, adjectives, demonstratives and numerals are easier to place in premodifier position than are the heavier modifiers, namely possessors and relative clauses. Similar suggestions have been made, based on empirical observations, by Greenberg (1963) (for relative clauses and adjectives only) and by Kuno (cited in Mallinson & Blake 1981:300ff.). In contrast, Tetun, though a prepositional language, requires that adjectives, demonstratives and numerals follow the head, while allowing possessors and a restricted class of relative clauses to occur in premodifier position.
In the following sections we consider only possible combinations and orderings of modifiers. The internal structure of all phrasal and clausal constituents, and the characteristics and closed-class membership of word classes, are discussed elsewhere, as outlined in the following table.

Table 7.1: Noun phrase constituents

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
<th>Discussion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premod</td>
<td>premodifying words</td>
<td>§7.4</td>
</tr>
<tr>
<td>Psr</td>
<td>possessor</td>
<td>§7.3</td>
</tr>
<tr>
<td>PreCl</td>
<td>premodifying relative clause</td>
<td>§7.4.4</td>
</tr>
<tr>
<td>Gen</td>
<td>genitive enclitic (-n, -r)</td>
<td>§7.3.3</td>
</tr>
<tr>
<td>NP</td>
<td>noun phrase</td>
<td>§7.5.1</td>
</tr>
<tr>
<td>AdjP</td>
<td>adjective phrase</td>
<td>§7.5.2</td>
</tr>
<tr>
<td>NumP1</td>
<td>numeral phrase: 'one'</td>
<td>§5.1</td>
</tr>
<tr>
<td>NumP2</td>
<td>numeral phrase: numerals greater than 'one'</td>
<td>§5.1</td>
</tr>
<tr>
<td>PP</td>
<td>prepositional phrase</td>
<td>§8.1</td>
</tr>
<tr>
<td>RelCl</td>
<td>postmodifying relative clause</td>
<td>§14.8</td>
</tr>
<tr>
<td>Det1</td>
<td>determiner: any determiner</td>
<td>§6</td>
</tr>
<tr>
<td>Det2</td>
<td>determiner: ne'e 'this', sia 'PL'</td>
<td></td>
</tr>
</tbody>
</table>

An NP can have only one possessor (Psr), either premodifying or postmodifying. The genitive marker (Gen) occurs with premodifying possessors. Only one of the two determiner slots can be filled. All types of modifiers are illustrated in the following examples, in which the third interlinear line indicates the start of each constituent within the NP.

7.1  Kal o  mana e  la  n-o'uk  bá  feto  nia  musti tate.
      if  male  not  3S-agree.to  to  female  3S  must  pay.fine
      N            N
      If the man doesn't agree to (marry) the woman, he must pay a fine.  (10.132)

7.2  feto  ki'ik  á
      woman  small  DEF
      N      AdjP    Det
      the small (or younger) woman  (F2.48)

7.3  ohin  nia-kan  ali-n  ne'e
      just.now  3S-POS  younger.sibling-GEN  this
      Premod  Psr  N-Gen  Det
      this his aforementioned younger brother  (P1.38)

7.4  ata  nia-k  rua  ne'e
      slave  3S-POS  two  this
      N      Psr  NumP  Det
      these her two slaves  (G4.51)

7.5  na  renu-n   iha  nia  ne'e  sia
      3S  commoner[Port]-GEN  LOC  3S  this  PL
      Psr  N-Gen  PP  Det  Det
      his commoners at that (place)  (Z3.272)
It is highly unusual to have more than two non-determiner modifiers within an NP, unless one of the modifiers is so closely bound to the head noun as to give the sequence the character of a compound (e.g. ai fuan ‘fruit’ in 7.6). A sequence of three adjectives followed by the indefinite article ida ‘one’ was given during elicitation about adjective sequences (ema mutin naruk bót ida ‘person white long big one’ = ‘a large tall white person’), but (with the exception of some non-restrictive relative clauses) there are no clear examples of such long sequences in the corpus.

Generic NPs have no determiners.

In the following sections various issues relating to combinations of modifiers will be considered. The first relates to various means of marking plurality and the extent to which these co-occur. This is followed by discussions of the relative ordering of various combinations of modifiers, namely numerals versus adjectives, prepositional phrases and relative clauses versus determiners, and sequences of determiners.

7.2.2.2 Marking plurality

An NP with plural referents usually has the plurality marked in some way. This may be indicated by the plural marker sia (7.5), by a numeral phrase (7.4), a numeric adjective (e.g. ruas ‘two’; §4.6.7, §4.9.6), a quantifying adjective (e.g. wa’in ‘many, much’, ruma(k) ‘others’), a plural pronoun head (e.g. emi ‘2P’), coordination (§7.6.1), repetition of the head noun (7.15; §4.8.4), or reduplication of the adjective (7.6, 7.16; §4.8.3).²

² According to Hull (1996b:37), the Tetun spoken in southern East Timor in the past had a number of plural noun forms ending in -n (e.g. belu-n ‘friends’) or -r (e.g. ema-r ‘persons’). I have no data supporting this for the present-day Fehan dialect, where belu-n is the genitive marked form of belu, and emar is an alternate form of ema which can refer to a single person.
The exception is *ema* ‘person’, which is frequently left unspecified for number when it is generic (7.11), or when it refers more generally to people whose identity (and often number) are not relevant (7.12). That is, from a pragmatic point of view it is non-referential, meaning ‘someone(s)’. In such instances subject marking on the verb, if any, is normally plural (7.12).

7.11 Ø bá m-usu *ema*, *ema* fó, la bele fó hotu.
2S go 2S-ask person person give not can give all
If you go and ask someone/people, and he/she/they give you (an explanation), (he/she/they) can’t give (it) all. (P4.89)

7.12 Ha’u k-o’i bá té *ema* r-aboba ha’u.
1S 1S-do.not want go because person 3P-hit 1S
I don’t want to go as (the) people (there) hit me. (R4.55)

Although *ema* is, in the corpus, the most common noun to be unspecified for number, it is possible to omit plural marking for other nouns also if the number of referents is not relevant.

7.13 Karau tama to’os. Ha bá daka.
buffalo enter garden 1S go look.after
(A/Some) buffalo entered the garden. I went to guard (the garden). (O2.61)

Most plural NPs have the plurality indicated only once. The only combinations of plural marking that frequently occur within a single NP are of a plural pronoun modified by a numeral or the attributive adjective *ruas* ‘two’ (e.g. *sia ruas* ‘they two’), and of the necessarily plural adjective *ruma(k)* ‘others’ followed by the plural marker *sia* (e.g. *ema rumak sia* ‘other people’).

Apart from this it is rare but possible to have more than one indication of plurality within the one NP (7.14). Such multiple marking occurs most readily when one of the plural indications has other meanings as well, namely specifying person for pronouns (hence *emi ne’e sia* ‘2P this PL’ = ‘you (plural) here’) or indicating variety for repeated nouns (7.15) and reduplicated adjectives (7.16).

7.14 ...To’o *isin* (é) *hira* *sia* má...
until time HES several PL and.then
N NumP Det
[She kept trying to waken them.] (This went) on several times, then...

7.15 *Hudi* hudi né *sí* *ita* la há.
banana banana this PL 1PI not eat
N Det Det
These (any) types of bananas we don’t eat (after giving birth). (J3.108)

7.16 ...*foho* bá-bót e’e *sia* be’o...
mountain RDP-big this PL shatter
N AdjP Det Det
[There was an earthquake,] these large mountains shattered...

---

3 Hull (1993:32) similarly observes that plural marking is often omitted from *ema*. The vagueness of *ema* is further shown by the fact that it is often followed by another NP which gives a more specific meaning (e.g. *ema feto* ‘woman’, *ema dōk* ‘shaman’, *ema to’os na’in* ‘garden owner’). This seems comparable to the use of the semantically general *ai* ‘plant’ and *na’an* ‘meat, fish’ before most plant and fish names (e.g. *ai kamelin* ‘sandalwood’, *na’an knase* ‘mugil fish’). *Ema* is also unusual syntactically in that it can take possessive marking either like a noun (*ema nia-kan* ‘person 3S-POS’) or like a personal pronoun (*ema-kan* ‘person-POS”).
7.2.2.3 Order of modifiers: numerals and adjectives

Numerals other than *ida* ‘one’ follow adjectives (11 examples).

7.17  

<table>
<thead>
<tr>
<th>Feto</th>
<th>kawa‘ik</th>
<th>na‘in</th>
<th>néén</th>
<th>né‘e</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>AdjP</td>
<td>NumP</td>
<td>Det</td>
<td></td>
</tr>
<tr>
<td>woman</td>
<td>older</td>
<td>CLS:human</td>
<td>six</td>
<td>this</td>
</tr>
</tbody>
</table>

In (Z3.74) these six elder sisters (lit. ‘women’).

However, *ida* ‘one, a’ can either follow or precede an adjective. Evidence from the corpus suggests that where *ida* follows the adjective, an interpretation as indefinite article is always possible, even if not required (7.18, 7.24; over 50 examples). Where *ida* precedes the adjective, a numeric interpretation is in each case forced, both by context and by the co-occurrence of a determiner (7.19; 6 examples).

7.18  

<table>
<thead>
<tr>
<th>Feto</th>
<th>kara‘u</th>
<th>aman</th>
<th>bót</th>
<th>ida</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>AdjP</td>
<td>AdjP</td>
<td>Det</td>
<td></td>
</tr>
<tr>
<td>buffalo</td>
<td>male</td>
<td>big</td>
<td>one</td>
<td></td>
</tr>
</tbody>
</table>

In (F0.162) a big bull.

7.19  

<table>
<thead>
<tr>
<th>Feto</th>
<th>ida</th>
<th>kawa‘ik</th>
<th>á.</th>
<th>Bano Eki Liurai, n-alai dalan seluk,</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>NumP</td>
<td>AdjP</td>
<td>Det</td>
<td></td>
</tr>
<tr>
<td>woman</td>
<td>one</td>
<td>older</td>
<td>DEF</td>
<td>Bano Eki ruler 3S-run road other</td>
</tr>
</tbody>
</table>

In (F2.43) (After the two wives argued,) the oldest woman, Bano Eki Liurai, ran (away) down one road, the youngest woman ran down another road.

The only generalisation in the data about adjective order is that specification of gender (e.g. *inan* ‘female (of animal)’) precedes specification of colour or size, as in 7.18.

7.2.2.4 Order of modifiers: prepositional phrases, relative clauses and determiners

Prepositional phrases and relative clauses follow adjective phrases (7.6, 7.20) and numeral phrases (7.21).

7.20  

<table>
<thead>
<tr>
<th>oan</th>
<th>ki‘ik</th>
<th>nú</th>
<th>né</th>
<th>né</th>
<th>sia</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>AdjP</td>
<td>PP</td>
<td>Det</td>
<td>Det</td>
<td></td>
</tr>
<tr>
<td>child</td>
<td>small</td>
<td>like</td>
<td>this</td>
<td>this</td>
<td>PL</td>
</tr>
</tbody>
</table>

In (J2.59) the little children like this (Said pointing.).

---

4 It is not surprising that the only numeral to precede (and thus fall within the scope of) the adjective is ‘one’. For higher numerals, the adjective must fall within the scope of the numeral. This is because the numeral indicates a collection, of which each of the members is individually described by the adjective. Thus, for instance ‘five red houses’ describes five houses which are red, as opposed to a red collection of five houses (as pointed out by Hurford (1987:221)). The order in which the numeral precedes the adjective contravenes Greenberg’s (1963) universal 20, but fits Hawkins’s (1983:119f.) weakened version of it.

5 This is consistent with the crosslinguistic preference for placing ‘heavy’ constituents towards the right within an NP, a fact for which Hawkins (1994) finds good processing-based reasons.
It is possible but rare for the one NP to be modified by both a prepositional phrase and a relative clause, or by two relative clauses.

Epistemic determiners and the indefinite article *ida* 'one' precede prepositional phrases and relative clauses (7.23, 7.24), while the definite determiner *ne'e* 'this' and the plural marker *sia* 'PL' usually follow them (7.20, 7.22). Nevertheless, non-restrictive and 'afterthought' relative clauses can follow what are normally NP-final markers, such as these determiners or a phrase-level pause (7.25; 5 examples).

Of the determiners, only *ne'e* 'this', *sia* 'PL' and *hotu-hotu* 'all' co-occur. The vast majority of examples in which *ne'e* and *sia* co-occur follow the order *ne'e sia* 'this PL' (7.20, 7.22; 205 examples); however, *sia ne'e* is also attested. The latter seems to occur only when the *ne'e* emphasises continuity with a preceding plural topic; that is, where the plurality has already been mentioned recently and the *ne'e* is used anaphorically (7.26; 7 examples).
Chapter 7

In the corpus hotu-hotu ‘all’ co-occurs with either a preceding sia ‘PL’ (7.27; 3 examples) or a following ne’e ‘this’ (7.28; 4 examples). Other ordering possibilities are not known.

7.27

Ita bolu gan sia hotu-hotu mai hamutu bá ne’e.

1PI call child PL RDP-all come together at this

We ask all the children to come together here. (V0.45)

7.28

...ema hotu-hotu ne’e la n-aré.

person RDP-all this not 3-see

[When the basket came down,] all these people didn’t see (it). (O6.96)

7.2.3 NP with deverbal noun head

Deverbal abstract nouns (§4.5.1) are found almost exclusively within possessive constructions, with the possessor corresponding to the subject of the verb (67 examples). The possessor usually precedes the nominalisation (7.29), although a following possessor was accepted in elicitation (7.30; 4 examples). Where present, the object of an underlying transitive verb is introduced by bá ‘to, go’ (7.31; 2 examples).

Deverbal.Noun (Psr) [Deverbal.Noun (ba NP) (Det)]

7.29

Feto Ikun ne’e nia-kan da-di’ak di’ak resi.

woman tail this 3S-POS NOM-good good too

This Youngest Sister, her beauty (lit. ‘goodness’) is exceptional. (Z6.45)

7.30

Fa-fiar ó-k á Sarani Protes.


Your religion is Protestant Christian. (X0.102 elicited)

7.31

Ó-kan fa-fiar bá ha’u la lós.

2S-POS NOM-believe to 1S not true

Your trust in me is not right. (Q0.115)

Such nominalisations cannot readily take modifiers other than possessors and determiners, with consultants accepting hahalok ái ‘behaviour bad’ = ‘bad behaviour’ but rejecting other attempts at modification by adjectives (e.g. ‘dirty talk’) and numerals (e.g. ‘two commands’). The only unelicited modifier in the corpus is a single instance of negation.

7.32

nia-kan di’ak ha’i

3S-POS good not

its bad aspects (lit. ‘its not-goodness’) (H7/5)

7.2.4 NP with pronominal head

Although pronouns are usually not modified, some (particularly personal pronouns) can take a considerable range of modifiers, as shown by the formula below.6 Adjectival modifiers are rare, with the exception of ruas ‘two’ (e.g. sia ruas ‘they two’), and the highly respectful formulaic ita bó’t (lit. ‘big you’) used to formally address God, spirits or very senior nobility.

---

6 Since combinations of lexical modifiers are rare for pronouns, the order of modifiers in the formula is tentatively presented as a simplified version of the formula in §7.2.2.1.
The noun phrase

7.2.5 NP with proper noun as head

7.2.5.1 Sociolinguistic rules for terms of address

Tetun offers a range of possibilities for addressing people, including pronouns, titles (a subclass of common nouns), and names ('proper nouns'). These options reflect the hierarchical nature of Fehan society, the importance of kinship, and the growing impact of Indonesian forms. In this section we consider names and titles, while the following section demonstrates the range of modifiers allowed for proper nouns.

Fehan people have several sets of names. The naran matabian (lit. ‘spirit.of.the.dead name’) is named after an ancestor, and as such is selected from a fairly small set of native Fehan names (e.g. Ho’ar or Luruk for girls, Seran or Klau for boys). Since many people are known by the same ancestral names, a modifier may be added as part of the name (e.g. ki’ik ‘small’, nurak ‘young’, klaran ‘middle’). The naran Sarani ‘Christian name’ is non-native, and generally does not follow Tetun phonology (e.g. Alfonsius, Yuliana). Only people with post-school education, such as schoolteachers, are known by this name, which for other people is used only in written documents. The naran baba’ur is a nickname which is often based on personal characteristics, such as Bo’uk ‘chubby’, Iku (lit. ‘tail’) ‘youngest’, Mutik ‘whitney’, Guru ‘teacher [Malay]’. Many people are known by these nicknames. Parents in addition acquire a teknonym based on the name of their first child, followed by ina ‘mother’ or ama ‘father’ (e.g. Ho’ar-ina ‘mother of Ho’ar’, Ser-ina ‘mother of Seran’, Luk-ama ‘father of Luruk’).\(^7\) Except when referring to children, names are normally preceded by a title.

Despite this profusion of names it is in many contexts considered rude to address a person by his or her personal name, especially if the addressee is of higher status than the speaker.

\(^7\) Whether all or part of the child’s name is used depends on the name.
Instead, titles are frequently used, both as term of address and in place of second person pronouns. For adults, various terms meaning ‘mother’ and ‘father’ are used. Village people with little formal education, and nobles regardless of their education, are respectfully addressed by the Tetun terms ina and ama, or in-bei and am-bei (lit. ‘father-grandparent/ancestor’). The most senior nobles are addressed as ina na’i or ama na’i (lit. ‘father noble’). People with advanced formal education, such as schoolteachers, are addressed using the Indonesian equivalents ibu and pāi (written ‘Pak’ and pronounced pā’ in Indonesian), as are most non-Tetun adults. Dialectal Malay mama and papa are also used for some outsiders, and by some children to their educated parents. For other older relatives, other kin terms may be used (e.g. tuak ‘mother’s brother’), or kin terms may be used as titles before the ancestral name or nickname (e.g. Mau Manek ‘elder.brother Manek’, Bí Iku ‘elder.sister tail/youngest’). Names without titles may be used for siblings (especially younger ones) and children.

In addition, the very common male name Manek and female name Bete can be used to honourably address people regardless of their personal name. Na’i (lit. ‘noble’) can be used for all nobles, and is also an affectionate term for people significantly younger than oneself (7.38).

When referring to oneself, some speakers similarly prefer to use the title by which the addressee should call them, rather than using first person pronouns.

7.37  

\[
\text{Ina } \text{dakar oan hawa'ı hanesan oan tε.}
\]

‘Mother’ (i.e. I whom you call mother) looks after adopted children like her children by birth.  

(D0.9)

7.2.5.2 Structure of a proper noun NP

In general, proper nouns are sufficiently restricted in their reference to require no further modification other than a title (7.38) or an adjective (7.39). Non-restrictive relative clauses may add information about the referent (7.40). Finally, discourse modifiers are accepted, including ohin ‘aforementioned’ (7.39) and determiners (7.39; Mama Funan sia ‘Mother Funan and associated people’).

NP.Proper.Noun → (ohn) (Title) Proper.Noun (AdjP) (RelCl) (Det*)

7.38  

\[
\begin{array}{llllllll}
\text{Mεo} & \text{Teti} & \text{Mεo} & \text{Bere} & \text{n-a'ak} & \text{Ami} & \text{hakoi ita} & \text{bå} & \text{ne'e} \\
\text{hero} & \text{Teti} & \text{hero} & \text{Bere} & \text{3-say} & \text{1PE} & \text{bury} & \text{2S.HON} & \text{at this} \\
\text{Title} & \text{PN/} & \text{Title} & \text{PN} \\
onan, & \text{na'i''}. \\
\text{IMM} & \text{noble} \\
\text{Hero Teti and Hero Bere said ‘We'll bury you here now, Sir/Dea'}. \\
\end{array}
\]

(K10.122)

---

8 Prepositional phrase modifiers are not attested in the corpus.
The noun phrase

7.39 \(\text{Liurai} \ Suri \ Nurak \ ne'e, \ nia \ mak \ taruka \ ohin \ liurai \ Suri\)
ruler Suri young this 3S REL replace just.now ruler Suri
Title PN AdjP Det Premod Title PN

Tuan.
old
AdjP
This Ruler Suri Junior, he was the one who replaced the aforementioned Ruler Suri Senior.

(F2.3)

7.40 \(\text{Bá futu manu nó Lakuleik, Liurai Lakuleik mak}\)
go tie bird with Lakuleik ruler Lakuleik REL
PN/Title PN RelCI

futu manu, laka manu.
tie bird shine bird
(They) went cockfighting (lit. ‘tying birds’) with Lakuleik, (the) Ruler Lakuleik who fought cocks (lit. ‘tied birds, shone birds’).

(G1.8)

7.2.6 NP without head nominal

An NP can readily occur without a head nominal if context makes the reference or denotation sufficiently clear. In this case constituents which are otherwise modifiers occur in the same order as if there were a common noun head. Only constituents that function as postmodifiers of nominal heads can occur in headless NPs.

The following examples illustrate NPs with an initial possessor (7.41, 7.42), adjective phrase (7.43), numeral phrase (7.44, 7.45), and relative clause (7.46).

7.41 \(\text{Ha'a-k mesa ruin.}\)
1S-POS solely bone
Psr
Mine (my roasted bird) is only bones.

(K12.40)

7.42 \(\text{N-ola ata nia-k rua ne'e, n-ák "M-asai ó-k}\)
3S-take slave 3S-POS two this 3S-say 2S-take.out 2S-POS
Psr

rua mai”.
two come
NumP
Taking her two slaves, (she) said “Bring out your two (slaves)”.

(G4.51)

7.43 \(\ldots \text{Naruk sì it la hatene.}\)
long PL 1PI not know
AdjP Det
[We know only the short poems.] Long ones we don’t know.

(J3.24)
7.44 ...na'in nén r-ameti tais, na'in ida n-ameti
    CLS:human six 3P-secure cloth CLS:human one 3S-secure
    NumP

karon.
sack

[Seven men came courting.] Six wore handwoven cloths, one wore a sack.

(U11.2)

7.45 ...lda iha Wehali né ... lda iha Wewiku å...
    one LOC Wehali this one LOC Wewiku DEF
    NumP PP Det NumP PP Det

[The three women were sent to different places.] The one here in Wehali [was
called Bano Eki Liurai...] The one in Wewiku [was called Aek Eki Liurai].

(F2.23)

7.46 Mais mak la r-atene adat sia...
    but REL not 3P-know tradition PL
    RelC Det

But (those) who don’t know tradition...

(110/7)

Note that many determiners (e.g. ne’e ‘this’, sá ‘what’, sia ‘PL’) have alternative
classification as pronouns. As a result they are analysed as pronominal heads (rather than as
modifiers within headless NPs) if there are no preceding constituents.

7.3 Possession

7.3.1 Overview

The Fehan dialect of Tetun allows the possessor to be specified either before the
possessum or after it. The preposed position is by far the more common, representing over
80% of textual examples in the corpus.9 This is in contrast to almost all other types of
modifiers, which follow the head noun. The part of the overall NP formula relevant to
possession is here rewritten for convenience, along with further formulae expanding the
constituents.

NP. Possessive → \{ ... Pre-possessor ... N (Gen) ... \}

{ ... N Post-possessor ...

Pre-possessor → \{ Pro (Pos) \}

NP (Pro3 (Pos))

9 The order in which possessor precedes possessum is typical of the so-called ‘reversed genitive’ languages of
eastern Indonesia east of the Brandes line, and is the only constituent order noted by Capell (1944:31) for
the Austronesian languages of Timor. It is also the only order mentioned in the brief Tetun descriptions by
Morris (1984b:xv) for East Timor, and by Mathijisen (1996:v) and Troeboes et al. (1987:77) for the Foho
dialect of northern Belu. According to Wortelboer (1955:177), the possessor-possessed order is found in
north Belu, with the reverse order being found in the Fehan dialect. Hull (Hull 1996b:59, 83), writing of Dili
Tetun, notes both orders for nominal possessors, with the possessor-final order being less common. While
the fact that only Hull and Wortelboer mention postposed possessors may reflect the brevity of most
accounts, it probably supports Wortelboer’s contention that postposed possessors, at least as a commonly
used option, have limited geographical distribution within the Tetun-speaking area.
As shown by the formulae, and illustrated in 7.47, preposed possessors are maximally realised either by a possessive-marked pronoun, or by an NP followed by a possessive-marked pronoun; the following possessed noun is marked by a genitive clitic if it is vowel-final and satisfies other conditions discussed in §7.3.3.2. Postposed possessors, illustrated in 7.48, similarly consist of either a possessive marked pronoun or an NP followed by such a pronoun, but there is no genitive clitic on the head noun, and no possibility of omitting the pronoun or the possessive marker.

### 7.47

<table>
<thead>
<tr>
<th>Possessor</th>
<th>Pro3</th>
<th>Pos</th>
<th>N</th>
<th>Gen</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ó</td>
<td>fé</td>
<td>-n</td>
<td>wife</td>
<td></td>
<td>your wife</td>
</tr>
<tr>
<td>2S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. sa</td>
<td>ata</td>
<td>-r</td>
<td>slave</td>
<td></td>
<td>(a) slave of theirs</td>
</tr>
<tr>
<td>3P</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. ha'ũ</td>
<td>-kan</td>
<td>tais</td>
<td>cloth</td>
<td></td>
<td>my cloth</td>
</tr>
<tr>
<td>1S</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. fahi</td>
<td>na'in</td>
<td></td>
<td>noble</td>
<td></td>
<td>pig's owner</td>
</tr>
<tr>
<td>pig</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e. Bui Hirik</td>
<td>ni</td>
<td>naran</td>
<td>name</td>
<td></td>
<td>Bui Hirik's name</td>
</tr>
<tr>
<td></td>
<td>3S</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. tamukun</td>
<td>nia</td>
<td>-kan</td>
<td>fé</td>
<td>-n</td>
<td>village head's wife</td>
</tr>
<tr>
<td>village.head</td>
<td>3S</td>
<td></td>
<td>wife</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g. ina-ama</td>
<td>sia</td>
<td>-kan</td>
<td>sasain</td>
<td></td>
<td>parents' wealth</td>
</tr>
<tr>
<td>mother-father</td>
<td>3P</td>
<td></td>
<td>wealth</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 7.48

<table>
<thead>
<tr>
<th>N</th>
<th>Possessor</th>
<th>Pro3</th>
<th>Pos</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>asu</td>
<td>ó</td>
<td>-k</td>
<td>your dog</td>
</tr>
<tr>
<td>dog</td>
<td>2S</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>inan-aman</td>
<td>feto</td>
<td>nia</td>
<td>parents of the girl</td>
</tr>
<tr>
<td>mother-father</td>
<td>woman</td>
<td>3S</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Predicative and headless possessors, illustrated in 7.49, have the same structure as postposed possessors.
The following table gives a preview of the major factors influencing the choice of possessive construction. These factors will be examined in the following sections. These sections deal first with factors affecting the form of the possessor, then with those influencing the presence and choice of genitive marking on the possessum. This is followed by a discussion of the differences between preposed and postposed possessors, and finally by a brief look at a construction in which the possessor slot is filled by an adjective.

<table>
<thead>
<tr>
<th>Options</th>
<th>Factors influencing choice</th>
<th>Person-number of possessor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Possessor is preposed/postposed</td>
<td>Preposed required for part-whole, usual for kin, optional for close and possessions</td>
<td></td>
</tr>
<tr>
<td>NP possessor has pronoun (indirect)/no pronoun (direct)</td>
<td>Direct usual for part-whole, optional for kin and close, disallowed for possessions</td>
<td></td>
</tr>
<tr>
<td>Possessor has possessive clitic/no clitic</td>
<td>Clitic required for postposed possessor, optional for preposed</td>
<td></td>
</tr>
<tr>
<td>Possessive clitic is -kan/-k</td>
<td>-kan strongly preferred for preposed possessor, -k for postposed</td>
<td></td>
</tr>
<tr>
<td>Possessum has genitive clitic/no clitic</td>
<td>Clitic allowed only if possessor is preposed</td>
<td>Clitic optional for 1S (even if required by alienability class)</td>
</tr>
<tr>
<td>Genitive clitic is -n/-r</td>
<td>Clitic required for part-whole and kin, usual for close, disallowed for possessions</td>
<td>-r allowed only for plural possessor, more common for 3P; -n allowed for all</td>
</tr>
</tbody>
</table>
7.3.2 Form of the possessor: pronouns and possessive clitics

7.3.2.1 Overview of possible forms

As demonstrated above, possessors can be expressed in a number of ways. A nominal possessor may consist of an NP juxtaposed directly to the possessum (7.47d), in what is known in the Oceanic literature as ‘direct possession’ (Lynch 1982:27). Alternatively, it may be followed by a third person pronoun that specifies the number of the possessor, namely nia ‘3S’ for singular (7.47e, 7.47f), or sia ‘3P’ for plural (7.47g). This use of a mediating pronoun is called ‘indirect possession’.

Pronouns allow the same range of options regardless of whether they follow an NP possessor or are themselves the possessor. If there is no following possessive marker, the pronoun can consist of either the short form of a pronoun (7.47b, 7.47e), or the full form (7.47a, 7.47c). However, if there is a following possessive marker, the full form of the pronoun is required. The possessive marker cliticises to the preceding possessive pronoun, formally marking it as possessive. It can take the form -kan (7.47c, 7.47f) or -k (7.48).

The choice between these various options for representing the possessor is affected by alienability, by the position of the possessor before or after the head noun, by stylistics, and by dialect.10

7.3.2.2 Alienability

A major determinant affecting the choice between direct and indirect possession is the alienability of the relationship between possessor and possessum. This relationship also affects whether postposing of the possessor is possible (§7.3.4.1), whether a genitive clitic is used with pronominal possessors (§7.3.3.2), and whether a final -n (possibly a fossilised genitive clitic) is used even when a noun is not in a possessive construction (§7.3.3.3).

Alienability in Tetun is a cline. At the one end are ‘clearly inalienable’ relations such as part-whole relationships. At the other are ‘clearly alienable’ relationships such as that of a person to temporary physical possessions. These extremes are expressed quite differently syntactically. In between it is convenient to distinguish two alienability classes which in some respects group syntactically with alienable, and in other respects with inalienable, as shown in Table 7.3 below. Semantically the four alienability classes may be characterised as follows:

1. Part-whole relations are inalienable. These include body parts (e.g. ha ulu-n ‘my head-GEN’, nia kabun ‘his/her stomach’) and spatial relations (e.g. laran ‘interior’, leten ‘top’). In fact body part terms are a major source of terms for spatial relations, as illustrated by words such as oin ‘face, front’, and sorin ‘side, beside’. For a fuller list of location terms see §8.3.1.

10 Possessive marking on pronouns appears to be an area of great dialectal variation. The corpus includes a few instances of a possessive clitic -n (e.g. ha’u-n ‘IS-POS’), which is recognised by consultants as being from outside the Fehan dialect area, and of -k on preposed possessors, which is similarly considered non-Fehan. Full pronouns without possessive marking (e.g. ha’u ‘1S’) are encountered in the corpus, but are in the Fehan dialect not as preferred as short pronouns (e.g. ha ‘IS’) or possessive-marked pronouns (e.g. ha’u-kan ‘1S-POS’).

For the Foho dialect of northern Belu, Mathijsen (1906:iv-v) lists -kan, -n and zero marking of the pronoun, while Troeboes et al. (1987:78) mention only the bare pronoun. For East Timorese Tetun Fernandes (1937:30f) and Morris (1984b:xv) list -nia ‘3S’ where the possessor is followed by a noun (e.g. ha’u-nia asu ‘my dog’), and nian, -k and -n where the possessor is used predicatively (ha’u nian, ha’u-k, ha’u-n ‘mine’). According to Hull (1996b:61) speakers of rural and literary East Timorese Tetun frequently use a bare pronoun as possessive (e.g. ha’u ‘my’).
2. Kin relations are expressed by terms such as *ama* ‘father’ and *oa* ‘child’. This category includes *na’in* ‘noble, owner’.

3. Close: This small category includes relationships that are close and significant, but which can be broken. Semantically it falls into two subclasses. The one involves important physical locative possessions, namely *rai* ‘earth, land’ and *uma* ‘house’. The other involves close non-kin social relationships, including *ata* ‘slave, servant’ and *belu* ‘friend’.

4. Possessions: Physical possessions other than house and land (e.g. *faru* ‘clothes’) are always expressed as alienable possession.

The table below summarises the interactions between these classes and the various syntactic and morphological factors mentioned above. Note that since genitive clitics (essential to two of the tests for class membership) can apply only to vowel-final words, the classification of some consonant-final words (e.g. *naran* ‘name’) is problematic. In addition, various tests give probabilistic rather than absolute results, further compounding the difficulty of classifying some words.

<table>
<thead>
<tr>
<th>Table 7.3: Syntactic correlates of alienability classes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Part-whole</strong></td>
</tr>
<tr>
<td>Kin</td>
</tr>
<tr>
<td>Close</td>
</tr>
<tr>
<td>Possessions</td>
</tr>
</tbody>
</table>

### 7.3.2.3 Effect of alienability on the form of the possessor

Alienability affects the expression of possessors in two ways. Firstly, alienability is the main determinant of the choice between direct and indirect possession. Secondly, it has some impact on whether possessor pronouns are formally marked as possessive (by *-k* or *-kan*).

Part-whole relations are usually expressed by direct possession (i.e. by juxtaposition of the possessor NP and the possesseum; 7.50). Many such expressions have the semantic character of a compound, being the conventional name for their referents. Part-whole relations can, however, be expressed by indirect possession as well (7.51), with indirect possession being usual where a body part is specified as belonging to a specific individual (7.52).13

---

11 This test of whether a noun retains final -n when there is no (explicit) possessor is of course not relevant to nouns which inherently end in *ni* (e.g. *nin* ‘wind’).

12 A percentage is not presented due to the difficulty of knowing the exact membership of the ‘close’ category and the relatively small number of examples. Examples for close possession are thus included in the percentage for ‘possessions’.

13 I suspect there are situations in which indirect possession is impossible, but this requires further testing. According to Capell (1944:31) part-whole relations are expressed by direct possession, while alienable possession is expressed using a pronoun (i.e. by indirect possession). This is a tendency but not a rule for the Fehan dialect.
7.50

<table>
<thead>
<tr>
<th>Direct possession</th>
<th>Gloss</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>faru kakaluk</td>
<td>clothes pocket</td>
<td>pocket in clothes</td>
</tr>
<tr>
<td>samodo né kakun</td>
<td>green tree snake</td>
<td>skin of this green tree snake</td>
</tr>
<tr>
<td>uma laran</td>
<td>house interior</td>
<td>inside the house</td>
</tr>
<tr>
<td>foho leten</td>
<td>mountain top</td>
<td>mountain top</td>
</tr>
<tr>
<td>foho bôt ida leten</td>
<td>mountain big one top</td>
<td>top of a big mountain</td>
</tr>
<tr>
<td>ai nahon</td>
<td>tree shade</td>
<td>shade of a tree</td>
</tr>
<tr>
<td>malae lia-n</td>
<td>Malay language-GEN</td>
<td>Malay language</td>
</tr>
<tr>
<td>ha’i kbon</td>
<td>fire smoke</td>
<td>smoke from a fire</td>
</tr>
</tbody>
</table>

7.51

mál nia-kan laran
mould[Mly] 3S-POS interior
inside the mould (AA3.20)

7.52

Bei Ulu Kmeik nia-kan ulu-n
Mr head sharp 3S-POS head-GEN
Mr Sharp Head’s head (K8.56)

Similarly, for kin, kin of a specified individual are expressed by indirect possession (e.g. Bei Lulik nia-kan fê-n ‘Mr Taboo’s wife’; 7.47f.), while kin of kin (i.e. complex kin terms) are usually identified by direct possession (e.g. ali la’e-n ‘younger sister’s husband, la’e ali-n ‘husband’s younger sibling’).

Close relations (particularly rai ‘earth, land’ and uma ‘house’) are often expressed by direct possession (7.53, ema rai-n ‘other people’s land’, na’in Lakuleik rai-n ‘noble Lakuleik’s land’), but freely occur with indirect possession as well (7.54).

7.53

feto Luk Morin ne’e uma-n
woman Luk sweet.smell this house-GEN
this girl Luk Morin’s house (U10/50)

7.54

feto Luk Morin ne’e nia-kan uma-n
woman Luk sweet.smell this 3S-POS house-GEN
this girl Luk Morin’s house (U10.21)

For relations to physical possessions, indirect possession seems to be required (e.g. ema nia-kan batar ‘person 3S-POS maize’ = ‘someone’s maize’).

Statistically, alienability also has some effect on the choice between marking a preposed possessor by a pronoun only, or by a pronoun which is formally marked as possessive by -k(an). In particular, kin relations are expressed marginally more often without possessive marking than with it, while all other relations are expressed much more often with a possessive clitic than without one.

---

14 The final /n/ on the other possessed nouns in this list are not glossed as separate genitive morphemes because those nouns end in /n/ even when they are not in a possessive construction (see §7.3.3.3).
Table 7.4: Effect of alienability on possessive marking of preposed possessors

<table>
<thead>
<tr>
<th></th>
<th>(NP) Pro-Pos</th>
<th>(NP) Pro</th>
<th>Number of examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part-whole</td>
<td>71%</td>
<td>29%</td>
<td>240</td>
</tr>
<tr>
<td>Kin</td>
<td>45%</td>
<td>55%</td>
<td>462</td>
</tr>
<tr>
<td>Close + Possessions</td>
<td>86%</td>
<td>14%</td>
<td>465</td>
</tr>
<tr>
<td>Overall</td>
<td>66%</td>
<td>34%</td>
<td>1,167</td>
</tr>
</tbody>
</table>

7.3.2.4 Effect of position of possessor

The position of the possessor within a phrase is the main factor determining the choice between a -k possessive clitic and the other possessive marking strategies.

A possessive clitic is required where the possessor does not have a following possessum, namely for postposed possessors (7.48), headless NPs (ita-kan in 7.49, 7.55), or predicative possessors (ema-k in 7.49, 7.56). In these positions the clitic -k is much more common than -kan, occurring in 90% of over 400 examples. Nevertheless, -kan is acceptable also (7.57). In these phrase-final contexts (and apparently only here) -kan can be phonologically stressed (7.58).15

7.55 ...Eeto ikun nia-k á n-ó ha’i racun.
   woman tail 3S-POS DEF 3S-have not poison[Mly]
   [Only Man poured wine for them.] Youngest Sister’s (drink) did not have poison.
   (Z6.74)

7.56 To’os ne’e ha’u-k.
   garden this 1S-POS
   This garden is mine.
   (K8.36)

7.57 Ha’u-kan nó lbu nia-kan manesak.
   1S-POS and mother 3S-POS same
   Mine and Ibu’s are the same.
   (D12/8)

7.58 Dadi Lakuleik Na’in mai, n-odi manu nia-kán, naran...
   so Lakuleik noble come 3S-bring bird 3S-POS name
   So the noble of Lakuleik came, bringing his cock (lit. ‘bird’) called [Falahok Loro Lakan.]
   (S2.30)

For preposed possessors, in contrast, the use of a possessive clitic is optional and the form -kan is much preferred to -k. In fact, the form -k accounts for only 2% of the over 800 preposed examples marked by -kan or -k in the corpus and is considered to be non-Fehan.

15 Mathijsen (1906:iv) lists the possessive marker as ‘kaan’ and Capell (1944:35) as ‘ka:n’. This suggests that they perceived the vowel as long (and hence stressed) as a general rule.
7.3.2.5 Effect of style

Style affects the choice of possessor marking for preposed pronominal possessors. A pronoun followed by -kan (e.g. ha'u-kan ina-n ‘1S-POS mother-GEN’ = ‘my mother’) is more likely in formal speech and almost compulsory in writing (presumably influenced by the northern Foho dialect). In contrast, the short form of a pronoun (e.g. ha ina-n ‘1S mother-GEN’ = ‘my mother’) is considered ‘friendlier’ in everyday speech.

7.3.3 Form of the possessum: genitive clitics

7.3.3.1 Distinction between -n and -r

Tetun has two genitive clitics, -n and -r, which are cliticised to the head of a vowel-final possessum NP if there is a preposed possessor and if the construction in addition satisfies other requirements (primarily relating to alienability) which are discussed in the next section. Consonant-final nouns (e.g. inur ‘nose’, hirus ‘chest’) cannot take genitive clitics, just as they cannot take consonantal suffixes.

The clitic -r is restricted to co-occurrence with plural possessors that are marked by a pronoun (7.59, 7.60, 7.66; 16 textual examples). It cannot co-occur with possessive marking on the pronoun (*sia-kan oa-r ‘3P-POS child-GEN,P’). As elsewhere with pronouns not formally marked as possessive, the short form of the pronoun is preferred over the full form.

7.59 Ami  
    o-a-r  
    é  
    sia i  
    uma am  
    dadobe.  
1PE child-GEN,P this PL LOC house 1PE pamper  
Our children—at home we pamper (them).  
(O5.152)

7.60 Mama  
    nó  
    Bapa  
    sa o-a-r  
    ulun  
    =  
    mother[Mly] and  
    father[Mly]  
    3P child-GEN,P head  
    sa  
    o-a-n  
    ulun  
3P child-GEN head  
Mama and Bapa’s oldest child  
(X0.101 elicited)

In contrast, -n is applicable with any type of preposed possessor. That is, it is applicable regardless of whether the possessor is singular (e.g. nia-kan fé-n ‘his wife’) or plural (ami-kan bei-n ‘our (exclusive) ancestor’, 7.61), and regardless of whether the relationship is expressed by direct possession (e.g. feto uma-n ‘(the) woman’s house’), contains a pronoun (e.g. ita o-a-n ‘our (inclusive) child’, 7.60), or is marked by a pronoun plus possessive clitic (7.61).

7.61 Ha’u-kan  
    ina-n  
    sia-kan  
    leo-n  
    á  
    ni  
    ne’e.  
1S-POS mother-GEN 3P-POS hamlet-GEN DEF be this  
My mother (and her people)’s hamlet is this one.  
(L0.115)

---

16 According to Capell (1944:32,35) genitive markers are used in Timorese languages spoken to the west of Tetun, but are not used in Tetun except in noun sequences, and are not used in the Austronesian languages spoken to the east of Tetun, nor in the neighbouring non-Austronesian languages. Genitive clitics are not mentioned in the brief discussions of Tetun possessives in Fernandes (1937:30f.), Hull (1996b:51f.), and Mathijsen (1906v), while Troeboes et al. (1987:78) mention only -n. Morris (1984b:xv) mentions both -n and -k but unfortunately does not illustrate the latter, for which I have no evidence in the Fehan dialect. While the lack of mention of genitive -r may reflect the brevity of the discussions, it could also reflect dialectal variation in genitive marking.
An additional difference between the two genitive clitics is that -r is used mainly by people with little or no formal education, with -n preferred by most younger secondary school educated consultants. Although one consultant considered -n to be incorrect in contexts where -r could be used (and so in complementary distribution with it), actual usage does not accord with this judgement. The fact that younger speakers prefer -n, and that -n can be used in all situations in which -r is applicable, suggests that -r is in the process of being taken over by an invariant clitic -n.\textsuperscript{17} Limited evidence from the texts suggests that the preference for -r over -n is stronger for third person plural possessors (for which about 2/3 of genitive clitics are -r) than for the other plural pronouns (for which about 1/3 are -r).

It is noteworthy that the singular/global -n and the plural -r are homophonous with the subject markers for third person singular and plural respectively. This pattern of homophony does not, however, extend to distinguishing person marking along the lines of subject marking on verbs (*ha ama-k 'my father', *ô ama-m 'your father'), as it does in the neighbouring language of Dawan.\textsuperscript{18}

### 7.3.3.2 Effect of alienability and 1S possessors

Genitive markers are obligatory for part-whole relations (*nia-kan ulu '3S-POS head' requires -n on ulu) and kin terms (nia-kan oa '3S-POS child' requires -n on oa), except where the possessor is first person singular. For first person singular possessors the genitive marker is optional (e.g. ha'u-kan bi-(n) 'my elder sister'); however, its omission is preferred, especially where the possessor is specified as h (e.g. h-oa 'my child', h-ina 'my mother'). In the corpus, close possession nearly always takes genitive marking (e.g. sia leo-(n) 'their hamlet', nia-kan rai-(n) 'his/her land'), although consultants differed in their judgments as to whether this was correct.

For relations of physical possession, in contrast, genitive marking is very rare (4 examples), with one consultant considering it unacceptable.\textsuperscript{19}

### 7.3.3.3 -n on non-possessed nouns

There is evidence that final -n is fossilised to various degrees on inalienably possessed nouns.

All location nouns (e.g. klaran 'middle', laran 'interior', kotuk 'back, behind') end in consonants, usually /n/, even when they are not in a possessive construction. Their phonology is thus consistent with an interpretation of fossilised genitive marking.

Body part terms (except those ending in non-/n/ consonants) usually, but don't necessarily, have a final /n/ even when not explicitly possessed. The exception is when they are

---

\textsuperscript{17} A similar collapse of person-number inflection to an invariant 3S -n has been proposed for the Masarete dialect of the Central Malayo-Polynesian language Buru (Grimes 1991a:283). There, however, person-number inflection is retained in other dialects.

\textsuperscript{18} In Dawan (also known as Timorese), the genitive marker on body parts inflects for both person and number of the possessor, and is identical with the subject marker for that person-number combination for all except 1S and 1PI (Middlekoop 1950:414f., 429). Galoli, spoken in East Timor, similarly has full person and number marking on the genitive (Capell 1944:36). The existence of such parallels is noted more generally by Anderson (1985:188), who says "Many languages...inflect nouns for the person and number categories of a possessor. This inflection is generally highly similar to, if not identical with, the way verbs in the same language are inflected for agreement with their subject (or whatever else they agree with).

\textsuperscript{19} Nichols (1986:77) notes that there are other languages too which use head-marking (in Tetun, genitive clitics) for inalienable possession but not for alienable possession.
themselves possessors, in which case the final /n/ is omitted (e.g. ai fuan, *ai-n fuan ‘foot/leg fruit’ = ‘toe’). Final /n/ is also omitted when the term is used as a person’s nickname (e.g. ulu ‘head’ for the oldest child in a family, and iku ‘tail’ for the youngest child). The citation forms for these body part nouns end in /n/. As one consultant expressed it, they are felt to ‘really’ (Indonesian sebenarnya) have a final /n/, even though it does not show up when the word is a possessor.20

Kin terms are frequently used with a final -n even when not possessed, unless they are used as vocative, title (Ina Funan), or possessor (such as ali in ali la’e-n ‘younger sibling husband-GEN’ = ‘younger sister’s husband), or unless they are modified by a postposed possessor (ina ha’u-k ‘mother 1S-POS’ = ‘my mother’). Apart from the constructions which either require a genitive or require its absence, it seems that the choice between a vowel-final and an /n/-final kin term is free, but that final -n is more likely if a possessor is implied in the context.21

In contrast, ‘close’ possession nouns and fully alienable physical possession nouns never have a final /n/ unless they are possessed with a preposed possessor (e.g. feto uma-n ‘woman house-GEN’ = ‘(the) woman’s house’), or inherently end in an /n/ (e.g. daun ‘needle’).

Since the apparently fossilised -n is, when it occurs on non-possessed nouns, not synchronically functioning as a genitive clitic (regardless of what it may have been diachronically), it is for convenience treated as part of the root in the interlinear examples.

7.3.4 Differences between preposed and postposed possessor

7.3.4.1 Effect of alienability

The choice between preposed and postposed possessors is influenced by the alienability of the relationship. Part-whole (including spatial) relations can only be expressed using preposed possessors (7.62; *sorin ha’u-k ‘side 1S-POS’ = ‘my side’, *kotuk uma nia-kan ‘back house 3S-POS’ = ‘the back of the house’). As is to be expected, body parts which have been dissociated from the body are no longer inalienable, and can have postposed possessors (7.63).

7.62 Tán ha’u-kan ulu-n kmeik tama iha ná kain, because 1S-POS head-GEN pointed enter LOC coconut stalk
[Help me.] Because my pointed head is stuck in (lit. ‘entered’) the coconut stalk. (K8.49)

7.63 Kawa’i nia-k á dadi bá bok kbetek oan ida... stomach 3S-POS DEF become go gourd squat small one [The fire devoured her whole body, and she died.] Her stomach became a little squat gourd...

(Z5.18)

20 Similar apparent fossilisation of 3S genitive -nl-na on terms for part-whole relations has been noted for the Central Malayo-Polynesian languages Buru (Grimes 1991b:179ff.), Larike and Tugun (Laidig 1993:342ff.).

21 Three kin terms are found as NP modifiers with an obligatory final /nl/ and a related, non-kin, meaning. These are oan ‘small’ from ooa ‘child’ (uma oan ida ‘house small one’ = ‘a small house’), inan ‘female (of animals)’ from ina ‘mother’, and aman ‘male (of animals)’ from ama ‘father’ (manu aman ‘bird male’ = ‘cock’).
There is a strong tendency to express kin relations using preposed possessors (7.64); however, postposed possessors are perfectly acceptable (7.65), making up 10% of all examples with possessed kin terms.

7.64 Nia má Bui Hirik ni ina-n nó na ama-n rona.
then Bui Hirik 3S mother-GEN and 3S father-GEN hear
Then Bui Hirik’s mother and her father heard (it).

7.65 “Ô, ama ó-k mai nó lale?” N-a’ak “Ama ha’u-k
2S father 2S-POS come also no 3S-say father 1S-POS
mai nó”.
come also
“You, did your father come too, or not?” (He) said “My father came too”.

Close possession and alienable physical possession are expressed by either preposed possessors (ha’u-kan tais ‘my cloth’, 7.66) or postposed ones (tais ha’u-k ‘my cloth’, 7.67), with postposed possessors making up a third (33%) of examples.

7.66 N-o’i tuku iha sa uma-r á.
3S-currently strike LOC 3P house-GEN.P DEF
(He) was metalworking (lit. ‘striking’) at their house.

7.67 ...rai ha’u-kan ó-k ti’an, rai ó-k bé ó-k.
earth 1S-POS 2S-POS already earth 2S-POS also 2S-POS
[The man who had lost his land in a gamble said:] my land is now already
yours, your land is also yours.

7.3.4.2 Other factors

It is not clear what determines the order of possessor and possessum when both orders are allowed. However, a number of correlations have been observed.

Postposed possessors are far more likely to co-occur with other modifiers (often with definite determiners) than preposed possessors are. Preliminary figures indicate that about 60% of postposed possessors co-occur with other modifiers, while only about 25% of preposed possessors do. This may be because NPs with preposed possessors are already specific (as noted for English by de Groot (1983:108)), and so need no anaphoric marker of definiteness.

Long possessors are somewhat more likely to be preposed than are short ones. In particular, nominal possessors are more likely to be preposed than are pronominal ones. Based on counts of possessor NPs which contain a possessive clitic -k or -kan, 25% of preposed possessors are headed by nouns, while only 10% of postposed possessors are. Furthermore, none of the 26 postposed nominal possessors in the corpus have modifiers within the possessor NP, suggesting that short NPs are strongly favoured in the postposed possessor slot.

7.3.5 Adjectival ‘possessors’

There is a construction in which an adjective or adverb acts semantically as a modifier within a clause or NP, but looks syntactically like a possessor, in that it is followed by nia-kan ‘3S-POS’ (16 examples). The possessor marking appears to signal an opposition,
emphasising that it is the preceding description which is valid, and not some other (possibly weaker) description. Note that, as elsewhere when it is phrase-final (§7.3.2.4), the -kan is optionally stressed (7.70).

7.68 \( \text{Wé n-akali tebe-tebes nia-kan.} \)
water 3S-boil/simmer RDP-true 3S-POS
The water is \textit{really} boiling (not just simmering). (I0.100)

7.69 \( \text{Lia fuan mak kwana nia-kan ní ne’e.} \)
word fruit REL right 3S-POS be this
The words which are correct (lit. ‘right-hand-side’) are here (in Fehan, as opposed to Foho). (H3/6)

7.70 \( \text{na’in tebe-tebes nia-kán} \)
noble RDP-true 3S-POS
really truly nobles (i.e. having the powers and bearing that nobles should have, as opposed to just being noble by birth) (V0.81)

7.4 Premodifiers

7.4.1 Introduction

Apart from possessors, there are only three types of premodifiers of NPs. These are \textit{ohin} ‘aforementioned’, \textit{sura} ‘every’ and a restricted type of relative clause.\textsuperscript{22}

7.4.2 \textit{ohin} ‘aforementioned’

The most common use of \textit{ohin} is as a time noun meaning ‘the very recent past, just now; today’. It also occurs as premodifier of NPs, with the related meaning of ‘aforementioned’. It is thus used as a participant tracking device (comparable to Indonesian \textit{tadi} ‘a moment ago’).\textsuperscript{23}

7.71 \( \ldots n\text{-afula ohin feto na’in tolu ne’e} \ldots \)
3S-spy.on just.now female CLS:human three this
[The shaman] spied on the aforementioned three women...
(G2.21)

7.4.3 \textit{sura} ‘every’

The premodifier \textit{sura} ‘every’ introduces either location NPs (7.72; 11 examples) or peripheral time NPs (7.73; 22 examples). It cannot postmodify NPs (‘\textit{fulan} \textit{sura} ‘month

\textsuperscript{22} Mathiesen (1906:ii) analyses a few additional words as premodifers. These are \textit{seluk} ‘other’, which I analyse as modifier of the preceding verb (\textit{buka seluk \textit{wé} ‘seek other water’}), and \textit{naran} ‘name’ which I analyse as a noun functioning as subject of a clausal predicate (§9.6).

\textsuperscript{23} This premodifying use of \textit{ohin} is relatively uncommon, with most of the 16 unelicited examples coming from a speaker who has spent much of his life amongst Suai speakers. Some speakers also use ohin \textit{é} in pre-NP position as some sort of ‘filler’ when storytelling (e.g. \textit{tōo ohin é dalan à ‘reach just.now HES road DEF} = ‘get to the road’). To what extent these two uses are related, or overlap, is not clear.

Time nouns (§3.3) such as \textit{ohin} are in most grammars analysed as adverbs. Adverbs which modify nouns are crosslinguistically rare, and fall outside the common definition of adverbs as “modifiers of constituents other than nouns” (Schachter 1985:20). Sasse (1993:664) notes that in languages which do allow adverbial modifiers of nouns, they “are limited to certain types (local and temporal) and are always distinguished from adjectival modifiers by their distributional characteristics”.

---

\( \text{The noun phrase} \quad 153 \)
every'). *Sura* is also a transitive verb meaning 'count' (e.g. *sura loit* 'count money'; 25 examples).24

7.72  *Kowá* mulai *semo, semo, sura* leó.
crow start[Mly] fly fly every hamlet
The crow started to fly, and flew, to each hamlet in turn.  (O4.20)

7.73  *Sura* kalan, *nia* tanis.
every night 3S cry
Every night, she cried.  (Z5.84)

Just as units of time follow each other, and items being counted are counted one at a time, so when *sura* is used for location (e.g. *sura uma* 'every house'), it can only be used when the locations are processed one at a time. For instance, *la'o sura rai* 'walk/go every earth' means 'travel to every place, one at a time'. Compare this with *hare sura uma* 'see every house', which is unacceptable because one can see all the houses at the same time.

7.4.4 Premodifying relative clause

There is only one head that takes premodifying clauses, namely *fatik* (sometimes *fatin*) 'place' (39 examples). Such NPs refer to the place where the action specified by the clause is (in most cases habitually) done. Premodifying clauses are syntactically restricted to 'small clauses' (§9.7.3), consisting of a single verb (*toba* *fatik* 'sleep place' = 'bed'; *haris* *fatik* 'bathe place'), a subject plus verb (*ro* *semo* *túr* *fatik* 'boat fly (= aeroplane) sit place' = 'airport') or a verb plus object (*jasi* *dai* *fatik* 'wash net place' = 'place where nets are washed').25

7.74  *Rai* né *rai* futu *manu* sia *fatik*.
earth this earth cockfight bird PL place
PremCl

This area is a cockfighting place.  (G1.110)

Premodifying clauses are non-finite, with no subject marking or adverbial modification.26 They are always restrictive. The resulting NP has the character of a compound, in that it often represents a conventional name, and allows no intervening modifiers. Nevertheless, there is no phonological evidence of compounding.

---

24 It seems likely that the premodifying position of *sura* 'every' reflects its diachronic relationship with the verb 'count'.

25 Premodifying clauses are relatively uncommon in prepositional languages (Keenan 1985:144; Mallinson & Blake 1981:285). The restriction that premodifying clauses be short is consistent with Hawkins's (1994) theory that constituent order is determined by processing constraints. A long premodifying clause would delay for a long time recognition of the fact that this is part of an NP, and so would be avoided.

26 Subject marking on the verb was accepted during elicitation (*Nia* n-ô *n-arls* *fatik* '3S 3S-have 3S-bathe place' = 'He has a bathing place.'). Similarly adverbial modification by *horisehik* 'yesterday' was accepted (*horisehik haris* *bá* *fatik* 'yesterday bathe at place' = 'place where (we) bathed yesterday' ), although the addition of postverbal *ti'an* 'already' was rejected. Nevertheless, I suspect these are forced constructions. The apparent limitations on subject marking and tense-aspect marking, as well as the absence of a relativiser, are common crosslinguistically for prenominal relative clauses (Keenan 1985:160).
7.5 Postmodifiers

7.5.1 Noun phrase

Nominal modifiers bear a wide range of semantic relations to the head. For convenience these modifiers are illustrated here under the set of semantic relations defined by Levi (1978:75ff.) for English 'complex nominals'. Where the modifier consists of more than one word, the whole is underlined, illustrating the fact that the modifier can itself be an NP, rather than just a single noun. The most common semantic relations are BE (including BE LIKE), HAVE, IN and FROM.

BE (essive/appositional)
- *oa ulun* 'child head' = 'oldest child'
- *ema to'os na'in* 'person garden noble/owner' = 'garden owner'
- *ai kamelin* 'wood/plant sandalwood' = 'sandalwood'
- *na'an knase* 'fish/meat mugil' = 'mugil fish'
- *feto Sikori* 'woman Sikori' = '(the) woman Sikori'

BE LIKE (essive/appositional; Levi (1978:107ff.) includes this under BE)
- *ai tebok* 'plant bowl' = 'k.o. water weed with bowl-shaped leaves'
- *rai manu matan* 'earth bird eye' = 'earth the size of a bird's eye'

HAVE (possessive/dative)
- *ai lolo mutin* 'plant trunk white' = 'k.o. bush with white stem'
- *ai funa meak* 'plant flower reddish' = 'k.o. lantana bush with orange flowers'
- *batar ai naruk* 'maize leg long' = 'sorghum'

IN (locative)
- *ai tasi* 'plant sea' = 'k.o. mangrove'
- *manu tasi* 'bird sea' = 'k.o. sea bird'
- *ai tahan rai né* 'plant leaf earth this' = 'leaves in this district'

FROM (source/ablative)
- *kabau malae* 'buffalo Malay/foreign' = 'horse'
- *loit Balanda* 'money Dutch/European' = 'Dutch money (i.e. old Dutch coins)'
- *huttun rai tolu* 'populace earth three' = 'populace from (the) three districts'
- *tebok ai* 'bowl wood' = 'wooden bowl'
- *ema loro-sa' e-n* 'person sun-ascend-GEN' = 'people from the east'
- *mina fahi* 'oil pig' = 'oil produced from pig fat'

ABOUT (topic)
- *ai.kanoik* 'Manu Kowa' 'story/folktale ' (The) Crow'
- *lia Ai Lotuk Laran* 'story (lit. 'word') (about) Ai Lotuk Laran'

USE (instrumental)
- *daun liman* 'needle hand' = 'needle for hand-sewing (not machine-sewing)'
- *ró ha'i* 'boat fire' = 'large ocean-going boat'

CAUSE (causative)
- *ai.kanoik laran moras* 'story/folktale interior sick' = 'sad story'
- *ai.kanoik surik bo'as* 'story/folktale side burst' = 'joke'
MAKE (productive)
  *fatuk ahu* 'rock lime' = 'limestone (k.o. rock used to make lime)'

FOR (purposive/benefactive)
  *tambak boek* 'pond[Mly] shrimp' = 'pond for raising shrimp'
  *tai s feto* 'handwoven.cloth woman' = 'handwoven cloth of the type worn by women'

### 7.5.2 Adjective phrase

Most adjective phrases functioning as noun modifiers consist of a single, unmodified, adjective. However, the following are also attested in the corpus.

\[
\text{AdjP.NP}_{\text{modifier}} \rightarrow \begin{cases} 
\text{Adj (Intensifier)} \\
\text{(mesa) Adj}
\end{cases}
\]

7.5.3 Relative clause

The internal structure of relative clauses is discussed in §14.8, and their order relative to other constituents in §7.2.2.4. Here we only point out that NPs modified by relative clauses can, like shorter NPs, occupy a wide range of positions within the matrix clause. These positions include subject, object (7.25), fronted object (7.77), indirect object, locative and topic.

7.77 *Ho tu, fatuk [mak ohin dadi] ne’e sia, [mak ohin cetak] ne’e sia, sia hawai bá loro.*
  finish rock REL just.now happen this PL REL just.now mould[Mly] this PL 3P dry(s.th.) at sun
  Then, the bricks which have just come into being, which (the workmen) have just cast, they (the workmen) dry in the sun. (AA3.33)

7.6 Complex NPs

7.6.1 Coordinate noun phrase

7.6.1.1 Regular coordination

NPs can be coordinated using the same means as coordination of other constituents such as clauses. Conjunction is expressed by listing of the conjuncts, or by means of *nó* 'and' before the final (or before all non-initial) conjuncts. Where subject NPs referring to humans are conjoined, however, the comitative verb *hó* 'accompany' is used more frequently than the
conjunction nó. Disjunction is expressed by ká ‘or’, or sometimes by lale ‘else’. Open disjunction, in which the disjuncts listed are not exhaustive of the possibilities, is expressed by ká ‘or’ after every disjunct, including the final one. These options are further discussed in §14.3 and illustrated there.

7.6.1.2 Dual pronoun construction

There is in addition a means of expressing conjunction which appears to be restricted to NPs that refer to two people.

NP.Coordinate → (NP1) Plural.Pronoun ruas NP2

This construction minimally consists of a plural personal pronoun followed by ruas ‘two’ and then an NP identifying one of the referents of the pronoun. The other referent is optionally specified before the pronoun (7.78; 17 examples).

\[
\text{7.78} \quad \text{Fu'a a ina-n sia ruas a ama-n bá.} \\
\text{waken} \quad 2\text{S mother-GEN} \quad 3\text{P two} \quad 2\text{S father-GEN} \quad \text{IMP}
\]

NP1 Pro ruas NP2

Wake up the two of them, your mother and your father. \((\text{Z4.122})\)

Where one referent is unspecified, limited evidence suggests that it is always the speaker if the pronoun is ami ‘1PE’ (7.79; 5 examples). For the pronoun sia ‘3P’ an unspecified referent need only be a salient one, such as the participant who is currently in focus (7.80; 4 examples). Since English lacks this construction, the translations for the examples in this section use apposition instead.

\[
\text{7.79} \quad \text{Ami ruas Pák Tóm bá Laran...} \\
\text{1PE two Mr[Mly] Tom go Laran}
\]

Pro ruas NP2

We two, Mr Tom (and I), went to Laran... \((\text{Q0.141})\)

\[
\text{7.80} \quad \text{Nia ti'a, si ruas na belu-n oan né} \\
\text{3S already 3P two 3S friend-GEN small this}
\]

Pro ruas NP2

\text{\quad r-asán...} \\
\text{3P-carry.on.shoulder}

Then the two of them, (the nobleman) and this his friend, carried [the woman home.] \((\text{R5.139})\)

Rarely (2 examples) a classifier plus numeral (na 'in rua) is used in place of the adjective ruas ‘two’.

\[
\text{7.81} \quad \text{Ami na'ìn rua Ibu há ti'an.} \\
\text{1PI CLS:human two mother eat already}
\]

We two, Ibu (and I), have eaten. \((\text{H6/6})\)

---

27 Schwartz (1988) reviews this construction from a crosslinguistic perspective, calling it a ‘plural pronoun construction’. She says it consists of a non-singular pronoun followed by an NP specifying one of the referents of this pronoun. If a language has dual pronouns, it uses dual rather than plural pronouns. Schwartz does not mention the use of a word meaning ‘two’ to force a dual interpretation, as in Tetun. Nor does she mention the possibility of a preceding NP specifying one of the referents. These two Tetun features are, however, also found in the Austronesian language Lewo, spoken on Vanuatu (Early 1994a:224).
7.6.2 Apposition of noun phrases

Apposition links two NPs which share a grammatical role, and which also share reference (at least to some extent). The second NP elaborates on the information given in the first by specifying the class of referents (*ita feto* 'we women'), a proper noun (*nia maun Suri Tuan* 'his elder brother Suri Tuan'), or some other description. Intonationally the construction ranges from falling under a single intonation contour to having a clear pause and 'comma' intonation between the two NPs.

7.82 \[ Fukun ida mama n-odi kose lawarik sia \]
elder one chew(usu.betel) 3S-COORD rub small.child PL

* feto mane.
girl boy

An elder chews betel and (uses it to) rub on the children — girls and boys.

(B2.36)

7.83 \[ Manu ne'e, manu kowá ne'e, rani iha hudi tahan. \]

bird this bird crow this perch LOC banana leaf

NP1 NP2

This bird — this crow — perched on the banana leaf.

(K12.57)

7.84 \[ Nó samea chin naruk bót ne'e, samea likusaen ne'e. \]

also snake just.now long big this snake python this

NP1 NP2

(They) also (put in) the aforementioned long thick snake, the python.

(K10.96)
8 Prepositions and the prepositional phrase

8.1 The prepositional phrase

A prepositional phrase consists minimally of a preposition followed by its complement NP. If the preposition is a locative one, such as general locative iha, a deictic particle (e.g. distal bá, proximal mai; §11.8) may follow the complement (8.4).

PP.Non-predicative → Preposition NP (Deictic)

All prepositional phrases can function as peripheral constituents within a clause (8.1, 8.10). Some prepositions can also head phrases functioning as modifiers within NPs (§7.2.2.1), as oblique arguments (particularly locative iha; 8.4), and as non-verbal predicates.

Prepositional phrases functioning as predicates accept adverbial modifiers which are not found in prepositional phrases in other functions. The justification for nevertheless analysing such predicates as non-verbal is discussed in §8.7.

8.2 Overview of prepositions

The defining characteristics of prepositions are discussed in §3.10.

Semantically most of the prepositions can be grouped into prepositions of location, time and similarity. The following table lists all prepositions in this semantic order, followed by those which don’t fit into these categories.

Some function words can introduce either NPs or clauses, and so are analysed as belonging to both the class of prepositions and the class of conjunctions. Since the meanings are the same regardless of complement type, both prepositional and conjunction uses will be discussed together. The column in the table marked ‘Cnj. too?’ shows whether the word is a conjunction as well as a preposition.

Remaining columns in the table list other word classes and glosses for the form in question (regardless of whether these appear to be semantically related to the preposition or not), whether the preposition is, in the corpus, found to head a predicate (as opposed to solely introducing peripheral phrases), and the section in which the preposition is discussed, if this is in another chapter. The various meanings of bá are discussed in the chapter on serial verbs on account of their grammaticisation from the verb ‘go’, while hosí as a conjunction is dealt with in the chapter on complementation because it only introduces complements.
In addition to the above prepositions, there is a complex preposition tán bá ‘because of’, which appears to be marginal to the Fehan dialect (§14.7.5.2).

8.3 Prepositions of location

8.3.1 iha ‘at (location)’

The locative preposition iha (or i, ia) is both very common (1,200 examples) and general in its meaning, being translatable according to context as ‘in, at, on, from, to’. It is thus assigned the general gloss ‘LOC’.

More exact location can be indicated by an inalienably possessed noun. As the list below shows, a number of these locative terms are also body part terms.1 There are several pairs of apparent synonyms.

---

1 The use of the same terms for body parts and locatives is of course common, as demonstrated by Bowden (1992) for Oceanic languages.
Table 8.2: Location nouns

<table>
<thead>
<tr>
<th>Word</th>
<th>Location meaning</th>
<th>Body part meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>oin</td>
<td>in front</td>
<td>face</td>
</tr>
<tr>
<td>kotuk</td>
<td>behind</td>
<td>back</td>
</tr>
<tr>
<td>sorin</td>
<td>beside</td>
<td>side</td>
</tr>
<tr>
<td>kbelan</td>
<td>beside</td>
<td>side</td>
</tr>
<tr>
<td>leten</td>
<td>at the top of, on top of</td>
<td></td>
</tr>
<tr>
<td>fohon</td>
<td>on top of</td>
<td></td>
</tr>
<tr>
<td>kidun</td>
<td>underneath, at the bottom of</td>
<td>buttocks, bottom</td>
</tr>
<tr>
<td>ohak</td>
<td>underneath, at the bottom of</td>
<td></td>
</tr>
<tr>
<td>laran</td>
<td>interior</td>
<td>character, emotion</td>
</tr>
<tr>
<td>klaran</td>
<td>middle</td>
<td></td>
</tr>
<tr>
<td>molin</td>
<td>outside</td>
<td></td>
</tr>
</tbody>
</table>

8.1 Nia ti’a Bapa bá n-usu ha’u iha ke’an laran
3S already father[Mly] go 3S-request 1S LOC room interior
á n-ák “...”
DEF 3S-say
Then Father went and asked me in the room, saying “...”. (F3.58)

8.2 Ai bilibala n-aklele iha we fohon.
plant k.o.water,plant 3S-float LOC water top
The bilibala plant floats on top of the water. (T0.139)

8.3 ...n-oku ti iha (é) uma kidun á, nia n-anono lai.
3S-lie already LOC HES house bottom DEF 3S 3S-listen
first
[Then the crocodile mother went out, and] having squatted under the house, she listened first (before departing). (Z4.38)

These inalienably possessed location nouns are not restricted to complements of iha, being found, for example, also in complements of the preposition hori ‘from’, in objects of verbs of motion (e.g. tama uma laran ‘enter house interior’ = ‘enter the house’), in NP modifiers (e.g. naha uma laran ‘baggage house interior’ = ‘household goods’), and as subject NPs (e.g. Kbonan laran é n-ó na’an ‘sarong interior this 3S-have fish/meat’ = ‘This sarong contains fish’). There is thus no reason to analyse these location terms as postpositions (or as forming complex prepositions with iha), as is the case in analyses of some other Austronesian languages (Bowden 1992:45).

The preposition iha frequently introduces a peripheral location (8.1). It also commonly introduces oblique locative arguments for motion verbs (8.4, 8.6) and transfer verbs (8.5), for which it in most cases introduces the goal of the motion (8.4, 8.5). However, for some verbs, such as hola ‘take, fetch’ and sai ‘exit’, it introduces the source location instead (8.6).

8.4 Mais ha’u la k-bá iha Suwai bá.
but 1S not 1S-go LOC Suai go
But I am not going over to Suai. (F3.21)
8.5 Ha’u-kan hudi nò ha’u-kan na’an ó rai iha nabez?
1S-POS banana and 1S-POS meat 2S lay.down LOC where
My bananas and my meat, where have you put (them)? (K12.29)

8.6 Sai i sekola tân kabuk.
exit LOC school[Mly] because pregnant
(Many girls) have left school because (they are) pregnant. (O5.202)

In addition to being a preposition, iha is an intransitive verb predicating presence or existence (§9.4.7.2). There is a close semantic relationship between the two, with the intransitive verb predicating presence at an unspecified location, and the preposition predicing presence at a specified location. This relationship is clearly shown in the following elicited example. The reasons for analysing transitive iha as a preposition rather than as a transitive verb in such examples are discussed in §8.7.

8.7 “Ama ó-k iha uma?”
father 2S-POS LOC house
“Is your father at home?”

“La iha.” = “Ama ha’u-k la iha uma.”
not be.present father 1S-POS not LOC house
“(He’s) not present.” = “My father is not at home.” (T0.101 elicited)

8.3.2 ré ‘at (location)’

The word ré ‘at’ introduces a location (6 examples). It may either head a clause predicate (8.8) or follow a verb such as hare ‘look’. In five of the six examples the complement is ne’e ‘this’, with ré ne’e meaning ‘here’. Data are insufficient to determine whether this is indeed a preposition or instead a transitive verb.

8.8 Nia ré ne’e.
3S at this
He is here. (Z5.98)

8.3.3 hori ‘from (location)’

As a locative preposition, hori ‘from’ is restricted to movement upwards and downwards (e.g. tûn hori leten ‘descend from above’). It is seldom used (3 textual and 6 elicited examples), and appears to be always replaceable by the semantically more general verb hosî ‘(come) from’ (§12.5.2.2). The following extracts from a keen attempt to convince me that the speaker’s ancestors came from America illustrate this, as the speaker spontaneously repeated the claim with hori twice, and later used a clause with hosî to make the same point.

8.9 Sa’e hori Amerika n-akur tasi wë-n kran hitu...
ascend from America 3-cross sea water-GEN CLS:layer seven
Sa’e hosî Amerika.
ascend from America
(Two ancestors) came up from America and crossed seven seas (to come to Timor)... (They) came up from America. (V0.97)
8.4 Prepositions of time

8.4.1 hori ‘since (time)’

In addition to being a locative preposition, hori is a temporal preposition (8.10; 9 examples) and conjunction (8.11; 2 examples) meaning ‘since’. In this capacity it introduces a constituent which in some way indicates time. As such it is found in the common expression hori bei sia ‘since ancestor PL’ = ‘since the time of our ancestors’.

8.10 Tetun n-ola ti’an hori uluk.
Tetun 3S-take already since former.times
Tetun has taken (this borrowed word kalák) a long time ago. (I4/7)

8.11 Ai kanoik né hori rai moris n-ó kedas.
story this since earth live 3-exist immediately
This story has existed since the earth came into being. (V0.15)

It also takes part in a number of fixed past-time expressions (e.g. hori.hirak ‘when (past), recently’), which are listed in §4.5.3. It appears to be reduplicated in hori.horin ‘quite a long time ago’.

8.4.2 to’o ‘until’

The form to’o ‘until’ is both a preposition (8.12, 8.13; 108 examples) and a conjunction (8.14, 8.15; 130 examples). It is presumably diachronically related to the verb tu’o ‘reach, arrive, suffice’. When it introduces time NPs or clauses, however, its non-verbal status is shown by its lack of a subject, the fact that it requires a complement, and the fact that it can introduce either an initial (8.13) or final (8.14) peripheral constituent.

8.12 Toba to’o sasawan, sasawan bá n-i’ka.
lie.down until morning morning go 3S-back
(They) lay until morning; in the morning (she) went back. (Z4.43)

8.13 Nia ti’an, to’o loron ida, nia loke.
3S already until day one 3S open
Then, one day, he opened (it). (R6.3)

8.14 Nia manán ni belu-n daudaun to’o ni belu-n
3S defeat 3S friend-GEN continue until 3S friend-GEN
osan la n-ó.
money not 3S-have
She kept defeating her friend (in gambling) until her friend had no money. (G4.43)

8.15 Kaban á turu kona karas á, to’o kawa’ik ti’a,
saliva DEF drip touch chest DEF until older already
karas n-ó rahun.
chest 3S-have hair
(If a baby’s) spittle drips onto the chest, then when (the child) is older (its)
chest has hair. (V0.83)
Usually the time introduced by *to'o* follows the main proposition, and presents a state, event or time until which that which is specified in the previous clause continues (8.12). The aspect of continuation can be emphasised by *daudaun* 'continue' (8.14). Less commonly, the constituent introduced by *to'o* gives a time setting for the following clause (8.13, 8.15).

### 8.4.3 natón bá, bá natón ‘at (time)’

The word *natón* is invariably in sequence with either a following (8.16) or a preceding (8.17) *bá*, with both resulting sequences meaning 'at (a specified time)'. These sequences are, in absence of further evidence, analysed as complex prepositions when they introduce NPs (8.16, 8.17), and complex conjunctions when they introduce clauses (8.18). *Natón* is quite uncommon, with only seven unelicited examples (all of prepositional *natón bá*) in the corpus. Examples 8.17 and 8.18 are from direct elicitation.

#### 8.4.3.1 natón bá

8.16 *Natón bá* loron ida, Beur mai bosok nia-kan ina-n.
One day, Trickster came and deceived his mother.

(8.16)

8.17 Ou, bele. Mais keta bá natón nia ha’u la iha.
Yes can but perhaps at (time) 3S 1S not be present
(8.17)

8.18 *Natón bá* ha’u sei ki’ik, ha’u la’o kód ha’u-kan
When I was still small, I walked with my mother and my father.

(8.18)

---

### 8.5 Prepositions of similarity

#### 8.5.1 nú, nu’u ‘like, just like’

The preposition *nú* (which native speakers write, but seldom pronounce, as *nu’u*) ‘like, just like’ (about 300 examples) indicates likeness in some respect, ranging from close physical likeness (8.19, 8.27) to likeness in action (8.20) or extent (8.24). Frequently the complement is a demonstrative pro-clause *ne’e* 'this' or *nia* 'that, 3S', in which case the comparison is with either a demonstrated action (8.20) or an event which has already been described (8.21) or is about to be talked about (8.22).

#### 8.5.1.1 nú

8.19 *Nia ikun ulun la hatene; ulun nú ikun, ikun nú ulun.*
(Concerning an *ai kaliruk* ‘throwing stick’ snake:) (We) can't tell its head from its tail; the head is like the tail, and the tail like the head.

(8.19)

8.20 *Tuir adat Wesei Wehali né, ita la haré nú ne’e.*
According to the tradition of this Wesei Wehali, we do not look (at the money in the basket) like this. (The speaker demonstrated by opening the lid of the basket and looking in.)
Prepositions and the prepositional phrase

8.21 Feto nia-k n-akotu, mane. Mane nia-k n-akotu, feto.  
woman 3S-POS 3S-finish man man 3S-POS 3S-finish woman

Simu malu nú nia.  
reply each other like 3S

(When) the women's (song) is finished, the men (sing). (When) the men's is finished, the women (sing). (They) answer each other like that.  
(J3.7)

8.22 Ai kanoik Mesak Oan ne'e nú ne'e Ibu. Mesak Oan  
story sole child this like this mother sole child

ne'e, nia...

this 3S

The story of this Only Child is like this, Ibu: This Only Child, he... (The speaker then tells the story.)  
(Z6.1)

This preposition is found in a number of standard expressions, including nú nabe (lit. 'like which/where'), a general question which frequently translates as 'how' (8.23). The expressions nú ne’e dei (lit. 'like this only') and nú nia dei (lit. 'like that/3S only') mean 'only so much, not very much' (8.24), and often signal the end of a monologue. As connectors in discourse, kalo nú nia (lit. 'if like that') translates as 'in that case', and masik nú nia (lit. 'despite like that') as 'nevertheless'.

8.23 N-akes n-ák "Nú nabe Iku? It at dadi ká lale?"  
3S-talk 3S-say like which tail 1PI IRR happen or no
(He) spoke saying, "How is it Iku? Will it work out for us or not?" (i.e. Will you accept my proposal of marriage?)  
(F3.18)

8.24 Nia soi nú nia dei.  
3S rich like 3S only
He is only somewhat rich (i.e. neither rich nor poor).  
(I0.159)

8.5.2 nudar 'like, be as if'

The preposition nudar (or nu'udar) 'like, be as if' introduces a simile, comparing one entity or situation with another which is in some respect similar. Unlike nú and hanesan, it is not used to express overall likeness or close physical resemblance. It is both a preposition (8.25; 26 examples) and a conjunction (8.26; 7 examples).

8.25 Karian nudar ata, há nudar na'in.  
work as if slave eat as if noble
(If you) work like a slave, (you will) eat like a noble.  
(D0.153)

8.26 Nanti á:isin nia n-etan susar iha na'i Maromak  
later [Mly] once 3S 3S-get difficult LOC noble God
nia-kan oin, nudar lilin n-a-wé.  
3S-POS face as if candle 3S-make-water
Later he will get into difficulty before Lord God, just like candles melt (before a fire).  
(C3.24)
8.5.3 hanesan ‘like, resemble’

The preposition hanesan ‘like, resemble’ introduces either an NP (8.27; 11 examples), or else a prepositional phrase introduced by the semantically overlapping prepositions nú ‘like’ (8.28; 12 examples) or nudar ‘like, be as if’ (8.29; 4 examples). It indicates resemblance in some respect, such as appearance (8.27) or behaviour (8.28), and is semantically close to nú, as illustrated by the spontaneous paraphrase in 8.27.²

8.27 Nia-kan katarak ne’e nú kawak larit, hanesan kawak 3S-POS serrated this like k.o.bird comb resemble k.o.bird
nia-kan larit ne’e. 3S-POS comb this
Its serrations are like a kawak bird comb—resemble this kawak bird’s comb. (U8.10)

8.28 Nia mak labu iha tasi we-n klaran hanesan nú na’an. 3S REL go.out LOC sea water-GEN middle resemble like fish
It was he who wandered around in the sea just like a fish. (P1.8)

8.29 ...Sia hanesan nudar anin. 3P resemble as.if wind
[The spirits of the dead are invisible.] They are like the wind. (E0.143)

8.6 Other prepositions

8.6.1 bodik ‘for (beneficiary); as for’

The preposition bodik ‘for, on behalf of’ can introduce a benefactive NP (8.30; 13 examples), in which capacity it appears to be synonymous with the more commonly used preposition bá (§12.5.3.6), with which it can occur in sequence (8.31). Alternatively, bodik can be a conjunction, with the complement clause specifying the purpose of the action for the beneficiary (8.32; 6 examples).

8.30 Bele n-ó rai maran oan ida iha tasi we-n can 3-exist earth dry small one LOC sea water-GEN
ne’e bodik ha’u, bele kâ lale? this for IS can or no
(Request from the first sea-dwelling man to Maromak:) Can there be a little bit of dry land in the sea for me; is that possible or not? (P1.22)

8.31 Ina soru tais bodik (bá) oan sia. mother weave cloth for for child PL
Mother weaves a cloth for the children. (D0.11 elicited)

² This preposition is homophonous with one variant of the causative verb hanesa(n) ‘to put in order, to make alike’. While the preposition is clearly related semantically and formally to both the verb and the adjectival root nesan ‘same’, it is unclear what the function of the normally causative prefix ha- is in the preposition. The non-verbal status of the preposition is reinforced by its inability to take subject marking.
8.32 Dadi Maromak n-alo fatin ida, bodik sia na’in
so God 3S-make place one for 3P CLS:human
rua tür bá.
two sit at
So Maromak made a place for the two of them to live in. (P1.60)

The word bodik extends to a more general meaning of ‘for, with regard to’ (8.33; 9 postverbal examples), and as such can also introduce a clause-initial topic (8.34; 7 examples).

8.33 Ne’e bodik mane dei, ‘ema klabuk’.
this for man only person gadabout
This is (applicable) only for males, (the term) ‘gadabout’. (Q0.174)

8.34 Mais bodik uma N... ne’e, mako’an sé sé dei,
but for house N... this historian who who just
Ibu n-usu...
mother 3S-ask
But regarding this house of N...: for any keeper of the myths, if you (‘Ibu’) ask (them) [, they would agree with what I have told you.] (P4.42)

8.6.2 hetak ‘as for’

The word hetak ‘as for, with regard to’ (4 examples) introduces an initial topic NP. There are insufficient data to determine whether hetak is indeed a preposition (like bodik, with which it overlaps semantically) or is instead a topic marker.

8.35 Hetak ami, ha-tama naran bat sekola i ne’e,
as for 1PE make-enter name so.that school LOC this
r-odi ami bá subar i Bolan bá.
3P-bring 1PI go hide LOC Bolan go
As for us, (when teachers) registered names for (children to start) school here, (our families) took us and hid us over in Bolan. (O5.148)

8.7 The prepositional phrase as predicate

When it heads a predicate a preposition can, like verbs, be modified by adverbs and auxiliaries. Certain adverbs precede the preposition while others follow it, with the position of each adverb relative to the preposition being the same as that relative to verbs.

PP.Predicative → (Adverb) (Auxiliary) Preposition (Adverb) NP (Deictic)

Postmodifying adverbs include kedan ‘immediately, already’ (8.36), hika(r) ‘back, again’ (8.37), kbesik ‘directly’, ta ‘already’ and ha’i ‘not’ (8.39), while la ‘not’ (8.7) is an example of a premodifiier. Auxiliaries which can precede prepositions include keta ‘do not’ (8.38), sei ‘still’, bele ‘can’ and irrealis atu.

8.36 Moris mai nú kedan nia.
live come like immediately 3S
(When he) was born (he) was already like that. (i.e. He was like that from birth.) (N0.28)
8.37  \( Oin \, nu \, n{\text{-}}ikar \, o \, dei. \)  
face like 3S-back 2S only  
(Her) face was just like yours.  

8.38  \( Ema \, ferik \, t\ur, \, keta \, iha \, oin. \)  
person mature.woman sit do.not LOC face  
(When) women are sitting (you children) must not be out in front.  

The only example in the corpus in which a non-predicative prepositional phrase contains a modifier other than a deictic particle is 8.39, where the phrase is a peripheral locative.

8.39  \( ...iha \, ha'{i} \, mota \, m\os \, 3-exist \, only \, LOC \, stream \, also \)  
\( n{-}\o \, dei. \)  
3-exist only  
[Does fafoek grass grow in the river?] Not at the river it exists, at the river it also exists. (i.e. It grows both away from and at the river.)  

When a prepositional phrase is predicative, the preposition looks rather like a transitive verb followed by an object NP. However, unlike transitive verbs, none of the predicative prepositions take subject marking, and none allow fronting of the complement NP.

The most verb-like of the predicative prepositions is locative \( iha, \) which is also an intransitive verb meaning 'be present' (§9.4.7.2), and which is, regardless of its part of speech, ineligible for subject marking by virtue of its vowel-initial phonology (§9.3.1). Predicative transitive \( iha \) could thus be analysed either as a preposition (aligning it with \( iha \) in non-predicative prepositional phrases) or as a transitive verb (aligning it with intransitive predicative \( iha \)). Unlike \( iha, \) the remainder of the prepositions do not have a corresponding intransitive form.

Since there are no factors compelling a verbal analysis of these words when they head predicates, and since they are all clearly prepositions when they are non-predicative, by Occam's razor the words under discussion are classed as prepositions (rather than as transitive verbs or prepositional verbs) even when they function predicatively.
9 The clause

9.1 Introduction

This chapter deals with the structure of the clause. A complete independent clause consists minimally of a ‘clause core’. The core in turn consists of the predicator plus any arguments which are included in the valency of that predicator, namely subject, object and oblique arguments.

The overall structure of the clause is presented in the formula below, where only the (underlined) clause core is obligatory. Formulae for the various types of clause core are given throughout this chapter, where they are labelled simply ‘clause’, with a subscript identifying the clause core type. The ‘topicalised object’ slot is for fronted objects of transitive clauses (§9.4.3.3). Note that the formula is simplified in that some peripheral elements can occur within the clause core.

\[
\text{Clause} \rightarrow \text{(Periphery)} \text{(Topicalised.object)} \text{Clause.core} \text{(Periphery)}
\]

The following section deals with properties of subjects and objects. Under certain circumstances verbs show person-number marking for the subject, a matter which is dealt with next.

It is the predicate type which determines the type of the clause. The predicate is the part of the clause which is obligatory, and which is retained in reduced complement clauses. All types of predicate follow the subject, except in some existential clauses. In terms of their internal structure, predicates fall into three overall categories. In the first, the predicate has a verbal head. In the second, the predicate consists of some non-verbal constituent, namely a prepositional phrase, numeral phrase, noun phrase, possessor or relative clause. The third category of predicate is that of ‘body-good’ expressions, in which the predicate consists of a clause which is itself composed of a subject and predicate. Justification of this analysis of body-good expressions is given in §9.6.4.

While all the clause types mentioned above can occur as main clauses, there are also some types which are restricted to dependent contexts. These are outlined in §9.7.

Discussion of complex predicates resulting from verb serialisation is deferred until Chapter 12, while auxiliaries are dealt with in Chapter 10, and adverbs and other modifiers in Chapter 11. Clause formulae in the present chapter are deliberately kept simple by not explicitly making allowance for verb serialisation or for modifiers.

Peripheral constituents, such as time and setting, tend to occur towards the periphery of the clause, although some can occur between the subject and predicate. The periphery of the clause is discussed in §9.8.

The final section in this chapter centres on non-declarative clauses. It distinguishes various types of imperatives and interrogatives, as well as greetings.
9.2 Subjects and objects

9.2.1 Properties of subjects and objects

Subjects and direct objects usually consist of NPs (including pronouns), although they can also be complement clauses (Chapter 13). Apart from subject and object, NPs can function within the clause only as predicative NPs (§9.5.3), time phrases (§9.8.2), and, marginally, indirect object recipients (§9.4.4). All other NPs within the clause are introduced by prepositions or by prepositional verbs.

Subjects are clearly distinguished from non-subject NPs in Tetun. Subjects always precede the predicate head, the one exception being presentative existential clauses, in which the subject can follow the predicate (§9.4.7.3). The only participant-marking on verbs is subject-marking. The subject controls reflexivisation (§11.4.2), and is obligatorily omitted in coordinations using hodi ‘and, while, in order to’ (§14.5.2), as well as in reduced complement clauses (§9.7.2).

Subjects have a strong tendency to have definite reference, with many being pronouns, proper nouns, or other expressions referring to entities already present in the discourse. This is, however, only a tendency, with indefinite subject NPs being possible also.

Direct object NPs characteristically follow the verb, but also occur topicalised in the pre-subject slot, or incorporated into the verb in a preverbal position. They always represent the ‘undergoer’ participant, in the sense of Foley and Van Valin (1984:29), that is, the object NP expresses the participant which “does not perform, initiate, or control any situation but rather is affected by it in some way”.

9.2.2 Ellipsis of subjects and objects

Both subject and object NPs are readily omissible in situations where the referent is either recoverable from context or irrelevant. This is demonstrated by preliminary counts. In a sub-sample of 250 clauses headed by intransitive verbs, 20% of subjects were omitted, while in a sub-sample of 130 clauses headed by transitive verbs (e.g. of cutting and hitting), 20% had no arguments. Of these clauses headed by transitive verbs, only 30% had both subject and object expressed. A further 15% had only a subject, while 35% had only an object.

Omitted subjects (marked ‘*S’) and objects (marked ‘*O’) are illustrated in the examples below.

9.1 Ami hodi kuda ba, (*S) sai, (*S) ba kesi (*O)
1PE bring horse[Mly] go exit go tie
iha hae...
LOC grass
We will take the horses, go out, and go and tie (them) up in the grass [, because there is no grass in the garden.] (K9.66)

Clauses with oblique arguments such as instrument or recipient were excluded from these counts, since the presence of an oblique argument further mitigates against expressing both subject and object.
Nia ti’a, (*S) n-a-sa’e fetoe ne ba.
3S already 3S-cause-ascend female this go
Then, (he) put the woman up there (so nobody could disturb her).

(Loro-) (*S) Bā n-afaho, (*S) n-afaho dadauk loro
sun go 3S-weed 3S-weed continue sun

manas, nia ba.
hot 3S go
(Stumble.) (He) went and weeded; (he) kept on weeding until the sun was hot,
and he went (to her).

Nia ba bolu. (*S) Bolu n-odi n-arí: (*S) n-arí
3S go call call 3S-COORD 3S-sing 3S-sing

ai kanoik.
story

He went and called out. (He) called by singing, singing (as in) a story.

Ne (*S) n-ák é “…”
this 3S-say HES
So (he) said [“Ririró, Taun Lawes, woman Taun Lawes. Lower the ladder, I’ll
rise up to you; lower the ladder, I’ll ascend up to you.”]

Nia ti’a, (*S) n-aroun odan mai, (*S) sa’e (*O)
3S already 3S-lower ladder come ascend

bā n-á.
go 3S-eat
Then, (she) lowered the ladder, and (he) went up (the ladder) and ate.

(*S) N-á ti’a, (*S) tūn n-ika mai.
3S-eat already descend 3S-back come
When (he) had eaten, (he) came back down.

(*S) Tūn n-ika mai, (ē) (*S) bā n-afaho.
descend 3S-back come HES go 3S-weed
(He) came back down, and went weeding. (P3.20)

9.3 ...sia mai r-ola Bita Nahak (*S) hō (*O) tama.
3P come 3P-take Bita Nahak accompany enter
...they came and fetched Bita Nahak, and came in with (her).

(*S) Hō (*O) tama ho’i tafatik bā...
accompany enter towards noble.house go
(They) accompanied (her) in to the noble house... (AA1.94)
In most cases the intended referents of omitted NPs must be inferred pragmatically, as shown by the fact that the actor shifts without warning from the man to the woman and back to the man again in the sixth line of 9.2. There are, however, two syntactic constraints. Firstly, subject marking, where it occurs, indicates the person and number of the subject. Note, however, that it is quite common to have neither subject NP nor subject marking, either because the verb’s phonology is such that it cannot take the appropriate subject marking (sai in 9.1) or because the speaker is not consistent in applying it (hó in 9.3).

And secondly, the elided subject of a reduced complement (§13.4) or of a non-initial verb in a serial verb construction (§12.1.2) is obligatorily interpreted as coreferential with either the subject or the object of the main verb (for complements) or the preceding verb (for serial verb constructions), depending on the semantics of that verb.

9.3 Subject marking

9.3.1 Overview

Subject marking is an inflectional category taking the form of an initial consonantal prefix.² It is primarily applicable to transitive and intransitive verbs; subject marking of other word classes is discussed in §9.3.3. The reasons for analysing these inflections as subject markers rather than as subjects are discussed in §9.3.5.

Whether a verb can take subject marking depends on the initial phoneme. /h/-initial verbs allow subject marking for a range of subjects, with the subject prefix replacing the /h/.³ (This /h/ is consequently omitted from transcriptions, as in k-azoru, the 1S inflection of the verb hasoru, in example 9.4 below.) All other consonant-initial verbs reference only first person singular subjects, by prefixing k- to the verb. Vowel-initial verbs take no subject marking at all (*Hau k-uru surat ‘1S 1S-re-collect paper’ = ‘I re-collect the playing cards’).⁴

9.4 Ha’u k-sai mai k-azoru nia.
1S 1S-exit come 1S-meet 3S
I came out and met him. (F3.15)

² Anderson (1985:194) points out that the borderline between genuine inflectional agreement and the attraction of clitics to the verb is extremely hard to draw in many languages. This is to be expected if clitic pronouns can develop diachronically into verbal inflections (Givón 1990:353). If subject marking is analysed as an inflection, as proposed here, then Tetun is a counterexample to Greenberg (1963:93) and Bybee’s (1985a:24) empirical observations that person and number marking on the verb implies that tense, aspect or mood inflection will also occur for that language, since none of the latter are inflectional categories in Tetun.

³ Morphologically this marking can be analysed either as a prefix or as fused verb-initial conjugation, the two types of subject-marking which are the norm in Central Malayo-Polynesian languages (Himmelmann 1996:131). The former analysis is preferred in that it extends more readily to inflection of consonant-initial verbs. It requires a phonological rule that deletes /h/ following /k/, a rule which is consistent with the general lack of /kh/ sequences in Tetun.

⁴ Subject marking is apparently an area of wide dialectal variation. In the Tetun of Dili (the capital of East Timor) there is no subject marking on verbs at all (das Dores 1907:17; Hull 1996b:285; Jonker 1911:282; Morris 1984b:xvi). Mathijsen (1906:x) and Troeboes et al. (1987:59), writing of the Foho dialect of northern Belu, state that subject markers occur on /h/-initial verbs only. For East Timorese Tetun outside Dili, consonant-initial verbs inflect for 1S (Morris 1984b:xvi), as they do in the Fehan dialect. In addition, Tetun spoken in southern East Timor is said to inflect /o/-initial verbs (Hull 1996b:285; Morris 1984b:xvi). The only example I have found in the literature is oho ‘kill’ (Hull 1996b:285), which in the Fehan dialect is /h/-initial ho’o.
The table below lists for each person-number combination the full form of the pronoun, the subject marker used for /h/-initial verbs, and the paradigm for the verb há 'eat'.

**Table 9.1: Subject markers**

<table>
<thead>
<tr>
<th>Person and number</th>
<th>Full pronoun</th>
<th>Subject marker</th>
<th>Inflection of há 'eat'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1S</td>
<td>ha 'u</td>
<td>k-</td>
<td>ká</td>
</tr>
<tr>
<td>2S</td>
<td>ó</td>
<td>m-</td>
<td>má</td>
</tr>
<tr>
<td>3S</td>
<td>nia</td>
<td>n-</td>
<td>ná</td>
</tr>
<tr>
<td>1PI, 2S.HON</td>
<td>ita</td>
<td>5</td>
<td>há</td>
</tr>
<tr>
<td>1PE</td>
<td>amí</td>
<td>-</td>
<td>há</td>
</tr>
<tr>
<td>2P</td>
<td>emí</td>
<td>-</td>
<td>há</td>
</tr>
<tr>
<td>3P</td>
<td>sia</td>
<td>r- (n-)</td>
<td>rá (ná)</td>
</tr>
</tbody>
</table>

In the neighbouring Suai sub-dialect 3P marking is the same as that for 3S, namely n- (9.10). Some Fehan speakers adopt this paradigm also, or alternate between it and the Fehan paradigm. (Note that when n- is used for 3P, it is glossed simply as '3'.)

All inflections are regular, with two exceptions. One exception concerns the 1S inflection of trisyllabic verbs beginning with hak- followed by a consonant. These are freely inflectable with either initial kak- or initial ka-. That is, the coda of the initial syllable is optionally omissible. Thus, for instance, the 1S inflection of hakdiuk 'play' is either kaksiuk or kadiuk, and of haklati 'topple' either kaklati or kalati. The second exception is irregular inflection of hosí when it means 'about' (§13.3.4.2) or when it introduces the source of a transfer verb (§12.5.2.3). In both contexts hosí is not fully verbal.

There is significant dialectal-internal evidence that the /h/-initial form is indeed the basic one. Firstly, where a verb does not have the subject marking that would be expected from its subject, it is always /h/-initial, suggesting that this is the uninflected form. Secondly, the /h/-initial form is the one used as a root in derivations with the ma- prefix (§4.4). And thirdly, the citation form of such verbs is in most cases /h/-initial. The exception is for verbs which, by virtue of their semantics, can only take third person subjects. For these verbs the citation form begins with the 3S marker n- (e.g. naksala 'out of joint', naklaik 'withered', which are both derived verbs incorporating the prefix hak-). In some such cases plural inflection using r- is disallowed, leaving an invariant /h/-initial verb.

Cross-dialectal evidence supports this conclusion, since for most of Fehan's inflectable verbs the form with an initial /h/ is used in the Tetun of Dili, which uses no subject marking (fn. 4). The remaining verbs are those which necessarily have third person subjects; these, as in Fehan, begin with n- in Hull's (1996b) word list for Dili Tetun (showing that Dili Tetun used to have subject marking).

---

5 The lack of subject marking for first and second person plural reflects a diachronic loss, since such marking has been reconstructed for Proto Central Malayo-Polynesian (Blust 1993:269; see also Jonker 1911:280), and is found in the closely related languages Dawan (Middelkoop 1950:429) and Rotinese (Fox & Grimes 1995:615).

6 According to Mathijsen (1906:x) there is, in addition, an invariant verb hierus 'hate, be angry with'. Fehan consultants recognised this as a non-inflecting verb used in East Timor. In Fehan hierus is only used as a noun meaning 'breast'. Since the breast is seen as the locus of the feelings of hate and anger, extension from a basic nominal to an extended verbal meaning may account for the lack of subject inflection.
As noted above, vowel-initial verbs do not allow subject marking at all. Such verbs are in any case rare. The only vowel-initial lexeme which one linguistically astute schoolteacher accepted as a verb was the gambling term *uru* ‘re-collect playing cards after an abortive round’. All other vowel-initial verbs in the corpus have a non-verbal meaning as their basic meaning, and were felt by this consultant to be not fully verbal on semantic grounds. These are intransitive *iha* ‘be present’ (§9.4.7.2) which is mostly used as a locative preposition, *udan* ‘rain’ and *ukun* ‘rule’ which are mostly used as nouns, and *uluk* ‘go first, go ahead’ which is also a time noun meaning ‘former times’.

Subject marking in Tetun does not strictly speaking ‘cross-reference’ (or ‘agree with’) the subject, since the subject itself need not be specified. Where the subject is unspecified (9.5), or is unspecified for number (as the semantically general noun *ema* ‘person’ frequently is; 9.6), the choice of subject marking shows the person and number of the understood subject.

### 9.5

**N-ák** “Ó, k-aré ti’an”.
3S-say oh 1S-see already
(He) said “Oh, (I) have seen (it)”.  
*(O4.25)*

### 9.6

*Tán luli wa’in, ema ikus é la r-o’uk luli.*
because taboo much person time this not 3P-agree.to taboo
Because the taboos are numerous, people these days don’t agree to (abide by)
the taboos.  
*(J3.81)*

When the subject is specified within the clause, the verb in most cases agrees with it. Thus, for instance, if an addressee is referred to by pronoun, the verb agrees with the pronoun (9.7), while if the addressee is referred to by title, the verb tends to take on a third person inflection (9.8). Second person agreement, however, is an alternative possibility when the subject is a non-pronominal term of address (9.9).

### 9.7

**Ó m-atene ká lale?**
2S 2S-know or no
Did you know (him) or not?  
*(H1.24)*

### 9.8

*Ibu n-obun dansa.*
mother 3S-watch dance[Mly]
You (‘Ibu’) will watch the dancing.  
*(D0.9)*

### 9.9

**Ai! Katuas m-á sá?**
EXCL mature.man 2S-eat what
Hey! What are you (‘old man’) eating?  
*(T0.72)*

### 9.3.2 Consistency of subject marking

There are a number of differences in the consistency of subject marking for /h/-initial verbs as compared to other consonant-initial verbs. Firstly, all /h/-initial verbs in a series take subject marking (9.10). In contrast, only the first consonant-initial verb in a clause-internal series of verbs will be marked (9.11).

### 9.10

**Sia n-alai onan, n-alai n-ola n-ikar loro-sa’e-n bá.**
3P 3-run IMM 3-run 3-take/via 3-back sun-ascend-GEN go
They ran, ran away further to the east.  
*(K10.12)*
9.11  **Ha’u**  **k-bá nono**  **wé á...**  
1S 1S-go heat(liquid) water DEF  
I went and boiled water...  (F3.55)

During elicitation some consultants allowed that inflecting two consecutive consonant-initial words might be possible (**Ha k-bá k-toba** ‘1S 1S-go 1S-lie.down’; **?Ha’u k-sei k-bá Barama** ‘1S 1S-still 1S-go Barama’). However, there is only one unelicited example of this in the entire corpus, an example which in any case has uncertain acceptability due to its inflection of the adverb **foin** ‘then’.

9.12  **Oras ida ama ó-k tún mai, ha’u k-foin**  
time one father 2S-POS descend come 1S 1S-then  
**k-fu’a ó.**  
1S-awaken 2S  
Soon (when) your father comes down, then I will awaken you.  (U1.23)

A second difference between the two types of inflection is the degree to which it is obligatory. Subject-marking of /h/-initial verbs appears to be obligatory in ‘correct’ written Tetun. However, it can be omitted in speech, and consultants frequently made no correction of uninflected (but inflectable) verbs when checking transcribed texts. Approximately 5% of /h/-initial verbs in the corpus for which inflection could be expected are in fact left uninflected. This is illustrated by examples 9.3 and 9.13, both from a speaker who is less consistent in applying subject marking than most.

9.13  **Mais nia hakés tuir ama nia-k dei.**  
but 3S talk follow father 3S-POS only  
But she just spoke following my (lit. ‘father’s’) (words).  (AA2.40)

In contrast, inflection of other consonant-initial verbs is much less frequent: 51% of the 486 such verbs with a full pronominal subject **ha’u** have no subject marking. Where the subject is the proclitic **ha** (42 examples), however, only 2% lack subject marking. Discussion of this discrepancy between **ha’u** and **ha** is deferred until §9.3.4.

The presence of subject marking for consonant-initial verbs is quite strongly correlated with the presence of an overt subject pronoun within the clause. This is not so for /h/-initial verbs. Based on over 300 examples of each type of inflection, the vast majority (82%) of inflected consonant-initial verbs immediately follow the subject pronoun (as in 9.11), compared to only 49% of inflected /h/-initial verbs. Only 11% of inflected consonant-initial verbs totally lack a subject within the clause, compared to 23% of inflected /h/-initial verbs.

A final difference between the two classes of verbs concerns writing conventions. In writing, subject-marking of /h/-initial verbs is required. In contrast, some consultants held that inflection of consonant-initial verbs should be pronounced but not written, since it is absent in the orthography used in official writing (such as the Catholic liturgy), which mainly comes from the Foho dialect of northern Belu.

The relative optionality of subject marking is presumably related to the fact that consonant-initial verbs are not inflected in some dialects, while in Dili Tetun there is no subject marking at all.

---

7 This is based on a sub-sample of 300 /h/-initial verbs with a range of subjects, and on an overlapping sub-sample of 250 clauses with 1S subjects. Matthijsen (1906:x), who worked in northern Belu at the turn of the century, similarly noted occasionally hearing uninflected verbs in the Foho dialect.
9.3.3 Inflection of adjectives and non-verbs

9.3.3.1 Introduction

There are two classes of predicate heads to which subject marking clearly is applicable, namely transitive and intransitive verbs. There are other predicate heads to which it is clearly not applicable, such as nouns, prepositions and numerals. However, in addition to these clear instances there are some non-verbs which take subject marking, as well as several word classes for which the applicability of subject marking is uncertain.

9.3.3.2 Adjectives

The major class for which there is uncertainty is predicative adjectives. For instance, for bót 'big' one consultant said inflection was possible, another (on a separate occasion) that it was not possible, and another changed her mind from one answer to the other. Similar responses were received for nurak 'young' and metan 'black, dark (in colour)'. This was in contrast to immediate agreement over the possibility of inflecting active verbs such as bá 'go'. Semantically, Tetun adjectives include at a minimum some words expressing value, colour, dimension and age (of things), which according to Dixon (1982:56) are crosslinguistically likely to belong to the adjective class, no matter how small that class is. No words belonging to these semantic classes are inflected in the corpus, except a single instance of di'ak 'good' (Ha'u k-di'ak basuk '1S 1S-good very' = 'I am very good'). The boundaries of the adjective class are, however, not yet clear (§3.4).

9.3.3.3 Other word classes

Auxiliaries constitute a word class for which there is diversity in subject marking potential, with one auxiliary (the only /h/-initial one) being consistently inflected, others never being inflected, and subject marking for yet others being a matter of disagreement. For details see §10.2.3.

The postverbal modifier hikar (or hika) 'back, return to an earlier location, state or activity' is syntactically an adverb, in that it cannot head a predicate on its own, but must always follow another verb (480 examples; §11.6). Nevertheless, it consistently takes subject marking (9.10). A word which is only occasionally heard inflected, and which consultants accepted both with and without inflection, is hotu 'finish(ed), complete'. It can be used as sole predicator (e.g. Lia hotu 'word finished' = 'The story is finished'), but usually follows a verb which identifies what is finished.

---

8 Mathijsen (1906:x) concludes from the inflection of hikar that non-verbs which are tied to verbs also get inflected. Lord (1993:221) notes a similar situation in the West African language Yoruba, which has a class of words with verb-like form and behaviour, adverbial meanings (e.g. 'again'), and which lack some of the syntactic capabilities of full verbs. The classification of hikar as an adverb is defended in §11.6.

9 Presumably the uncertain status of subject marking for hotu indicates that it is in the process of grammaticisation from serialised verb to adverb. Similar grammaticalisation of verbs meaning 'finish' into completive markers is found in many languages (Bybee, Perkins & Pagliuca 1994:59; Lord 1993:228ff.). There may be dialectal or diachronic variation in the extent of inflection of hotu, since it is regularly inflected in Mathijsen's (1967) Bible stories, written earlier this century in the Foho dialect.
9.14  
*Sia* r-á *hotu* ti’a...  
3P 3P-eat finish already  
When they had finished eating...  

(S2.12)

9.15  
*Sia* r-á *r-otu* ti’an. *Nia* n-á *n-otu* ti’an.  
3P 3P-eat 3P-finish already 3S 3S-eat 3S-finish already  
They have finished eating. He has finished eating.  

(V0.45 elicited)

Although most preverbal adverbs cannot take subject marking (e.g. *hetak* ‘increasingly’, *la* ‘not’), the preverbal construction marker *na’i* ‘just’ does (§4.8.1). The adverb *foin* ‘then’ is occasionally heard inflected (9.12); however, consultants differ as to whether they accept this.

Prepositional verbs have some characteristics of verbs, and others of prepositions. The prepositional verbs *hó* ‘accompany, with (person)’, and *hodi* ‘use (instrument); by (time)’ take subject marking just as verbs do. For discussion of these words see §12.5.4–§12.5.5.

A final non-verb which takes subject marking, albeit with uncertain rules, is the preposition and conjunction *hosi* ‘about’, which is presumably grammaticised from the verb *hosi* ‘(originate) from’. It is discussed in §13.3.4.2.

### 9.3.4 A developing pronoun hak-?

Earlier it was noted that consonant-initial verbs are inflected almost without fail when the subject is an immediately preceding proclitic *ha* ‘1S’ (98% of 42 examples), whereas subject-marking is far less common when the subject is the full pronoun *ha’u* ‘1S’ (49% of 486 examples). This discrepancy suggests that the combination of the subject proclitic *ha* and a following subject marker *k-* is in the process of being reinterpreted as a single unit.

Such a reanalysis would align the lexical boundary with the syllable boundary, since *k-* is phonologically the coda to the preceding *ha*, even though it is syntactically a prefix to the following verb.

It is noteworthy that the only example in the texts of an apparently inflected noun follows the proclitic *ha*. Consultants accepted such inflection for other nouns also (e.g. *Ha k-feto* ‘1S 1S-woman’ = ‘I am a woman’).

9.16  
*Ha’u* ha *k-busa* ne’e, at *ha* k-bá kaer wé k-án sa?  
1S 1S 1S-cat this IRR 1S 1S-go grasp water 1S-do.like what  
I being a cat, how would I go and hold water?  

(Z2.112)

In addition, the preverbal adverb *na’in* ‘self’ is in the corpus inflected only following *ha* (3 examples). These facts, and the fact that the sequence *ha-k* usually occurs without a preceding pronoun *ha’u* (unlike example 9.16), suggests that *ha-k* behaves like a subject pronoun rather than a subject marker.

9.17  
*Ha* *k-na’in* du’uk bá-n.  
1S 1S-self self go-IMM  
I’ll go myself (not send someone else).  

(T0.37)

### 9.3.5 Subject marker versus subject

In the preceding discussion it was taken as given that the subject marker is not the actual subject of the clause. An alternative analysis is that the subject marker is in fact the subject,
while a preceding NP (if any) is coreferential with it, either in apposition to it or as some sort of topic. There are two main arguments for the analysis adopted here.

In the first place, if the subject marker were in fact the subject, one would expect that free pronouns preceding the subject would be rare, except in emphatic contexts. This is not the case in Tetun, where independent subject pronouns are common, as Text 2 in Appendix A demonstrates. Similarly, one would not expect a sequence of topic NP, co-referential independent subject pronoun, and subject marking on the verb, such as is found in left-dislocation constructions in Tetun.

9.18  
Maun, nia n-odi kahūk.  
elder.brother 3S 3S-bring blowgun

<table>
<thead>
<tr>
<th>Topic</th>
<th>Subject</th>
<th>Subject-marker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elder brother—he brought a blowgun. (K12.9)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Secondly, if the subject marker were analysed as the subject, the possibility of a clause having a subject would depend on the phonology of the verb. This would be a most unattractive dependency. In fact, as noted above, verbs which begin with consonants other than /h/ are much more likely to take 1S subject marking if there is an immediately preceding 1S subject pronoun than if such a subject is omitted. This is the reverse of what one would expect.

9.4 Verbal predicates

9.4.1 Introduction

Verbal predicates are predicates headed by a verb. Their subtypes are distinguished by the valency of the verb.

Tetun appears to have no verbs with a valency of zero; that is, all verbs take at least one argument. Weather verbs, which in some languages have zero valency, in Tetun allow rai ‘earth’ as subject.

Intransitive verbs and adjectives take a single argument. Existential verbs are discussed separately from other intransitive verbs, both because the single argument can follow the verb, and to allow the close relationship between intransitive existential clauses and transitive possessive clauses to be made explicit.

Verbs which take two arguments are either transitive verbs or copulas. The verb kona ‘touch, undergo’ is discussed separately because it is unique in the types of subjects and complements it can take. Note that complementation, in which the subject or object slot is filled by a clause, is dealt with in Chapter 13.

Ditransitive clauses, in which there are three NP arguments, are rare.

Certain transitive and intransitive verbs take arguments in addition to subject and object, such as recipient, or locative goal. These always follow the verb and the object NP (if any), and are introduced by prepositions, prepositional verbs or serialised verbs. For more details see Chapter 8 on prepositions and Chapter 12 on prepositional verbs and serial verb constructions.

---

10 Bresnan and Mchombo (1987:743ff.) show an alternative possibility in their analysis of Chichewâ, in that they analyse subject marking as the subject if there is no NP to fill that slot, but analyse it as grammatical agreement if there is a subject NP. Nevertheless, the details of that language are rather different.
9.4.2 Intransitive predicates

Intransitive predicates allow only one argument, namely the subject. The subject always precedes the predicate. There is no difference between actor (9.19) and undergoer (9.20) subjects.

Clause.Intransitive → (S) \{ \begin{align} \text{Vi} \\ \text{Adj} \end{align} \}

9.19 Nia n-alai ti’an.
3S 3S-run already
She has run away.  (G4.131)

9.20 Monu kona bá kursi á, liman n-ak-tesi.
fall touch go chair[Mly] DEF arm 3S-INTR-break
(When I) fell and hit the chair, (my) arm broke.  (I0.75)

9.21 Feto né sei k‘ik...
girl this still small
This girl was still little...  (G1.4)

9.4.3 Transitive predicates

9.4.3.1 Introduction

Transitive verbs allow for two nominal arguments, namely subject and direct object. Both are syntactically omissible.

The unmarked order for transitive clauses is SVO. However, the object may be fronted to before the subject to give a topicalised OSV construction. Under very restricted conditions an apparent SOV order is also possible; this is analysed as object incorporation.

Clause.Transitive → \{ \begin{align} \text{(S) Vi} \\ \text{(O)} \end{align} \}

Clause.Incorporated.Object → (S) O.Incorporated V

For all transitive verbs the subject is the semantic ‘actor’ (in the sense of Foley and Van Valin (1984:29); e.g. agent, experiencer), and the object is the ‘undergoer’ (e.g. patient, locative, recipient). That is, the subject is higher in agentivity than the object.

9.4.3.2 SVO constituent order

Transitive clauses with unmarked SVO order are illustrated here.

9.22 Tan nia n-ak-lelek ha’u / join ha’u / fight nia.
because 3S 3S-speak.abuse 1S then 1S hit 3S
Because she verbally abused me, then I hit her.  (J4.37)

9.4.3.3 OSV constituent order: topicalisation

The object of a transitive verb may be fronted to immediately before the subject, if any (213 examples). This includes the objects of some serial verb constructions (9.24). The
fronted object (highlighted by double underlining in the examples) follows any clause-initial
adverbs (e.g. *dadi* ‘so’), or periphery such as time setting.

9.23 Kalo besi, nia rotus. Mai ai aksisi ne'e, rai
as for iron 3S rust but plant k.0 hard.wood this earth
la n-á. not 3S-eat
As for iron, it rusts. But this *aksisi* wood, earth does not devour (it; i.e. it
doesn’t rot). (L0.109)

9.24 Nia tuir leten dei. Hadak e'e nia tún ha'i bá.
3S follow top only floor this 3S descend not go
She (flew) around only up high. This floor, she didn’t go down to (it).
(S2.166)

9.25 ...n-a'ak "Oa ne'e ó m-atene lale?"; N-a'ak "Oa
3S say child this 2S 2S know no 3S say child
ne'e k-atene".
this 1S know
[The nobleman] said “This child, do you know (him) or not?”. (He) said “This
child (I) know”. (Z5.99)

Fronted object NPs are always definite.11 Fronted personal pronouns consist of the full,
stressed, variant. The subject has lowered prominence.

Givón (1990:706) characterises topicalisation constructions as contrasting with normative
expectations which have been set up by the speaker listing various members of a group which
are expected to be similar. Prince (1981a) too notes that topicalisation constructions select
one member from a set of specified or evoked entities. This characterisation is true of a
significant proportion of Tetun examples (9.23, 9.24), and may account for the fact that
21% of topicalised clauses in the corpus are negated. In addition, topicalisation is a
precondition for relativisation on the object, and hence on focusing it in a cleft construction
(9.26; §9.5.5).

9.26 ...“Ida ne'e mak ha'u atu k-ó bá iha
one this REL 1S IRR 1S accompany go LOC
rai klaran”.
earth middle
[The shaman, who was choosing a wife for his master, selected one. He said]
“This one is (the one) that I will take to earth”. (G2.25)

9.4.3.4 ‘SOV’ constituent order: object incorporation

Apparent ‘SOV’ constituent order is restricted to irrealis clauses. In the corpus, most of
the examples are negative (9.27; 9.52), while the remainder are questions (9.28). SOV order
cannot be used for positive statements (*Ami kabau hó. ‘1PE buffalo have’ = ‘We have

11 This is a common condition crosslinguistically, since fronting applies primarily to highly topical, anaphoric
arguments (Givón 1990:709).
The clause

This constituent order is relatively rare, with the corpus containing only 15 indisputable examples of it apart from the examples for the verb *lalek* 'lack' (for which SOV order is obligatory; §9.4.7.7). Note that preverbal position also occurs, under the same conditions, for 'small clause' object complements (§13.4.3.3). Objects in this construction are always non-referential, and consist simply of a single noun. The limited data thus suggest that they are incorporated into the following verb.

9.27 $\text{Ha}'u\ kopi\ k-emu\ ha'i,\ \text{kangkung}\ k\hat{a}\ ha'i.$
1S coffee[Mly] 1S-drink not k.o.vegetable[Mly] 1S-eat not
I don't drink coffee and don't eat the vegetable *kangkung* (for health reasons).
(T0.45)

9.28 $\text{Sia}\ \text{boli}tik\ \text{r-}\text{atene}?!$
3P politics[Mly] 3P-know
Were they politically aware?! (Rhetorical question—of course they weren’t!)
(H3/6)

9.4.4 Ditransitive predicates

Ditransitive clauses have three NP arguments, namely the subject and two objects. Such clauses are rare in the corpus. There are only four unelicited ditransitive examples in the corpus, all for the verb *f6* 'give'. The recipient (indirect object) NP follows the verb, while the patient (direct object) NP either follows the recipient (9.29) or is topicalised in pre-subject position.

\[
\text{Clause.Ditransitive} \rightarrow \begin{cases} (S)\ VfO\ (IO.\text{Recipient}) \ (O) \\ (O)\ VfO\ (IO.\text{Recipient}) \end{cases}
\]

9.29 $\text{In-bei,}\ \text{f6}\ \text{ha}'u\ \text{w}\ \text{k-emu}\ \text{lai.}$
mother-ancestor give 1S water 1S-drink first

Old lady, give me water to drink.
(O2.6)

Several consultants agreed that such ditransitive constructions were acceptable, although one expressed reservations. It is in practice far more common for the recipient to be introduced by *b\hat{a} to* (§12.5.3.3). Clauses with two postverbal NPs were rejected by consultants for other exchange verbs such as *fa'en* 'sell' (for which the recipient is introduced

---

12 Note that this ungrammatical example would be grammatical if the verb *h\hat{a} to* 'have' were given third person subject marking to agree with *kabau* 'buffalo' instead of *ami* '1PE'. For details of this topic-comment construction see §9.4.7.5. In addition to the clear cases of SOV word order, there are 8 examples with the verb *n-\hat{o} to* '3-have'/'3-exist', which can in principle be analysed either as having SOV order (with the verb 'have') or as having this topic-comment construction (with the verb 'exist').

13 Irrealis mood is not a common crosslinguistic trigger for constituent order variation within the clause. Givón (1979:124) notes that negation and a number of other modal environments retain an older SOV order in the now otherwise SVO Benue-Kwa language, Kru. However, he hypothesises that this is because they all represent historically embedded structures, which are conservative when it comes to word order change. Siewierska (1988:92) points out one other example of negation triggering word order variation, namely the SOV Nigerian language, Gwari. It has SVO order in negative perfective clauses as well as in perfective affirmatives with contrastive emphasis on the object (Lord 1982:295).
by bá ‘(to)’ and lók ‘offer betel’ (for which the recipient is introduced by bá only if there is also a patient NP).14

9.4.5 Copular constructions

9.4.5.1 Introduction

Copulas are two-place predicates which take a preceding subject NP and a following predicative complement. Unlike transitive clauses, in which there are in general selectional restrictions between verb and object, copular constructions have selectional restrictions between subject and complement. The complement cannot be fronted.

‘True copulas’ are semantically empty. Tetun has only one of these, ní. In addition, there are two ‘semi-copulas’ (Hengeveld 1992:35f.) which contribute meaning, namely tu’an ‘grow, become more’ and dadi ‘become’.

9.4.5.2 ní ‘be’

The copula ní indicates a relationship of unique identity, in which the referent of the postverbal NP is presented as uniquely satisfying the description given in the preverbal NP (47 examples).

Clause. Be → (S) ní Complement. Pronoun

The preverbal NP is descriptive, giving the basis of the unique identity. It is usually possessive, but the uniqueness may also come about via a description such as mak kwana nia-kan ‘REL right 3S-POS’ = ‘which is the correct one’. It is presupposed that one member of a group satisfies this identity. The postverbal NP is headed by a pronoun (usually ne’e ‘this’, nia ‘that, 3S’, but also other personal pronouns or interrogatives such as nabé ‘which’ and sé ‘who, which person’), and specifies which member of the group satisfies the description.

9.30...

...N-ák “Lale. Tais ó-k ní nia...”

3S-say no cloth 2S-POS be 3S

[The man denied owning the sarong.] (The girl) said “No. That is your sarong. [Take it and wrap it around you.]”

N-ák “Lale. Tais ha’u-k á ní ha’i ne’e”.

3S-say no cloth 1S-POS DEF be not this

(He) said “No. This is not my sarong”. (R5.27)

---

14 There are two verbs of naming which marginally accept both an object NP which refers to a person (and which can be fronted via topicalisation) and a following object complement specifying a name. These are the apparent synonyms hanaran and hamed ‘give a name’, which are both illustrated in the rather stilted elicited example below. Another consultant, however, disapproved of a similar clause (*Oan ne’e ha’u kanaran Manek ‘child this I name Manek’). In the texts either the object or the object complement is always omitted and/or the name is introduced by ha’ák ‘say’. In these ways a clause with two non-subject NPs is avoided.

Ω ha’u k-a-namí Abraham, Ha’u k-a-naran ó Abraham.

2S 1S 1S-make-name Abraham, 1S 1S-make-name 2S Abraham

You I name Abraham. I name you Abraham. (Q6.133 elicited)
9.31 *Na fé-n kbesik á ní ha’u.*
3S wife-GEN direct DEF be 1S
His true wife is me (as opposed to the other girl who wants to marry him).

(V0.41)

Clauses with *nì* thus present constituents in the reverse order to cleft constructions (§9.5.5). However, *nì* can be negated (second line in 9.30; 5 examples), showing that it is predicative, and not just a non-predicative word marking the following constituent as the focus.

The requirement that the postverbal NP be pronominal is much more natural for a subject than for a complement. Nevertheless, the fact that the initial NP, in addition to being in the usual Tetun subject position, can be omitted from the clause (3 examples) provides syntactic support for its analysis as subject.

9.32 *M-ola tais ó-k á tê nì nia.*
2S-take cloth 2S-POS that because which 3S
Take your sarong as (yours) is that one.

(R5.29)

9.4.5.3 *tu’an ‘grow’*

The copula *tu’an ‘grow, become more’* takes an adjectival complement denoting the quality with respect to which the growth takes place (10 examples).15

Clause.Grow → (S) *tu’an* Complement.Adjective

9.33 *Sura kalan nía tu’an bokar.*
every night 3S grow wide
Every night it (the moon) grows wider.

(F4.14)

9.4.5.4 *dadi ‘become’*

The verb *dadi ‘become’* can take a range of predicative complements, namely NPs (9.34; 20 examples), adjective phrases (9.35; 13 examples), and prepositional phrases (9.36; 2 examples).16 The complement is frequently introduced by *bâ ‘go, to’* (9.37; 54 examples).

Clause.Become → (S) *dadi* (*bâ*) Complement

Complement → \{NP, AdjP, PP\}

9.34 *Ô madinas sekola, ó *dadi* ibu, *dadi* ba’a.*
2S diligent school[Mly] 2S become mother[Mly] become Mr[Mly]
(If) you are diligent at school you’ll become ‘Ibu’ (or) become ‘Pak’. (These are Indonesian terms of address used for schoolteachers and others with comparable levels of education.)

(Q0.22)

15 The form *tu’an* is also an active transitive verb meaning ‘add’, which takes nominal or numeric complements.

16 In addition, *dadi* is a complement-taking verb meaning ‘happen, work out’.
9.35  
_Nia dadi ha' i ktomak._
3S become not whole
It won't become complete.  

(I22/7)

9.36  
_Nia dadi nú ita ti'an é._
3S become like 1PI already TAG
He had become like us, hadn't he.  

(U4.70)

9.37  
_Nia dadi bá fahi._
3S become go pig
He turned into a pig.  

(Z6.8)

Unlike object NPs, complement NPs rarely refer. Most complement NPs are not modified
by determiners; the indefinite article _ida_ is not unusual (8 examples), but definite determiners
are (the one example being a reiteration: 'Having become that beautiful man...'). Complement
NPs cannot be fronted in topicalisation constructions.

9.4.6 _kona_ ‘touch, undergo’

In the physical realm, the transitive verb _kona_ ‘come into contact, touch, undergo’ refers
to two entities unintentionally coming into contact with each other.17 Usually, though not
necessarily, this touching has an adverse effect on one of the two participants. The moving
participant is coded as subject regardless of which of the participants is adversely affected.
Thus in 9.38 it is the object referent who is adversely affected by acquiring a skin blemish,
while in 9.39 it is the subject referent that is caught on a spear and as a result starts bleeding.

9.38  
_Ita ho'o manu, rán kona ita, manu rán é._
1PI kill bird blood touch 1PI bird blood this
dadi bá kfuti.
become go mole
(If) we kill a chicken, (and) blood touches us, the chicken blood becomes a
mole.  

(Q0.41)

9.39  
...nia-kan kelen né _kona_ diman á...
3S-POS thigh this touch spear DEF
...his thigh caught on the spear [, and blood dripped onto the floor.]  

(E0.132)

By extension, events and conditions can touch a participant. The event or condition is
coded using a ‘small clause’ (§9.7.3), whose structure is discussed below. Either the event or
condition is specified as subject and the experiencer as object (9.40) or the two occur in the
reverse order (9.41). Any difference in meaning between the two orders is not apparent to
me.18

Clause._kona_ → [(S.NP) _kona_ (O.NP)]
{(S.NP) _kona_ (O.Small.Clause)}
{(S.Small.Clause) _kona_ (O.NP)}

17 In addition to the meanings discussed in this section, _kona_ is an intransitive verb meaning ‘correct, true’. It is
also commonly used in referring to people, words or prices ‘touching each other’ (_kona malu_), that is, to
reaching agreement.

18 Fronting of the object of _kona_ was accepted during elicitation (Nia moras _kona_ ‘3S sick touch’ = ‘He got
sick’), confirming analysis of _kona_ as a fully transitive verb.
Small clause complements of *kona* nearly always refer to things which adversely affect the participant, such as sickness, death, cold or work. Positive effects are, however, possible when they are specified in contrast to negative ones, as in the following example.

9.42 *Emi bá kona di’ak, ami kona át. Emi kona soi...*
2P go touch good 1PE touch bad 2P touch rich
You (Europeans) went and received good, we (Timorese) received hardship.
You got rich...

In nearly all examples the small clause consists of a single adjective (9.42), intransitive verb (e.g. *monu* ‘fall’, *mate* ‘die’) or transitive verb (e.g. *kohi* ‘catch’, 9.44). However, one example has a verb plus object (9.43), while subject plus object was accepted during elicitation (*Ô kona ema ta’e* ‘2S touch person hit’ = ‘You are hit by people’).19

9.43 *Ami kona hafaho rai, emi kona Malae.*
1PE touch weed earth 2P touch non.native
We (Timorese) inherited manually weeding the ground, you got the foreign way (of working).

Where the complement is headed by a transitive verb, the affected participant may be either the actor (9.43; 5 examples) or the undergoer (9.44; 16 examples) of that verb, so long as it is a participant which is adversely affected in some way.

9.44 *Ami loron kona fotá.*
1PE day touch hit
We were daily hit (by him).

While it is in most cases clear what the semantic role of the participant is, it seems that ambiguity is possible. Thus *Nia kona hana’o* ‘3S touch steal’ was held by one consultant to mean ‘He was robbed’, while *Nia kona n-ana’o* ‘3S touch 3S-steal’ was said by another to mean ‘He stole’. The difference between these is only in the subject marking; the verb cross-references the subject of *kona* in the active interpretation but is left uninfl acted in the undergoer interpretation. Presumably true ambiguity is possible if subject marking is not applicable (e.g. for *Ami kona hana’o* ‘1PE touch steal’).20

---

19 A longer object clause consisting of subject, verb and complements was spontaneously used by one consultant who was strongly influenced by Indonesian, but it proved difficult for others to interpret. It appeared to be an attempt to use *kona* to derive a passive construction.

20 Baxter (1988) notes a similar, although syntactically more restricted, construction in Portuguese creole. He analyses it as serialisation of the verb *taka* ‘touch’ and a transitive verb. To analyse the Tetun construction in this way would be to obscure the close relationship between ‘touching’ a physical object (with NP complement) and ‘touching’ a condition (with non-NP complement). It would also result in a unique serial verb construction in which the subject of *kona* could be coreferential with either the subject or (where applicable) the object of the following verb.
9.4.7 Existential and possessive constructions

9.4.7.1 Introduction

Semantic relations akin to both existence and possession can be predicated with the verbs *iha* 'be present', *n-ó* '3-exist' and *hó* 'have'. The verb *n-ó* '3-exist' is idiosyncratic syntactically, in that its single argument can either precede or follow the verb, and in that existential clauses headed by *n-ó* take part in topic-comment constructions. The verb *iha* is given special attention on account of its close semantic relationship with *n-ó*. The verb *lalek* 'lack' predicates lack of possession, and is unique in requiring SOV word order. There is also a minor, verbless, presentational construction, which is discussed in §9.4.7.8.

9.4.7.2 Intransitive *iha* 'be present'

The word *iha* as general locative preposition has already been discussed in §8.3.1. As an intransitive verb it indicates that something is present or available. Mostly it is negated (9.45; 69% of 42 examples). Where it is not negated it tends to be used only if the presence of the participant is surprising (9.46). This is because most positive statements that something is present specify the location as well, thus using *iha* transitively (as a preposition). As Lyons (1967:391) puts it, “whatever is, is somewhere”.

9.45  
*K-anoin ni ina iha ha'i.*
1S-think 3S mother be.present not
I think her mother isn’t present.  
(K6.16)

9.46  
*Ema iha kedan.*
person be.present immediately
People were already here (even before I had my siesta).  
(T0.20)

The subject of *iha* always precedes the verb. Since *iha* is vowel-initial, there is no possibility of subject marking.

9.4.7.3 Intransitive *n-ó* '3-exist'

As an intransitive existential verb, *n-ó* '3-exist' is used with both singular and plural third person subjects. It is not clear whether this verb is synchronically a third person inflection of *hó* (with *n-* inflection used for both singular and plural), or whether it is an invariant verb *nó*. Third person plural marking is not possible (*Kabau la r-ó* 'buffalo not 3P-exist'). The fact that *hó* also means 'have', and that many languages have a close relationship between possessive and existential constructions (Bateman 1982; Lyons 1967; Omar 1974; Schachter 1985:57) supports at least a diachronic analysis as a 3S inflection of *hó*. This is the analysis that is assumed in the example glosses, although nothing else hinges on it analytically.

Clause. Existential *nó*  \[ \{ (S) \ nó \ \} \]

---

21 Note that *nó* is also a coordinator meaning ‘and’ (9.50), and an adverb meaning ‘also’.

22 Since existence is not readily predicated of first and second person subjects, I do not know how these would be inflected.
The single NP in an existential clause can either precede the verb (9.47, 9.48; 40 examples) or follow it (9.49, 9.50; 78 examples).

9.47  Uluk kantung sei la n-ö.
      former.times bag[My] still not 3-exist
   In former times there weren't any bags yet. (They are a modern introduction.)
   (O5.38)

9.48  Hofonin mane sia mai labu iha h oo sia né,
      last.night man PL come go.out LOC 1S child PL this
      mama n-ö.
      betel 3-exist
   Last night men came courting my children (or nieces ...), (and so) there will be
   betel (at the house).
   (O5.32)

9.49  N-ö ha‘i ema r-ö ita mai.
      3-exist not person 3P-accompany 2S.HON come
   There was no-one who accompanied you here. (lit. ‘There weren’t people...’)
   (T0.37)

9.50  ...“Ei! N-ö feto ida. N-ö feto ida nô oo”.
      EXCL 3-exist woman one 3-exist woman one and child
   [Looking out at the tiny offshore island, the man said:] “Hey! There is a
   woman! There is a woman and child”.
   (R5.89)

The preverbal subject position is used only for NPs which are already topical. As such none of the preverbal NPs are marked as indefinite. All examples in the corpus indicate contrast with some other (earlier, later or expected) state of affairs. 75% are negative.

In contrast, the postverbal NP position is mostly used to introduce new, albeit often very minor, participants or props into the discourse. This is consistent with the strong crosslinguistic tendency for presentational constructions to be verb-initial, thus cancelling the expectation that the single argument NP is ‘given’ (Clark 1978; Givón 1978:295). Unlike preverbal NPs, none of the postverbal existents in the corpus are marked as definite. The lack of definiteness and the postverbal position both point to such NPs being highly unusual subjects, if indeed they are best labelled as such at all. 41% of clauses with this word order are negative.

9.4.7.4 Transitive hó ‘have’

The transitive verb hó ‘have’ can indicate a wide range of relationships between the two clause participants. These are the same relationships that are encoded within possessive NPs. They include the relation of a person to his or her kin (9.51; e.g. wife, child), social relations (e.g. companion, servant), body parts (e.g. arm, leg), physical possessions (9.52; e.g. buffalo, shoes), time, power, work and good fortune. They also include the relationship of a whole to a part (e.g. a house has rooms), or of an entity to things which are associated with it (e.g. clothing has fleas, a drink contains poison, and an item has an associated poem).

Two main orders of clause constituents are found for hó ‘have’. As for other transitive verbs, the most common is SVO (over 180 examples for (S)VO).

23 Hó is also a transitive verb meaning ‘accompany’ (9.49).
Chapter 9

9.51 "Ha'u k-ó ama ká ha'u la k-ó ama?" ... "Ó m-ó ama".

2S have father or 1S not 1S have father 2S

"Do I have a father, or don't I have a father?" [Abuk replied] "You have a father". (U1.12)

The most common alternative order is SOV, in which the object is incorporated into the verb. This occurs under the same conditions as other clauses with object incorporation (i.e. particularly with negation; but see §9.4.3.4 for details; 6 unambiguous examples).

9.52 Kalo belu ó osan la m-ó, it taru ata.

if friend 2S money not 2S have IP I gamble slave

If, friend, you have no money, we'll gamble for slaves. (G4.45)

In addition there is an uncommon OSV order (2 examples), in which the object is topicalised.

9.53 Kalo adat é ó m-ó, ó m-ó oan.

if tradition this 2S 2S have 2S 2S have child

If you have (i.e. follow) tradition, you will have children. (T0.132)

9.4.7.5 Overlap for nó '3S-haves/exist'

To say that someone has something (a possessive construction) implies that as far as that person is concerned the thing exists or is present (an existential construction). This semantic fact is reflected in an alternative means of expressing possession. In this construction, the possessor NP is followed by an intransitive existential clause, with an invariant verb nó. The subject of nó denotes the possessum, and can either precede nó (9.54), or follow it (9.55).

Initial possessor NPs in this construction are what Chafe (1976:50f.) calls 'Chinese-style' topics, in that they "limit the applicability of the main predication to a certain restricted domain". In this case, the possessor indicates the one for whom the existence of the possessum is relevant.

\[
\text{Clause:Possessive,nó} \rightarrow \begin{cases} 
\text{(Topic.Psr) (S) nó} \\
\text{(Topic.Psr) nó (S)}
\end{cases}
\]

9.54 Lale. Ha'u buat e'e sia n-ó.

no 1S thing this PL 3-exist

No. These things I have. (So no need for you to give them to me.) (S2.176)

9.55 Isin di'ak ita n-ó ha'i kók.

body good 1PI 3-exist not swollen.spleen

(When we are) healthy, we don't have a swollen spleen. (It only appears when we have malaria.) (I0.124)

The proposed analysis introduces a topic-comment construction, which is not found elsewhere in the grammar. An alternative analysis of many of these examples is that there is a single verb hó 'have', which is in a state of flux between being a regularly inflectable verb (as in 9.51) and an invariant nó (as in 9.55). Nevertheless, there are certain circumstances under which the topic-comment analysis is preferred. The first is where a verb-final clause does not meet the conditions for incorporating the possessum into the verb. This is so if the
possessum NP is definite (9.54; 1 example), or if the clause is used to make a positive, non-emphatic, statement (possibly 9.54, 9.56). A topic-comment analysis is furthermore encouraged, although not required, if there is pause after the initial possessor NP (9.56).

\[ \text{Ah! Lalika fó. Ami, osan n-ó, mortén n-ó.} \]
EXCL need.not give 1PE money 3-exist beads 3-exist
Ah! No need to give (anything). We have money and beads. (or: As for us, there is money, and there are beads.) (S2.173)

Where the possessor is 3S, there is usually no basis for choosing between a ‘have’ analysis (in which nó agrees with the possessor) or an ‘exist’ analysis (in which the possessor is a topic NP). Both options are reflected in the alternative translations for 9.57 (which has an explicit possessor) and 9.58 (which does not).

\[ \text{Nia fini la n-ó!} \]
3S seed not 3S-have/exist
She has no seed. (or: ‘As for her, there is no seed.’) (This is said of childless women.) (T0.132)

\[ \text{Feto ne’e kiak: la n-ó ina, la n-ó} \]
girl this orphan not 3S-have/exist mother not 3S-have/exist
father
This girl was an orphan. (She) had no mother, had no father. (or: There was no mother, and there was no father.) (R4.2)

The use of an invariant nó form in possessive clauses is common. If one considers only examples with non-3S possessors, nó is found in 30% of the 94 examples with Possessor-Verb-Possessum (i.e. apparent SVO) word order, and in 68% of the 19 examples with Possessor-Possessum-Verb (i.e. apparent SOV) order. Elicitation confirmed that both an inflected hó and an invariant nó are acceptable, and that inflection for the possessor is considered better in SVO clauses, while an invariant nó is more natural for SOV.

Regardless of the analysis accepted, there is a correlation between word order and negation, with 33% of the 204 verb-medial (i.e. Possessor-Verb-Possessum) clauses being negative, and 85% of the verb-final clauses being negative. This is comparable to the situation for intransitive nó clauses in which there is no possessor, for which 41% of verb-initial clauses and 75% of verb-final ones are negative.

\[ \text{9.4.7.6 Comparison of iha and nó} \]

Of the two intransitive verbs discussed above, iha ‘be present’ specifies current (perhaps temporary) presence of something at a certain place, while n-ó ‘3-exist’ comes closer to predicating absolute existence or association with an understood person or location. Consultants explained this contrast to me using examples such as the following. Papa la iha ‘Papa not be.present’ is appropriate if Papa (the father) has temporarily gone out. In contrast, Papa la nó is appropriate if Papa has died, or has abandoned the family; that is, if he is more or less permanently gone. This distinction is consistent with the fact that iha is also a locative preposition, used to predicate location of an entity, while hó is also a transitive verb predicating the less accidental relationship known in the literature as ‘possession’.

Nevertheless, there is significant overlap between the two. This was illustrated by an occasion when two people were explaining to me that a certain sacred place was uninhabited.
One used *Ema la iha* ‘person not be.present’, while the other, almost simultaneously, used *Ema la nó*.

In East Timor *iha* covers the semantic domain of Fehan * nó* as well as *iha*. This statement is supported by the dictionary of Morris (1984b) and the word list of Hull (1996b), as well as by the observations of Fehan consultants. Influence from other dialects may explain the rare use of *iha* in the corpus to predicate absolute existence, rather than merely current (perhaps temporary) presence. The following example from an origin myth illustrates this usage.

9.59 ...*loro sei dauk la iha*. *Fulan sei dauk la iha.*

sun still not not be.present moon still not not be.present

*But it had not yet become light;* the sun did not yet exist. The moon did not yet exist. (B1.10)

**9.4.7.7 lalek ‘lack’**

The verb *lalek* ‘lack’ predicates the lack of possession of something (38 examples). This is a unique verb in Tetun in that it compulsorily follows the constituent identifying that which is lacking (*’lalek oa ‘lack child’*). The complement of *lalek* is indicated by double underlining in the examples.

Clause.*lalek* → (S) O.Incorporated *lalek*

O.Incorporated →

\[
\begin{align*}
\text{N} & \\
\text{Adj} & \\
\text{Vi} &
\end{align*}
\]

9.60 ...*Nia inan lalek, aman lalek*...

3S mother lack father lack

[This girl is a real orphan.] She has no mother/sisters, no father/uncles [, no brothers or sisters.]

(T0.41)

Evidence that *lalek* is nevertheless a verb comes from its ability to be intensified by *basuk*.

9.61 *Emi ne‘e ukun lalek basuk.*

2P this rule lack very

You are really without rule. (i.e. You are very disobedient.)

(Q0.73)

*Lalek* can be reduplicated with no apparent change in meaning.

9.62 *Nia kbokur ilas la-lalek.*

3S fat magnificent RDP-lack

He is fat without beauty (i.e. fat and ugly).

(V0.37)

The constituent specifying the possessor is clearly the subject. This is indicated by its initial position, by the fact that it is the semantically shared constituent in clausal coordinations using *hodi ‘COORD’* (9.63; §14.5.2), and by the fact that it is the possessor which is the object of the derived causative verb *halalek* ‘cause to lack’ (9.64).

24 This word is probably morphologically related to the negative pro-clause *lale* ‘no’.
The constituent specifying what is possessed can be a noun (9.60, 9.61; 34 examples), adjective (9.62, 9.63; 2 examples) or intransitive verb (9.65; 2 examples). Although karian ‘work’ can be a noun, its verbal status in 9.65 is reinforced by the following clitic ta ‘already’, which only follows verbs. There are syntactic restrictions on this constituent. While compound nouns are accepted (ina-ama lalek ‘mother–father lack’ = ‘lacking parents’, maun–alin lalek ‘elder.brother–younger.sibling lack’ = ‘lacking brothers’), it appears the noun cannot be modified (*uma bót lalek ‘house big lack’ = ‘lacking a big house’). This suggests that the object is incorporated into the verb phrase headed by lalek.

9.65 It karian ta lalek mane sia soe kananuk nó ita.
1PI work already lack man PL throw k.o.poem with 1PI
If we don’t do work the men ‘throw’ a poem at us (which accuses us of laziness). (V0.13)

Although lalek is unique in requiring object incorporation (i.e. apparent SOV word order), incorporation is optional in some other situations, most of which involve negation. It thus seems that what is grammatically optional for these other, inherently positive, verbs has been fully grammaticised for this inherently negative one.

9.4.7.8 Presentational NP

An NP with no accompanying verb can specify the identity of a new referent (10 examples).25 In the corpus each presentational NP is intonationally linked to a preceding clause, which sets the context.

9.66 ó moris mai, mesa ó dei.
2S live come solely 2S only
(When) you were born, (there was) only you. (You had no brothers or sisters.)
(G1.20)

Some such NPs which identify new referents could be analysed as equative clauses with omitted subjects. For instance, the subject of the NP ‘a young male bird’ in 9.67 would be coreferential with that of the preceding clause, namely nia ‘3S’. Nevertheless, such an analysis is highly unlikely for 9.66, for which the insertion of a subject would result in a clause saying ‘you were only you’. It thus seems that 9.66 represents a minor construction type in Tetun.

25 According to Durie (1994:518) the introduction of new referents is in Acehnese conversation a common function for NPs not connected to clauses.
9.67 \textit{Nia moris, manu a\textit{man} oan ida. Oa m\textit{os} n-a-horis,}  
3S live bird male small one child also 3S-make-live  
\textit{mane oan ida.}  
boy small one  
It hatched, (and it was) a young male bird. The young woman (lit. ‘child’) also gave birth, and (the baby was) a boy.  
(R5.80)

9.5 Non-verbal predicates

9.5.1 Prepositional phrases

A number of prepositions can introduce prepositional phrase predicates. These are listed in §8.2, and are further discussed and exemplified in that chapter.

\textit{Clause.PP} \rightarrow (S) Predicative.Complement.PP

9.68 “...Uma \textit{iha nabe?” N-\textit{ak} “Uma \textit{iha hali bot ida...}”}  
house LOC where 3S-say house LOC banyan.tree big one  
“...Where is (your) house?” (She) said “The house is at a large banyan tree...”  
(O4.16)

9.5.2 Numeral phrases

Numeral phrases can function as predicate, with a variety of semantic interpretations.

\textit{Clause.Numeral} \rightarrow (S) Predicative.Complement.NumP

They frequently indicate the number of the subject.

9.69 \textit{Nia-kan ulu-n a hitu.}  
3S-POS head-GEN DEF seven  
It (this snake) has seven heads. (lit. ‘Its heads are seven.’)  
(Q0.95)

9.70 \textit{Kbu ne’e fuan rua dei.}  
rice.parcel this CLS:round two only  
There were only two rice parcels. (lit. ‘The rice parcels were only two items.’)  
(Z3.251)

Alternatively, predicative numerals (without classifiers) may indicate price or some measure of size.

9.71 \textit{Inan tulan mak sei ki’ik-ki’ik oan ne’e sia}  
female immature(.bird) REL still RDP-small small this PL  
m\textit{os rihun h\textit{at}.}  
also thousand four  
Even immature female (chickens) that are still small are 4,000 (rupiahs).  
(K0.184)

Where the numeral is \textit{ida} ‘one’, it can indicate a unity amongst the referents of the subject, showing that they are considered as one.
9.72  Adat nó Sarani ida dei.
tradition and Christian[Mly] one only
Tradition and Christianity are one. (This was in reply to a question as to whether the Christian Maromak 'God' was different to the traditional one.)

(B0.122)

9.5.3 Noun phrases

In a clause consisting of two juxtaposed NPs, the first (regarded here as the subject) usually represents given information, and is in the vast majority of cases marked as definite. The second, representing new information, indicates the identity of the first. This includes membership of established groups (such as by sex, age, type of entity (9.73), or class (9.74)), and proper nouns.

\[
\text{Clause.Nominal} \rightarrow (S) \text{ Predicative.Complement.NP}
\]

9.73  Buat nia emar iha laran. Nia bók ha'i.
thing that person LOC interior 3S gourd not
(Referring to a gourd that had become human:) That thing was a person inside. It wasn’t a gourd.

(Z5.25)

9.74  ...mais, Buku Lasak ne’e ata. Na’im mak Bita Nahak.
but Buku Lasak this slave noble REL Bita Nahak
[The nobleman had taken a liking to Buku Lasak.] However, this Buku Lasak was a slave. The noble(woman) was Bita Nahak.

(AA1.48)

The order of the NPs in an equational clause is reversible. The less common order in which the new information comes first is found in the corpus under only two conditions. The first is where the initial NP is focused by a cleft construction (second clause in 9.74; §9.5.5), and the second is where it is focused by being questioned (9.75; 2 examples).

9.75  Sé ida ema tür ne’e?
who one person sit this
Who is this person sitting here?

(K9.93)

Negation of nominal predicates is by means of ha’i (9.73) or the contrastive negator lahós ‘indeed not’. Negation by premodifying la was not well accepted by consultants.

The predicate of an equational clause can be a verb phrase (4 examples). Given the paucity of nominalisation strategies in Tetun, it is not surprising that this verb phrase need receive no nominal marking.

9.76  Feto nia-k knawar, so’i batar, taman hare.
woman 3S-POS work harvest(.maize) maize plant rice
Women’s work is harvesting maize, planting rice.

(H0.72)

9.5.4 Possessors

The possessor may be presented as a predicate following the subject. It has the same internal structure as a possessor that follows the nominal head within an NP (‘post-possessor’ in §7.3.1).

---

26 Van den Berg (1989:170) observes that the Muna language of Sulawesi similarly allows both orders.
Possessive predicates are best analysed as a subtype of nominal equative predicates, as suggested by de Groot (1983:111f.) for Hungarian and subsequently by Bugenhagen (1986:141) for the Austronesian language Mangap-Mbula. In other words, the predicate is an NP without a head noun. This is semantically motivated, in that a possessive construction necessarily implies, even if it does not state, that which is possessed. Thus, for instance, 'The coconuts are ours' can be paraphrased as 'The coconuts are our coconuts'. Syntactic support for this analysis comes from the fact that the structure of the predicative possessor is the same as that found in headless NPs, in contrast to the different structure found in preposed modifiers within NPs.

9.5.5 Relative clause NPs: clefts

9.5.5.1 Structure

Cleft clauses are equative clauses with an NP subject (indicated by double underlining in the examples) and a subject complement headed by a relative clause. The relative clause is introduced by the relative clause marker mak (sometimes ma’ak). For further details on the internal structure of relative clauses see §14.8.

Example 9.79, with the determiner ne’e ‘this’ following the relative clause, shows that the complement is an NP headed by a relative clause, and not a relative clause directly.

9.79 Tuan bot se mak n-á uluk ne’e?
important.man big who REL 3S-eat go.first this
[Riddle: Even before I start eating, an important man eats before me.] Which important man is it who eats first? (The answer is ‘a fly’.)

The subject of the main clause is equated with the relativised constituent in the relative clause. In the corpus the vast majority of relativised constituents are subjects (9.79, 9.81; 139 examples). Others are object (9.80, 9.26; 8 examples), subject complement (second clause in 9.74; 28 examples), reason (first clause in 9.90; 2 examples) or time (1 example).

---

27 Close syntactic and pragmatic parallels between relative clauses and it-cleft constructions have frequently been observed in the literature (Andrews 1985:86; Keenan & Hull 1973; Steinhauer 1992).

28 This presumably does not exhaust the possibilities, since there are more constituents which can be relativised.
9.80  
Na’an  ida nabé mak at  há  di’ak?  
meat one which REL IRR eat good  
Which meat is it that would be good to eat?  

(Q0.160)

The object of the complement clause can be fronted to before the main clause subject (second clause in 9.81; 8 examples). Object fronting from within a relative clause predicate thus functions exactly like object fronting from within a verbal predicate.

9.81  “Sé mak lita faru ó-k ne’e?”  “Faru ha’u-k ne’e,  
who REL sew clothes 2S-POS this clothes 1S-POS this  
ha’u mak lita.”  
1S REL sew  
“Who is it who sewed your dress?” “This dress, it is I who sewed (it).”  
(H0.77)

9.5.5.2 Function

In the vast majority of cleft examples in the corpus the relative clause presents a situation with which the hearer is presumed to be familiar (although it need not yet have come up in the discourse) while the focused subject NP is questioned, presents an answer to a question, or presents contrastive information.  

Only a minority (about 18%) of information (‘wh’) questions have this cleft construction, those that do nearly always have the questioned constituent as the focused subject of the cleft clause (9.79, 9.80, 9.81; 23 examples). Similarly, answers to information questions can present the new information in the focus slot (second clause in 9.81; 7 examples).

Apart from questions and answers, the focused constituent frequently explicitly contrasts its referent with other referents with which it forms a set (67 examples). This is done either by explicitly specifying that another referent is not meant (9.82), or by stating that other members of the set are otherwise disposed (e.g. ‘I sat in the room. It was my parents who went to the discussion.’).

9.82  ...“Hai!  Buik.  Ó  la  bele  todak  ha  ulu-n  nú  nia”.  
EXCL Buik 2S not can yank 1S hair-GEN like 3S  
[The noblewoman said] “Hey! Buik. You can’t pull my hair like that”.

Nia.má  n-ák  “Lahós  ha’u  mak  todak.  Ama  Na’i  
then 3S-say indeed.not 1S REL yank father noble  
mak  todak”.  
REL yank  
Then (Buik) said “It is not I who am pulling. It is Noble Father who is pulling (it)”.

(O2.39)

Another use for cleft constructions is to present further background information on a participant who was mentioned in the immediately preceding clause (10 examples). The
information often has no bearing on the rest of the story, serving purely to help identify the referent (9.83). This presenting of background information is a function shared with postmodifying relative clauses, and has also been noted elsewhere for cleft constructions (Givón 1990:717; Prince 1978:902).

9.83 To'os na'in nia-kan naran N...
garden noble 3S-POS name N...
The garden owner's name was N...

Nia mak n-odi ulun iha leo K...
3S REL 3S-hold.office.of head LOC hamlet K...
He is the one who was head of K... hamlet.

Finally, cleft constructions are often used to mark the end of a narration (12 examples). Most are variations on two fairly standardised formulae, with the focused constituent referring either to the speaker's words (9.84) or to finishing (9.85).^31

9.84 Lia mak ne'e dei.
word REL this only
That is all the story. ('The end.') (R1.62)

9.85 ...ne'e mos mak nia Ibu.
this/now finished REL 3S mother
[When he had married the woman,] (the tale) is finished, Ibu. (U5.87)

9.6 'body-good' predicates

9.6.1 Basic description

Character, emotions and physical attributes are often expressed by 'body-good expressions' which are based on a noun followed by a (normally) single-word predicate (e.g. isin di'ak 'body good' = 'healthy'). The noun usually denotes a body part, and will hence be loosely referred to as a 'body part noun' throughout the discussion. However, it can also name some other inalienable characteristic, such as naran 'name' or folin 'value, price'. The predicate is usually an intransitive verb or adjective, but may also be a transitive verb (sa'e 'ascend'), noun (fatuk 'rock'), numeral (hira 'how many, how much'), adverb derived from a reduplicated adjective (di'a-di'ak 'well') or serial verb construction (la'o ba mai 'walk go come' = 'walk to and fro'; 9.95).

Predicate.Body-good → S.Body.part Predicate

Some examples of such expressions are given below. Note that most expressions are restricted to describing people and animals. The only ones in the corpus which are not normally applicable to people are those dealing with cost (folin 'value, price').

---

31 These three uses (questions and answers, contrast, end of narration formulae) together account for 85% of the 140 examples for which sufficient context is recorded. It is quite possible that many of the remaining 15% can be interpreted contrastively; however, unlike the majority of cases, there is nothing in the context to force such an interpretation.

32 This syntactic analysis is defended in §9.6.4.
Table 9.2: Examples of body-good expressions

<table>
<thead>
<tr>
<th>Expression</th>
<th>Gloss</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>abut naruk</td>
<td>root long</td>
<td>have a long life</td>
</tr>
<tr>
<td>folin hira?</td>
<td>value how much?</td>
<td>cost how much?</td>
</tr>
<tr>
<td>folin ktodan</td>
<td>value heavy</td>
<td>expensive</td>
</tr>
<tr>
<td>ibun luan</td>
<td>mouth wide</td>
<td>habitually talkative</td>
</tr>
<tr>
<td>ibun naruk</td>
<td>mouth long</td>
<td>habitually speak badly of people</td>
</tr>
<tr>
<td>inur busa</td>
<td>nose cat</td>
<td>have a sensitive sense of smell</td>
</tr>
<tr>
<td>isin di'ak</td>
<td>body good</td>
<td>healthy</td>
</tr>
<tr>
<td>isi karuk</td>
<td>body left</td>
<td>left-handed</td>
</tr>
<tr>
<td>kulit talas tahan</td>
<td>skin yam leaf</td>
<td>thin-skinned, easily angered or shamed</td>
</tr>
<tr>
<td>laran di'ak</td>
<td>interior good</td>
<td>have good character</td>
</tr>
<tr>
<td>laran kle'an</td>
<td>interior deep</td>
<td>have a strong stomach, able to eat anything</td>
</tr>
<tr>
<td>laran moras</td>
<td>interior sick</td>
<td>sad, upset</td>
</tr>
<tr>
<td>lian kmeik</td>
<td>voice sharp</td>
<td>have a squeaky voice</td>
</tr>
<tr>
<td>lisam at</td>
<td>character bad</td>
<td>have bad character</td>
</tr>
<tr>
<td>matan delek</td>
<td>eye blind</td>
<td>blind</td>
</tr>
<tr>
<td>matan wa'i</td>
<td>eye grow</td>
<td>wide awake</td>
</tr>
<tr>
<td>meti sa'e/tasi sa'e 33</td>
<td>sea ascend</td>
<td>(get) angry (of nobles)</td>
</tr>
<tr>
<td>meti maran/tasi maran</td>
<td>sea dry</td>
<td>dead (of nobles)</td>
</tr>
<tr>
<td>naran X</td>
<td>name X</td>
<td>have the name X</td>
</tr>
<tr>
<td>nawan åt</td>
<td>breath bad</td>
<td>bad-tempered, quick to anger</td>
</tr>
<tr>
<td>nawan n-akraik</td>
<td>breath 3S-lower</td>
<td>calmed down, no longer angry</td>
</tr>
<tr>
<td>nawan naruk</td>
<td>breath long</td>
<td>able to hold one's breath for a long time</td>
</tr>
<tr>
<td>neon di'ak</td>
<td>emotion good</td>
<td>content, happy</td>
</tr>
<tr>
<td>oin åt</td>
<td>face bad</td>
<td>ugly-faced</td>
</tr>
<tr>
<td>oin n-alai</td>
<td>face 3S-run.away</td>
<td>dizzy</td>
</tr>
<tr>
<td>otas badak</td>
<td>age short</td>
<td>short-lived (will die young)</td>
</tr>
<tr>
<td>ulun fatuk</td>
<td>head stone</td>
<td>stubborn</td>
</tr>
<tr>
<td>ulun moras</td>
<td>head sick</td>
<td>have a headache</td>
</tr>
</tbody>
</table>

9.86 Ema  ne'e  lian  kwana.
This person has a lovely voice.  

9.87 Sia  bá  r-usu  ferik  uan  na'in  rua:  ida  matan
3P  go  3P-ask  mature.woman  small  CLS:human  two  one  eye
át, ida  tilun  diuk.
bad  one  ear  deaf
They went and asked two old women, one blind and one deaf.  

9.88 Fatin  ne'e  naran  To'ös  Lalawar.
place  this  name  garden  house  garden
This place was called Garden of Lalawar.  

33 Both meti sa'e, tasi sa'e 'get angry' and meti maran, tasi maran 'dead' are synonymous pairs used in ritual or poetic language when speaking of nobles.
Mr goat tooth DEF dry die already

[Once the fire roasting him had died down,] Mr Goat’s teeth were dry—(he) had died.

(J1.39)

Some such expressions have a common metaphorical meaning, but can also be used literally, as illustrated in the following examples.

Table 9.3: Examples of literal and metaphorical interpretations of body-good expressions

<table>
<thead>
<tr>
<th>Expression</th>
<th>Gloss</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>isin sa'e</td>
<td>body ascend</td>
<td>1. grow physically 2. fevered</td>
</tr>
<tr>
<td>kulit karau</td>
<td>skin buffalo</td>
<td>1. have buffalo-like skin 2. thick-skinned, do not know fear or shame</td>
</tr>
<tr>
<td>liman kmetis</td>
<td>hand tight</td>
<td>1. have arms tied up tightly 2. tight-fisted, stingy</td>
</tr>
</tbody>
</table>

Others appear unable to be used literally. Thus laran moras (lit. ‘interior sick’) means ‘sad, upset’, and cannot be used to speak of physical pain inside one’s body.

9.6.2 Lexicalisation and productivity

Many body-good expressions are the usual means of describing the particular condition. The question arises as to whether they are lexicalised. Pawley (1986), using examples from English, proposes a number of tests to determine to what extent an expression is lexicalised. Tetun body-good expressions pass some of these tests, particularly those relating to their resemblance to single-word lexemes. In particular, they belong to a terminology set along with single-word members (adjectives), and can have single-word synonyms and antonyms. This is demonstrated by the examples in the following table.

Table 9.4: Examples of near synonyms: body-good expressions and adjectives

<table>
<thead>
<tr>
<th>Expression</th>
<th>Gloss</th>
<th>Translation</th>
<th>Adjective</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>isin manas</td>
<td>body hot</td>
<td>fevered</td>
<td>moras</td>
<td>sick</td>
</tr>
<tr>
<td>matan dukur</td>
<td>eye close</td>
<td>sleepy</td>
<td>hakati</td>
<td>nodding off</td>
</tr>
<tr>
<td>mata metan</td>
<td>eye dark</td>
<td>dead (of commoner)</td>
<td>mate</td>
<td>dead, die</td>
</tr>
<tr>
<td>isin kreon</td>
<td>body thin</td>
<td>thin</td>
<td>kreon</td>
<td>thin</td>
</tr>
<tr>
<td>nawan sa'e (X)</td>
<td>breath ascend</td>
<td>get angry (at X)</td>
<td>krakat (X)</td>
<td>angry (at X)</td>
</tr>
<tr>
<td>neon kiki</td>
<td>emotion tremble</td>
<td>be afraid</td>
<td>hata'uk (X)</td>
<td>fear (X)</td>
</tr>
</tbody>
</table>

Pauses do not occur within these expressions in fluent speech, although as we shall see below certain closed classes of words (such as negators and auxiliaries) can intervene between the two parts of the expression.

A syntactic similarity to single words is that no part of the expression can be omitted; in particular the body part noun cannot be omitted, in contrast to subjects in other contexts which often are omissible. When modifying an NP, these expressions usually are not introduced by the relative clause marker *mak* (9.102). In this they are like single-word adjectives and unlike clausal modifiers.
Semantically, too, many body-good expressions are lexicalised to a significant degree, since they represent the standard means of expressing certain concepts, and in certain instances it is a metaphorical rather than literal interpretation which is the conventional (or only) meaning attributed to the expression.

While many body-good expressions are standardised, this construction is productive. For instance, there is a story in which a suitor thinks he is in the company of the noblewoman whom he wants to court. When she doesn’t reply to his conversation, he taps her on the head. Since the head is in fact a statue and not the real person, it shatters and the red palm juice which had been placed inside spills out. He tastes the juice and exclaims that the blood is sweet. Presumably ‘head shatter’ and ‘blood sweet’, which both fit this construction and which are both interpreted literally, are not institutionalised expressions.

Evidence that such expressions are not syntactically compounds comes from the constituents that can intervene between the body part noun and the predicate (see below), as well as from causatives. Just as for single-word adjectives and verbs, causatives can be expressed either periphrastically using the verb halo ‘make, do’ (9.91) or (for some expressions) by the causative prefix ha- (9.92). In the latter case the word order of the expression is reversed (contrary to what one would expect of compounds), with the causative prefix ha- attaching to the verb, and the body part noun following the verb as its object.

We turn now to the grammatical status of the two NPs in body-good constructions. For convenience these will be referred to by the loose semantic labels of ‘experiencer’ and ‘body part’. There are several properties which ambivalently point to either the body part as subject and the following word(s) as predicate, or alternatively point to the experiencer as subject and the whole body-good expression as predicate.

In the first place, there are a number of words which elsewhere must be placed before the predicate head. In these constructions they can be placed either before the body-good expression as a whole or before its predicate. These words include the auxiliaries atu ‘IRR’ (9.93), keta ‘do not’ (9.94, 9.95), and sei ‘still’, and the adverbs hetak ‘increasingly’ (9.96).
and bei ‘also’. The difference in position does not affect the basic meaning; it is unclear what difference, if any, it does make.

9.93  
\[\text{Nia at nawan so'e onan.} \approx \text{Nia nawan at so'e onan.}\]
\[3S \text{IRR breath ascend IMM} \quad 3S \text{breath IRR ascend IMM}\]
He is about to get angry.  
(X0.123 elicited)

9.94  
\[\text{Keta neon kadolik.}\]
\[\text{do.not emotion tremble}\]
Don't (let your) heart tremble.  
(C2.64)

9.95  
\[\text{Emi neon keta la'o bâ ma.}\]
\[2P \text{emotion do.not walk go come}\]
Don't (let) your heart walk to and fro. (i.e. Don't be afraid.)  
(C2.12)

9.96  
\[\text{Nia hetak isin kreon.} \approx \text{Nia isin hetak kreon.}\]
\[3S \text{increasingly body thin} \quad 3S \text{body increasingly thin}\]
He got thinner.  
(V0.36, X0.122 elicited)

Secondly, subject marking, where applicable, usually agrees with the body part noun (9.97; 16 examples). However, there is one example in the corpus in which it agrees with the experiencer (9.98).

9.97  
\[\text{Ha'u isin n-anusa, tân malirin.}\]
\[1S \text{body 3S-spotty because cold}\]
I have goose bumps, as it is cold.  
(L0.34)

9.98  
\[\text{Ha'u, ha tilun diuk, ha matan k-aré má, mane}\]
\[1S \quad 1S \text{ear deaf 1S eye 1S-see and.then man}\]
\[\text{ida di'ak basuk.}\]
on one good very
I—I am deaf, (but) my eyes see, and (I saw) a wonderful man.  
(U4.39)

Thirdly, the floating quantifier hotu ‘all’ (§6.4) can quantify either the body part noun (9.99; 3 examples, all with a plural marker on the body part) or the experiencer (9.100; 1 example).

9.99  
\[\text{Nia tanis daudaun to'o lian sia bûs hotu.}\]
\[3S \text{cry continue until voice PL hoarse all}\]
He cried on and on until his voice was all hoarse.  
(X0.15)

9.100  
\[\text{Feto mane sei laran moras hotu.}\]
\[\text{woman man still interior sick all}\]
(We) women and men are all still sad.  
(IO.131)

The ability to relativise on the experiencer supports analysing the experiencer NP as subject, since subjects can be relativised but topic NPs which precede subjects can not.  

---

34 These could alternatively be analysed as relativisation on a possessor constituent within the subject (‘people whose legs are crippled’). However, since the only clear examples of this construction involve body-good expressions, there is, in the absence of further supporting evidence, no justification for introducing this category of relativisation.
9.101 *ema mak aîn ât, matan ât.*
person REL leg bad eye bad
people who are crippled (or) blind

9.102 *Ema hûn di’ak tuir adat.*
person origin good follow tradition
People who have good backgrounds follow tradition.

One word which is restricted to occurring within the body-good expression is the negator *la.* It cannot precede the expression as a whole (*ita la abut naruk* ‘we not root long’ = ‘We (are) not long-lived’).35

9.103 *Ô la bele fiar Pâk, tán nia lisan la di’ak.*
2S not can believe Mr[Mly] because 3S character not good
You cannot trust Pak, as he is of bad character.

9.6.4 Analysis

9.6.4.1 Analysis as clausal predicate

In accordance with these facts I propose analysing body-good expressions as clauses consisting of body part subjects and simple predicates. Such a clause may itself function as predicate in a larger clause, whose subject refers to the experiencer (who is the possessor of the body part).36 This analysis is consistent with the ambivalence of the properties discussed above, which sometimes point to the body part as subject and sometimes the experiencer. It is also consistent with the lexicalised nature of many body-good expressions, in that a complex lexeme forms a single syntactic constituent. The proposed analysis is schematised below.

A small minority of examples which use these body-good expressions have a clearly different structure, in which the experiencer NP has possessive marking and so forms a single NP constituent with the following body part noun.

---

35 Note that *la* does not readily negate nominal predicates either.

36 This analysis was suggested to me for Tetun by Robert Early (pers. comm.). He first proposed it in his analysis of the Vanuatu language Lewo (Early 1994b:70).
There are many examples for which either of these analyses is possible. That is, the body part could be analysed either as forming a constituent with the preceding possessor (as represented below in 9.106), or as a separate NP representing the subject of the overall clause (as depicted above in 9.104).

This structural ambiguity comes about because inalienable possession can be indicated simply by juxtaposition of a pronominal possessor and a possessed NP, without any overt possessive marking. In addition, genitive -n (the final consonant of ne’an ‘tooth’), which marks possessed NPs, is frequently found on body part terms even when they are not in a possessive construction (§7.3.3.3), and so does not conclusively mark these as possessive constructions.

Despite this structural ambiguity several factors point to a normal reading of most such examples as involving two separate NPs. In the first place, intonation frequently marks the two NPs as separate. Secondly, although full pronouns are acceptable as syntactically juxtaposed possessors, short forms of the pronoun or pronouns followed by a possessive clitic are statistically much more common for possessor NPs. Most examples with pronouns followed by body-good expressions, in contrast, use the full form of the pronoun. Finally, while indisputably possessed body parts (as in 9.105) are quite uncommon for body-good expressions, there are a significant number of examples in which the two NPs are indisputably separate. This comes about when an auxiliary (9.100) or adverb (9.107) intervenes between the two. In conclusion, while recognising the structural ambiguity of many examples, analysis of the construction as incorporating two separate NPs is favoured unless there is evidence to the contrary.

Note that the ambivalent distribution of subject properties amongst the experiencer and body part NPs is unsurprising if this construction is undergoing grammaticisation, with subject properties in the process of being shifted from one NP to the other.

### 9.6.4.2 Arguments against analysis as subject incorporation

The analysis proposed above for the body-good construction is of a clause functioning as predicate within another clause. An alternative analysis for this construction is that the body part term is incorporated into the subject. Most examples in the corpus fit Mithun’s (1984) ‘type 2’ noun incorporation. In this type an argument is incorporated into the predicate. The position which it vacates is then filled by another argument which would otherwise not be a core argument. Most of her examples involve object NPs, such as the following from

---

37 Klamer (1994:297ff.) records somewhat similar expressions for the Central Malayo-Polynesian language Kambera, and proposes an incorporation analysis for some of the data. Hyman (1977) analyses comparable data in the Bantu language Haya in terms of possessor raising to subject.
The clause

Blackfoot: ‘I-break-back-him that man’ = ‘I broke the man’s back’. Her one example of a subject being incorporated is also from Blackfoot: ‘I-DUR-pain-back’ = ‘I have a backache’. As illustrated by these examples, a high proportion of such incorporated nouns are body part terms, whose possessors are given core argument status. Mithun (1984:856) describes the features common to such incorporation:

In all of them, a V stem and a N stem are combined to form an intransitive predicate denoting a unitary concept. The compound is more than a description; it is the name of an institutionalized activity or state. The [incorporated noun] loses its individual salience both semantically and syntactically. It no longer refers to a specific entity; instead, it simply narrows the scope of the V. It is thus unaccompanied by markers of definiteness or number, or by demonstratives. Although it may function semantically as a patient, location, or instrument, it has no independent syntactic role in the sentence as a whole, and so is unmarked for case.

Many Tetun examples clearly fit this description. Syntactically this is so in that the experiencer has many subject properties, and the body part is usually not marked for definiteness or number. Semantically this analysis fits in that the construction is used primarily for standardised expressions while also being productive.

Nevertheless, the fit between this construction and an incorporation analysis is not perfect. The body part term has more subject properties than would be expected for subject incorporation, in that subject marking normally identifies it as the subject and the behaviour of a number of auxiliaries and adverbs is ambivalent as to whether the body part is subject.

In addition, it is possible, though not very common, for the body part term to be modified by a determiner (9.89, 9.99, 9.108; 19 examples). While some of these could be analysed as having experiencer and body part together forming a single subject NP, others (such as 9.108, with a determiner on the experiencer NP) cannot readily be analysed as such.

The fact that the body part noun is usually not modified (e.g. by demonstratives, plural markers or adjectives) is consistent with a crosslinguistic tendency for nouns referring to ‘attached’ body parts to lack such modification, and so to be less prototypically nominal. Hopper and Thompson (1984:726) argue that this is because in discourse body parts are not in general autonomous, discourse-salient entities. Since they are physically undifferentiated from their ‘possessors’, body parts are treated in grammar and discourse as dependent, non-individuated entities. Typically they are not participants in the discourse as distinct from their ‘owners’.

The one expression for which a subject incorporation analysis is in principle not feasible is that which attributes a name. Here the subject of the expression is the noun naran ‘name’, and the predicate is a proper noun. It is hard to see how naran could be incorporated into a proper noun.

9.108 Feto á naran á Bita Nahak.
woman DEF name DEF Bita Nahak
The girl’s name was Bita Nahak. (O7.38)

9.6.4.3 Arguments against analysis as topic-comment

A second alternative analysis to consider for this construction is as a topic-comment construction, in which the comment is itself a clause. Andrews (1985:77f.) gives a standard definition of such a construction as follows: “In topic-comment articulation there is usually one NP, the topic, which indicates what the sentence is about. The remainder of the sentence, the comment, provides information about the topic. By definition, topics are given”. 
Li and Thompson (1976) distinguish between 'subject-prominent' and 'topic-prominent' languages. In the former, subject-predicate constructions are dominant (as in English), while in the latter, topic-comment constructions are basic (as in Chinese). All languages which they investigated have a topic-comment construction. The difference is that in subject-prominent languages the topic is highly marked and set off from the rest of the sentence, while in topic-prominent languages it is integrated into the basic syntax.

Tetun does have a number of topic-comment constructions typical of subject-prominent languages. That is, the topics are pragmatically marked, often set off phonologically from what follows, and syntactically quite restricted. They are in nearly all cases topicalised objects (§9.4.3.3) or coreferential with subject pronouns (left dislocation). Only seldom does a topic NP precede a clause without relating to it in one of these two ways (§14.2.2).

The initial experiencer NP in body-good constructions is clearly not a pragmatically marked 'topic'. Is it then a topic such as is found in topic-prominent languages? While Tetun is not fully topic-prominent, it does have a number of characteristics which Li and Thompson associate with topic-prominent languages. In particular, passivisation is absent, there are no dummy subjects, and there are so-called 'double subject' constructions, such as this body-good construction and clausal possessives (§9.4.7.5). Such constructions were not found in any pure subject-prominent languages.

If such an analysis were accepted, the experiencer would be a topic NP, and the body part a subject. The experiencer does in fact fit some of the criteria listed by Li and Thompson as diagnostic of topics. In particular, the topic is definite (the corpus containing no indefinite examples) and occurs in initial position, while the subject (usually) controls verb agreement. The topic is not usually set off from the comment by a major intonation break. Nevertheless, we have seen that a number of grammatical rules (such as positioning of auxiliaries and adverbs) are able to treat either the experiencer or the body part as subject.

In short, then, the experiencer NP clearly has a grammatical role in the sentence other than what can be expected of grammatical topics. While body-good constructions can of course be analysed as topic-comment constructions on a pragmatic level (as for instance in Halliday's (1985:36ff.) Theme-Rheme analysis of clause as 'message'), there is no justification for analysing it this way syntactically.

9.7 Dependent clauses

9.7.1 Introduction

There are two clause types which are restricted to syntactically dependent contexts, and which have reduced syntactic possibilities relative to the types of clauses discussed so far. These two types are labelled 'reduced complement clauses' and 'small clauses'.

Necessarily dependent clauses which are less limited in their core structure, and which are discussed elsewhere, are clauses introduced by subordinating conjunctions (§14.7), relative clauses (§14.8), and nominalised clauses (§14.6).

9.7.2 Reduced complement clauses

Reduced complement clauses are found only as complements of certain predicates. They consist simply of a predicate, with their tense-aspect options restricted to irrealis 

\[ au \] by virtue of the fact that their time reference is dependent on that of the main clause. Reduced complements are further discussed and exemplified in §13.4.
9.109  *Tama ti’a, na’in mane n-akara atu sukai sira.*
enter already lord male 3S-like IRR measure 3P
(When she) had come in, the nobleman wanted to test them.  (AA1.96)

9.110  ...*Só m-aruka ó ata-n sia bá r-ola h ina*
unless 2S-order 2S slave-GEN PL go 3P-take 1S mother
*knusuk á...*
noodle DEF
[You can’t kill me.] Unless you order your servants to go and fetch my
mother’s weaving needle [; then if you use it to kill me, I will die.]  (F0.120)

9.7.3 Small clauses

A small clause consists minimally of a single verb (9.111). It appears to consist maximally
of a subject and verb (9.112), or a verb and following object (9.113).\(^ {38} \) In small clauses
functioning as fronted complements the verb takes subject marking (9.111). Data are
insufficient to be certain about subject marking in other small clauses. There is no adverbial
modification, and the subject and object NPs nearly always consist of only a single word.

\[
\text{Clause. Small} \rightarrow \begin{cases} (S) & V \\ V & (O) \end{cases}
\]

There are four constructions in which embedded clauses must have the structure of a small
clause. In two the small clause functions as complement, namely a complement which
precedes the complement-taking verb (9.111; §13.4.3.3), and a clause functioning as subject
or object of the verb *kona* ‘undergo, touch’ (§9.4.6).

9.111  *Ha’u k-akes la k-atene.*
1S 1S-talk not 1S-know
I don’t know how to talk (about this topic).  (10.123)

In the remaining two constructions the small clause functions as modifier within an NP. In
particular they occur as premodifying relative clauses (9.112; §7.4.4), and as clauses which
postmodify NPs, but which (unlike most relative clauses) have no syntactic slot within the
clause with which the head noun can be interpreted as coreferential (9.113; §14.8.3.3). In
both these constructions the resulting NP normally functions as a conventional name for the
class of referents it describes.

9.112  *loro ga’e fatik*
sun rise place
place where the sun rises

9.113  *besi tá rai*
iron/tool chop earth
broad hoe (tool used for hoeing)

\(^ {38} \) Note that these ‘small clauses’ are rather different to English constructions which go under the same name
(Cardinaletti & Guasti 1995; Matthews 1981:324).
9.8 Peripheral constituents

9.8.1 Overview

Peripheral constituents are those constituents (other than verbs and adverbs) which are not determined by the valency of the verb. A peripheral constituent can thus be added to any sentence in which it makes sense, is always optional, and has no collocational restrictions with the predicator (Matthews 1981:130).

Peripheral constituents include time, location, beneficiary (§8.6.1, §12.5.3.6), expressions of similarity (§8.5), purpose (§14.7.6) and reason (§14.7.5). Of these, time and location are the most common; since other types of peripheral constituents are discussed elsewhere, only these two will be considered here.

Peripheral time and location expressions tend to occur clause-initially when they provide information about setting. This tendency is stronger for time than for location. Other peripheral constituents are normally located after the clause core.

Initial peripheral expressions can be, though often aren’t, separated from the rest of the clause by a pause. When they follow the clause core there is no separating pause.

9.8.2 Time

Time expressions consist of an NP with a time noun as head\(^{39}\) (e.g. \textit{loron ida} ‘day one’ = ‘one day’, \textit{sura kalan} ‘every night’, \textit{awan} ‘tomorrow’), a clause, or an NP introduced by a preposition (\textit{bd loron ida} in 9.118; §8.4) or prepositional verb (§12.5.5.5). They are always peripheral.\(^{40}\)

Most time expressions provide information about the setting and occur clause-initially, following any clause-joining expressions.

\begin{verbatim}
9.114  Hotu ma sesawan na’in mane né ba. 
       finish and.then morning noble man this go 
       Then in the morning the nobleman went. (O2.53)
\end{verbatim}

Where time is itself focused on as new information, limited evidence suggests that it tends to, but need not (9.116), occur clause-finally. This tendency holds for questions about time (9.115), answers to such questions (9.115), and other instances in which the whole clause except the time expression is already known information.

\begin{verbatim}
9.115  “Ni ina-n nó na ama-n mate hori.hirak?”
       3S mother-GEN and 3S father-GEN die when(past)
       “When did his mother and his father die?”

       “Mate kleur ti’an. Mate tinan ida ti’an.”
       die long.time already die year one already
       “(They) died a long time ago. (They) died a year ago.” (N0.184)
\end{verbatim}

\(^{39}\) The justification for treating these time words as a subclass of nouns is presented in §3.3.

\(^{40}\) That is to say, the time expression as a whole is peripheral within the clause; of course a time NP introduced by a prepositional verb is itself object of the prepositional verb. In addition to peripheral constituents, time may be indicated by temporal adverbs (e.g. \textit{kala-kalan} ‘RDP-night’ = ‘nightly’).
The third slot for time expressions is between the subject and the verb. The great majority of time expressions in this position are single nouns or short NPs (e.g. *sura loron* ‘every day’) which indicate frequency.

The clause (Z2.33)

9.116 *Wa hira* kтомak?

When will (it) be complete?

Daily he doesn’t work; (he) just hangs around the house. (L0.79)

9.8.3 Location

9.8.3.1 Difference between core and peripheral location

Location can be part of either the periphery of a clause or its core (as an ‘oblique argument’). Core location is part of the valency of the verb, being in some way essential to its meaning. It always follows the verb (and object, if any). It often identifies a direction of motion rather than a static location, and can give the location/direction for either the actor (e.g. for motion verbs like *ba* ‘go’) or the undergoer (e.g. for *tau* ‘put’).

In contrast, peripheral location specifies a static setting. As such the actor is normally entirely situated within that location. Peripheral location can occur either initially or finally within the clause.41

9.8.3.2 Peripheral location

Peripheral locatives are introduced by the locative preposition *iha*. They mostly follow the clause core (9.118), and appear to always do so if it is the location that is questioned (by *iha nabe* ‘LOC where’ = ‘where’). Locatives can, however, be placed clause-initially if they provide information about the setting (9.119) or are contrastive (second clause in 9.120). Occasionally they are situated between the subject and the predicate (first clause in 9.120); however, in this slot it is not always clear whether the locative modifies the subject NP, or is peripheral to the clause as a whole.

The clause (K12.11)

9.118 *Bá loron ida sia r-afaho mo’at *iha to’os.

One day they were weeding in the garden.

9.119 *Iha vai Uma Katahan ne’e, ema mak soi, ema kmesak basuk.*

Here in Uma Katahan, people who are rich are very scarce.

41 These differences are succinctly spelled out as crosslinguistic generalisations by Pike and Pike (1982:45f.).
208 Chapter 9

9.120 Gaja iha Java bā n-akiak ti’an. Iha alas laran sta krakat.

Elephants over in Java have been domesticated. In the jungle they are wild.

(D0.163)

9.9 Non-declarative clauses

9.9.1 Introduction

Non-declarative clauses are clauses which are formally marked for non-declarative mood. As main clauses, their primary use is for performing speech acts other than making statements. In Tetun one can formally distinguish three types of imperatives, as well as information interrogatives, disjunctive interrogatives (including polar ones), and greetings.42 I have no evidence for exclamative mood or other minor sentence types.

9.9.2 Imperatives

Imperative clauses are characteristically used for issuing commands and requests. There are three distinct markers of imperative mood in Tetun, none of which occur in dependent clauses in the corpus. One marker of imperatives is bā ‘IMP’, which is discussed in §11.9. The auxiliary keta ‘do not’ is used in prohibitions (i.e. negative imperatives), as well as in negative purpose clauses (§10.4.4).

A third marker, found in hortatives, is an initial mai ita ‘come 1PI’ = ‘let us’ before the predicate (8 examples). Such sentences are used for suggestions that the addressee do something together with the speaker.

9.121 Ema tene ita r-āk “Mai ita bā nebā”.

People invite us saying “Let’s go over there”.

(J2.22)

However, in Tetun most clauses used for commands and requests have no overt marker of imperative mood, being formally declarative. If the actor is the addressee(s), and any specified tense/aspect is compatible with interpretation as a command (e.g. by not indicating past time), then the pragmatic interpretation is that this is some sort of request. The force may vary from strong commands to polite invitations. Just as in other clauses, the subject is optional.

9.122 Mane Kmesak. Ō m-ādēr onan.

Only Man. You get up now.

(P1.57)

42 As is common crosslinguistically (Sadock & Zwicky 1985:183f.), there is no formal basis for treating polar and disjunctive interrogatives as belonging to a superordinate ‘interrogative’ class along with information interrogatives. They do, however, have in common the fact that the same markers are used both for interrogatives and for indefinite statements. In particular, epistememes (e.g. sā ‘what, anything, something’) are used both in information questions and as indefinite constituents in statements, while kā ‘or’ is used in both interrogative and non-interrogative disjunctives.
9.123  *Te'in m-ola kbú fuan rua lai.*
      cook 2S-take riceparcel CLS:round two first
Cook two rice parcels now.

Degrees of urgency can be suggested by the adverbs *lai* ‘first’ (9.123; §11.5.6) and *onan* ‘imminent, now’ (9.122; §11.5.4). There are no morphological markers of politeness corresponding to English ‘please’.

9.9.3 Information interrogatives

9.9.3.1 Description of information interrogatives

Information (‘wh’) interrogatives make use of words from the class of epistememes (e.g. *sé* ‘who’), which are listed and discussed in §6.6. These fill the same syntactic slot in the question as their answer would in the corresponding statement. They can be placed in the focus slot in a cleft construction, but need not be. In the corpus such focus slots account for only 18% of the 67 questions for which the question word is subject, and for only four non-subject question words. Information questions do not take tags.43

All major slots in a sentence can be questioned. This includes all core arguments (9.124), as well as object of preposition, location (e.g. *iha nabe* ‘LOC where’ = ‘where’), possessor (9.126), time (horihirak ‘when (past)’, wainhira ‘when (future)’), numeral and quantity (hira ‘how many, how much’), reason (tán sá, bá sá ‘why’), the second NP in a coordination (... Nona n s é ... Nona and who’), predicate (halosá ‘do what’; 9.127), and type (oin sá ‘face what’ = ‘what’). In addition, there is a general question ná nabé ‘like which/where’ = ‘how’.

9.124  “Ai. Katuas m-á sá?”  N-ák  “Ou, k-á asu
      EXCL mature.man 2S-eat what 3S-say EXCL 1S-eat dog
ha'ú-k  ten  ne”.
      1S-POS excrement this
[Then the Chinese said] “Hey. What are you (old man) eating?”. He said “Oh,
I’m eating my dog’s excrement”.
      (T0.72)

9.125  “Lakon bá loron sá? Lakon bá loron nábé?”
      disappear at(time) day what disappear at(time) day which
      “On what day did (the buffalo) disappear? On which day did (it) disappear?”
      “Lakon bá loron basar Betun. basar Betun kalan.”
      lose go day market Betun market Betun night
      “(It) disappeared on the Betun market day (i.e. Monday)—on the night of the
      Betun market day.”
      (H1.8)

43 There is one sentence in the corpus which is syntactically an information interrogative, but which also has a
tag. Its function, however, is as a rebuke rather than as an information-seeking question. It thus
demonstrates the oft-emphasised distinction between the grammatical notion of interrogative clause (with
which a tag is possible), and the pragmatic notion of information question (with which a tag appears
impossible) (Huddleston 1994; Searle 1975).
Eman n-ata “Kabau mak mak-lakon ne’e, se person that 3S-answer buffalo REL who-disappear this who nia-k?”.

The person replies “Whose is this buffalo that disappeared?”. (H1.19)

Polisi n-ata “Kabau ne’e Si Klau nia-k”.

The policeman replies “The buffalo is Si Klau’s”. (H1.20)

“Manek bá alas ne’e bá n-alo sá?” “Nia bá tá ai.”

Manek go forest this go 3S-make what 3S go chop plant “Manek going to the forest—what is (he) going to do?” “He is going to chop wood.” (H0.163)

Consultants did not approve my attempts at questions with two question words (‘Who stole what?’, ‘Who went where?’). However, this could be a result of pragmatic rather than syntactic constraints, since the corpus contains four unelicited examples in which separate question words were used to simultaneously ask the identities of two opposing competitors.

Sia kalera sita; sé at n-odi sé?

They have a horserace; who will defeat who? (H15/5)

9.9.3.2 Answers to information questions

Answers to information questions can consist of either a single NP (9.129) or a fuller reply (9.81, 9.124).

“Sé ida ema tür ne’e?” … “Ha’u, na’i”.

“Who is this person sitting here?” [The woman replied,] “Me, Sir”. (K9.93)

When the answer is not known, the respondent can echo the question, adding ruma after the question word (9.130; 7 examples). This implies ‘How would I know?’. Idle questions, which air one’s wondering without expecting an answer, can similarly place ruma after the question word (9.131; 5 examples).

“Rai Ulun iha nabé?” “Rai Ulun ne’e at iha earth head LOC where earth head this IRR LOC nabé ruma.”

where unknown

“When is Earth Head?” “I don’t know where this Earth Head would be.” (P0.110)

Nona to’o nabé ruma?

Nona reach where unknown

Where could Nona have got to? (Said in idle wondering.) (T0.123)
9.9.4 Disjunctive and polar interrogatives

9.9.4.1 Description of interrogatives

Disjunctive interrogatives consist of two or more alternatives, usually separated by the conjunction ká ‘or’. An open disjunctive, in which only some of the alternatives are listed, has ká after the final disjunct as well. These constructions are discussed and exemplified in §14.3.3.

Polar (‘yes-no’) interrogatives are semantically and syntactically a special subclass of disjunctive interrogatives, in which one alternative is the negation of the other. The two alternatives can be fully specified (9.132). More commonly the negative alternative is shortened to ká lale ‘or not’ (9.133), or simply lale ‘no, not’ (9.134), or is left open by a final ká ‘or’ (e.g. Tebes ká? ‘Truly?’).

\[9.132\] ...at sei suk {n}a lai: nia na’ \(\text{in} \) ká na’ \(\text{in} \) ha’i?
IRR still measure 3S first 3S noble or noble not
(I) would like to test him first: is he a noble, or isn’t he a noble? (S2.170)

\[9.133\] N-ák “Bele? Ita ruas suk {t}a malu bele ká lale?”.
3S-say can 1PI two measure each other can or no
N-ák “Bele”.
3S-say can
(He) said “Can (we)? Can we two test each other, or not?”. (She) said “(We) can”.

\[9.134\] Ó m-akara lale?
2S 2S-like no
Do you like (him or) not?

Finally, the alternative can be totally unspecified, with final rising intonation alone signalling that this is a question (9.135, Bele? in 9.133). This is relatively uncommon for unbiased information-seeking questions, being found more often in biased or rhetorical questions (9.139) and complaints (9.141).

\[9.135\] At hanawa-n?
IRR stop-IMM
Do (you) want to stop?

9.9.4.2 Answers to polar questions

Answers to polar questions normally include an echo of the question or of some part of it (including at least the auxiliary or predicator; 9.133, 9.136).

\[9.136\] ...N-á’ak “Lale. Ha’u k-subar ha’i”.
3S-say no 1S 1S-hide not
[She said “Did you take and hide my thing?”:] (He) said “No. I didn’t hide (it)”.

For questions of positive polarity, it is possible to preface a positive answer with hou (é), a response also used to agree to suggestions or statements. Negative answers to positive

\[44\] The final type of polar question is thus syntactically a declarative clause.
questions can begin with *lale 'no' (9.136), a reply also used to contradict a previous statement, to reply negatively to a suggestion, or to contradict the assumptions underlying a preceding statement or question. Both *hou and *lale always initiate a speaker's turn. The expressions *hou 'yes' and *lale 'no' cannot be used in replies to questions with negative polarity (9.137).

9.137 "Ó la bá sekola ká?" "K-bá ha'i."
2S not go school[Mly] or 1S-go not no
(Said to child playing:) “Didn’t you go to school?” “(I) didn’t go.”

(L0.36 elicited)

9.9.5 Speech acts using polar interrogatives

9.9.5.1 Question bias

One function of polar interrogatives is to genuinely ask a question without indicating a bias as to which answer is expected. Sentences which specify the positive alternative followed by the negative one (whether fully or by means of *ka *lale 'or not' or *lale 'no') appear to always be such genuine, unbiased, questions. In addition, when interrogatives with positive polarity and no explicit negative alternative are used to ask genuine questions (as opposed to being used for rhetorical questions or indirect speech acts), they are always unbiased (e.g. *Bele? ‘can’ = ‘Is it possible?’).

Two syntactic factors bias a question in favour of an answer agreeing with the polarity of the question. These are use of the tag to (§14.2.5), and negative polarity in the question (9.138; 6 examples).

9.138 N-ó ha'i husu?
3-exist not request
Aren’t there any questions? (Said when it appeared that I had none, contrary to the usual pattern.)

(L0.99)

9.9.5.2 Rhetorical questions

When one asks a rhetorical question (whether positive or negative), one does not expect a reply, but rather one makes a point which has the opposite polarity to the question itself (15 examples without tag, 7 with tag ê).

9.139 ...*kabau kulit é *dade’o be’o? Be’o ha’i *Ibu.
buffalo skin this trample shatter shatter not mother
...buffalo hide (when) trampled, will it shatter? It won’t shatter, Ibu. (U1.49)

---

45 Li and Thompson (1979:202, 205) note that disjunctive polar interrogatives in Chinese are similarly restricted to 'neutral' contexts, and hypothesise that this may be universally true, since the disjunctive explicitly presents both the positive and the negative alternative.

46 Negative polarity questions are nearly always biased in English also, although the bias can be towards either a positive or a negative answer (Ladd 1981:164ff.). This is not surprising since, as Bublitz (1981) points out, the negative is “the marked form of the polarity in yes-no questions and markedness may trigger off inferences and implications on the part of the hearer”.
Like the people in former times—why wouldn’t they be rich, Ibu?

Did they have schools (which keep us poor)? [... They were rich.] (O5.125)

9.9.5.3 Other indirect speech acts

Polar interrogatives (of either positive or negative polarity) can function as complaints, presenting a proposition which is true but which is not approved of (9.141; 11 examples).

Then (Mr Turtle, who was given only one piece of banana while expecting a more generous share) said “Ah! Is mine only one piece?”.

(Mr Monkey) said “Be quiet!”.

Such indirect uses of interrogatives are of course widely recognised crosslinguistically (Anzilotti 1982:290; Brown & Levinson 1978:137ff.; Searle 1975), and it can be anticipated that many more indirect uses still remain to be found for Tetun interrogatives.

9.9.6 Greetings

Tetun greetings have the intonation of polar questions, but do not have any syntactic characteristics of polar interrogatives. In particular, tags and disjuncts are not possible on greetings. Unlike polar interrogatives, greetings express self-evident fact (e.g. Mai ti’an? ‘Have you come?’), and the only acceptable response is the acknowledgment He’e.

47 Bolinger (1978:88ff) similarly points out that tags are not possible on English interrogatives that state self-evident fact.
10 Auxiliaries

10.1 Introduction

Auxiliaries are markers of aspect or modality whose usual position is immediately preceding the verb or other predicate head.

Although auxiliaries constitute a diverse class syntactically, the various ones all share some characteristics with verbs; in particular, they have much in common with complement-taking verbs with reduced same-subject complements, such as *hakara* 'like, want' (§13.4.3). In the following sections we consider first characteristics which are shared by auxiliaries and complement-taking verbs, then ways in which they differ, and finally ways in which auxiliaries differ from one another.

Following this discussion of the class of auxiliaries as a whole, the various aspectual and deontic modal auxiliaries will be discussed and exemplified individually.

Note that although auxiliaries are here analysed as a separate word class from verbs, they could alternatively be considered a non-prototypical subclass of verbs. From a diachronic perspective, auxiliaries in many languages are understood to have developed diachronically from full verbs, and to in turn be a source for further development into tense-aspect-mood inflections on the verb (Heine 1993:86). If Tetun auxiliaries have similarly developed from full verbs, and are still undergoing grammaticisation, then it is not surprising that they are such a mixed bag with respect to the properties they share with verbs, and in particular with complement-taking verbs.

10.2 Characteristics of auxiliaries

10.2.1 Similarities to complement-taking verbs

Auxiliaries share the following characteristics with verbs which take same-subject reduced complements. Note that for purposes of comparison, the term ‘complement’ will be used to refer to that part of the clause which follows the auxiliary, as well as to the complement of complement-taking verbs.¹

1. The usual position for both is immediately before the head of the complement (auxiliary *bele* ‘can’ in 10.1, complement-taking verb *ho*i* ‘do.want’ (to)’ in 10.2).

```
10.1 ...ha’u  la  bele  k-ola  tenik  liurai  ne’e.
       1S not can 1S-take again ruler this
[Now,] I can no longer marry this ruler.  (S2.17)
```

¹ This temporary use of the term ‘complement’ is a convenience only, and does not imply commitment to an analysis of auxiliary as head of the construction.
10.2

\[ \text{Lale ha'u k-ö'i k-ola ó.} \]
else 1S 1S-do.not.want 1S-take 2S
[If you kill your new wife and children, I'll be your wife again.] Otherwise I refuse to take you (back).
(Y0.35)

2. Those which can occur without a complement (as many can), can do so only when the context is understood (auxiliary bele in 10.3, complement-taking verb ho'i in 10.4). That is, clauses with an auxiliary or complement-taking predicate are elliptical without a complement.

10.3

\[ \text{"Ö, Maromak, ó bele tulun ha'ú ká lale?" N-ák "Bele..."} \]
oh god 2S can help 1S or not 3S-say can
"Oh, Maromak, can you help me or not?" He said "(I) can..."
(P1.19)

10.4

\[ \text{"Feto Ikun, ita hamutu mane ó-k bá."} \]
female tail 1PI together man 2S-POS go
[They said] "Youngest Sister, we'll share your man."
\[ \text{Ni Feto Ikun, nia la n-o'i.} \]
so female tail 3S not 3S-do.not.want
Youngest Sister, she refused.
(U4.75)

3. The auxiliary bele 'can' shares with the complement-taking verb hatene 'know' the ability to follow the complement under certain conditions (§10.4.2, §13.4.3.3).

10.5

\[ \text{Ita tahan la bele.} \]
1PI endure not can
We can't endure (it).
(F2.66)

10.6

\[ \text{...Ema ne'e n-akneter ema la n-atene.} \]
person this 3S-honour person not 3S-know
[This person greeted me rudely:] He doesn't know how to honour people.
(H0.134)

10.2.2 Differences from complement-taking verbs

There are, however, also a range of characteristics which distinguish auxiliaries from complement-taking predicates:

1. All complement-taking verbs which take reduced clausal complements (i.e. complements which have the same form as complements of auxiliaries) also accept NP objects (10.7). Auxiliaries, in contrast, cannot take NP objects. Some auxiliaries can be immediately followed by NPs (10.8), such an NP is, however, the complement in an equative clause, rather than object of the auxiliary.

10.7

\[ \text{Ha'ú k-ö'i mane ne'e.} \]
1S 1S-do.not.want man this
(Describing how to rudely reject a courting man:) I don't want this man.
(V0.13)
10.8 Rai sei manu matan, sei bua klaras.
earth still bird eye still betel.nut slice
The earth was still a bird eye, still a slice of betel nut. (i.e. The newly created earth was still very small.) (P1.28)

2. Complement-taking verbs can be negated by either preverbal la (10.9) or postverbal ha'i (10.10), and allow independent negation of the matrix (10.6) and the complement clause (10.11). Of the auxiliaries, it appears that only bele ‘can’ allows independent negation of the auxiliary (10.1) and the complement (10.12; see Table 10.1). When bele is followed by a complement, it can be negated only by preverbal la (10.1; *bele ha'i bá ‘can not go’).

10.9 Nia la  n-akara kahur ema Timor.
3S not 3S-like mix person Timor
She doesn’t like mixing with Timorese people. (L0.47)

10.10 ...hanoin ha'i toba.
consider not lie.down
[The busily playing child] doesn’t think to sleep. (I14/7)

10.11 Ó  m-akara at bá bele, m-akara at la bá
2S 2S-like IRR go can 2S-like IRR not go
mós  bele.
also can
If you want to go, you may, if you want to not go, that is also fine. (T0.23 elicited)

10.12 Ó buka fahi. Buka bele hetan, bele la hetan.
2S seek pig seek can get can not get
(Explaining that buka ‘seek’ doesn’t imply success:) You look for a pig—(you) look and can get (one), or not get (one). (X0.60)

3. Many of the auxiliaries can occur in pre-subject position (e.g. irrealis atu in 10.31, bele ‘can’ in 10.46, see Table 10.1). This is not possible for complement-taking verbs, nor for any other verbs apart from in existential constructions.

4. Complement-taking verbs take subject marking just like other verbs. Amongst auxiliaries, only the /h/-initial progressive auxiliary ho'i is consistently inflected, while the remaining auxiliaries vary in their subject marking possibilities, as shown in the following section.

5. Auxiliaries specify aspect or deontic modality. Verbs which allow the relevant type of complement include verbs of thinking (e.g. hanoin ‘think to, remember’), liking and wanting (e.g. bér ‘desire’), and attention (horan ‘sense, feel’). There is, however, some semantic overlap in that the irrealis auxiliary atu can have the sense ‘want to’, so overlapping in meaning with the complement-taking verb hakara ‘like, want’.

6. A semantic difference between the two classes concerns the acceptability of the irrealis auxiliary atu in the complement. The reduced complements of complement-taking verbs are frequently introduced by atu. This is not surprising in view of the fact that the time reference of reduced complements is necessarily the same as or future to the time reference of the main verb. It is highly unusual (and at least in some cases unacceptable) for atu to follow other auxiliaries. An exception is the sequence foin atu ‘only, just IRR’, which means ‘just about to’.
10.2.3 Distinctions amongst auxiliaries

Auxiliaries constitute a diverse class syntactically and semantically.

The following table shows a number of areas in which auxiliaries vary, namely subject marking, the possibility of negating the auxiliary itself and of negating the complement, whether the auxiliary places selectional restrictions on the subject in some or all of its uses, whether it can function as a single-word utterance (given appropriate context), and whether it can precede the subject. Two of these words, *fain* and *keta* are analysed as adverbs when they precede the subject; although the meanings of the auxiliary and adverb are related, the two uses are sufficiently separate in syntax and meaning to warrant dual classification. (A question mark in the table indicates uncertainty, either due to insufficient confirmation, or to conflicts between elicitation and corpus data.)

<table>
<thead>
<tr>
<th>Subject marking?</th>
<th>Can negate auxiliary</th>
<th>Can negate complement</th>
<th>Selection restrictions on subject (for some meanings)</th>
<th>Selection restrictions on subject (for all meanings)</th>
<th>Can function as one word utterance</th>
<th>Can precede subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>ho'i currently</td>
<td>Y</td>
<td>N</td>
<td>N?</td>
<td>N</td>
<td>N</td>
<td>N?</td>
</tr>
<tr>
<td>fain only just</td>
<td>?</td>
<td>N</td>
<td>N?</td>
<td>N</td>
<td>Y</td>
<td>Y: then (§11.11)</td>
</tr>
<tr>
<td>seti</td>
<td>Y?</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y: not yet</td>
<td>N?</td>
</tr>
<tr>
<td>IRR</td>
<td>-</td>
<td>N</td>
<td>Y: person</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>bele can</td>
<td>?</td>
<td>Y</td>
<td>Y: person</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>lalika(n) need not</td>
<td>N</td>
<td>Inherently N?</td>
<td>Y: person</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>keta do not</td>
<td>-</td>
<td>Inherently negative</td>
<td>Y: person</td>
<td>N</td>
<td>Y</td>
<td>Y: perhaps (§11.12)</td>
</tr>
<tr>
<td>musti must</td>
<td>N</td>
<td>N?</td>
<td>N: person</td>
<td>N</td>
<td>N</td>
<td>Y?</td>
</tr>
</tbody>
</table>

In the first six columns of the table (concerning subject marking, negation, selectional restrictions, and ability to function as a single-word utterance), a positive ('Y') answer means that the auxiliary behaves in this respect like complement-taking verbs, while in the final column, a negative ('N') answer has the same implication. In each case answers indicating verbal characteristics are highlighted by shading. As can be seen in the table, the auxiliaries range from having few properties of complement-taking verbs (*ho'i* 'currently', which is, however, the most verb-like with respect to subject marking) to having much more in common with them (*bele* 'can').

One area in which auxiliaries differ is in the ability to take subject marking. Only /h/-initial *ho'i* 'currently' consistently takes subject marking. The remaining auxiliaries do not begin with /h/, so would, if they were verbs, be expected to inflect only for 1S subjects.
Nevertheless, apart from *ho‘i*, the only other indisputably inflectable auxiliary is *sei* ‘still’. It is, however, inflected in only 20% of the 34 examples in which it immediately follows the subject pronoun *ha‘u* ‘1S’ (this position being the one in which non-/h/-initial verbs are statistically most likely to be inflected). This compares to a 49% inflection rate for non-/h/-initial verbs in this position.

There was disagreement amongst consultants over whether *foin* ‘only just’ and *bele* ‘can’ could be inflected, and this uncertainty is borne out in the corpus: three of the seven unelicited examples in which *foin* immediately follows *ha‘u* ‘1S’ are inflected, while only one of the 25 examples in which *bele* immediately follows *ha‘u (la) ‘1S (not)’ carry subject marking.

/k/-initial *keta* ‘do not’ and vowel-initial irrealis *atu* are on account of their phonology unable to take subject marking. Being trisyllabic could account for lack of inflection on *lalika(n) ‘need not’* (since consonant clusters on trisyllabic words are uncommon), while being recognised as a Malay borrowing could account for the alleged impossibility of inflecting *musti* ‘must’.

It is possible for two auxiliaries to occur in sequence. Sequences found in the corpus include *sei bele n-akes* ‘still can 3S-talk’ = ‘still able to talk’, *foin atu n-ú* ‘only just 3S-blow’ = ‘just about to blow’, and *sei n-o‘i tatiti* ‘still 3S-currently rock’ = ‘still rocking (the baby)’. Co-occurrence and ordering constraints for auxiliary sequences are not yet known.

10.3 Aspectual auxiliaries

10.3.1 Overview

Aspectual auxiliaries are syntactically optional. Aspect can alternatively be expressed by adverbs, or be left implicit (§11.5).

In the following sections the four aspectual auxiliaries are considered in turn. These are *ho‘i*, which marks progressive aspect, *foin* ‘only just’, which indicates the very recent past, *sei* ‘still’, and *atu*, which marks both imminent future and irrealis in general. The auxiliary *sei* ‘still’ participates in a number of expressions meaning ‘not yet’; these instances are discussed separately from those where *sei* occurs with positive polarity to mean ‘still’.

10.3.2 *ho‘i*: progressive

The auxiliary *ho‘i* ‘currently’ (42 examples) indicates progressive aspect. The time frame for the activity may be either the time of speaking (if one is talking about the present; 10.13) or the time of some reported simultaneous event (10.14). It is used relatively infrequently, and its absence does not indicate non-progressive aspect.

10.13  
*M-aré. Manu oan n-o‘i kokorék.*

2S-see bird small 3S-currently crow

Look. The cock (lit. ‘small bird’) is crowing. (R5.116)

10.14  
*Sikori tama bá. Té! Ema t-o‘i tanis Rainahak é.*

Sikori enter go EXCL person 3P-currently cry Rainahak this

Sikori went in. Oh no! People were crying for this Rainahak (her husband, whom she was unaware had died). (D1.56)

It has the same phonological form as a complement-taking verb meaning ‘do not want’ and a verb meaning ‘go to’.
10.3.3 foin ‘only just’

Aspectual foin ‘only just’ (46 examples) means that the situation described in the clause occurred in the very recent past. This is relative to either the time of speaking (10.15) or some other reference time (10.16).

10.15  
Ha’u  k-subar  ha’i  té  ha’u  k-foin  mai.  
1S 1S-hide not because 1S 1S-only just come  
I didn’t hide (it), because I have only just come.  (Z6.23)

10.16  
Moras  ne’e  foin  kona,  ladún  bót.  
sick  this  only just  touch  not  very  big  
(When) this illness first struck, (it) was not very serious.  (I0.102)

The form foin is also used clause-initially as an adverb to mean ‘then’ (§11.11).

10.3.4 sei with positive polarity: ‘still’

The auxiliary sei ‘still’ usually indicates that a proposition that was true in the past continues to be true. There is an expectation that it will stop being true in the future (152 positive polarity examples).

10.17  
Ket  saseni  ha’u  lai  té  ha’u  sei  k-akés.  
do not hinder 1S first because 1S still 1S-talk  
Don’t interrupt me now, because I am still talking.  (X0.137)

10.18  
Sīlū  sei  kalan.  Tuku  hát,  sīlū.  
snap  still night  hour  four  snap  
(We) pick (mung beans when it) is still night. At four o’clock (in the morning) we pick.  (U8.28)

In a minority of cases sei indicates some sort of immediate future (20 examples, of which half are from a speaker heavily influenced by the Suai sub-dialect). It tends to involve either interrupting a main activity to do another, or repeating an activity (10.19). More data are required before this description can be firmed up.

10.19  
...kberan  n-ô  ti’a  ami  sei  harík,  la’o  hika  tenik.  
power 3-exist already 1PE still stand walk back again  
[When we had rested,] (and we) had energy, we stood back up, (and) walked on again.  (Z1.18)

Sei can precede a wide range of predicates, including transitive verbs (second clause in 10.21), intransitive verbs (10.17), adjectives (first clause in 10.21), prepositional phrases, body-good expressions, and noun phrases (10.8), including time phrases (10.18).

10.3.5 sei with negative polarity: ‘not yet’

The auxiliary sei appears unable to be negated by the standard negators la or ha’i. Rather, negation falls within the scope of sei, with the resultant meaning ‘still not, not yet (but expected to come true later)’. Such negation is usually expressed by dauk (or its variant daun), which occurs only in expressions for this concept ‘not yet’, and which either immediately follows sei or follows the predicator (and object NP, if any). Various expressions meaning ‘not yet’ are listed below, where ‘V’ represents the predicator (which is
usually but not necessarily verbal), and la and ha'i are both general negators (§11.2.2). Note that where dauk co-occurs with a general negator (la or ha'i), sei can be omitted.

\[
\begin{align*}
\text{sei dauk} & \quad \text{V} & (60 \text{ examples}; 10.20) \\
\text{sei la} & \quad \text{V} & (25 \text{ examples}) \\
\text{sei dauk la} & \quad \text{V} & (10 \text{ examples}; 10.21) \\
\text{sei la} & \quad \text{dauk V} & (1 \text{ example}) \\
\text{sei la} & \quad \text{V(O) dauk} & (3 \text{ examples}) \\
\text{la} & \quad \text{V(O) dauk} & (5 \text{ examples}; 10.22) \\
\text{V} & \quad \text{ha'i daun} & (3 \text{ examples}) \\
\text{dauk} & \quad \text{la V} & (1 \text{ elicited example only})
\end{align*}
\]

10.20  
\text{Ha'\text{u} k-sei dauk k-\text{\text{"a}}, tuan b\text{\text{"o}t ida n-\text{\text{"a}}}  \\
1S 1S-still not 1S-eat important.man big one 3S-eat  \\
uluk ti'an.  \\
go.first already  \\
(Riddle:) (When) I hadn't yet eaten, an important man had already eaten first. (M0.33)

10.21  
\text{Feto né sei ki'ik, nia sei dauk la n-atene sá ida...}  \\
woman this still small 3S still not not 3S-know what one  \\
(When) the girl was still little, (and) she didn't yet know anything... (G1.4)

10.22  
\text{Nia la n-ô liras dauk.}  \\
3S not 3S-have wing not  \\
She didn't yet have wings. (O6.5)

Even without formal negation, sei has the negative meaning 'not yet' when it has no following predicator. Most such instances in the corpus have sei as the second alternative in disjunctive questions in which the first alternative contains its antonym ti'an 'already' (10.23; 15 examples), or less commonly, onan 'imminent, now'. Single word answers to such questions also mean 'not yet' (10.23).

10.23  
\text{"Oa ne'e keta ti'an ká sei?" "Sei."}  \\
child this wean already or not.yet not.yet  \\
"Has this child been weaned already or not yet?" "Not yet." (H0.28)

10.3.6 atu: irrealis

10.3.6.1 Introduction

The word atu (or at in about 30\% of occurrences, rarely ata; over 250 examples) is found in two separate constructions, in which it has different, but overlapping, irrealis meanings. In preverbal position it means 'about to, want to, intend to, in order to', while elsewhere it is a more general irrealis marker.\(^3\) The two uses will be discussed separately.

\(^3\) Thus atu is an aspectual marker in one context, and a mood marker in another. The conflation of the two is not surprising. As Lyons (1977:815) points out, "we are seldom in a position to lay claim to knowledge of the future; and it is no doubt for this reason that reference to future world-states is grammaticalized in the category of mood, rather than tense, in many languages".

The term atu appears not to be used for the following functions which Bugenhagen (1993:35–37) identifies for irrealis markers in various Austronesian languages of Papua New Guinea: hypothetical
10.3.6.2 *atu* in preverbal position: ‘about to, want to, intend to, in order to’

When it occurs in preverbal position, *atu* indicates proximal future aspect (or ‘prospective aspect’ as Comrie (1976:64) calls it), and can be variously translated as ‘about to, want to, intend to, in order to’. Within narratives the futurity is with respect to the time within the narrative, and not to the time of speaking (10.24). It is well described by Comrie’s (1976:65) comments on English ‘be going to’, as indicating “the present seeds of some future situation which might well be prevented from coming about by intervening factors”. This present relevance of an anticipated future situation, as well as the fact that the reference time for *atu* is not necessarily the time of speaking, shows that *atu* marks aspect rather than tense.

In the vast majority of instances the subject referent for a clause containing *atu* is human, and the action specified by the head is intended or desired for the near future (10.24). As such it is common (but not required) in complements of verbs with meanings such as ‘desire’ and ‘consider’ (10.25), but allegedly cannot occur in the complement of *ho’i* ‘do not want, refuse’.

10.24  

*Sia at bá r-afaho r-akawak.*

3P IRR go 3P-weed 3P-mutual.aid  
They were going to go and help each other weed.  (F5.3)

10.25  

*Na’in mane n-akara *atu* sukat sira.*

noble man 3S-like IRR measure 3P  
The nobleman wanted to test them.  (AA1.96)

Another common context for *atu* is in the second of two clauses, where the first indicates something that is done in order to achieve the second (10.26). The presence of *atu* shows that the second clause represents a purpose, which was not necessarily achieved (10.27).

10.26  

*Sia mai at ukun rai Timor.*

3P come IRR rule earth Timor  
They (the various colonisers) came in order to rule Timor.  (H0.87)

10.27  

*Nia lolo liman *atu* kohi, semo.*

3S stretch arm IRR catch fly  
He stretched out his arms to catch (the bird, but it) flew.  (K12.9)

The word *atu* is not found in negative purpose clauses, for which *batu...keta* ‘in order that...do not’ is used.

Although most clauses with *atu* involve human agents acting intentionally, it is also fully acceptable for *atu* to co-occur with inanimate subjects (10.28) and/or predicates which specify an undesirable future (10.29). In such instances there is clearly no element of intention, and *atu* purely marks proximal futurity.4

---

4 Heine (1992:341) notes that “grammaticalization of a verb of volition (‘want’) to a kind of aspectual marker expressing the notion ‘nearly, almost’ can be observed in a number of languages”. Similar observations on the development of future or prediction terms from verbs of desire are made by Bybee et al. (1994:256) and Lord (1993:215ff.), amongst others.
Chapter 10

10.28 Kfitun na’in sa’e. Rai at n-aroma onan.
Venus (lit. ‘star lord’) has risen. The earth is about to lighten. (i.e. It is almost dawn.) (N0.149)

10.29 M-ót ha’u lai, té ha’u atu kiki.
Squeeze up against me, as I am about to shiver (from sickness). (V0.73)

The auxiliary atu cannot be negated by the standard negators la and ha’i (e.g. to mean ‘don’t want to, don’t intend to’). Instead, any negation of the clause must follow atu and fall within its scope, so meaning ‘want not to, intend not to’, as in the elicited example below.

10.30 Ha’u atu la k-bá bei, k-adomi.
I’d like to not go, but pity (them, so will go). (Q0.132 elicited)

Unlike the auxiliary sei, atu cannot stand alone within the predicate (as in English ‘I will’).

When atu has the meanings ‘about to, want to, intend to, in order to’ discussed above, it must be followed by a verb, which is non-stative in nearly all instances. It cannot, for instance, be used of wanting to be, or going to be, a teacher (*Nia atu guru ‘3S IRR teacher [Mly]’).

10.3.6.3 atu: irrealis mood

There is a separate construction in which atu can occur in pre-subject position (10.31, 10.33, 10.36, 10.38; 43 examples), or be followed by the full range of predicate types. In addition to non-stative verbs, possible predicates include stative verbs and adjectives (10.32, 10.35; 9 examples), NPs (10.34; 3 examples), and prepositional phrases (4 examples). In this construction atu is a general marker of irrealis, and need not indicate futurity of any kind. While many examples do refer to a time which is in the future relative to the reference time (10.33, 10.36), some are not specifically future (10.35), while others have past reference (10.32).

This construction is restricted to clauses which are already irrealis on grounds other than future reference. In the corpus the most common context is in rhetorical questions (10.31, 10.32). In these a gloss of ‘would’ rather than ‘will’ is in most cases appropriate. The construction is uncommon, though possible, in genuine information-seeking questions (10.33).

10.31 Hai! Ó ā’-át e’e, at ó bá m-ó na’in sia
EXCL 2S RDP-bad this IRR 2S go 2S-with noble PL
futu manu?!
fight.(cocks) bird
Hey! You being bad, would you go and cockfight with nobles?! (Z5.93)

10.32 Ná ema uluk á, at la soi bá sá Ibu?
like person former.times DEF IRR not rich for what mother
Like the people in former times, why wouldn’t they be rich, Ibu? (O5.125)

10.33 Ne’e at há.n sá, belu? ... At it tá ró ká?
this IRR do.like what friend IRR 1PI chop boat or
Now what are (we) to do, friend?...Shall we chop (i.e. make) a boat? (R5.117)
Embedded questions, for which there is some indication that the answer to the question is unknown or irrelevant, favour this construction.

10.34 Mane ne' e at sé ruma, sé n-atene dei?
man this IRR who unknown who 3S-know only
Who this man would be, who knows? (Z2.108)

10.35 Sia r-odi mai at wa'in ká ki'ik, bá rai nónók
3P 3P-bring come IRR much or small go lay.down be.silent
oan dei.
small only
They bring (the bride price and) whether (it) is much or little, (they) just go and
lay it down quietly (without saying how much money there is). (F1.48)

Another context for it is the (indirect speech) complement of verbs of ordering (§13.4.2), of which about 6% contain atu.

10.36 Tán Wahali na' in n-a' ak at ó bá m-ein
because Wehali noble 3S-say IRR 2S go 2S-wait
tora, m-ein hare.
foxtail.millet 2S-wait rice
Because the noble of Wehali said you were to go and watch millet, and watch
rice. (lit. ‘wait for’ it, that is watch it, and keep the birds away.) (S2.88)

10.37 Nia n-a' ak é “Oa. Man-tolun ne' e, ha' u katak ó at
3S 3S-say HES child bird-egg this 1S tell 2S IRR
há hola...”
eat take
She said “Child. This egg I told you to eat. [But you hatched it.]” (U10.11)

Limited evidence suggests that hypothetical conditionals too can use this construction.

10.38 Ata lear e'e at ó m-odi m-o'i ne' e mai mós
slave many this IRR 2S 2S-bring 2S-go.to this come also
la bele.
not can
These many slaves—even if you were to bring them over here, (you) wouldn’t
be able to. (Z3.302)

10.4 Deontic modal auxiliaries (permission and obligation)

10.4.1 Overview

Deontic modality is the modality of permission and obligation. It is expressible by the auxiliaries bele ‘can’, lalika(n) ‘need not’, keta ‘do not’ and the Malay loan musti ‘must’.

Since permission and obligation can only apply to “morally responsible agents” (Lyons 1977:823), it is not surprising that each of these auxiliaries in at least one of its meanings

---

5 Foley and Van Valin (1984:214) call this simply ‘modality’, subsuming under it obligation, intention and ability. However, the modifier ‘deontic’ helps distinguish this concept from others which have also been called ‘modality’.
requires that the subject refer to persons (e.g. humans, spirits, personified animals, or personified entities such as the sun and moon in folktales).

### 10.4.2 bele ‘can’

The auxiliary bele ‘can’ (over 360 examples) has three interrelated meanings, relating to permission, ability and future expectation. It frequently occurs without a following verb if the context is understood, such as in a question or answer about one’s ability to do something (first instance in 10.46), or in a permission-giving ‘OK’ response to a suggestion.

Syntactically bele is like the complement-taking verb hatene ‘know’ (§13.4.3.3) in that in certain irrealis contexts, usually where bele is negated, bele can follow its complement (10.39). Such postposing of bele accounts for 25% of the 167 negative examples in the corpus for which bele has a complement.

10.39  
\[
\text{Ha‘u k-la‘o bele ha‘i ti‘an.}
\]
1S 1S-walk can not already
I can’t walk any more (due to fatigue).  

Fronted complements are nearly always small clauses (§9.7.3), consisting simply of a verb (10.39) or a verb and generic object (e.g. foti ai la bele ‘lift wood not can’ = ‘unable to lift wood’). However, they need not be, and the corpus contains examples in which the complement is negated (second clause in 10.52), and longer examples such as fasi bá ne’e ‘wash at this’ = ‘wash here’, and 10.40.

10.40  
\[
\text{...Hodi sa‘e liu bá leten la bele.}
\]
bring ascend further go above not can
[They flew with her to the tree top, then came down again.] They couldn’t carry her further up to heaven.  

A further similarity to complement-taking verbs is that it is possible to negate either bele (10.39, 10.42) or the rest of the clause (e.g. bele la bá ‘allowed to not go’), or both independently (10.41, second sentence in 10.52).

10.41  
\[
\text{...Ita iha uma laran, saseni la n-ó. Bá ha‘i}
\]
1PI LOC house interior hindrance not 3S-exist go not

\[
\text{la bele.}
\]
not can
[If we give excuses for not attending weddings and the like, it is to strangers.] Within our household there are no excuses. We can’t not go.  

When used for the deontic modality of permission to mean ‘be allowed to’, the source of the permission is unspecified, normally being some socially recognised authorities (such as parents or elders), tradition, or even commonsense. This is the more common interpretation for negative clauses in which bele precedes the verb (10.42), accounting for 76% of 118 such examples.

10.42  
\[
\text{Fulan mate, la bele tiha, la bele halo lia...}
\]
moon die not can lay.warp not can make festival
(When) the moon is dead (i.e. at new moon), (one) may not lay the warp threads for weaving, (and) may not hold a festival... (according to tradition).
10.43  
_Talain_ foin _bele_ kawen._

cross.cousin then can marry
(If people are) cross-cousins, only then may (they) marry (according to
tradition).  

(N0.4)

In addition, _bele_ is used for the dynamic modality of ability (often considered marginal to
modality proper) to mean ‘be able to’. This is the most common interpretation for clauses
with positive polarity (10.44), and also for negative clauses where _bele_ is postposed (10.39).

10.44  
_Sia_ _bele_ r-aré _ita; ita_ _la_ _bele_ haré _sia._

3P  can  3P-see  1PI  1PI  not  can  see  3P
They (spirits of the dead) can see us; we can’t see them.  

(E0.143)

Finally, _bele_ is used for future expectation (something akin to ‘will’), especially in purpose
clauses where the main clause has already indicated that the conditions for ability and/or
permission are satisfied.

10.45  
_Ha’u_ _k-adoko_ nia _k-odi_ _k-ananu, bat_ _oan_ _la_ _bele_ _tanis._

1S  1S-rock  3S  1S-COORD  1S-sing  so.that  child  not
can  cry
I rock him (the baby) while singing, so that the child won’t cry.  

(N0.95)

The corpus contains two examples of _bele_ in pre-subject position. Both are questions (one
of the triggering conditions for pre-subject position of irrealis _atu_); however, it is not known
whether this is coincidental.

10.46  
_Bele_ _lale?_ _Bele_ _it_ _hakletek_ _bá?_

can  no  can  1PI  walk.along.top  go
Can (we) or not? Can we walk there along the top (of the log)?  

(R5.123)

10.4.3  _lalika(n) ‘need not’_

The auxiliary _lalika_ (or _lalikan_) means ‘need not, unnecessary’ (23 examples). If
someone (especially a child) is told that they ‘need not’ do something, the clear inference is
that he or she needs to not do it. As a result _lalika_ can often be translated as a prohibition
(‘don’t’). The actor of clauses marked by _lalika_ is nearly always the addressee. This word can
be used in single-word utterances, particularly as a reasonably polite means of refusing offers
(10.48).

10.47  
_Lalikan_ _m-ata’uk._  _= _ _Ó_ _lalika_ _m-ata’uk._

need.not  2S-fear  2S  need.not  2S-fear
You don’t need to be afraid. (i.e. Don’t be afraid.)  

(G4.90, C2.79)

10.48  
_“Ita_ _há_ _la.”  _N-a’ak_  _“Lalika…”_

2S.HON  eat  first  3S-say  need.not
“You eat first (before you depart).” (He) said “No need. [We’ll leave, as the
day is getting on.]”  

(K9.114)

---

6 According to Mathijsen (1906) and Morris (1984b) it consists of _la_ ‘not’ and _lika_ ‘necessary’; however,
Fehan consultants were not familiar with the root _lika(n)_.
10.4.4 *keta ‘do not’*

The auxiliary *keta* (or *ket*) ‘do not’ indicates that the condition specified by the clause should be avoided. It is found in two contexts.

The first is in prohibitions, where the second person subject is optionally specified (over 150 examples). It can be used without a following verb, usually in the expression *Keta lai* (lit. ‘do not first’), which is translatable as ‘Don’t!’ or ‘Wait a second!’. The clause-final adverb *oli* is restricted to such prohibitive clauses (10.49); its meaning, however, is unclear.

10.49  *Keta* titu *oli* té *luli.*  
*do not look*  
Don’t look (at her) because it is taboo.  
(Y0.108)

10.50  *Belu, ó keta bolu ema hudi na’an sia.*  
friend 2S do.not call person banana owner PL  
Friend, don’t you call the owners of the bananas.  
(U2.15)

The second context for *keta* is in negative purpose clauses (38 examples). In this case, in contrast to prohibitions, the subject is not restricted to referring to the addressee, and can readily have an animal or inanimate entity as referent.

10.51  *Ita semprut saka, bat kamerun ket tama.*  
1PI spray[Mly] hurry so.that k.o.rice.pest do.not enter  
We spray immediately, so that rice pests don’t get in (to the crop).  
(V0.174)

10.4.5 *musti ‘must’*

The word *musti* (also *mesti*; 72 examples) is a commonly used borrowing from Malay *mesti*, retaining the Malay meaning ‘have to, be obliged to’. The subject must be personal. This concept is alternatively, but less precisely, expressible by the adverb *dei* ‘only’.

10.52  *Luli ita musti tuir. La tuir la bele.*  
taboo 1PI must follow not follow not can  
Taboos we must go along with. We can’t not follow (taboos).  
(V0.97)

Three examples in the corpus have *musti* in pre-subject position, a construction which was disallowed in elicitation.

Limited evidence suggests that *musti* can also be used for epistemic necessity, to mean ‘must, inevitably’. The subject is in this case not restricted to persons.

10.53  ... *kalo haroun bá wé, udan musti mai.*  
if immerse go water rain must come  
[People who don’t believe this taboo.] if (they) immerse (the sago) in water, rain must come. (Rain is the inevitable result of breaking the taboo.)  
(R2.16)

Clause-initially *keta* is an adverb meaning ‘perhaps’ (§11.12).

It appears that, as pointed out by one consultant, *oli* adds politeness to a prohibition and would not be used in anger. According to Morris (1984b) it adds force to a prohibition. For Mathijsen (1906) the meaning is unknown.
Adverbs and verbal modifiers

11.1 Overview

Adverbs function as modifiers of constituents other than nouns (Schachter 1985:20). They cannot themselves head predicates, have no valency, and cannot be modified (e.g. negated or intensified). There are very few instances of morphological derivation from adverbial bases. Adverbs do not in general take subject marking; there are, however, a few exceptions, discussed in §9.3.3.3.

In certain situations, verbs (in which category is included adjectives) can function as modifiers to other verbs, filling the same function as adverbs. In particular, some modifiers of degree and manner are verbs. Because of the similarity in function, they are discussed in this chapter along with adverbs.

Tetun adverbs constitute a very mixed group, both semantically and syntactically. As a result of this heterogeneity, the use of a single cover term ‘adverb’ to encompass them all is problematic, as is commonly recognised for English also (e.g. Huddleston 1984:96f.). Nevertheless, it is both traditional and convenient to discuss various types of adverbs together.

There are four basic slots available for adverbs within a clause, with the final two slots collapsing into a single one for clauses with no object or oblique arguments. These positions are clause-initial, pre-predicate (between subject and predicate), postverbal and clause-final. Although some adverbs (e.g. degree modifiers and *ta* ‘already’) are restricted to verbal clauses, others (e.g. *ha’i* ‘not’) are not. For these adverbs ‘V’ in the formula below should be interpreted as the predicate head, regardless of whether that head is a verb.

\[
\text{Clause} \rightarrow (\text{Adv}_1) (\text{S}) (\text{Adv}_2) (\text{Aux}) V (\text{Adv}_3) (O) (\text{Oblique}) (\text{Adv}_4)
\]

Many adverbs have some freedom of distribution, which is at least partly determined by issues of scope. This relative freedom, and the fact that semantic classes do not necessarily correlate well with preferred syntactic positions, makes adverbs hard to subclassify. The following subclasses of adverbs have been selected for special attention in this chapter.

1. Negators
2. Degree modifiers
3. Participant coreference markers: reflexive and reciprocal
4. Aspectual and temporal markers
5. Adverbs of repetition and frequency
6. Manner modifiers
7. Clause-final modifiers homophonous with *bá* ‘go’
   a) Deictic particles
   b) Imperative *bá*
   c) *bá* ‘so be it’
8. Connective adverbs
9. Sentence adverbs of possibility

11.2 Negators

11.2.1 Overview of negation

Tetun has four words which are classed as negators. The general clause negators are preverbal la and postverbal ha'i 'not'. A negator with much more restricted distribution is dauk, which is found only in expressions for the concept 'not yet' (e.g. sei dauk, lit. 'still not'), and is for this reason discussed in §10.3.5. Finally, lahós 'indeed not' indicates contrastive negation of the following constituent.

Apart from these, there are a number of other means of expressing negation. These include the degree adverb ladún 'not very' (§11.3.1) and the modal auxiliary lalika 'need not' (§10.4.3). The modal auxiliary keta 'do not' provides for negation in prohibitions as well as in purpose clauses (§10.4.4). The pro-clause lale 'no' is used as a single-word negative response to positive polarity questions and suggestions, as a contradiction of a previous statement, as a negative alternative in a polar interrogative, or as an emphatic reiteration of a previous negative statement (11.7). Finally, there are some inherently negative verbs, namely lalek 'lack' (§9.4.7.7), and the complement-taking verb ho'í 'do not want, refuse' (11.4).

11.2.2 Clausal negators: la, ha'i 'not'

11.2.2.1 Shared features

Clauses with verbal (11.1) or prepositional phrase (11.5) predicates can be negated by the general negators la and ha'i.1 The former precedes the head of the predicate (the verb or preposition),2 while ha'i follows the head.

11.1 N-ó ha'i ema dakar nia, nia la karian. 
3-exist not person look.after 3S 3S not work 
(If) there isn't someone watching over him, he doesn't work. (MO.77)

11.2 Sei we ne, hola ha'i na'an, hola ha'i boek, 
collect(liquid) water this take not meat take not shrimp 
hola ha'i ni'is. 
take not k.o.crab 
When (they) were net-fishing (lit. 'collecting water'), they didn't catch fish, didn't catch shrimp, didn't catch crab. (U4.3)

Although the negators are normally immediately contiguous with the head, there are exceptions. The auxiliary bele 'can' is the only preverbal constituent that can follow the negator la, resulting in the preverbal sequence la bele 'not can' = 'cannot'. There are also a

---

1 This is what Klima calls 'sentential negation' and Payne 'standard negation' (Payne 1985:198).
2 Thus la has the same distribution as auxiliaries. Since auxiliaries constitute a diverse class, la could be added to the class without extending the range of auxiliaries too much. Nevertheless, la has no verbal properties: it cannot take subject marking, a verb following la does take normal subject marking (just as it does for the irrealis auxiliary au; 11.3), it places no selectional restrictions on the subject, and it cannot occur without a following predicate.
few postverbal modifiers which precede postmodifying *ha'i* if they are to be specifically negated; these include the adverbs *hika(r)* ‘back, again’ (*fila n-ikar ha'i* ‘return 3S-back not’ = *not return*), *tenik* ‘again’ (*mai tenik ha'i* ‘come again not’ = *not come again*), and *resik* ‘too, overly’ (11.15). There is no difference between negation of main clauses and that of subordinate clauses.

Where *la* and *ha'i* co-occur with each other the negators reinforce each other (final clause in 11.3; 10 examples), resulting in more emphatic negation.

11.3

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ha'u</td>
<td>k-bá.</td>
<td></td>
<td>Ha'u</td>
<td>la</td>
<td>k-bá.</td>
</tr>
<tr>
<td>1S</td>
<td>1S-go</td>
<td>not</td>
<td>1S</td>
<td>not</td>
<td>1S</td>
</tr>
</tbody>
</table>

*ha'i*.

1S-go not

I'm going. I'm not going. I'm not going (stronger). I'm certainly not going.

(X0.106 elicited)

The same is true if these negators co-occur with the inherently negative verb *ho'i* ‘do not want’. Thus negating this verb strengthens the claim, rather than reversing the polarity (first clause of 11.4; 29 examples). The only counterexample in the corpus, for which two negatives result in a positive interpretation, is the following elicited one. In the second clause of this example the negative *la k-o'i* (‘not 1S-do.not.want’) is a single, repeated, unit which is negated as a whole by the following *ha'i*.

11.4

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Ö</td>
<td><em>la</em></td>
<td>m-o'i!&quot;</td>
<td>&quot;Ha'u</td>
<td><em>la</em></td>
<td>k-o'i&quot;</td>
</tr>
<tr>
<td>2S</td>
<td>not</td>
<td>2S-do.not.want</td>
<td>1S</td>
<td>not</td>
<td>1S-do.not.want</td>
</tr>
</tbody>
</table>

Ha'u  

k-ahoo'uk!"

1S  

1S-agree

"You don't want to!" "I don't 'not want to!' I agree to!"  (H0.33 elicited)

Indefinite subjects are included within the scope of negation, as illustrated by the following example, where the negation is of the entire clause *ida iha nia* ‘one LOC 3S’ = ‘someone is there’ (8 examples). Such a combination of indefinite NP with negator is the standard means of expressing what in English is expressed using negative quantifiers.

11.5

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Ita</em></td>
<td>kaur</td>
<td>mós</td>
<td><em>ida</em></td>
</tr>
<tr>
<td>1PI</td>
<td>call.loudly</td>
<td>also</td>
<td>one</td>
</tr>
</tbody>
</table>

Even if we call out, there is no one there (to hear and help us).  (V0.44)

### 11.2.2.2 Differences between la and ha'i

Apart from the difference in position, there are a number of other differences between the two negators *la* and *ha'i*. In terms of frequency, *la* is much more common, being found in over 75% of the more than 1,400 negated clauses in the corpus.

Some constituents can be negated by only one or other of these terms. Nominal (including possessive) predicates can be negated by *ha'i* (11.6) or by a combination of *la* and *ha'i*. Negation by *la* alone was considered of dubious acceptability, a judgment supported by the lack of such negations in the corpus. The more usual negator for nominal predicates is the contrastive negator *lahós*, discussed below.

11.6

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Ö!</em></td>
<td><em>Buat</em></td>
<td><em>e'e</em></td>
<td><em>Beur</em></td>
<td><em>ha'i</em>!</td>
<td></td>
</tr>
<tr>
<td>oh!</td>
<td>thing this</td>
<td>Mr</td>
<td>deceive</td>
<td>not</td>
<td></td>
</tr>
</tbody>
</table>

Oh! This thing (actually a person) isn't Mr Trickster!  (T0.82)
On the other hand, there are some non-verbal expressions which can be negated by *la* but not by *ha'i*. Such negative expressions include *la oras* 'not time' = ‘soon’ and *la sâ ida* 'not what one' = ‘it’s nothing; it doesn’t matter’. In addition, the auxiliary *bele* ‘can’, when it precedes the verb, can be negated only by premodifying *la* (e.g. *la bele la'o* 'not can walk' = 'cannot walk'), and not by postmodifying *ha'i* (e.g. *bele ha'i la'o*). Where *bele* follows the verb, however, either negator is grammatical (e.g. *la'o la bele, la'o bele ha'i* 'cannot walk').

Phonological stress is different for the two negators, since *ha'i* contains a stressed syllable, but *la* is always unstressed (except in poetry, where normal stress rules do not apply). Thus *la* is phonologically a clitic, and is written attached to the following word by some Fehan people. In addition to this phonological stress *ha'i* is more readily used than *la* when the fact of negation is stressed (11.3); this is, however, only a tendency.

Finally, there is a dialect difference, in that *ha'i* is restricted to the region of southern Belu (i.e. primarily to the Fehan dialect), while *la* is used throughout the Tetun-speaking region. This fact, noted by Wortelboer (1955:177), is recognised by some consultants, and is supported by the mention of *la* but not *ha'i* in Hull’s (1996b) and Morris’s (1984b) dictionaries for East Timorese Tetun, and Mathijsen’s (1906) one for the Foho dialect.

### 11.2.3 Contrastive constituent negator: *lahós* ‘indeed not’

The word *lahós* (or *lahos*) negates the constituent that follows it. All of the 49 examples in the corpus are explicitly contrastive, being accompanied either by the contrasting positive statement (which may immediately precede or follow the negative one), or by a following emphatic *laLe* ‘no’, or both. A wide range of constituents can be negated by *lahós*. These include verbal predicates (11.7), nominal predicates (11.8), object NPs, locative prepositional phrases (11.9), and entire clauses (11.10).

#### 11.7

| Ha’u lahós foi.wa’i. LaLe. Ha’u k-ferik. |
|-----------------|------------------------|
| 1S indeed.not 1S 1S-mature.woman |
| I am not single. No. I am married. |

(Z5.79)

#### 11.8

| Keta rona manu lian; ne’e lahós ema lian. Manu lian. |
|-----------------|------------------------|
| do.not hear bird voice this indeed.not person voice bird voice |
| Don’t listen to the bird’s voice. This isn’t a person’s voice. (It’s) a bird’s voice. |

(AA3.13)

#### 11.9

| Ami hola malu lahós iha uma; iha hae laran. |
|-----------------|------------------------|
| 1PE take each.other indeed.not LOC house LOC grass interior |
| (Describing illicit sex:) We have sex (lit. ‘take each other’) not in the house, (but) outside (lit. ‘in the grass’). |

(10.29)

Unlike the clausal negators, *lahós* negates, rather than reinforces, a negative that falls within its scope (4 examples).

#### 11.10

| Ha’u k-ô la’e. Lahós ha’u la k-ô la’e, lale! |
|-----------------|------------------------|
| 1S 1S-have husband indeed.not 1S not 1S-have husband no |
| I have a husband. It is not true that I don’t have a husband; no! |

(T0.39)

The word *lahós* is phonologically unique in that the stress usually falls on the second syllable but can, for some speakers, fall on the initial one. This variability is consistent with analysis as a sequence of two words, with the first being the clausal negator *la*. Nevertheless, consultants did not recognise *hós* (or *ös*) as a word. Similarly, while both Mathijsen (1906)
and Morris (1984b) list *lahós* with either a hyphen or a space after *la*, neither of them list the second part as a lexeme.

11.3 Degree modifiers

11.3.1 Downtoner: *ladún* 'not very'

The ‘downtoner’ (Quirk & Greenbaum 1973:218) *ladún* ‘not very, not very much’ (17 examples) lowers the force of the verb. This adverb can either precede (11.11) or follow (11.12) the constituent it modifies.\(^3\)

11.11  
*Nia* *ladún* *bót* = *la* *ki'ik* *resi*, *la* *bót* *resi.*

3S not.very big not small too not big too

He’s not very big = not overly small (and) not overly big.  \(\text{(Q0.76)}\)

11.12  
*Nia* *n-adomi* *ha'ú* *ladún* *sá.*

3S 3S-love 1S not.very just

She loves me only a bit.  \(\text{(T0.143)}\)

11.3.2 Intensifiers

There are two intensifying adverbs, namely *basuk* ‘very’ (*dois basuk* ‘very smelly’, 11.13; 114 examples) and *resi* (or *resik*) ‘too, overly’ (11.11, 11.14, 11.15; 35 examples). They immediately follow the constituent they modify, which may be an adjective or stative verb but not an active verb such as *karian* ‘work’. Neither adverb can occur as the head of a clause, indicating that neither has alternative analysis as a verb.

11.13  
*Kalo* *m-ó* *ha'í* *basuk*...

if 2S-have not very

If you really don’t have (anything)... \(\text{(V0.104)}\)

11.14  
*Ita* *loron* *keta* *haneo* *resik* *Pák.*

2SHON day do.not ponder too Mr

You mustn’t think too much about Mister (your husband) every day. \(\text{(M0.17)}\)

The term *basuk* can modify the negator *ha'í* (‘really not...’; 11.13). In contrast, *resi* is found within the scope of negation (‘not too...’; 11.11, 11.15).

11.15  
*Manas* *resik* *ha'í*.

hot too not

(Betun is) not too hot (but still hot). \(\text{(H29/4)}\)

There are, in addition, four words which can function as intensifiers of a preceding verb, but which are also independent verbs with meanings vaguely related to intensification. These are *liu* ‘go further’ (11.16; 104 examples), *to’o* ‘arrive, reach’ (e.g. *ás* *to’o* ‘very tall’; 34 examples), *mate* ‘die’ (which as an intensifier is informal; e.g. *karian* *mate* ‘work very hard’; 7 examples), and *bót* ‘big’ (which as an intensifier is considered non-Fehan; 11.17; 14 examples). These intensifying verbs are unlike adverbial intensifiers in three respects. They can modify active as well as stative verbs, they follow the object NP, if any (11.16), and they

---

\(^3\) Although *ladún* appears to incorporate the negator *la*, the remainder of the expression, *dún*, is not recognised as an independent lexeme.
can themselves be modified by an intensifying adverb (11.17). As intensifier mate (lit. ‘die’) retains the additional verbal feature of being able to follow the irrealis auxiliary at(u), as in susar at mate ‘very difficult’. Although these intensifiers could be analysed as being adverbs as well as verbs, there seems no advantage in doing so, especially in the light of the syntactic differences between them and the intensifying adverbs basuk and resi(k).

11.16 Nia ne’e, ha’u k-adomi nia liu.
3S this 1S 1S-love 3S further
This one, I really pity her. (I5/7)

11.17 Ita soi bôt basuk!
1PI rich big very
We’d be very very rich! (Q0.99)

The intensifying verb liu (lit. ‘go further’) is used to form comparatives, by making the object of comparison the syntactic object of liu (22 examples).

11.18 Mina fahi mak kmidar liu mina nú.
oil pig REL tasty further oil coconut
It is lard that is tastier than coconut oil. (N0.118)

In addition to the intensifiers, there is the adjective kaliuk ‘above all’, which can be used as a superlative (3 examples).

11.19 Manas kaliuk mak fulan sanulu resin ida.
hot above.all REL moon ten extra one
The hottest (month) is the eleventh month (November). (N0.147)

11.4 Participant coreference markers

11.4.1 Introduction

There are two markers of participant coreference, namely reflexive án and reciprocal malu. Both follow the predicate head.

11.4.2 Reflexive án

Coreference of the subject and object of transitive verbs is optionally indicated by the postverbal reflexive marker án or its variants á, ón, ô or oan (over 80 examples). Coreference with other syntactic roles (such as with the object of the prepositional verb bá ‘to (recipient)’) cannot be expressed by án.

11.20 Sia na’in ukun án.
3P self rule REFL
They themselves ruled themselves. (i.e. They were not ruled over by others.) (V0.15)

Reflexive án is found in some fixed polite expressions used when talking to or about nobles, including hatodan án ‘sit down’ (lit. ‘make-heavy REFL’) and halolo án ‘die’ (lit. ‘stretch.out REFL’).
Adverbs and verbal modifiers

11.21  ...rai  n-aroma  ti'a,  Wahali  na'in  n-a-sai  ǵ.  
earth 3S-become.light already Wehali noble 3S-make-exit REFL  
[When...] it was dawn, the noble of Wehali brought himself out (i.e. revealed himself).  
(S2.152)

As illustrated in 11.20, án with a plural subject still gives a reflexive meaning, with reciprocal meanings instead being expressed by malu.

Alternatively, coreference of subject and object can be indicated by the postverbal adverb hikar (or hika) 'back, return to earlier location, state or activity', in conjunction with either an object pronoun (11.22; 6 examples) or reflexive án (11.23; 4 examples).

11.22  Ni  la'en  á  sona  n-o'o  n-ika  nia.  
3S husband DEF pierce 3S-kill 3S-back 3S  
Her husband stabbed and killed himself. (He actually died of the wounds some weeks after the stabbing.)  
(V0.38)

11.23  Nia  fota  n-ika  án.  
3S hit 3S-back REFL  
He hit himself.  
(H0.125 elicited)

Syntactically reflexive án fills the slot normally filled by object NPs, with two differences. Firstly, án cannot be fronted (a position which would in any case render it cataphoric rather than anaphoric). Secondly, limited evidence suggests that it precedes rather than follows the postverbal negator ha'i (11.24; 1 example). The latter observation suggests that rather than being a reflexive pronoun which constitutes an object NP, án is an adverb in a necessarily intransitive clause (an analysis noted for other languages in Anderson (1985:192)).

11.24  Hadak  e'e  babidu  ǵ  ha'i  ti'an.  
slat.floor this k.o.move REFL not already  
The slat floor doesn't move about any more (as the slats are tied down firmly).  
(P0.135)

Analysis of án as adverb is also consistent with its non-reflexive uses, which are possible with both transitive and intransitive verbs. In particular, án can indicate that the actor intentionally did the action specified by the verb, as illustrated by the contrast in 11.25. Non-reflexive án also has other uses which are not yet clear to me.

11.25  Nia  la'o  lakon  án.  
3S walk disappear REFL  
He walked to go missing (on purpose).  
Nia  la'o  lakon.  
3S walk disappear  
He walked and went missing (either intentionally or not).  
(K0.159)

11.4.3 Reciprocal malu

Reciprocity is indicated by either the word malu ‘each other’ (390 examples) or, in restricted circumstances, the circumfix hak- -k. The latter is discussed in §4.3.4.

Clauses with malu require that their subjects (whether specified or implied) have plural referents. How that plurality is marked is irrelevant; it may be indicated by any of the usual markers of plurality, such as coordination of NPs (11.26), plural pronouns (11.27), and...
enumerated or plural-marked NPs. Alternatively, the plurality may be left implicit (e.g. Folin kona malu ‘price touch each other’ = ‘They reached agreement on the prices’).

11.26 Asu nó busa funu malu.
don and cat enemy each other
Dog(s) and cat(s) are enemies to one another. (L0.93)

11.27 Bá oras sia n-á, sia mós n-usu malu.
at(time) time 3P 3-eat 3P also 3-request each other
While they were eating, they asked each other (questions). (K9.46)

As Lichtenberk (1985) points out in his overview of reciprocal constructions, such constructions involve a low degree of individuation of participants. The reciprocal indicates that some participants are actors, and some (often an overlapping group) are undergoers, but does not specify which individuals fall into which group.

Within this general observation, Tetun reciprocals fall into two of the semantic categories outlined by Lichtenberk. In the first, each of the participants stands in the same relation to the others as the others do to him or her. These relations may be simultaneous (11.26) or not (11.27), with the temporal order of the relations being unspecified.

The second semantic category is a variation of what Lichtenberk calls ‘chaining’. If there are two participants, one of them stands in the specified relation to the other (11.28). If there are more than two participants, an unspecified number of them stand in that relation to others within the group. In the neatest case, the one described by Lichtenberk, each participant except the last stands in a certain relation to the next. This is illustrated by tuir malu ‘follow each other’, which could describe a line of people, each following the one in front (11.29). However, the relationship need not be that neat (11.30). For instance, one can say that a group of people bolu malu ‘call each other’, which need not imply that each individual is calling to another.

11.28 Botir nó blék r-atán malu.
bottle[Mly] and tin[Mly] 3P-on.top.of each other
The bottle and the tin: one is on top of the other. (V0.64)

11.29 Sia lolo fulan bá malu. Sia mate tuir malu.
3P pass month to each other 3P die follow each other
They (three people) ‘handed each other months’. (i.e. Each died approximately a month after the other.) They died one after the other. (H2/5)

11.30 ...hawa’i hikar malu dei.
adopt back each other only
[We] just adopt each other. (i.e. We adopt within our group, not from strangers.) (N0.49)

The word malu is not used for reflexive plurals (for which reflexive án is used), nor for collective or cooperative activities (for which the verb hamutu(k) ‘together’ is available; §12.6).

With transitive verbs malu fills the same postverbal slot as object NPs. However, unlike object NPs, it cannot be fronted, and can follow nominal (11.26) and intransitive (11.31) predicates. This is evidence that malu should, like reflexive án, be analysed as an adverb which requires that the clause not have an object NP. Unlike reflexive án, malu readily fills the recipient slot introduced by the prepositional verb bá ‘to’ (11.29).
Betun and Kletek far each other Betun far Kletek
Betun and Kletek are far apart. (N0.7)

There are a number of verbs which allow participants to be specified as subject in the
reciprocal which could not be subjects in non-reciprocal clauses. These include the transitive
verbs fa'e ‘divide, separate’, sasa ‘separate’ and kari ‘scatter’. When used non-reciprocally,
the subject specifies an actor who divides or separates the undergoer participants from one
another, as in Ha'u bele fa'e emi na'in rua 'I can separate you CLS:human two'. The subject
cannot refer to (one of) the individuals being separated; hence *Ha'u fa'e ó 'I separate
(from) you'. Nevertheless, this is precisely what occurs in a reciprocal construction, where
fa'e malu means 'separate from each other'.

11.32 ...emi na'in rua sasa malu; emi fa'e malu.
2P CLS:human two separate each other 2P divide each other
[If we tell you, husband and wife, the news,] the two of you will separate; you
will part from one another. (E0.156)

11.5 Aspectual and temporal markers

11.5.1 Introduction

There is no obligatory marking of tense or aspect in Tetun, with these notions being
largely implicit. Thus, for instance, Nia bá ‘3S go’ can be interpreted according to context as
'She went', 'She is going' or 'She will go'. Nevertheless, temporal relationships can be made
explicit through aspectual auxiliaries (§10.3), through time nouns and temporal adverbs
which refer to situation-external time (e.g. awan ‘tomorrow’, kala-kalan ‘nightly’), and
through the aspectual and temporal markers which will be surveyed in this section.

While many of these aspectual and temporal terms have some freedom of position within
the clause, all tend to occur towards the end of the clause, following the predicate and any
object NP. The one exception is the proclitic ta, which immediately follows the verb.

11.5.2 ti'an, ti'a, ta ‘already’

The three terms ti'an (over 800 examples), ti'a (over 700 examples) and ta (140
examples) all have the same basic meaning of 'already' and are phonologically related, but
have different, though overlapping, syntactic distributions.

The word ti'an ‘already’ indicates perfective aspect (in the sense of Comrie (1976:16ff.)),
in that it denotes a complete situation which occurred in the past, but “involves lack of
explicit reference to the internal temporal constituency of a situation” (p.21). It also involves
perfect aspect in that the past situation so described still has present relevance. Although the
form ti'a is sometimes found in this context too, such usage is not considered correct.

mother further already mother further already
Ibu has already gone past. (Q0.145)

11.34 Mai to'o uma mai, mai, ha'u k-toba ti'an.
come reach house come come 1S 1S-lie.down already
(They) came to the house here; (when they) came, I was already lying down
(in bed). (F3.51)
11.35 Nia n-akfunin ti’an iha nia.
3S 3S-hide already LOC 3S
He had hidden (himself) there. (S0.26)

When clausal negation (by the negators la or ha’i) co-occurs with ti’an, the resulting meaning is nearly always ‘no longer, not any more’ (11.36; 89 examples). In only one example in the corpus is the resulting meaning ‘already not-Predicate’ (11.37).

11.36 Ikus é tau ha’i faru baban ti’an.
final this put not clothes patch already
Now (people) don’t wear patched clothes any more. (Q0.164)

11.37 Tán nia la bá uma loron hitu ti’an.
because 3S not go house day seven already
Because he had already not gone home in seven days. (O1.23)

The form ti’a, like ti’an, means ‘already’; however, in contrast to ti’an it is primarily used in clause-sequencing constructions. It indicates that the event described in the following clause will occur after that referred to in the ti’a-marked clause is complete. The first clause normally has rising, non-final, intonation. The initial clause may be a backgrounded reiteration of an earlier clause (i.e. tail-head linkage, as in 11.38 and 11.44; §14.4), but need not be (11.39). Frequently the time setting is reduced to nia ti’a ‘3S already’ = ‘after that; then’ (11.40). As illustrated by 11.39, the time referred to can be in the future relative to the time of speaking, so long as it is in the past relative to the time of the following clause. The form ti’an is acceptable but less common than ti’a in this clause-sequencing context.

11.38 ...it hâ onan, é. Há ti’a toba.
1P I eat IMM TAG eat already lie.down
[When dinner is ready,] we eat, don’t we! Having eaten, (we) sleep. (J2.17)

11.39 ...Kawen ti’a, tûr iha ne’e dei.
marry already sit LOC this only
[I agree to marry you so long as we will live near my family.] Once (we) are married, (we) must live here. (F3.25)

11.40 Nia ti’a, kalan ti’an r-anawa.
3S already night already 3P-stop
Then, it was already evening, and they stopped. (Z2.37)

The third related form is ta (usually pronounced [tɔ]). Syntactically it immediately follows the verb. Phonologically it is an unstressable proclitic, attaching to a following serial verb or adverb (11.42), or to the first word of a following object NP (11.43). The forms ti’a and (rarely) ti’an may also be used in this syntactic position, but are phonologically not clitics. The form ta is used like ti’an (and often together with it) for past time (11.41, 11.42), and like ti’a (and often together with it) in clause-sequencing constructions to mean ‘after’ (11.43, 11.44).

---

5 Mathijsen (1906) reaches the same conclusion in his dictionary of the Foho dialect of northern Belu, saying that ti’an forms past tense, while ti’a translates conjunctions such as ‘when’ and ‘after’ for past time as well as forming commands (for which see following discussion on ta).

6 Some consultants considered ta to be a short form of both ti’an and ti’a. Others stated, like Mathijsen (1906), that it has no meaning.
11.41 ...Busa Oan á mai n-ikar ti’an; n-alai uluk
cat small DEF come 3S-back already 3S-run go.first

\( \text{ta} \text{ mai} \text{ ti’} \text{an}. \)
already come already

[When they returned,] Kitten had already come back: (he) had already run home ahead (of them). 
\( \text{(Z2.90)} \)

11.42 Sasawan n-odi n-ika bá, n-a-be’o ta dei.
morning 3S-bring 3S-back go 3S-make-shatter already only

In the morning (she) took (the water pitcher) back there, (and) just shattered (it). 
\( \text{(Z4.7)} \)

11.43 Nia.má ku’u n-ola ta fuka tahan á, beni
then pick 3S-take already k.o.plant leaf DEF cover

\( \text{ta} \ bá... \)
already go

Then having picked the \( fuka \) leaves, (she) laid (them) over (the crocodile’s back) [and climbed up on top.] 
\( \text{(Z4.23)} \)

11.44 N-asusu oa. N-asusu ta oa ti’a,
3S-breastfeed child 3S-breastfeed already child already

\( n-ák “...” \)
3S-say

(He) nursed the child. After (she) had nursed the child, (she) said “...”
\( \text{(O4.61)} \)

In addition to past time reference, \( ta \) and (less commonly) \( ti’\) are used for future time (11.45), particularly in commands (11.46). This future perfective aspect use of \( ta \) and \( ti’\) is often in conjunction with \( lai ‘first, beforehand’ (\$11.5.6). \)

11.45 Emi fó we ida mai, ha k-emu ta lai.
2P give water one come already 1S-drink first

You give me some water, and I’ll drink.
\( \text{(Z5.65)} \)

11.46 Há ta bá lai. = Há ti’a bá lai.
eat already IMP first eat already IMP first

Eat first. (i.e. Eat now, before doing something else.)
\( \text{(123/6)} \)

11.5.3 kedan ‘immediately’

The word \( kedan \) (or \( kedas, kedak; 85 examples) indicates immediacy or promptness, and can often be translated ‘immediately’ or ‘in advance’. Thus it can indicate that the event in the clause marked by \( kedan \) occurred immediately after the preceding event (11.47) or is expected to occur immediately after the time of speaking (11.48). In a request such as 11.49 it again adds immediacy, reinforcing the fact that the addressee must bring the salt with him when he comes, as opposed to fetching it at some later stage.

11.47 Taru bá, futu, manu nia-k á n-o’o kedan.
gamble go fight(.cocks) bird 3S-POS DEF 3S-kill immediately

(They) gambled, fought cocks, (and) his cock immediately killed (the other).
\( \text{(Z3.300)} \)
11.48 Katak kedak bá.
tell immediately IMP
Tell (us) immediately. (Q0.23)

11.49 Ó mai, m-odi kedas masin.
2S come 2S-bring immediately salt
(When) you come, bring salt. (U6.33)

The word *kedan* can also apply to past events which were done in preparation for the present (11.50). In this case the sense of immediacy is retained, since the action was forward-looking.

11.50 Mane sia at mai onan ... am ina-n sia bobi
man PL IRR come IMM IPE mother-GEN PL pound
kedan badut, buti bá kesak.
immediately k.o.nut massage go palm.midrib
(When) men were about to come (courting us)...our mothers and aunts would pound *badut* nuts beforehand, (and) massage (the pounded nut) onto the midribs of palm leaves (to make wick lamps for the courtship). (O5.27)

### 11.5.4 onan ‘imminent’

The word *onan* presents the event specified by the clause as imminent and inevitable. The reference time may be the time of speaking or the time within the narration. Within this general meaning, most of the 340 examples in the corpus fall into three broad categories.

Firstly, the reference point for the soonness can be the time of speaking. In this case *onan* indicates a time in the very near future. The time may be further indicated by a time phrase, such as *awan* ‘tomorrow’ (11.51) or *oras ida* ‘time one’ = ‘soon’. The word *onan* is often used in combination with the irrealis auxiliary *atu* ‘be about to, intend to’, in which case *onan* adds inevitability to the statement (11.53). It can also attach directly to a time phrase (11.54).

11.51 N-a'ak “Ha’u k-tanis tán *awan* ema r-o’o ha’u
3S-say 1S 1S-cry because tomorrow person 3P-kill 1S
*onan*.
IMM
(The child whom the witch was fattening up to eat) said “I am crying because tomorrow people will kill me”. (U6.31)

11.52 Ó mai ha’i tür onan?
2S come not sit IMM
Aren’t you coming to sit now? (Greeting to passer-by who was clearly going further.) (Q0.77)

11.53 Kalan böt ti’an, Ibu at fila *onan*.
night big already mother IRR return IMM
The night is already late, (and) you (‘Ibu’) are about to return (home). (P4.130)

---

7 In other dialects, such as those of Suai, northern Belu (Mathijsen 1906) and East Timor (Morris 1984b), this is *ona*. However, judging by its occurrence in texts from Suai and from the aforementioned dictionaries, the usage of *ona* in these other dialects is somewhat different to that of *onan* in the Fehan dialect.
Kalan onan, semo mai, tün mai rai klaran.

(As) night fell, (they) came flying, coming down to earth.  

The above examples are all of statements and questions. A second context in which onan is common is in commands, invitations and statements of intent. In commands and invitations it suggests insistence (‘do it now!’), as in 11.55. The ubiquitous invitation to eat, Há onan ‘eat IMM’, for instance, implies that the food is standing ready to eat. Culturally one is required to announce to people whenever one is parting from them or about to eat or drink. Such statements too can use onan to indicate that the action is about to take place (11.56).

A lakon! Bā m-ika onan.

You disappear! Go back now.  

Ami há onan.

We are going to eat now. (Said to those who won’t be eating.)  

In the third broad category of use of onan, the soonness is relative to the time within the narrative rather than to the time of speaking. This is frequently indicated by placing onan in the second of two intonationally-linked clauses, where the first specifies one event or condition, and the second specifies something which is presented as an inevitable consequence or an immediately following event.

Te’in tasa ti’a, it há onan, é.

After cooking (until the food) is cooked, we eat, don’t we!  

The enclitic -n (91 examples) is placed on the final word of a clause (so before tags and vocatives) if that word is vowel-final. Otherwise it is not used.

Like onan, from which some consultants believe it to be abbreviated, it indicates that the event is expected to occur in the very near future. Unlike onan, it is usually (86% of examples) used in conversation when the speaker refers to something that he or she is about to do (11.58), or thinks the addressee (or speaker and addressee together) is or should be about to do (11.59).

Hou. Ne’e ha’u k-bá k-ika-n...

OK. Now I’m going back. [Tomorrow I’ll bring it here.]  

As such it is polite in greetings. For instance, an approaching visitor can be greeted with Mai-n? ‘(Have you) come?’, while a passer-by who is addressed by the Malay term Ibu ‘mother’ can be greeted with Ibu liu-n? ‘(Is) Ibu going further?’ Similarly, this clitic is polite in the statements of intended action which are culturally required before one eats, drinks or takes leave of people (e.g. Ami liu-n ‘We (are) going further.’).
Only occasionally is it used outside this face-to-face context. The following example shows that it can, like onan, be used with third person subjects, with the time reference being within the narrative rather than the time of speaking.

11.60  

\[
\text{Niamá si la'o-n é.}
\]
then 3P walk-IMM TAG
Then they went (lit. ‘walk, go’), didn’t they. (D1.22)

11.5.6 *lai* ‘first, beforehand’

The term *lai* ‘first, beforehand’ (over 300 examples) indicates that that which is described by the clause must happen first, before some other (often unspecified) event takes place.

As such it is a polite word to use in requests for immediate action, since it implies that, once having complied with the request, the addressee can return to what he or she is currently doing, if so desired. Thus, for instance, *Hanawa lai* ‘stop first’ could be used to ask someone to temporarily stop what he is doing. In contrast, *Hanawa onan* ‘stop IMM’ has no such implication that the addressee can later return to the halted activity. Well over half of the corpus examples of *lai* are in fact in requests, commands or invitations.

11.61  

\[
\text{N-ák “Iku mai lai” .}
\]
3S-say Iku come first
(He) said “Iku, come here”. (F3.39)

Apart from such contexts, *lai* can be used for sequencing of events (in procedural, or less commonly narrative, texts), with *lai* marking the initial clause and *foin* ‘then’ often introducing the subsequent one. The subsequent event can also be left unspecified, as in 11.63, where the speaker agrees to a request but says she will get ready first. What she will do next, namely work to fulfil the request, is not stated but understood.

11.62  

\[
\text{Tunu fuan ida-.ida ita mama lai foin hodi hahán lawarik sia.}
\]
bake fruit RDP-one 1PI chew first then use feed small.child PL
(We) baked the (banana) fruits one at a time, we chewed them first, then used (the chewed banana) to feed the babies. (J3.132)

11.63  

\[
\text{Hou ina. Ha k-aris ta lai: luku ta lai.}
\]
yes mother 1S 1S-bathe already first massage.hair already first
Yes, mother. (I’ll go as you ask, but) I’ll go and bathe first, massage my hair first. (R4.93)

11.6 Adverbs of repetition and frequency

Most adverbs of repetition and frequency have some freedom of position within the clause. Some such adverbs are listed below, along with the position in which they are most often found.
Table 11.1: Adverbs of repetition and frequency

<table>
<thead>
<tr>
<th>Common position</th>
<th>Adverb</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postverbal</td>
<td>osak</td>
<td>in turn</td>
</tr>
<tr>
<td></td>
<td>fali, falik</td>
<td>in turn</td>
</tr>
<tr>
<td></td>
<td>tenik, teni</td>
<td>again</td>
</tr>
<tr>
<td></td>
<td>hikar, hika</td>
<td>back, return to an earlier location, state or activity</td>
</tr>
<tr>
<td></td>
<td>beibeik</td>
<td>constantly, habitually</td>
</tr>
<tr>
<td></td>
<td>fatik</td>
<td>usually, habitually</td>
</tr>
<tr>
<td></td>
<td>nimanimak</td>
<td>continually, forever</td>
</tr>
<tr>
<td>Pre-predicate</td>
<td>fudik</td>
<td>sporadically</td>
</tr>
<tr>
<td></td>
<td>biasa</td>
<td>usually, habitually [Malay]</td>
</tr>
<tr>
<td>Clause-final</td>
<td>terús</td>
<td>incessantly, continuously [Malay]</td>
</tr>
</tbody>
</table>

11.64 ...nia tama uluk bá mí, ... n-ák “Belu, ó tama 3S enter go first go urinate 3S-say friend 2S enter

osak bá”.
in turn go

[They had a urinating competition.] She went in first to urinate, and when she had finished, she came out, saying “Friend, you in turn enter now”. (G4.21)

11.65 ...Desi Malae n-á’ak “Ha’u bá fali”.

Desi non-native 3S-say 1S go in turn

[The previous emissary died.] Desi Malae said “I will go this time”.

(AA4.114)

11.66 N-ásai n-ola tenik rua né mai futu tenik, 3S-cause-exit 3S-take again two this come fight (cocks) again

nia manán tenik, mós ti’an...

3S win again finished already

Having brought out two more (slaves, to gamble for), (and) fought cocks again, she won again, (and it) was finished. [She had won all his slaves.]

(G4.63)

11.67 Kalo ita isin manas beibeik, fudik kiki, if 1PL body hot constantly sporadically tremble

fudik kiki...
sporadically tremble

If we continually have fever, and shiver on and off [, then Tetun-speaking people say of it “This is a swollen spleen”].

(K2.2)

The adverb hikar(r) ‘back’ (480 examples) is unusual in that it takes subject marking. The fact that hikar(r) inflects suggests that it may historically have been a non-initial verb in a serial verb construction rather than being an adverb. Synchronically, however, it cannot be so analysed. The ‘returning’ can be true of the preceding subject referent (11.68), or the object referent (11.69), or the truth of the proposition as a whole (11.70). Such behaviour is characteristic of modifiers rather than of non-initial verbs in serialisation.
11.7 Manner modifiers

Manner can be expressed by a phrase headed by an adjective (11.71, 11.72), or by an adverb which morphologically has the form of a full reduplication (11.73). Adjectival but not adverbial modifiers can be negated (11.71) or intensified (11.72). The manner phrase follows any postverbal object NP (11.72, 11.73).

11.71 ...túr la kmetis.
sit not tight
[She] doesn’t sit still. (X0.80)

11.72 Nia n-ú manu lós basuk.
3S 3S-blow.on bird straight very
He shot birds very accurately (with a blowpipe). (K12.10)

11.73 Ó kolu faru nainaik.
2S take.off clothes slowly
You (must) take off your clothes slowly. (G4.92)

Alternatively, manner modifiers can be introduced by the verb halo ‘make’, which takes subject marking agreeing with the subject of the clause. In the corpus halo introduces two types of manner complement. The first is adverbial complements, in which the adverb semantically modifies the predicate (11.74; 8 examples). The second type of complement is the adjective mós ‘finished’, with the combination halo mós meaning ‘completely, all’ and semantically modifying the subject (11.75; 12 examples).8

11.74 Rai n-aktidin n-alo ūi-ūit.
earth 3S-cave.in 3S-make RDP-little
The earth (around the well) has caved in a little. (P0.113)

11.75 Sia ne’e mate t-alo mós.
3P this die 3P-make finished
They all died. (G5.6)

---

8 In the corpus all manner modifiers introduced by halo modify an intransitive predicate; it is not known whether this is an accident of the data.
11.8 Deictic particles

Tetun has two commonly used deictic particles, *mai* (11.76) and *bá* (11.77), which indicate closeness to or movement towards the speaker and distance from or movement not towards the speaker respectively (over 100 examples each).[^9]

11.76 *Mane fila n-ika mai.*
man return 3S-back come

The man returned back (home/here). (U4.56)

11.77 *Sia r-odi to’o sia rai-n bá, iha Lakaluta Rai.*
3P 3P-bring reach 3P earth-GEN go LOC Lakaluta earth

*Lór bá.*
Lór go

They took (it) over to their country, over in Lakaluta Rai Lór. (K10.51)

In addition, there is the less common tone, which always indicates movement away from the speaker (16 examples). It is used mainly for movement towards the addressee (11.78) or towards some place associated with the addressee, such as his or her home or destination (11.79).

11.78 *Emi sës, lale ai n-aklati tone, kona emi.*
2P go.away else plant 3S-topple go.(to.you) touch 2P

You move away, else (when) the tree falls towards you, (it) will hit you. (This would be an appropriate warning when felling a tree.) (V2.117 elicited)

11.79 *Ibu at liu tone.*
mother IRR further go.(to.you)

Ibu is about to go on (to your house). (V0.112)

All three deictics are grammaticised from motion verbs which incorporate the same directional meanings as the corresponding deictics. Nevertheless, unlike the verbs, the deictic particles do not of themselves imply motion. When they follow an intransitive verb of motion (11.76, 11.78), the deictics indicate the direction of motion, whether towards or not towards the speaker. When they follow the object of a verb or preposition, however, their interpretation depends on context. In particular, if the modified clause or prepositional phrase contains no motion component, the deictic particle indicates the relative distance of the specified location from the speaker (11.80). In contrast, when it follows a motion verb and an NP specifying a distant source, as in 11.81, the deictic particle *mai* must be interpreted as designating the direction of motion towards the speaker, and not the proximity of the source. Where the NP gives the goal of motion, as in 11.77, the distance and direction of motion interpretations are in practice equivalent.^[10]

[^9]: In a preliminary count of 200 clauses involving the motion verbs *halai* ‘run’ and *fila* ‘return’, over 25% of the clauses included one of these deictic particles.

[^10]: This semantic bleaching supports the analysis of the deictics as particles and not serial verbs, as they are in some other Austronesian languages such as Paamese (Crowley 1987:50) and Numbami (Bradshaw 1993:148). According to Foley and Olson (1985:41, 48), ‘come’ and ‘go’ are in fact crosslinguistically the most favoured verbs for serialising. The crucial evidence for the particle analysis is, however, the fact that the deictics can be part of peripheral locative prepositional phrases (11.80).
Over in the east, Lakaluta Rai Lór Na’in really liked this message (idea).

(K9.13)

By far the most common use of the deictic particles is within locative prepositional phrases (second instance in 11.77, 11.80), or clauses specifying motion (11.79). The deictic follows any object NP (11.77, 11.81) and is usually clause-final. The sequence bá mai means ‘to and fro’.

The deictic particles are also found apart from contexts involving movement or location. All can indicate the ‘direction’ of speech or giving (e.g. Katak mai ‘tell come’ = ‘tell me’; fó tone ‘give to.’(addressee)’ = ‘give to you’). The basic deictic meaning is also retained in expressions involving perception, such as m-anono mai ‘2S-listen come’ = ‘listen here, listen to me’, haré bá ‘see go’ = ‘look (out)’, and titu bá ‘look outwards’.

Apart from these contexts, proximal mai (at least) is used in contexts which indicate some kind of change of state, such as moris mai ‘live come’ = ‘be born’, haktekir mai ‘aware come’ = ‘wake up’, hadér mai ‘arise come’ = ‘get up’. The limited data are consistent with observations made for both English (Clark 1974) and Thai (Gandour 1978:383) that ‘come’ (to which deictic mai is clearly related) indicates change to some ‘normal’ state.

11.9 Imperative bá

As an imperative marker, bá can be added to commands and invitations, particularly when the speaker is not intending to participate in the action (85 examples). As such it can only be used with second-person subjects (including terms of address, such as Malay Ibu ‘mother’).

The imperative marker follows the object NP, if any (11.83), but precedes clause-final adverbs such as lai ‘first’ (11.46) and dei ‘only’.

---

11 The deictic particles thus differ slightly in distribution from the deictic determiners/pronouns which incorporate them, namely nemai ‘here’ and nebá ‘there’. The latter can, though rarely do, occur without a preceding verb or preposition (§6.3.5).

12 Imperative bá shares some of the basic ‘motion away from speaker’ meaning of the verb bá ‘go’, in that the speaker will not participate; however, examples such as 11.83 show that it is perfectly compatible with requests to come towards the speaker. Note that invitations to join the speaker in doing something are introduced by mai ‘come’ (11.59; §9.9.2).
11.10 *bá* ‘so be it’

When it means ‘so be it, nevertheless’ *bá* characteristically occurs in the second of two clauses, where the initial clause describes a condition, while the second specifies acceptance of something despite that condition (5 examples). This acceptance is often of the continuance of the condition itself.

11.84 *Ha‘u k-ala‘a, k-ala‘a bá-n...*  
1S 1S-hungry 1S-hungry so.be.it-IMM  
(Even if) I were hungry, so be it, I would stay hungry [, because I don’t want to eat this very dirty food.]  
(V0.35)

11.85 *Emi hola di‘ak, hola bá. Ha‘u k-ola át, át*  
2P take good take IMP-so.be.it 1S 1S-take bad bad  
bá-n.  
so.be.it-IMM  
If you take good (men as husbands), do so (or ‘so be it’). If I take a bad (one), so be it.  
(U11.6)

11.11 Connective adverbs

Connective adverbs link two consecutive clauses. Some appear to be restricted to starting new sentences (complete with phonological sentence boundary and optional topic position), and some to linking clauses in multi-clausal constructions, while others are found in both contexts. It is likely that sentence-initial connective adverbs play a role in the broader structuring of discourse; this, however, has not been tested. Some connective adverbs (e.g. *mais* ‘but’) are glossed in other languages by what are traditionally analysed as coordinating conjunctions. At this stage, however, no syntactic distinctions have been observed which would justify recognising connective adverbs and clause-coordinating conjunctions as separate word classes.

Most connective adverbs are phonologically part of the following clause. Some common connective adverbs of this kind are listed below, with approximate glosses.

**Table 11.2: Connective adverbs which are phonologically clause-initial**

<table>
<thead>
<tr>
<th>Introduces Sentence</th>
<th>Adverb</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Introduces</strong></td>
<td><em>nia</em>má</td>
<td>then (temporal)</td>
</tr>
<tr>
<td><strong>Introduces</strong></td>
<td><em>nia</em></td>
<td>so, then</td>
</tr>
<tr>
<td><strong>Introduces</strong></td>
<td><em>ne‘e</em></td>
<td>so, then</td>
</tr>
<tr>
<td><strong>Sentence</strong></td>
<td><em>dadi</em></td>
<td>so, then</td>
</tr>
<tr>
<td><strong>Sentence</strong></td>
<td><em>huti</em></td>
<td>after that</td>
</tr>
<tr>
<td><strong>Sentence</strong></td>
<td><em>lale</em></td>
<td>else, otherwise</td>
</tr>
<tr>
<td><strong>Sentence</strong></td>
<td><em>foin</em></td>
<td>then (temporal), only then (conditional)</td>
</tr>
<tr>
<td><strong>Sentence</strong></td>
<td><em>mais</em></td>
<td>but</td>
</tr>
<tr>
<td><strong>Sentence</strong></td>
<td><em>má</em></td>
<td>however, but</td>
</tr>
<tr>
<td><strong>Sentence</strong></td>
<td><em>tapi</em></td>
<td>but</td>
</tr>
</tbody>
</table>
Several of these are apparently reduced adverbial clauses, which is not surprising since both connective adverbs and connective adverbial clauses fill the same clause-initial syntactic slot. In fact it is not always clear whether the word is functioning as an adverbial clause, or is fully grammaticised into a connective adverb. This uncertainty is evident for *hotu* ‘after that’, whose interpretation clearly derives from an adverbial clause meaning ‘(when that is) finished’.

Similarly, clause-initial *lale* ‘else, otherwise’ (76 examples) can often be understood as an implicitly conditional pro-clause meaning ‘(if) not’. Nevertheless, its frequency in clause-initial position, as well as the fact that it can follow the conjunction *ká* ‘or, maybe’ to mean ‘or else’ (6 examples, mostly from Indonesian-influenced people), suggests that *lale* has been at least somewhat grammaticised into a connective adverb.

11.86  
*ita*  *tula*  *na’an*  *halo*  *án*,  *lale*  *asu*  *n-á*  *n-ola*.

1PI put.high meat make high else dog 3S-eat 3S-take

We put meat up high, otherwise (or: ‘if not’) the dog eats (it) up.  

(Q0.70)

The connective adverb *foin* ‘then’ is unusual in that although it is usually clause-initial (11.62), it can alternatively follow the subject (11.87). It only rarely begins a new phonological sentence.

11.87  
*ema*  *rúa*  *ne’e*  *mate*  *lai*,  *ita*  *foin*  *mate*  *hatene*...

person two this die now 1PI then die know

These two people died first, and only then did we know death. [If they hadn’t died, there would be no death.]  

(W0.30)

The three connective adverbs meaning ‘but, however’ all start new phonological sentences (i.e. follow sentence-final punctuation in the transcriptions) in approximately 50% of the examples.

11.88  
*Há*,  *bele*  *há*.  *Má*  *la*  *toba*,  *é*.

eat can eat however not lie.down  

(With regard to) eating, (we) can eat. But (we) don’t sleep.  

(J3.31)

A second group of connective adverbs characteristically connect two clauses within the one sentence, phonologically linking with the preceding and/or the following clause. Some common examples are listed below.

<table>
<thead>
<tr>
<th>Introduces Clause</th>
<th>Adverb</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>bei</em></td>
<td><em>má</em></td>
<td><em>daudaun</em> (to’o) however, but and then continue until</td>
</tr>
<tr>
<td><em>be</em></td>
<td></td>
<td>Also connective adverb: ‘however’ to’o is also preposition/conjunction</td>
</tr>
</tbody>
</table>

The connective adverb *bei* ‘however, but’ is always phonologically part of the preceding clause, in contrast to the semantically similar connective adverbs *mais* ‘but’, *má* ‘however, but’ and *tapi* ‘but’.

11.89  
*Ha’u*  *at*  *la’o*  *ta*  *bá*  *bei*,  *uma*  *sé*  *dakar*?

1S IRR walk already go however house who look.after

I wanted to go there but who would look after the house?  

(X0.105)
The connective adverb má ‘and then’ can group phonologically with either the preceding or the following clause. In addition to joining full clauses (11.90; 62 examples), it is common in complex connective adverbs which derive from reduced clauses and which mean ‘then’. These include nia má lit. ‘3S and.then’ (193 examples), nia ti’a má lit. ‘3S already and.then’ (8 examples), and hotu má ‘finish and.then’ (19 examples).

11.90  Tama alas laran á má manu kokorék.
      enter forest interior DEF and.then bird crow
      (They) entered the middle of the forest, and then a cock crowed.  (D1.25)

The term daudaun (or daaun, daudauk, dadau) indicates that the action described in the preceding clause has commenced but is not yet completed. This initial clause is nearly always followed by an NP introduced by the preposition to’o ‘until’ (11.91), or by another clause which is often but not necessarily introduced by to’o (11.92). This second constituent indicates a time, place or event until which the activity mentioned in the first clause continues.

11.91  Ke’e utu daudaun to’o kalan...
      dig louse continue until night
      (He) kept picking out lice until it was night...
      (Z5.16)

11.92  ...Tunu daudaun tasa ti’an...
      bake continue cooked already
      [He baked the pigeon and bananas.] (He) baked until (they) were cooked
      [, then ...]
      (K12.19)

While daudaun is normally phonologically part of the preceding clause (77 examples), it can also start a new phonological sentence or clause (11.93; 26 examples). In this case it is understood that a situation described in a previous clause or clauses is continuing. 13

11.93  Mai r-ikar r-alo, r-alo sobu r-alo sobu
      come 3P-back 3P-make 3P-make demolish 3P-make demolish
      r-alo sobu, nú nia. Dadaun to’o loron ida...
      3P-make demolish like 3S continue until day one
      (They) came back to build, and built and demolished and built and demolished and built and demolished, like that. And so it went until one day [this Kitten came out again.]
      (Z2.31)

11.12  Sentence adverbs of possibility

There are a number of sentence adverbs which specify epistemic modality, with meanings similar to ‘perhaps, maybe’. I was unable to elicit comparable words to indicate certainty, with the nearest being the adverb dei ‘only, just’ (which can be used to indicate insistence), and the tag lâ, which adds certainty (§14.2.5). This situation parallels that for deontic modality, where there are native Tetun auxiliaries indicating permission (corresponding to epistemic possibility) but none indicating obligation (corresponding to epistemic necessity).

Apart from ruma ‘perhaps’, which is clause-final (11.97), these words are usually found clause-initially (11.95), although they can also follow the subject of the clause (11.98) and

13 Some speakers allow daudaun to be repeated, to emphasise iteration (daudaun daudaun ‘and so on, and so on’). This is a characteristic of verbs. Nevertheless, unlike verbs, daudaun has no subject.
can even interrupt a phrase (11.94). They are only loosely integrated into the clause, being what Quirk and Greenbaum (1973:126) call ‘disjuncts’, and what Foley and Van Valin (1984:215) call ‘peripheral operators’.

Several of these epistemic modality words can occur in combination, including de’ik nambé, keta aruma and keta...ruma (11.97).

The words de’ik (11.94; 18 examples) and aruma ‘perhaps’ (11.95; 16 examples, mostly elicited) both mean ‘perhaps, approximately’.

11.94 ...“Matan de’ik hát ká lima.”
CLS:animal perhaps four or five
[“How many buffalo?”] “About four or five head.” (Q0.100)

11.95 Aruma la n-ó.
perhaps not 3-exist
Perhaps (vowel-initial verbs, for which we were unsuccessfullly searching)
don’t exist. (Q0.112)

The word keta (20 examples) similarly means ‘perhaps’. In questions it is polite and could be translated ‘by any chance, perhaps’. It can be repeated (keta keta) with no clear change in meaning.14

11.96 N-ák “Hó, keta ha’u k-ó namane”.
3S-say Oh perhaps 1S 1S-have woman’s.brother
(He) said “Oh, I thought perhaps I had brother(s)”.

11.97 Lalika rekam! Keta luli ruma!
need.not tape.record[Mly] perhaps taboo perhaps
No need to record (what I say about my sacred house)! What if/Perhaps it is
taboo! (T0.47)

The adverb nambé ‘I suppose’ (21 examples) indicates more certainty than de’ik or aruma.15 Almost all examples in the corpus mention some sort of evidence, however loose, on which the assumption is made. For instance, in 11.98 the speaker supposes that a child in a photo must have been afraid, because she looked like she was about to cry.

11.98 Ida ki’ik oan ne’e nambé n-ata’uk ai.laran bót é, one small small this I.suppose 3S-fear forest big this
nia at tanis.
3S IRR cry
This little child is I suppose frightened of the big forest; she is about to cry.

K6.19

14 The form keta is also an auxiliary meaning ‘do not’ (§10.4.4).

15 The adverb ruma has a second, related, use in rhetorical and reflected questions, where it immediately follows the question word, and implies ‘How would I know?!’ (§9.9.3.2). For instance, the question Ne’e sá? ‘this what’ = ‘What is this?’ could elicit the reply Sá ruma?! ‘How would I know?!’.

16 The adverb nambé has unusual phonology for Tetun, suggesting that it is either a loan or a compound. However, I have not found a convincing source for it. It is (like de’ik) missing from the dictionaries of Mathijsen (1906) for the Foho dialect, and Hull (1996b) and Morris (1984b) for East Timorese Tetun.
The word *k-anoin* ‘1S-think’ can, as a complement-taking verb, be followed by a complement clause which specifies what is thought (§13.3.2). However, it can also be placed within such a clause (11.99; 3 examples). This suggests that, in addition to being an inflected verb, it is grammaticised as an adverb (presumably with origins as a parenthetical clause).

11.99 *Ema rai klaran ida k-anoin mate.*

person earth middle one 1S-think die

An earth person has I think died. (This was said by a heavenly being who smelt a dead body while visiting earth.) (Q0.189)
12 Serial verb constructions and prepositional verbs

12.1 Overview of serial verbs

12.1.1 Introduction

Constructions containing two or more verbs fall into a few broad categories. Those in which the verbs belong to separate clauses are discussed in Chapter 13 (on complementation) and Chapter 14 (on other inter-clausal relations). Those in which one verb functions as modifier to another have already been dealt with in Chapter 11.

The present chapter focuses on the remaining broad category of relationships between verbs, namely that found in serial verb constructions. In these constructions, two or more verbs come under a single clause, without any one functioning as modifier to another.

The chapter begins with a survey of the characteristics shared by all serial verb constructions, and an overview of the constructions to be discussed in this chapter. Following this, the various types of serial verb construction are discussed in turn. Where the constructions bear resemblances to other, non-serialised, constructions, comparisons between the two are discussed also.

Verbs in certain types of serial verb construction have a diachronic tendency to develop into prepositions, via the intermediate stage of prepositional verbs. This diachronic development results in close synchronic relationships between serialised verbs, prepositional verbs and prepositions which are presumed to be grammaticised from verbs. The latter two classes of words are therefore dealt with in this chapter also (§12.5).

12.1.2 Characteristics of serial verbs

In Tetun serial verb constructions, a single clause contains two or more verbs, which share the following features.1

1. No verb is subordinate to another, or modifies another.
2. Verbs in serialisation fall under a single intonation contour, and are not separated by the types of pauses and intonation patterns found at clause boundaries. However, the length and types of pause used in speech vary enormously, and depend not only on syntactic structure, but also on such matters as overall speed of speaking, hesitancy and style.

1 These features correspond to those pointed out for serial verb constructions in many publications, including Crowley (1987:38), Foley and Olson (1985) and Schachter (1974).
Thus, until further study of Tetun pause and intonation patterns has been done, the use of intonation as a guide remains impressionistic.  

3. There is no syntactic marking of a clause boundary between the verbs, such as complementisers, conjunctions and connective adverbs (e.g. dadi ‘so’, niama ‘then’).

4. There are no intervening peripheral constituents (such as time or location phrases) or sentence adverbs (e.g. aruma ‘perhaps’).

5. Each verb shares one or two core arguments with the neighbouring verb. In particular, the subject of a second verb is always interpreted as being identical to either the subject or the object of the preceding verb. In some constructions two verbs share both subject and object.  

6. The verbs share tense, aspect, deontic modality and negation. Syntactically, this has the following implications.
   a) Preverbal auxiliaries (e.g. bele ‘can’, atu ‘IRR’) and the preverbal negator la precede the first verb.
   b) Clause-final aspectual markers (e.g. ti’an ‘already’) come after the final verb.
   c) The postverbal negator ha’i and other postverbal modifiers such as hotu ‘all’ follow either the first or a subsequent verb, depending on the subtype of serial verb construction. For details see Table 12.1.

7. A serial verb construction has a single illocutionary force. It cannot, for instance, contain both a conditional and an assertion.

8. The first verb takes subject marking. Subsequent verbs do not take subject marking in nuclear serialisation, but do in core layer serialisation if the phonology allows. Phonology is relevant because all /h/-initial verbs in core layer serialisation take subject marking, while verbs beginning in other consonants take subject marking only if they are the first non-/h/-initial verb in the series (e.g. k-ba toba ‘IS-go lie.down’).

9. In serial verb constructions, the sequence of verbs is semantically interpretable as referring to sub-parts of a single overall event. This statement (which is made in many published discussions of serial verbs) is of course open to the charge of circularity, since ‘event’ tends to be defined in linguistic terms as that which is expressed by a single unit (Givón 1991:140). Nevertheless, despite its problems, the statement retains intuitive appeal.

10. Where there is an iconic order of the sub-events depicted by two verbs, that iconic order must be followed. The only construction in which both orders appear to be identical, and

---

2 Pause in actual spoken data can probably never be more than a useful guide. One reason is that pause length inherently involves continua rather than discrete cut-off points (Matthews 1981:33). Another is that the likelihood of pause is probabilistic. Both the usefulness and the probabilistic nature of pause were shown by Givón (1991), who measured pauses in consultants’ descriptions of a film. He found that pauses within serial verb constructions occurred with a similar probability to pauses associated with lexical words within the clause. This was dramatically lower than the probability of inter-clausal pauses.

3 I have no clear evidence for the ‘ambient serialisation’ proposed by Crowley (1987:40, 49), in which the second verb has no NP subject, but rather predicates something about the preceding predicate. The functions of ambient serialisation are in Tetun performed by adverbs or by verbs functioning as modifiers (§11.7).
for which there is in fact no iconic order, is that which uses *mesan* 'alone' to indicate that the actor is acting alone (§12.6).4

12.1.3 Comparison with paratactic clauses

The boundaries of what constitute serial verb constructions in Tetun are somewhat fuzzy. This is in part due to the paucity of formal grammatical clues which indicate the boundaries of either clauses or verb phrases. Subject marking is the only obligatory grammatical marker in clauses, and even it is only obligatory under certain phonological conditions.

In particular, it can be difficult to determine whether a sequence of verbs (with or without following object NPs) constitute a serial verb construction, or whether the verbs are in separate clauses which are related by parataxis.5 Contributing to this difficulty is the fact that the patterns of argument sharing in serial verb constructions also constitute common patterns of zero anaphora across clause boundaries. Thus one statistically common pattern of zero anaphora is for an omitted subject to be interpreted as being identical to the subject of the previous clause, while the omitted object is interpreted as identical to the object of the previous clause (12.1). Alternatively, objects of transitive verbs are readily interpreted as coreferential with subjects of intransitive undergoer-subject verbs (12.2). In the examples below, omitted subjects are indicated by ‘(*S)’ and omitted objects by ‘(*O)’.

12.1  

<table>
<thead>
<tr>
<th>Nia n-ú n-ola lakateu ida.</th>
<th>He successfully shot a pigeon.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3S 3S-take pigeon one</td>
<td></td>
</tr>
</tbody>
</table>


\[(*S)\] N-odi (*O) mai, sia bá iha laen, sia tunu (*O).

| 3S 3S-bring pigeon come 3P go LOC hut 3P bake pigeon |
| (When he) brought (it) back, they went to the hut, (and) they baked (it). |

(K12.12)

---

4 In addition, *hó* ‘accompany’ (§12.5.4) and *hodi* ‘bring, use’ (§12.5.5) can either precede or follow another verb; however, when these two words occur initially they are truly verbs, while when they follow the other verb, they are classed as prepositional verbs (so that the construction is no longer one of serialisation).

5 This problem is recognised by Early (1994b:370) for the Vanuatu language Lewo, and more generally by Baker (1989:546–549). Baker notes that serial verb constructions tend to be perceived by native speakers as representing single events, while coordinate clauses report distinct events. In addition he notes that coordinations allow intervening pause (as in Tetun), and provides a series of syntactic tests for determining whether a sequence of NP VI NP l V2 NP2 is a serial verb construction or a ‘covert coordination’. Those which I have been unable to apply to Tetun are:

a) In serial verb constructions, either object NP can be fronted by *wh*-extraction. In coordination, neither NP can. (This test is also used by Bradshaw (1993:145). The closest Tetun equivalent to *wh*-extraction is topicalisation, which fronts an object NP to before the subject (§9.4.3.3). The object of the second verb in serialisation can be fronted in some serialisation types, but not in comitative or instrumental ones. (See Table 12.1 in §12.1.5 for a summary.) I have found no instances of fronting from an indisputably coordinate structure of the type Baker mentions.

b) In serialisation, the two object pronouns (NP1 and NP2) cannot be coreferential. In coordination, they can. I have no data for this in Tetun.

c) Predicate clefting (which Tetun does not have) distinguishes the two constructions.
Serial verb constructions and prepositional verbs

12.2 (*S) Hawai (*O) ti’a, (*S) maran;
1PI dry(s.th.) cassava already cassava be.dry
[We dry the cassava.] (When we) have dried (the cassava), (it) is dry.

(*S) Maran ti’a, ita ra’ut (*O) (*S) hodi (*O)
cassava be.dry already 1PI gather cassava 1PI bring cassava
mai uma...
(Once it) is dry, we gather (it and) carry (it) home...

Some of the problems in recognising serial verb constructions are illustrated in the following example, where ‘/’ represents a clause boundary.

12.3 Ita bá haré iha to’os, tá hola / hodi mai /? hisa.
1PI go look LOC garden chop take bring come suspend
We go and look in the garden, chop (a bunch of bananas), bring (it home and) hang (it) up.

In this example, the comma marks a clear clause boundary, one which in any case would be recognisable by the preceding peripheral locative phrase ‘in the garden’.

The following sequence of two verbs, tá hola ‘chop take’, represents a common pattern of serialising a transitive verb and hola ‘take’ (§12.3), with an object NP, if any, preceding or following the entire sequence. Since tá hola is also followed by a very short pause (intuitively too short to be recorded as a comma), and does not form a common sequence with the following verb hodi, I tentatively interpret this two-verb sequence as forming a clause of its own. A tertiary-educated consultant agreed with this positioning of a clause break. Further support comes from the possibility of placing an object NP before the following sequence hodi mai, as in the following example.

12.4 ...bá te’in nono we manas / hodi mai /
go cook heat(.liquid) water hot bring come
sia r-emu...
3P 3P-drink
[Then] (I) went and boiled water and brought (it and) they drank [, and they ate.]

The final sequence of three words in 12.3, hodi mai hisa ‘bring come suspend’, also fits a common pattern of a verb of motion followed by mai ‘come’ or bá ‘go’, followed in turn by another verb which shares the same subject. However, this sequence of three words in turn consists of two very common patterns. The first is a verb of motion followed by mai or bá, which in this context are analysed as deictic particles rather than verbs (§11.8). The second consists of mai or bá followed by any other verb (§12.4). It is not clear whether the three words constitute one serialisation within another, or instead a single serialisation (either of the two main verbs with the deictic analysed as an intervening particle, or of three verbs with the deictic interpreted as a verb). In any case it intuitively seems to me that the deictic is more closely associated with the preceding than the following verb. This is, impressionistically, supported by the intonation, since if there is any perceivable break at all, it follows rather than precedes the deictic.
Serial verb constructions can be negated only once. The preverbal negator *la* necessarily precedes the first verb in the series (12.15, 12.62, 12.67). The postverbal negator *ha'ì*, however, follows the first verb in some constructions (12.5, 12.6, 12.25), and the second verb in others (12.16). In each case it fits the same slot as other postverbal modifiers such as *hika(r)* 'back, again'. (For details see Table 12.1 in §12.1.5.)

The scope of negation is indeterminate. It can be the entire sequence or only the second verb (with its object), but it cannot be only the first verb (and object). Thus, for instance, example 12.5 represents the words of a heavenly woman who could no longer fly, and who was therefore negating the whole proposition 'I will ascend to the sky'. In contrast, 12.6 explicitly says that the subject referent was going out (so affirming the first verb), but that he was not going out to the garden (so denying the second). Both examples use the same motion-direction serialisation construction (§12.5.2.2).6

12.5  
\[ \text{Ha'u so'e ha'ì bā lāle'an ti'ān.} \]
\[ 1S \text{ ascend not go sky already} \]
I will no longer ascend to the sky (as I have lost my wings). (00.63)

12.6  
\[ \text{Nia n-ā ti'a, saì bā, saì ha'ì n-o'ì to'os bā...} \]
\[ 3S \text{ 3S-eat already exit go 3S exit not 3S-go.to garden go} \]
(When) he had eaten, (he) went out; he didn't go out to the garden [, but crouched at the base of a banana tree, to watch.] (P3.15)

12.1.5 Overview of construction types

In this chapter I distinguish seven types of serial verb construction, which each have characteristic syntactic constraints and semantic ranges. Although these types do not represent a full list of what is possible in Tetun, they do give some indication of the diversity that is found within verb serialisation.

While verb sequences can consist of more than two verbs, it appears that each verb in such a series relates to at least one other verb in the series in the same way as if the construction contained only those two verbs. This point is illustrated in the discussion above of the sequence *hodi mai hisa* 'bring come suspend' (§12.1.3), in which *mai* 'come' relates to both the preceding and the following verb, thus mediating between the two. For this reason, the discussion in the rest of this chapter will normally make the simplifying assumption that constructions consist of only two verbs.

In differentiating between various types of serial verb construction, it is useful to draw a distinction between 'nuclear', 'core' and 'peripheral' layers of clause structure. These distinctions and terms were proposed within Role and Reference Grammar by Foley and Van Valin (1984:77ff.), and further developed by Foley and Olson (1985) and Van Valin (1993:106ff.). They have been found helpful for the analysis of serial verb constructions in many other Austronesian languages also (Bugenhagen 1995; Crowley 1987; Early 1993; Grimes 1991a:391, 399).

---

6 This same indeterminacy of the scope of negation has been noted for serial verbs by Bruce (1988:27) and Stahlke (1974:274), and is well known in English also for certain adverbs and prepositional phrases, amongst others (Giv6n 1979:105ff.; Huddleston 1984:428ff.). The indeterminacy is true of at least motion-action serialisation (type 4) and motion-transitive direction serialisation (type 5). For the remaining construction types there are as yet insufficient data concerning scope of negation.
The innermost layer of the clause is the nucleus. In Tetun it consists of a verb without its subject marker. The next layer out is the core. It consists of the nucleus, the subject marker and the core arguments of the verb, where core arguments are those which are determined by the valency of the verb. The outermost clause layer is the periphery. This includes constituents such as time, spatial setting and beneficiary phrases, which are not determined by the valency of the verb, as well as most adverbial modifiers. The clause as a whole consists of the core plus the periphery.

A token of any layer can be joined to another token of that same layer. Thus, in Tetun, a clause nucleus can form a serial verb construction with another clause nucleus. The entire construction makes up a single nucleus, and so shares all arguments (since these are in the core, outside the nucleus), and all modifiers. Similarly, a clause core can serialise to another core. In this case the single resulting core shares all of the periphery. However, each core can have its own arguments, subject to the rule that in serialisation constructions the subject of the second must be identical with one of the arguments of the first verb.

In addition to the distinction between nuclear and core layer serialisation, there are a number of other parameters on which serial verb constructions differ. The overall list of parameters is presented below.

1. Some serialisation is at the nuclear layer, while some is at the core layer.

2. In some constructions, the verbs are necessarily contiguous. In others, adverbs or object NPs can intervene between the two verbs.

3. In some constructions, postverbal modifiers (such as ha'i 'not', hika(r) 'back') must follow the first verb, in others they follow the second verb, and in yet others it appears that some modifiers follow the first verb while others follow the second.

4. The object of the first verb (if any) may either immediately follow that verb, or instead follow the second verb.

5. The subject of the second verb may be coreferential with either the subject or the object of the first verb.

6. The object of the second verb may be fronted in a topicalisation construction for some serialisation types but not for others.

7. In most constructions one of the verbs is selected from a closed class. This is usually the second verb but may be the first.

8. In tail-head constructions (§14.4) it may be either the entire serialisation or just one part of it that is repeated.

These differences are summarised in the table below, where the numbered headings in the left-hand column correspond to the parameters listed above.

---

7 Tetun, like the Oceanic language Lewo (Early 1993), amongst others, is thus a counterexample to Foley and Olson's (1985) generalisation that verb medial languages "decidedly disfavor" nuclear serialisation.

8 Note that some of the information on tail-head repetition is tentative due to insufficient data.
Table 12.1: Serial verb construction types

<table>
<thead>
<tr>
<th>Type Description</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclear causative</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Motion-direction</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Core</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Motion-action</td>
<td>V2</td>
<td>V2</td>
<td>V2</td>
<td>V1</td>
<td>V1/V2</td>
<td>V1/V2?</td>
</tr>
<tr>
<td>Motion-direction</td>
<td>-</td>
<td>V1</td>
<td>V1</td>
<td>V1</td>
<td>V1</td>
<td>V1</td>
</tr>
<tr>
<td>Instrument</td>
<td>both</td>
<td>both</td>
<td>both</td>
<td>V2</td>
<td>V2</td>
<td>V1</td>
</tr>
<tr>
<td>7. Closed class</td>
<td>V2?</td>
<td>V2</td>
<td>V2?</td>
<td>V1</td>
<td>V2</td>
<td>V1</td>
</tr>
<tr>
<td>8. Tail-head repeats</td>
<td>both</td>
<td>both</td>
<td>both</td>
<td>both</td>
<td>both</td>
<td>both?</td>
</tr>
</tbody>
</table>

Participant cooperation (type 7) serialisation (using hamutu(k) ‘together’ or mesan ‘alone’) is omitted from the table on account of insufficient data.

In the remainder of this chapter, the seven recognised construction types are discussed in the order in which they are presented in the table. Thus nuclear serialisation is considered first, followed by core layer serialisation. The formulae presented for serial verb constructions are simplified in two ways. Firstly, ‘(...)’ is used to represent optional modifiers and (preceding the first verb) auxiliaries. Secondly, the formulae present all objects in postverbal position, whereas it is possible for objects in some serial verb constructions to be fronted to pre-subject position.

12.2 Nuclear serialisation

12.2.1 Overview

The distinguishing features of nuclear serialisation all follow from the distinction between nucleus and core. They are:

1. In nuclear serialisation the two verbs constitute a single clause nucleus. They share a single set of arguments, which precede or follow the sequence as a whole. The transitivity of the construction is determined by the transitivity of the initial verb. In contrast, verbs in core layer serialisation each retain their own transitivity.

2. Verbs in nuclear serialisation are inseparable. Thus any modifiers precede or follow the sequence as a whole.

3. Only the first verb in a nuclear serial verb construction is eligible for subject marking. Both verbs are eligible for subject marking in core layer serialisation.

9 This statement is based on limited data, but accords with what one would expect for nuclear serialisation. There are two types of nuclear serialisation. None of the second-slot verbs in type 2 serialisation are /hi/-initial, and so none are by virtue of their phonology eligible for subject marking. In nuclear causative (type 1) serialisation there are some textual examples in which the second word is /hi/-initial (horas ‘sick’, hat ‘bad’), but since these particular forms cannot occur predicatively they are adjectives and as such ineligible.
Two types of nuclear serial verb constructions are distinguished, namely causative serialisation (type 1) and motion-intransitive direction (type 2).

12.2.2 Nuclear causative (type 1)

12.2.2.1 Description

A situation in which an action causes a resulting condition may be described by a transitive clause in which the verb slot is occupied by two consecutive verbs. The first is a transitive verb which describes how the situation was caused (although the general verb halo 'make' is allowed also; 12.9), while the second is an intransitive verb or adjective describing the resulting condition of the undergoer (32 unambiguous examples). It is unclear whether the lexical items that can fill the second, intransitive, slot of this construction constitute an open or a moderately large closed class.

Predicate.Nuclear.Causative $\rightarrow \{ Vt \} \{ VI, Adj \}$

12.7 N-itak be'o ta lolo ã...
3S-throw.on.ground shatter already earthen.water.pot DEF
Having thrown the earthen water pot to the ground such that it shattered...

12.8 Ó mai, m-a ñit nù kain ne'e lai.
2S come 2S-eat be.cut.off coconut stalk this first
You come, and chew through this coconut stalk now.

12.9 Rai na’in halo horas ita.
earth noble make sick IPI
Spirits (lit. 'lords of the earth') make us sick.

This construction may only be used if the causation was successful. This is reflected in the fact that the second verb may not be independently negated. Although all textual examples involve intentional causation, consultants insisted that this need not be the case.

A unique feature of this construction is that the subject of the second verb follows that verb. This follows from it being the object of the verb series as a whole. Nowhere else in Tetun (except in existential clauses) can subjects follow the verb.

Where the object NP is either omitted or fronted (12.10), only intonation assists one in determining whether a semantically appropriate sequence of transitive and intransitive verbs constitutes a single causative predicate or instead consists of an action clause followed by a separate result clause.

12.10 Na kakun ema buti be'o ti'an.
3S shell person massage shatter already
His shell someone had squashed to pieces.

The biclausal interpretation is forced if the second verb is independently negated.

for subject marking. The only relevant evidence thus comes from elicitation, in which subject marking was disallowed within a nuclear causative serialisation for the otherwise inflectable verb hakesir 'be weary'.
12.11 Kabau kulit e'e dade’o / la be’o.
buffalo skin this trample not shatter
Buffalo skin—if one) tramples (it, it) doesn’t shatter. (U0.31)

12.2.2.2 Comparison with morphological causatives

This nuclear serialisation construction has much in common with morphological causatives, in which a causative prefix ha- is followed by an intransitive result verb or adjective (§4.2.2). Both share the same constituent order, in having the subject precede the prefix/transitive verb, and the (non-topicalised) object follow the intransitive result.

Both are largely consistent with Givón’s (1980) binding hierarchy for manipulative verbs. Syntactically they reflect a high degree of binding, and semantically they involve an actor exercising strong influence on an undergoer with a high degree of success.

Most verbs in the result slot of nuclear causative constructions can also function as sole verb in a clause; however, some can not. It appears that these verbs with restricted distribution can all also function as base in a causative derivation. One such verb is rahu ‘to pieces, ruined’ (12.12; ha-rahu ‘break into pieces’), which cannot function predicatively.

Another is horas ‘sick’, which functions attributively (e.g. ulu horas ‘head sick’ = ‘a disease in the ‘head’ of fruit that causes it to fall from the tree prematurely’), as a result verb in nuclear causatives (12.9) and as a base in morphological causatives (ha-horas ‘make sick’); in predicative constructions the form moras ‘sick’ is used instead.

12.12 Ó fota rahu kfui né, ita rua fa’e malu onan.
2S hit to.pieces flute this 1P I two divide each.other IMM
(If) you smash this flute to pieces, we two will separate. (F2.35)

The adjective at ‘bad’ can in this serial verb construction (but not when functioning predicatively or attributively) optionally have an initial /h/, resulting in the sequence halo (h)at ‘make bad’ = ‘ruin’. In morphological causatives /h/-insertion is a compulsory rule which applies to all vowel-initial bases (§2.10.3.1), resulting in derivations such as ha-h.at ‘make bad, ruin’. I was, however, unable to find other vowel-initial words which acquired initial /hl within verb sequences. For instance, as ‘high, tall’ and ilas ‘beautiful’ cannot take initial /hl in nuclear causatives.

12.2.3 Motion-intransitive direction (type 2)

12.2.3.1 Description

The second class of nuclear serial verb construction in Tetun involves an initial verb which is typically a verb of motion (although uluk ‘go first’ can serialise to non-motion verbs also). This is immediately followed by a verb which gives more information about that motion, often about its direction.

Predicate. Motion-Intr. Direction \( \rightarrow \) (...) V.motion V1 (... O)

12.13 Sia n-alai sai.
3P 3-run exit
They ran outside. (K10.109)

Only six verbs have been noted to fill the second slot in this construction; these are listed below, with nuclear serialisation examples of each.
**Table 12.2: Direction verbs in nuclear serialisation**

<table>
<thead>
<tr>
<th>Verb</th>
<th>Example</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>sa'e</code></td>
<td><code>samo sa'e</code></td>
<td><code>fly up</code></td>
</tr>
<tr>
<td><code>tün</code></td>
<td><code>la'oi tün</code></td>
<td><code>walk down</code></td>
</tr>
<tr>
<td><code>tama</code></td>
<td><code>hamán tama mai</code></td>
<td><code>walk [no noble] enter = come in</code></td>
</tr>
<tr>
<td><code>sai</code></td>
<td><code>n-akfút sai</code></td>
<td><code>3S-spurt out</code></td>
</tr>
<tr>
<td><code>liu</code></td>
<td><code>monu liu</code></td>
<td><code>fall further</code></td>
</tr>
<tr>
<td><code>uluk</code></td>
<td><code>sa'e uluk</code></td>
<td><code>ascend go.first = go up first</code></td>
</tr>
</tbody>
</table>

The first four predicate intrinsic direction, and occur in serialisation with preceding verbs specifying manner of motion.\(^{12}\) The final two, `liu` 'go further' and `uluk` 'go first', serialise with a wider range of motion verbs, including the four intrinsic direction verbs (e.g. `tama liu` 'enter further' in 12.15). The verb `uluk` 'go first' can, in addition, serialise with verbs which involve no motion (e.g. `mate uluk` 'die go.first' = 'die first (before someone else dies)'.

In this construction, the two verbs are immediately contiguous, with any postverbal modifiers following the second verb (e.g. 'already' in 12.14; `ha'i` 'not' in 12.16; `halai sai ha'i` 'run exit not' = 'not run outside'; *halai ha'i sai*). If there is a locative complement (12.15) or the verb series is in core layer serialisation with another verb (bá in 12.14, to'o in 12.28), the complement or subsequent serial verb also follows the second verb. (Note that clause-final `bá` in 12.15 is a deictic particle rather than a verb.)

12.14  *Fahi n-atai uluk ta bá rai bön ti'an.*

The pig had run ahead to the cave.

12.15  *...Tudik e'e la tama liu laran bá, to'o lerek kulit dei.*

[He stuck a knife into himself.] The knife didn't enter right into the body, only reaching to the skin.

12.16  *Feto Ikun n-adér uluk ha'i té matan dukur ne'e...*

Youngest Sister didn't get up first (i.e. before her husband) because being sleepy [. she slept in.]

---

\(^{10}\) Note that `liu` is also a verbal modifier of degree and is used in comparatives (§11.3).

\(^{11}\) The word `uluk` is also a time noun meaning 'former times, formerly'.

\(^{12}\) In his survey of ten Oceanic languages, Durie (1988:11) found that all had a well-defined class of intrinsically orientated verbs which behave in much the same manner as those of Tetun. According to Baker (1989:533), serialisations of manner of motion and direction of motion verbs are very common crosslinguistically, and invariably occur in this order. The fact that these directional lexemes precede the negator `ha'i` fits nicely with Foley and Van Valin's (1984:208ff.) analysis of intrinsic directionalson (i.e. terms which do not make reference to speaker and addressee) as 'operators' whose scope is the nucleus. Note that directionaldirections which do use the speaker or addressee as reference point (bá 'go (usually from speaker)', mai 'come (to speaker)', tone 'go (usually towards addressee)', §11.8) follow the negator `ha'i`, and so are core operators.
12.2.3.2 Comparison with verb-modifier constructions

Motion-intransitive direction serialisation has some features in common with head-modifier constructions in which both head and modifier are verbs. The latter include verbal modifiers of degree (§11.3.2) and manner (§11.7). In both the serialisation and the head-modifier constructions, transitivity is determined by the first verb, the second verb can be interpreted as providing information about the first, and the second verb cannot take subject marking.13

Nevertheless, there are significant differences. The second verb in serialisation predicates something of the subject, while a modifying verb modifies the predication of the preceding verb in some way. As a consequence, while verbs in serialisation can be active verbs (e.g. sa 'e 'enter'), active verbs are not readily interpretable as modifiers.

Syntactically, two verbs in nuclear serialisation must be contiguous, while verbs functioning as degree or manner modifiers follow the object NP, if any. A final difference is that manner modifiers can be explicitly negated (e.g. la'o la di'ak 'walk not good' = 'walk badly'), while the second verb in serialisation can not.

12.3 Transitive verbs sharing subject and object (type 3)

We turn now to core layer serialisation. In type 3 serialisation, two consecutive transitive verbs share both subject and object. The shared object, if present, occurs either after the second verb (12.17) or topicalised in clause-initial position (12.19). Modifiers too either precede or follow the verb sequence as a whole (12.20, 12.21). Both verbs in this core layer serialisation construction take subject marking if they meet the appropriate phonological conditions (12.20).

\[ \text{Predicate.Core.Transitive} \rightarrow (...) \text{Vt} \text{Vt} (...) \text{O} \]

The second verb slot appears to be reserved for a closed class of verbs, although the extent of this class is not known. The most common semantic relation is one in which the second verb represents the result of the first.

By far the most common verb for the second slot is hola 'take, fetch, hold' (over 300 examples). In most verb sequences it retains its literal meaning, so that the result of the first verb is that the actor holds or acquires the undergoer (12.19). This is the case, for instance, in the idiomatic expression té hola 'defecate hold' = 'give birth'. However, the verb sequence can also take on a more general meaning of success, such as in kahi hola 'dissuade take' = 'dissuade successfully', or haré hola 'see take' = 'intentionally see, take sight of' (which alternatively has a biclausal interpretation as 'see, then grasp').

12.17 Nia simu n-ola sala ne'e.
3S receive 3S-take fine this
He accepted this fine. (J4.39)

---

13 The similarity between adverbs and certain verbs in serialisation (particularly nuclear serialisation) is also noted by Crowley (1987:77–82), Early (1994a:365) and Sperlich (1993:97) in their discussions of various languages of Vanuatu, and more generally by Lynch et al. (in press) for some Oceanic languages. Lord (1993:215ff.) gives examples (primarily from West Africa) of adverbs and auxiliaries which putatively derive diachronically from serial verbs, further supporting a close link between the two.

Note that Crowley’s (1987) concept of ‘ambient serialisation’ covers the semantic ground of manner modifiers. However, the differences between manner modifiers and verb serialisation in Tetun are such that the former cannot be analysed as serialisation.
Serial verb constructions and prepositional verbs

12.18 Tasi n-akali sa’e mai, taka n-o-la Suri Tuan nia
sea 3S-boil ascend come cover 3S-take Suri Tuan 3S
ulu-n.
head-GEN
The sea bubbled up, (and) completely covered Suri Senior’s (severed) head.
(Note that his head was still there after the sea retreated.) (K10.59)

12.19 Kabir mak ohin nia tol-an n-o-la ne’e sia, sai hotu
betel.basket REL just.now 3S swallow 3S-take this PL exit all
bá rai.
go earth
The betel nut baskets which she had earlier swallowed all (spilled) out on the
ground. (AA1.105)

Other verbs that can fill the second slot of this construction include ho’o ‘kill’ (12.20) and
hela ‘leave’ (12.21).

12.20 ...tan hori. fonin ha’u k-song k-o’o ti’an fahi inan
because last.night 1S 1S-spear 1S-kill already pig female
ida...
one
...because last night I speared to death a sow...
(J4.16)

12.21 ...mane n-a-toba n-ela ta feto á, nia la’o.
man 3S-make-lie.down 3S-leave already woman DEF 3S walk
...(and when) the man had laid the (sleeping) woman down and left her, he
walked away.
(Z5.16)

Another sequence found within this construction type consists of an initial halo ‘make,
cause’ and a following transitive verb beginning with the causative prefix ha- (11 examples).
It is unclear why a causative verb should be used as well as a causative prefix.

12.22 Ai tahan n-al-o n-a-di’ak moras.
plant leaf 3S-make 3S-make-good sick
The medicine healed the sickness.
(F0.49)

Clear evidence that the sequences of transitive verbs discussed in this section involve
serialisation, and not clause parataxis, is that the second verb cannot be negated within this
construction. Instead, if the second verb is to be negated, a biclausal construction is required
(12.23; 27 examples). In particular, the shared object (if not omitted, topicalised or right-
dislocated) must then follow the first verb and not the second (*kahi la hola oan ne’e
dissuade not take child this’ = ‘this child cannot be dissuaded’). In addition, a phonological
break is required between the two verb phrases. This phonological clause boundary need
consist only of rising intonation and a very short pause; in any case it is often not significant
enough to have been recorded by a comma in my transcriptions, and so is marked by '/' in the
example below.

12.23 Oan ne’e kahi kahi / la hola...
child this dissuade dissuade not take
Continually reproving this child doesn’t work [, because he is very naughty.]
(L0.81)
12.4 Motion-action (type 4)

12.4.1 Description

In motion-action serialisation an intransitive verb of motion is in construction with a following action verb (300 examples). This construction is very common for the motion verbs bá ‘go’ and mai ‘come’. Other verbs which can fill the initial slot in this serialisation type include la’o ‘walk, go, travel’ and tone ‘go (usually towards addressee)’.

Predicate.Motion-action ➔ (…) Vi.motion V.action (...)

12.24 ...mai fai. Fai ti’a ita bá hola we.
    come pound pound already 1PI go fetch water
    [We] come and pound (grain). Having pounded, we go and fetch water.
    (J2.13)

12.25 Nia k-tone k-á. Lale ha’u k-tone
     then 1S-go,(usu.to.you) 1S-eat else 1S 1S-go,(usu.to.you)
     ha’i k-á.
     not 1S-eat
Then (if you do as I request), (I’ll) go and eat. Otherwise I won’t go and eat.
     (Y0.8)

The unmarked interpretation for such a sequence is that the two events follow one another, with the first usually being done in order to do the second. Consultants differed as to whether this construction could be used to indicate purpose, or whether it actually stated that the action described by the second verb was carried out.14 A consultant who favoured the latter interpretation pointed out that if the purpose was not achieved, the second verb should be introduced by irrealis atu (e.g. bá atu hola... ‘go to fetch...’), and that in the absence of atu (or some other explicit indication) one would assume that there was no hindrance to the purpose being carried out.15 The following example shows that such an expectation can, however, be explicitly overridden.

12.26 Ema Lakaluta Rai Lór bá foti atu hodi bá iha
    person Lakaluta Rai Lór go lift IRR bring go LOC
    sia rai-n bá, bá foti mós foti la hola.
    3P earth-GEN go go lift also lift not take
The people of Lakaluta Rai Lór went to lift (his head) in order to take (it) to their country; although (they) went to lift (it), they didn’t succeed in lifting.
    (K10.46)

In contrast to other serialisation types, only the second verb of this construction is repeated in tail-head linkage (12.24).

---

14 Such inter-speaker differences are found for serial verb constructions elsewhere too. Sebba (1987:102f.) conducted interviews on this topic for his analysis of serial verbs in the creole Sranan. He found that verbs following ‘come’ and ‘go’ were always interpreted as indicating purpose. However, consultants differed as to whether a verb following ‘take Instrument’ expressed a purpose or necessarily expressed a completed result.

15 This is of course a nice formulation of Grice’s (1975:45) cooperative principle, in particular as it relates to his maxim of quantity (‘make your contribution as informative as is required’).
12.4.2 Comparison with auxiliaries

This motion-action serial verb construction is somewhat similar to an auxiliary-verb construction, in that both have an initial closed-class element followed by the semantically more important open-class verb. Nevertheless, there are significant differences.

Semantically, the serialisation represents two sub-activities which occur in sequence; auxiliaries instead modify the clause in some way, in terms of aspect or deontic modality. Syntactically, there are three main differences. Firstly, postverbal modifiers follow the initial motion verb in serialisations (ha'i ‘not’ in 12.25), but follow the main verb in auxiliary constructions. Secondly, some auxiliaries can precede the subject, but the motion verb in a serialisation can not. Finally, the initial verb in serialisations takes subject marking for 1S subjects (12.25). Although some auxiliaries can take subject marking, such marking is uncommon except for the one /h/-initial auxiliary, progressive ho'i.

12.5 Introducing oblique arguments

12.5.1 Introduction: serial verbs, prepositional verbs and prepositions

The following sections deal with constructions in which one verb in a series introduces an argument role which has oblique status for the other. That is, the verb introduces a role such as source or goal of motion, instrument or co-actor. Such verbs constitute a restricted class, since there are only a limited number of distinct roles.

As pointed out by many scholars (Durie 1988; Foley & Van Valin 1984:207f.; Li & Thompson 1974; Lord 1993) there is a crosslinguistic tendency for verbs of this type to develop into prepositions. At an intermediate stage of diachronic development the erstwhile serial verb will be neither fully a verb nor fully a preposition. Such in-between forms have gone by a variety of names, including ‘prepositional verb’ in Oceanic studies (Pawley 1973), ‘co-verb’ in studies of Chinese (Li & Thompson 1974), and ‘verbal preposition’ (Durie 1988).

When diachronic change occurs in stages, and earlier uses are retained while newer ones are added, the result is what Claudi and Heine have termed a ‘grammaticalisation chain’ (Heine 1992). Since diachronic change is in the direction of increasing grammaticisation, the process of change results synchronically in a chain of related uses, each more grammaticised than the one before. In such a chain, each member shares characteristics with adjacent members, although the members at the two ends of the chain may have little in common. In this they are like a linear version of Wittgenstein’s ‘family of resemblances’, with the members of the chain not all sharing a ‘prototypical’ meaning. A morpheme may develop in more than one direction at once, resulting in more than one chain.

Tetun has one verb for which there is a reasonably long grammaticalisation chain, namely bá ‘go, to, at (location), for, at (time)’, which runs the gamut from verb to prepositional verb to preposition. Its various uses are summarised in §12.5.3.8. Shorter chains are found for hó ‘accompany’ (§12.5.4), hodî ‘use, …’ (§12.5.5) and to’o ‘reach, until’ (§8.4.2).

Although constructions involving prepositional verbs cannot strictly speaking be called ‘serial verb constructions’ since they include only one verb, prepositional verbs, nevertheless, have so much in common with serialised verbs that they are discussed in this chapter. Fully prepositional uses of bá are discussed here too, rather than in Chapter 8 (on prepositions), to facilitate a clearer overview of the range of meanings of bá.

To determine whether a form is a transitive verb used in serialisation, a preposition or a prepositional verb, the following differences between transitive verbs and prepositions are relevant. Firstly, the complement of a preposition cannot be omitted or fronted, while the
object of a transitive verb can, under the appropriate discourse conditions, do either. Related to this, when a clause is relativised on the complement of a preposition, a resumptive pronoun is retained following the preposition; no such resumptive pronoun is used when relativising on the object of a transitive verb. Verbs have subjects, which are either specified or understood. In contrast, although prepositional phrases can have subjects (when used predicatively) they need not. Finally, /h/-initial serial verbs take subject marking while prepositions (e.g. hori ‘up/down from’) do not. Nevertheless, the subject marking test has pitfalls, since some non-verbs which are presumed to derive diachronically from verbs (including the preposition hosí ‘about’ ($13.3.4.2)) do take subject marking. In addition to the above general differences, there is a verb/preposition distinction which applies specifically to bí ‘go, to, at (location), for, at (time)’. This is that when bí is fully verbal, it contrasts with its deictic opposite mai ‘come’, but when it is prepositional, no such contrast is operational (since mai is not a preposition). These differences between transitive verbs and prepositions are summarised in the table below.

Table 12.3: Contrast between transitive verbs and prepositions

<table>
<thead>
<tr>
<th>Transitive verb</th>
<th>Preposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complement can be omitted</td>
<td>Complement is required</td>
</tr>
<tr>
<td>Complement can be fronted</td>
<td>Complement must follow preposition</td>
</tr>
<tr>
<td>Relativise object without resumptive pronoun</td>
<td>Relative clause requires resumptive pronoun</td>
</tr>
<tr>
<td>Has a (stated or implied) subject</td>
<td>Need not have an implied subject</td>
</tr>
<tr>
<td>Serialised verbs must be in iconic order</td>
<td>Some PP can either precede or follow the verb</td>
</tr>
<tr>
<td>Takes subject marking if /h/-initial</td>
<td>Does not take subject marking</td>
</tr>
<tr>
<td>bí ‘go’ contrasts with mai ‘come’</td>
<td>bí is preposition; mai is not</td>
</tr>
</tbody>
</table>

Note that anything which has a mixture of verbal and prepositional properties is classed as a prepositional verb, with the result that prepositional verbs have quite diverse combinations of properties.

The following sections deal firstly with serialisations in which the second verb introduces the source or goal for a preceding motion or transfer verb, and then with various ways in which the direction verb bí ‘go’ has been grammaticised. This is followed by a discussion of constructions in which hö ‘accompany’ introduces co-actors, and finally an overview of the uses of hodi ‘use’, which primarily introduces instrument NPs.

12.5.2 Motion-transitive direction (type 5)

12.5.2.1 Introduction

In motion-transitive direction serialisation, a motion or transfer verb is followed by a transitive direction verb, whose object specifies the source or goal of the motion. It is possible to have two direction phrases in a row, in which case they must follow iconic order, with the first introducing source, and the second goal (12.38). Often the serialisation is followed by a deictic particle indicating the direction of motion towards (mai) or not towards (bí) the speaker (12.30; §11.8). Modifiers such as the negator ha‘i typically follow the first verb (12.5, 12.6), although it is possible for adverbs to follow the direction verb instead (fali in 12.31).
There are two subtypes of this construction. In one, the subject of the direction verb is interpreted as coreferential with the subject of a preceding verb of motion, while in the other, it is interpreted as identical with the object of a preceding transfer verb.

### 12.5.2.2 Same-subject motion-direction

In same-subject motion-direction serialisation, it is the subject referent that moves in the direction indicated by the direction phrase. Note that the initial motion verb can itself be constructed via nuclear verb serialisation (monu tun ‘fall descend’ in 12.28, tun uluk ‘descend go.first’ in 12.29).

**Predicate: Motion-direction** → (... V.motion (...) [Vt.direction (...) (O)]* (Deictic)

12.27

...nia mós n-alai n-atutuk fahi maten.
3S also 3S-run 3S-go.direct pig corpse
[Having heard the news,] he ran straight to the dead pig.  (J4.19)

12.28

Nia monu n-osi ai leten, monu tun to’o rai,
3S fall 3S-from plant top fall descend reach earth
nia n-aktesi.
3S 3S-break
He fell from a tree top; (when he) fell down to the ground, he broke (bones). (K3.1)

12.29

Belu, ó tun uluk bá wé bá.
friend 2S descend go.first go water go/IMP
Friend, you go down into the water first (i.e. before I do).  (F5.10)

12.30

Oras é sé at tama uluk n-o’i rai bón bá?
time this who IRR enter go.first 3S-go.to earth burrow go
Now who will go into the burrow first?  (J1.20)

12.31

Mais mai, sia la’o n-ola fali foho.
but come 3P walk 3S-via in.turn mountain
But coming (home), they walked this time via the mountains. (K10.119)

12.32

Sira fasi, at fila r-ika mai uma, mane
3P wash IRR return 3P-back come house man
n-o’i.
3S-do.not.want
They washed, and were about to return home, (when) the man refused (to go).  (Z5.8)

Transitive direction verbs commonly used in this same-subject construction are listed below. Although all can occur as sole verb, most are used more frequently in serial verb constructions than as sole verbs. An exception is hola; however, its most common use apart from this direction serialisation is with the meaning ‘take, fetch’.
Table 12.4: Direction verbs in core layer serialisation

<table>
<thead>
<tr>
<th>Verb</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>hosi</em></td>
<td>originate from, go via</td>
</tr>
<tr>
<td><em>bá</em> (<em>iha</em>)</td>
<td>go, to (usually away from speaker)</td>
</tr>
<tr>
<td><em>mai</em> (<em>iha</em>)</td>
<td>come, to (in direction of speaker)</td>
</tr>
<tr>
<td>ho'í</td>
<td>go to (reaching destination)</td>
</tr>
<tr>
<td><em>hatutuk</em></td>
<td>go directly to (without detour or delay)</td>
</tr>
<tr>
<td><em>hola</em></td>
<td>via, go past</td>
</tr>
<tr>
<td><em>to’o</em></td>
<td>reach, arrive at</td>
</tr>
</tbody>
</table>

The object of some of the direction verbs is omissible if it is contextually understood (e.g. *hosi* ‘from, via’ (12.33) and *to’o* ‘reach, arrive (at)’). When there is no following object for *bá* ‘go’ or *mai* ‘come’, these two words are instead interpreted as deictic particles (§11.8). Note that the prepositions *hori* ‘up/down from’ and locative *iha* (12.34) can introduce direction NPs in the same way as directional verbs can.

12.33  *Oras* ida foin *mai* *hikar* *hosi*.  
Soon (she’ll) come back past (here). (i.e. She’ll visit us.)  
(Q0.121)

12.34  *Mais* ha’u *la* k-*bá* *iha* Suwai *bá*.  
But 1S not 1S-go LOC Suai go  
But I am not going over to Suai.  
(F3.21)

The locative complements of *bá* ‘go’ and *mai* ‘come’ can be either NPs or prepositional phrases introduced by the locative preposition *iha*. Whereas *bá* ‘go’ often has an NP complement, for *mai* ‘come’ the prepositional phrase option is much more common than the NP one. These NP and prepositional phrase options are available both when *bá* or *mai* is the sole verb in the clause (12.34), and when they introduce the direction phrase in a serialisation (*hodi* *bá* *iha* in 12.26, 12.38).

12.5.2.3 Switch-subject motion-direction

In switch-subject motion-direction serialisation, it is the object referent of the first verb that moves. Initial verbs in this construction thus include transfer verbs such as *tau* ‘put’, *hasai* ‘remove’ and *hasa*’e ‘raise’, in which the subject referent causes the object referent to move.

**Predicate.**

\[
\text{Put-direction} \quad \rightarrow \quad (...) \mathbf{V}_{\text{transfer}} (...) (O) \quad [\mathbf{Vt}_{\text{direction}} (...) (O)]^* \quad (\text{Deictic})
\]

The choice of direction verb appears to be much more limited in these switch-subject constructions than in the same-subject serialisations discussed above.\(^\text{16}\) In fact the only goal phrases attested in the corpus are introduced by *bá* ‘go’ (12.35), *bá* *iha* ‘go LOC’ (12.38), or directly by the locative preposition *iha*. In addition, elicitation showed that *mai* *iha* ‘come LOC’ can be used. Source NPs can be introduced by *hosi* ‘from’ (12.38).

\(^{16}\) It is possible that this is an accident of the data.
12.35 ...*solok* surat *bá* Australia.
send letter [Mly] go Australia
...send a letter to Australia.  

That *bá* is a serial verb in this construction (and not a preposition like *iha*) is shown by the fact that its object is readily omissible (although when it is omitted I synchronically analyse *bá* as a deictic particle instead (§11.8)), and that the object can be fronted (12.36; 3 examples).

12.36 *Ema bulan tau dadur *bá*, bat nia keta la’o.
person idiot put handcuff go so that 3S do not walk
A crazy person has handcuffs put on (him), so that he won’t walk (i.e. wander around). (This can be done by handcuffing an ankle to a wrist.)  

The possibility that direction verbs in this construction are not 100% verbal is raised by the subject marking found on *hosi* ‘from’. Although all examples for *hosi* as a main verb and in same-subject motion-direction serialisation show regular subject marking (or absence of subject marking, as happens periodically for other verbs also), there is some confusion concerning its subject marking in switch-subject serialisation. If one ignores examples for which the subject and object of the initial transitive verb both trigger the same person-number agreement on verbs, or in which one of them triggers zero agreement (e.g. *ami* ‘1PE’), the corpus contains four examples in which *hosi* agrees with the subject of the motion verb (12.37), three in which it agrees with the object (12.38), and eight in which it remains uninflected. A similar ambivalence about subject marking is found for *hosi* when it is a preposition/conjunction meaning ‘about’ (§13.3.4.2).

12.37 *Mai m-a-sai ha’u m-osi rai ne’e...*
come 2S-make-exit 1S 2S-from earth this
Come and get me out of the ground [, then I will reward you.]  

12.38 *Ha’u k-a-sai kabau baka sia n-osi lalu’an*
1S 1S-make-exit buffalo k.o.cattle[Port] PL 3-from animal.pen
*bá iha hae molik bá.*
go LOC grass bare go
I take the cattle out from the pen into the grass.  

12.5.3 Other functions of *bá*

12.5.3.1 Introduction

In the previous section it was shown that *bá* ‘go’ can function in serialisation with a preceding verb to introduce the goal for a verb of motion. The following sections deal with various other functions of *bá*, in which it introduces other semantic roles. In these functions *bá* ranges from prepositional verb to preposition.
12.5.3.2 bá ‘to (addressee)—prepositional verb

The NP referring to the addressee of a speech act is often introduced by bá (90 examples). It always follows the direct object, if any.

In this construction bá is semantically bleached of its verbal meaning ‘movement away from (or at least not towards) speaker’. This is shown by the fact that it cannot be replaced by its deictic opposite mai ‘come’ (*Katak mai ha’u ‘tell come 1S’), being used even if it is the speaker who is presented as addressee.

12.39 Mais Ibu katak bá ha’u, n-a’ak “Tone, ó uluk
but mother tell to 1S 3S-say go.(usu.to.you) 2S go.first
bá”.
go
But Ibu tells me, “Go; you go ahead”. (L0.117)

Addressee bá cannot be interpreted as a serial verb. Most verbs of speaking seldom take an object NP. As a result a switch-subject serial verb interpretation is often not possible, since there is no NP (stated or implied) to function as ‘subject’ of bá. On the other hand, a same-subject serialisation analysis is excluded on semantic grounds. (The only serialisation analysis possible would thus be one of ‘ambient’ serialisation, for which there is no evidence.)

According to consultants, the addressee NP can be fronted (leaving bá in clause-final position), and can also be relativised without a resumptive pronoun. If these judgements are correct, they show that addressee bá still has some verbal characteristics even though it cannot be analysed as a serial verb.

12.5.3.3 bá ‘to (recipient)—prepositional verb

The NP referring to the recipient of an act of giving is usually introduced by bá (over 60 examples; see fn. 18 for alternatives). The semantic bleaching evident in addressee bá occurs here also, in that bá is also used where the speaker is himself the recipient (12.40). Nevertheless, one consultant accepted that mai ‘come’ could introduce first person recipients, just as, according to Hull (1996b:28), it normally does in Dili Tetun.

12.40 Ó m-ola ktodan bá ó, fó kmán bá ha’u.
2S 2S-take heavy to 2S give light.(weight) to 1S
You take my heaviness (problems) upon you, and give lightness (ease) to me.
(F4.10)

Although the recipient phrase normally follows any direct object NPs (12.40), it is possible for long objects (but not short ones) to follow the recipient phrase (12.41; 2 examples). This is in accordance with a crosslinguistically common tendency to shift ‘heavy’ constituents to the right (Hawkins 1983:91).

---

18 Addressee and recipient can alternatively be expressed as direct object NPs for some verbs (e.g. katak ha’u ‘tell me’, fó ema ‘give (to) people’). In addition, fó tone ‘give go.(usually.towards.addressee)’ can be used to specify a second-person addressee or recipient (e.g. katak fó tone ‘tell you’, fó tone ‘give to you’). The deictic particles mai ‘come’ and bá ‘go’ can be used instead of (but not in addition to) addressee or recipient phrases to specify speaking/giving towards or not towards the speaker (e.g. katak mai ‘tell me’, fó mai ‘give (it) to me’).
12.41 Nia tate n-ika bá mane tais ida, fahi ida,
3S pay.fine 3S-back to man cloth one pig one

rihun lima-nulu.
thousand five-tens
(If a girl breaks her engagement,) she pays to the man one cloth, one pig, and
50,000 (rupiahs). (F1.6)

A switch-subject serial verb analysis is on semantic grounds possible for recipient bá, since
the object NP of the verb of giving is interpretable as subject of bá. Such an analysis would
require that this construction be one of very few which allows a subject to follow the verb (in
this case on condition that the subject be long; the other constructions allowing postverbal
subjects being nuclear causative constructions and existential clauses). Nevertheless, the
recipient NP cannot be fronted in a topicalisation construction, which shows that bá is not
fully verbal in this construction. It is therefore analysed as a prepositional verb.

12.5.3.4 bá ‘at (location)’—prepositional verb

The word bá can introduce a locative adjunct for a preceding verb (60 examples). While
the preceding verb often specifies posture (e.g. toba ‘lie down’, harfik ‘stand’; 12.42, 12.44),
this need not be the case (12.43).

12.42 Uluk tur bá uma mesa á-’át.
former.times sit at house solely RDP-bad
In former times (people) lived in houses that were all of poor quality.
(N0.71)

12.43 Nia monu, kidan n-aktesi. Nia n-aktesi bá ne’e.
3S fall backside.bones 3S-break 3S 3S-break at this
(When) he fell, the bone(s) of his backside broke. He broke his bones here (lit.
‘at this’). (This was said pointing.) (I0.75)

Locative bá has the syntactic features of a verb. In particular, its complement NP is
omissible (12.44; 5 examples) and can allegedly be relativised without use of a resumptive
pronoun (12.45; 2 examples).

12.44 Fahi n-anua rai; foin n-oku bá.
pig 3S-root earth then 3S-lie at
The pig roots up the dirt using its snout, then lies in (it). (Q0.93)

12.45 fatik mak sia r-arís bá
place REL 3P 3P-bathe at
a place where they bathe
(X0.80 elicited)

Nevertheless, it has some non-verbal features also, including the bleaching of a
‘movement’ meaning from bá, and the fact that it introduces a peripheral NP. In addition, a
same-subject serial verb analysis is semantically possible for posture verbs, in which the

19 Locative bá is semantically similar to the locative preposition iha (§8.3.1), and to the verb hosí when it
means ‘at’ (e.g. tur n-osí labsí ‘sit 3S-at verandah’ = ‘sit on the verandah’; 13 examples). However, unlike
bá, hosí can also occur as sole verb in this sense (Ami hosí ne’e ‘1PE at this’ = ‘We are here’), and takes
regular subject marking. There are no data concerning whether the complement of locative hosí can be
omitted or fronted.
subject referent of the posture verb is also the one whose location is being given (12.42); however, such an analysis is not possible in example 12.43, where the location is of the break, and not of the subject referent.

Locative ba thus has the mixed characteristics of a prepositional verb.

12.5.3.5 ba ‘to (factive result)—preposition?

In addition to its other functions, ba can introduce a result for the verb halo ‘make, cause’. This verb allows for two arguments other than the subject, namely the undergoer of the causation (the direct object ‘causee’), and the result of the causation.

It is uncommon for both causee and result to be specified within the one clause. However, when they are both specified the causee is direct object and the result follows it, being introduced by ba (7 examples). The causee NP can be fronted just like any other object NP.

12.46 Ibu n-alo ema ida ba ai.kalete.

mother 3S-make person one to intermediary

Ibu makes someone an intermediary. (L0.130)

A much more common pattern for expressing causation with an NP result is for the causee to be understood from context. Usually there is a preceding clause which indicates what the causor does to the causee. The verb halo then simply introduces the result, either expressed as an NP (12.47; 14 examples) or introduced by ba (12.48; 21 examples).

12.47 Batar fulin sanulu hatali libur halo liman ida.

maize head(.grain) ten tie assemble make arm one

(When) ten cobs of maize are tied together (we) make one ‘hand’. (H0.39)

12.48 Aka ne’e r-otus n-alo ba klinun hât...

sago this 3P-cut.trunk.into.lengths 3P-make to trunk.section four

This sago palm (they) cut into lengths, making four lengths...

(K4.5)

Where the result is a bare NP, halo is best interpreted as (non-causative) transitive ‘make’. This possibility is reflected in my translation of example 12.47.

Data are insufficient to determine whether result ba is a preposition or a prepositional verb.

12.5.3.6 ba ‘for (beneficiary, purpose)—preposition

Where ba introduces a beneficiary (12.49; 20 examples) or purpose (12.50; 52 examples), it has none of the verbal characteristics of ba ‘go’, and is totally grammaticised into a preposition. The beneficiary or purpose phrase follows the postverbal object NP, if any. In keeping with its analysis as preposition, the complement of ba is not omissible and cannot be fronted, and ba cannot be replaced by mai ‘come’ when it is the speaker who is the beneficiary (12.49). Most of the purpose examples involve the expression ba sâ ‘for what’ = ‘why’.

---

20 Verbal complements of halo are discussed in §13.3.3.

21 Both beneficiary and purpose NPs can alternatively be introduced by the preposition bodik ‘for’ (§8.6.1). Other alternative expressions for beneficiary include fô ba NP ‘give go NP’ = ‘for NP’, and fô tone ‘give go.(usually.towards.addressee)’ = ‘for you’ (e.g. loke fô tone ‘open for you’).
12.49 Ina! M-anana oda.matat bá ha’u lai.
mother 2S-open door for 1S first
Mother! Open the door for me now. (Z4.63)

12.50 Ô mai bá sá, Bolan Rano’is?
2S come for what Bolan Rano’is
What have you come for, Bolan Rano’is? (G2.14)

12.5.3.7 bá ‘at (time)’—preposition

When bá introduces peripheral time NPs (60 examples), it is fully prepositional, as shown by the lack of any (logical or implied) subject, and by the ability to place the prepositional phrase either initially (12.51) or finally (12.52) within the clause.

12.51 Bá loron ida, sia na’in rua bá to’os, atu
at.(time) day one 3P CLS:human two go garden IRR
r-afaho mo’at.
3P-weed weed
One day, the two of them went to the garden, to remove weeds. (K12.8)

12.52 Nia lakon bá tuku hira?
3S disappear at.(time) hour how.many
At what time did it disappear? (H1.5)

Over half of the 60 examples of this preposition are from texts by two well-educated people, suggesting possible Indonesian influence. In part this reflects a greater tendency for such speakers to mention time at all. In fact the specification of exact time, which uses concepts incorporated via Indonesian (e.g. bá tuku nen ‘at hour six’ = ‘at six o’clock’, bá loron Selasa ‘at day Tuesday,[Malay]’ = ‘on Tuesday’), seems to be largely limited to educated speakers, such as schoolteachers. A related construction which seems largely limited to educated speakers is the specification of a simultaneous event using bá oras ‘at time’ followed by a clause specifying the event, such as bá oras ita toba ‘at time 1PI lie.down’ = ‘when we lie down’ ($\S 14.8.3.2$).

12.5.3.8 Summary of functions of bá

The table below summarises the characteristics of bá in its various transitive uses as serial verb, prepositional verb and preposition. These functions are introducing direction for a verb of motion (‘go Dir’) or a transfer verb (‘put Dir’), addressee (‘Add’), recipient (‘Rec’), location (‘Loc’), factitive result (‘Res’), beneficiary/purpose (‘Ben Pur’), and time (‘Tim’).

In the table ‘?’ marks information which has been derived by elicitation but which has not been confirmed in texts, and blanks reflect missing information. In each case a positive answer (‘Y’, emphasised by shading) indicates that bá behaves just as verbs do with respect to that character. Some of the characteristics, however, are not unique to verbs. That which specifies that the phrase must follow any object NP is true of many prepositional phrases as well as of all serial verbs, while that of introducing oblique arguments (as opposed to peripheral NPs), is shared by the locative preposition iha (e.g. for transfer verbs).
Table 12.5: Verbal characteristics of *ba* in various functions

<table>
<thead>
<tr>
<th>Verbal characteristic</th>
<th>Verb</th>
<th>Prep. verb</th>
<th>Preposition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>go</td>
<td>put</td>
<td>Add Rec</td>
</tr>
<tr>
<td>Complement of <em>ba</em> can be omitted</td>
<td>Y</td>
<td>Y</td>
<td>Y Y Y Y N N N</td>
</tr>
<tr>
<td>Complement can be fronted</td>
<td>Y</td>
<td>Y</td>
<td>Y N Y Y N N</td>
</tr>
<tr>
<td>Relativise object without resumptive pronoun</td>
<td>Y</td>
<td>Y?</td>
<td>Y Y Y Y N N</td>
</tr>
<tr>
<td>Has a 'logical' subject</td>
<td>Y</td>
<td>Y (N)</td>
<td>Y Y N N N N</td>
</tr>
<tr>
<td>Subject of <em>ba</em> = first verb's S/O</td>
<td>S</td>
<td>O</td>
<td>O - - - - -</td>
</tr>
<tr>
<td><em>ba</em> phrase must follow object NP</td>
<td>Y</td>
<td>Y Y Y Y Y Y</td>
<td></td>
</tr>
<tr>
<td>Introduces oblique argument</td>
<td>Y</td>
<td>Y Y Y Y Y Y</td>
<td></td>
</tr>
<tr>
<td>Contrasts with <em>mai (ihā)</em> NP</td>
<td>Y</td>
<td>Y N N N N N</td>
<td></td>
</tr>
<tr>
<td>Contrasts with intransitive <em>bāmai</em></td>
<td>Y</td>
<td>Y Y Y Y N N</td>
<td></td>
</tr>
</tbody>
</table>

From the table it is clear that the uses of *ba* fall into three categories. Those introducing direction (goal) are fully verbal. When *ba* introduces addressee, recipient or location it has a mixture of verbal and prepositional characteristics, and so is classed as a prepositional verb. And when it introduces the peripheral roles of benefactive, purpose and time, *ba* has no uniquely verbal characteristics at all. Result *ba* is tentatively classified as a preposition, although more information is required for a definitive classification.

12.5.4 *hō* ‘accompany’—verb, prepositional verb (type 6)

We turn now to two other words which occur both as verbs in serialisation and as prepositional verbs, namely *hō* ‘accompany’ and *hodi* ‘use’.

The very *hō* ‘accompany, be with’ frequently occurs in construction with either a previous (12.53) or a following (12.54, 12.56) verb, which specifies what the co-participants are doing.

12.53 *Ha’u bā k-ō lōs Am Bo’uk dei.*
1S go 1S-accompany just father Bo’uk only
I will go with only Am Bo’uk. (i.e. No-one else will go.) (V0.102)

12.54 *Feto na’in rua ne’e r-ō malu r-akāt.*
woman CLShuman two this 3P-accompany each other 3P-fight
These two women fought with each other. (F2.40)

The entire comitative serialisation may in turn serialise with a preceding verb of motion, forming a motion-action serialisation.

12.55 ... *ha k-bā k-ō feto sia k-adiuk...*
1S 1S-go 1S-accompany woman PL 1S-play
[Every evening,) I go and play with (i.e. court) the girls...

(U5.13)

In contrast to the situation with most verbs, the object of *hō* ‘accompany’ is very rarely fronted. This is so regardless of whether it is used in serialisation (for which the corpus contains no examples of fronting) or is the sole verb in a clause.

As is the case for instrumental *hodi* (see below), the object of *hō* is omissible if *hō* is the initial verb in the serialisation (12.56), but not if it is the final one. Instead, a final adverb *nō*
'also, too' (homophonous with ‘3S-accompany’) can be used to indicate that an activity is done with unspecified co-actors (e.g. Ha’u k-bá nó ‘1S 1S-go too’ = ‘I am going too’). Series-final hó is thus a prepositional verb, taking subject marking, but being preposition-like in not allowing omission or fronting of the complement.

12.56 At n-á laho bei, k-ó k-á.
IRR 3S-eat rat also 1S-accompany 1S-eat
Even if (he) eats rats, (I) will eat with (him).  (Z2.74)

12.5.5 hodi 'bring, use' (type 6)

12.5.5.1 Overview

When it occurs as main verb, hodi usually means ‘bring, take’, and involves bringing (usually carrying) something from one location to another (12.3, 12.26). Generally the items to be brought are physical things; however, hodi is also used of bringing non-physical things such as messages. It can also be used of bringing animals (e.g. hunting dogs) and humans (e.g. slaves), where these move independently but without free will. For accompaniment of free individuals, hó ‘accompany’ is used instead.

Occasionally hodi as sole verb takes on the meaning of ‘hold, use’ without accompanying change of location. It is this meaning which is predominant when hodi is used in serialisation with other verbs, to introduce the instrument phrase. The other verb may either follow or precede hodi. There are differences between the two orders, which will therefore be considered separately. Other uses of hodi, as a serialised comitative verb and as a prepositional verb introducing time, will subsequently be dealt with briefly.

12.5.5.2 ‘use (instrument)’—verb

Where hodi ‘use’ precedes the verb with which it is in serialisation (104 examples), the series follows the iconic order in which taking hold of the instrument precedes doing something with it. This construction is found in many verb serialising languages (Foley & Olson 1985:53; Lefebvre 1991; Lord 1993:65). It is a classic serial verb construction in that it involves a typical collocation of ideas (instrument and activity), no significant pause and a shared subject.

12.57 Nia n-odi tudik e’e ko’a sít ti’a...
3S 3S-use knife this cut be.cut.off already
He took the knife and cut through (the umbilical cord)...  (P1.38)

In this initial position within a serialisation hodi is fully verbal. In particular, it readily allows omission of the object NP under conditions of anaphora (12.58; 37% of examples) as well as fronting of the object (12.59; 25% of examples). Where the object is fronted, as in 12.59, the construction gives the purpose or intended use for the item in question, rather than the actual use on a particular occasion.

---

22 Other meanings of main verb hodi are ‘hold the office of’ (e.g. hodi liurai ‘hold the office of ruler’), ‘win, defeat’, and ‘using language’ (e.g. (hakés) hodi Dawan ‘(talk) in Dawan’). It also means ‘it’s up to; that depends on’ (e.g. Hodi ô dei ‘up.to 2S only’ = ‘It’s up to you; do as you please.’).
12.58 M-atali m-ola (é) kabas hitu, bat ha k-odi
2S-tie 2S-take HES cotton.thread seven so.that 1S 1S-use
ko'us nia.
3S cradle Tie together seven strands of cotton, so that I will use (them) to cradle her.
(U1.34)

12.59 Tais ne'e hodi ha-di'a nia-kan naran.
cloth this use make-good 3S-POS name
This cloth (paid by a man to a lover he refuses to marry) is to clear her name.
(T0.131)

12.5.5.3 ‘use (instrument)’—prepositional verb

When hodi ‘use’ follows the other verb, it always has a following object (71 examples). This object is in the corpus nearly always non-specific.\(^{23}\) Not only are the objects thus presented as non-typical participants in discourse, but none of the examples report discourse events. That is, none perform the typically verbal function of answering the question ‘What happened?’ (Hopper & Thompson 1984:726). In fact a very high percentage of examples of this construction come from sentences which people gave me to teach me the meaning of new words. Example 12.60, for instance, was given to explain the term kalete beni ‘covered bridge’.

12.60 kalete beni = kalete mak tate hodi rai
bridge protect bridge REL scatter use earth
a covered bridge = a bridge which is covered with dirt (i.e. not just consisting of logs)
(N0.19)

12.61 ...dadi at n-aroun odan n-odi sá?
so IR 3S-lower ladder 3S-use what
[Her legs and arms were firmly tied,] so what could (she) lower the ladder with?
(P3.56)

12.62 Ha' u la k-atene tán ha' u la k-aré k-odi matan.
1S not 1S-know because 1S not 1S-see 1S-use eye
I don’t know as I didn’t see (it) with (my) eyes.
(D0.38)

Series-final hodi (like final hó ‘accompany’) is like a preposition in that its object is never omitted and cannot be fronted. A further characteristic not shared with serial verbs is that it can be independently negated (12.63). Nevertheless, hodi takes subject marking (12.61, 12.62), giving it the characteristically mixed properties of prepositional verbs.\(^{24}\)

---

\(^{23}\) The discourse-function contrast in the two constituent orders is reminiscent of the difference noted in Foley and Van Valin (1984:207f.) for verbal versus prepositional instrument phrases in Thai. An instrument introduced by a preposition is out of focus, while one introduced by a serial verb is highlighted and of central importance.

\(^{24}\) Instrumental prepositional verbs are common in Oceanic languages (e.g. Paamese: Crowley 1987:54f.), and one is also reported for the Central Malayo-Polynesian language Kambera (Klamer 1994:278). If series-final hodi were fully verbal, such sequences would run counter to the general principle that serial verb constructions follow iconic order.
12.63 Nia sita n-odi ha’i taha.
3S compete 3S-use not big.knife
She competed not with a knife (but with a weaving stick). (V0.54)

12.5.5.4 ‘with (comitative)’—verb

In addition to meaning ‘bring, use’, and being used to link clauses (§14.5.2), hodi can occur in serialisation with a following verb to introduce a comitative malu ‘each other’ (12 examples). This is a function for which hō ‘accompany’ is usually used.

12.64 ...r-odi malu de’an; de’an Feto Ikun.
3P-with each other speak.angrily speak.angrily woman tail
[The sisters...] together spoke angrily: spoke angrily to Youngest Sister. (R4.128)

12.5.5.5 ‘by (time)’—prepositional verb

In addition, hodi can introduce a time phrase for a preceding verb (4 examples). Note that hodi ‘by (time)’ has the verbal property of taking subject marking (12.65), but the non-serialisation property of introducing a peripheral constituent and of being able to be negated (12.66). Temporal hodi is thus analysed as a prepositional verb.

12.65 Kakaluk na’in mai n-odi kalan dei.
spirit.house lord come 3S-by.(time) night only
The lord of the spirit-house (a spirit) comes only by night. (I0.183)

12.66 Hu’an hodi ha’i loron; Hodi kalan bá oras
tuku tolu.
fell(tree) by.(time) not day by.(time) night at.(time) time
hour three
(One) cannot fell (akaria sago palms) during the day. (One must do it) at night, at three o’clock. (N0.166)

12.6 Participant cooperation: ‘together’/‘alone’ (type 7)

A final type of same-subject verb serialisation involves one of the verbs indicating whether the subject referent of the other verb acted alone or in concert with others. These verb sequences are tentatively analysed as serialisation at the core layer; however, more data are required on the position of object NPs and adverbial modifiers.

The verbs hamutu and hamutuk are semantically very similar, and when used intransitively mean ‘(be) together’. In serialisation with a preceding verb they indicate that the subject referents (or subject plus co-actors introduced by hō ‘accompany’) act together.

12.67 Ibu la la’o n-amutuk n-ó ema nia...
mother not walk 3S-together 3S-accompany person that
Ibu doesn’t walk/travel together with that person [because he is not trustworthy.]
(M0.43)

The lexemes mutu and mutuk behave similarly to the above verbs in serialisation (hā mutu ‘eat together’), but appear unable to function as sole verbs. If this observation is correct, mutu and mutuk are better analysed as adverbs.
Chapter 12

The word mesan (sometimes mesa) 'be alone' can co-occur with a preceding (12.68) or following (12.69) verb to indicate that the subject referent is alone in doing that which is specified by the predicate.

12.68  *Nía toba mesan dei.*
      3S lie.down alone only
  He sleeps alone.  \(\text{(X0.90)}\)

12.69  ...*ni mesan mai.*
      3S alone come
  [The nobleman left his dogs at home, and] he came alone.  \(\text{(O2.34)}\)
13 Complementation

13.1 Classes of complement-taking verbs

This chapter deals with constructions in which clauses or sentences function as arguments. I will for convenience follow Dixon's (1991) semantic classification of such verbs, which proves illuminating for Tetun. However, I will use the term 'complement-taking predicate' (Noonan 1985:43), instead of Dixon's more opaque 'Primary-B verb', to distinguish this verb from the verb in the complement.

Complement-taking predicates include mainly verbs of speaking, thinking, liking and attention. Amongst verbs of speaking, those which signify ordering the addressee to do something have different syntactic possibilities to other verbs of speaking, and so are presented as a separate subclass. Some common verbs in each of these categories are listed below.

**Speaking:**
- bolu: call
- haruka: order
- katak: tell, inform

**Ordering**
- call
- order
- tell, inform

**Speaking:**
- hakés: talk, converse
- ha’ak: say
- halón: plead
- hameno: arrange that speaker/addressee will do something
- meno: arrange that speaker/addressee will do something
- haheno: arrange that a third person will do something
- simu: reply, receive

**Thinking (including knowing)**
- fiar: believe
- ha’ak.fali: believe incorrectly (lit. 'say in.turn')
- halu’a: forget
- haneo: ponder, consider
- hanoin: think, ponder, remember
- hatene: know, understand
Chapter 13

Liking/Wanting

bér desire, like
hakara like, want
hoʻi do not want
hahoʻuk agree to, approve of
hoʻuk agree to, approve of
hataʻuk fear

Attention

haré see
hobun watch
hora smell
horan sense, feel
rona hear

Complement-taking verbs in other categories include the following: the verb of making halo ‘make’, used in analytic causatives; the phasal verbs hahú ‘start’ and hanawa ‘finish’; and the verb of happening dadi ‘happen’. Another verb of happening, kona ‘undergo, touch’, takes very restricted complements and is discussed in §9.4.6.2

13.2 Overview of complement types

There are two types of complement clause in Tetun, namely ‘sentential’ (or ‘sentence-like’) and ‘reduced’ (Noonan 1985:49, 73). The former has the full range of capabilities of a clause, including independent mood (e.g. interrogative, declarative), negation and subject. The time reference may be independent of that of the complement-taking predicate (e.g. ‘say’), or dependent on it (e.g. ‘see’, ‘hear’), depending on semantics.

Reduced complements, in contrast, cannot have an independently specified subject, cannot independently be interrogative, and have a time reference which is restricted (usually future or co-temporal) relative to that of the complement-taking predicate (e.g. ‘want to’). They can, however, be independently negated. Reduced clauses in turn have two subtypes, namely normal reduced clauses, which are found in clause-final position, and ‘small clauses’, which are placed before the complement-taking predicate.

1 There is overlap between the verbs of liking (which express a feeling about something) and verbs of wanting (which describe an attitude to an as yet unrealised state or event). It seems that hataʻuk comes under the category of verbs of liking only, while the remaining verbs listed fall into both categories.

2 I have little or no data on complements for Dixon’s remaining categories, such as deciding, annoying, acting, relating and seeming. Some categories which in some languages are handled by complementation make use of different strategies in Tetun. Comparison uses liu ‘go past, further’ in a non-complementation construction (§11.3.2). The word ‘try’ is expressed by the postverbal modifier kokon (e.g. hā kokon ‘eat try’ = ‘try eating some, have a taste’). Modality is expressed by preverbal auxiliaries (§10.4). The phrase ‘not matter’ is expressible by la sā ida ‘not what one’ = ‘it’s nothing’ following a concessive clause (e.g. Biar haʻu mate, la sā ida ‘Even if [Mly] I die, it doesn’t matter’).

3 Complements for nouns (e.g. kakotun ‘decision’) do not occur in the corpus, and attempts to elicit them proved unsuccessful.

4 There is no evidence for nominalisation as a complementation strategy in Tetun.

5 The primary split largely fits Noonan’s (1985:133) generalisation that complement systems with two members tend to make their primary break at the distinction between dependent and independent time reference.
The following table summarises the types of complements that are potentially found for members of each of the larger semantic classes of complement-taking verbs. In the table ‘Comp’ stands for ‘complement’, ‘Comp.S’ for subject-control reduced complements, and ‘Comp.O’ for object-control complements. Quotes, particularly direct quotes, are frequently introduced by the verb *ha’ak* ‘say’; although such constructions are analysed as coordinate-dependent constructions rather than complementation, they are listed in the table for purposes of comparison. Also for purposes of comparison, the bottom line of the table shows whether verbs of this semantic class typically allow an object (or indirect object) NP in the same postverbal slot as the complement clause.

| Table 13.1: Complement types found for each class of complement-taking predicate |
|-----------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                  | Speak:          | Speak:          | Think           | Like            | Atten.          | Make            | Phasal;        | Happen          |
|                                  | Order           | Other           |                  |                  |                  |                  |                  |                  |
| **Sentential complement**        |                 |                 |                 |                 |                 |                 |                 |                 |
| $S\, V\, Comp$                   | §13.3.2,        | Y               | Y               | Y               | (Y)             | Y               | Y               | Y               |
|                                  | §13.3.3         |                 |                 |                 |                  |                 |                 |                 |
| $S\, V\, hosí-Comp$              | §13.3.4.2       | Y               | Y               |                 |                 |                 |                 |                 |
| $S\, V\, batu-Comp$              | §13.3.4.3       | Y               | Y               | Y               |                 |                 |                 |                 |
| Question S V                     | §13.3.5.2       |                 |                 |                 |                  |                 |                 |                 |
| **Reduced complement**           |                 |                 |                 |                 |                 |                 |                 |                 |
| $S\, V\, O\, Comp,O$             | §13.4.2         | Y               |                 |                 |                 |                 |                 |                 |
| $S\, V\, Comp,S$                 | §13.4.3.2       | Y               | Y               | Y               | Y               |                 |                 |                 |
| $S\, Comp,S\, V$                 | §13.4.3.3       | Y               |                 |                 |                 |                 |                 |                 |
| *ha’ak* + Quote                  | §14.5.3         | Y               | Y               |                 |                 |                 |                 |                 |
| NP complement                    |                 |                 |                 |                 |                 |                 |                 |                 |

In the examples in this chapter, complements are enclosed in quotation marks where they are clearly direct quotes, and in square brackets otherwise. The complement-taking predicate is underlined. Sentential complements will be discussed first, followed by reduced complements.

### 13.3 Sentential complements

#### 13.3.1 Introduction

A sentential complement has the potential to consist of a full clause, or, in some cases, a full sentence.

In this section we consider a range of clauses which are semantically interpretable as complements, although in some cases alternative syntactic analyses are considered. Such clauses are found either finally or initially within the matrix clause. If they are final, they can be introduced by various complementisers. The following subsections deal first with complements (or apparent complements) functioning as objects, namely clause-final complements without and with a complementiser, and initial clauses. This is followed by a brief discussion of complements functioning as subject in the matrix clause.
Note that quotes, particularly direct quotes, for verbs of speaking are very often introduced by *ha‘ak* ‘say’; discussion of such constructions is postponed to the next chapter (§14.5.3).

### 13.3.2 Clause-final complements: non-causatives

Clause-final sentential complements without complementisers are found for verbs of speaking, thinking (13.1, 13.2; 30 examples), liking (13.3; 6 examples) and attention (13.4, 13.5; 45 examples). Complements for verbs of making are somewhat different, and are discussed separately below (§13.3.3). For verbs of speaking a sentential complement usually represents a direct quote (13.9, discussed below; over 1,000 examples), although it can also be indirect (13.6, 13.7; about 100 examples).

13.1 *Mais ha‘u laka* *katene* [sé at n-ola].
but 1S not 1S-know who IRR 3S-take
But I don’t know who would take (it). (L0.141)

13.2 ... *ha‘u katene* [ó m-ó buat di‘ak ida].
1S 1S-know 2S 2S-have thing good one
...I know you have a good thing. (O2.28)

13.3 *Sia r-aquo‘uk* [feto nó mane bele kawen].
3P 3P-agree.to woman and man can marry
They agree to the girl and boy marrying. (I0.85 elicited)

13.4 *Bei Kidu Kmeik, nia n-aré* [fahi ida mai]...
Mr bottom pointed 3S 3S-see pig one come
Mr Pointed Bottom, he saw a pig coming...

13.5 *Ita rona* [ema hakés] dei, *ita ladún kohi*.
1PI hear person talk only 1PI not.very catch
If we just hear people talk (about sacred things, instead of dreaming them like
da *mako‘an* does), we don’t catch (i.e. comprehend) much. (Q0.141)

13.6 *Ket m-a‘ak* [ó túr iha rat maran].
do.not 2S-say 2S sit LOC earth dry
Don’t say you live in dry land. (C2.68)

13.7 *Pák Tóm meno* [ha‘u dakar Ibu].
Mr[Mly] Tom arrange 1S look.after mother
Mr Tom arranged (that) I look after you (‘Ibu’). (C1/8)

The complement is the final element in the matrix clause. It thus follows addressee NPs and prepositional phrases, as well as usually following adverbs (13.5 being an exception). In this it is unlike NP objects, which precede prepositional phrases.

---

5 For verbs of liking there are no examples of sentential complements from natural discourse, with five of the examples coming from direct elicitation and the remaining one from poetry.
Complements have considerable syntactic freedom. They can have independent negation (13.1, 13.8) and aspect with respect to the matrix clause, and can be interrogative (13.1).

13.8  \[\text{N... n-atene [nia-kan lia fuan ne'e la kona].}\]
\[\text{N... 3S-know 3S-POS word fruit this not touch}\]
\[\text{N... knows these words of his are not true.}\]

The corpus contains no naturally occurring examples of complements having pragmatic topics (such as left-dislocation) or topicalised objects, but these were accepted in elicitation.

Indirect quotes, used with verbs of speaking, tend to be preceded by comma intonation (13.7). Such a pause is not common before complements of thinking, liking or attention verbs.

As Noonan (1985:130) points out, attention verbs, when these are used as verbs of immediate perception (e.g. ‘see X’ rather than ‘see that X’), take complements with determined time reference, and so have limited aspectual possibilities. This follows naturally from the fact that one can only perceive what is happening at that particular time.

Direct quotes (13.9) occur in the same position in the sentence as the complements discussed above, but, unlike them, are not integrated into the matrix clause in terms of deixis. Syntactically, too, direct quotes are free. For instance, they can be of any length, from a single word (e.g. a vocative) up to multiple sentences, and can be any speech act. They freely have initial exclamatives, vocatives and topics, none of which are found in indirect quotes. Intonationally, the matrix clause may carry clause-final intonation or be followed by pause, and the quote may start on a new, raised, pitch level.\(^6\)

13.9  \[\text{Nia'má kataba Feto Ikun: “Ei! Feto Ikun. Ó}\]
\[\text{then tell to woman tail EXCL woman tail 2S}\]
\[\text{bi-n sia r-a-sasar ha’u ti’an;}\]
\[\text{elder.sister-GEN PL 3P-make-be.in.difficulty 1S already}\]
\[\text{ne’e ha’u k-la’o-n’}.\]
\[\text{so 1S 1S-walk-IMM}\]
Then (he) said to Youngest Sister “Hey! Youngest Sister. Your older sisters have put me in difficulty; so now I am leaving”.\(^{13.48}\)

13.3.3 Clause-final complements: analytic causatives

13.3.3.1 Syntax

Analytic causatives make use of the verb halo (sometimes ha) ‘make’.\(^7\) Syntactically there are two subclasses of analytic causatives, namely those in which the complement has a verbal predicate (e.g. ‘make him go’), and those with a nominal result (e.g. ‘make her his wife’). Nominal results are introduced by ba ‘to’; this results in a monoclausal construction which is discussed in §12.5.3.5.

\(^6\) It is for such reasons that direct quotes are widely regarded as not being syntactically part of the matrix sentence (de Vries 1990:294), although some linguists dispute this view (Haiman & Thompson 1984).

\(^7\) As an intransitive verb halo means ‘act, behave’, while as a transitive verb it primarily means ‘make, prepare’.

The verb hamamar ‘let’ (lit. ‘make-soft’) was also given for this construction during elicitation, but does not occur in this construction in the texts.
Clause.Causative $\rightarrow$ S $\{\text{Complement.causative} \} \hspace{1cm} \{\text{(NP.causee) ba NP.result} \}$

Complement.causative $\rightarrow$ (NP.causee) V...

Verbal complements are usually intransitive (13.10; 184 examples), but can be transitive (13.11; 6 examples).

13.10 Anin dadurus $n$-alo [hare kabelak hotu].
wind hurricane 3S-make rice flat all
The hurricane made the rice all flattened. (Q0.47)

13.11 Hakim n-akotu, $n$-alo [nia (s) tate
judge[Mly] 3S-conclude.case 3S-make 3S HES pay.fine
juta ida...].
million[Mly] one
The judge concluded the court case, (and) made him pay one million [316,000 rupiahs, a horse, a woven cloth and a pig]. (AA2.43)

The biclausal nature of this construction is shown by the fact that the two verbs can be independently negated, with negation being either of $halo$ (13.12; 1 example) or of the complement (13.13; 8 examples). In addition, the second verb can be preceded by the irrealis auxiliary atu, something which cannot be done in monoclausal serial verb constructions. Finally, it is the causee NP, and not the subject of $halo$ 'make', that controls subject marking (13.13), reflexivisation, and the interpretation of reciprocal $malu$ in the complement, as well as the interpretation of $hotu$ 'all' in intransitive complements (13.10).

13.12 Aman sia la r-alo [oan sia karian]. Habusik lós dei.
father PL not 3P-make child PL work leave.free just only
The fathers don't make the children work. (They) just leave them alone. (V0.130)

13.13 ...Ó bilan bá bilan mai, $m$-alo [ha'u la k-arej.
2S change go change come 2S-make 1S not 1S-see
[Where is my book?] You shift (it) to and fro, making me not see (i.e. find it).
(V0.71)

On the other hand this construction is like a single clause in that it is possible for modifiers of the construction as a whole to follow the complement (few examples). For instance, clause-final $oli$ is a modifier of unknown meaning which only co-occurs with the prohibitive auxiliary $keta$ (13.14). Similarly, the two instances of $ti'$an 'already' in 13.15 modify the whole preceding propositions and not just the adjectival complements. Example 13.15 also illustrates the fact that the causee NP can be fronted in a topicalisation construction.

13.14 Keta $m$-alo [buat e'e n-akfore] $oli$.
do not 2S-make thing this 3S-come.undone ?
Don't make this thing (a pile of bundled cane) come untied. (X0.6)

13.15 Ha ulu-n ne'e, ha'u k-sisi $halo$ [di'ak] $ti'$an,
1S head-GEN this 1S 1S-comb make good already
My hair, I had combed (it) making (it) tidy,
Complementation

13.3.3.2 Comparison with morphological and nuclear serial verb causatives

Apart from analytic causatives, causation can also be expressed by morphological causatives (§4.2.2) and nuclear serial verb causatives (§12.2.2). The latter two generally involve a causer who is a conscious being acting intentionally, and a causee which is usually inanimate, or animate but not able to resist (such as a baby or a rape victim). 

In contrast, analytic causatives, while usually having animate causers, also allow inanimate ones such as wind (13.10, 13.15), words and fire. Causees in analytic causatives can readily be human (13.13; 30% of examples), and may be either willing or unwilling participants. The causation can be intended (13.11) or unintended (13.13).

13.3.4 Complementisers

13.3.4.1 Introduction

Tetun has few complementisers, and even those which do exist are not frequently used. The words hosí ‘about’, batu ‘so that’ and possibly ne’ebé ‘so that’ function as complementisers for clause-final complements.

There is no complementiser for direct quotes, which are instead very frequently introduced by the verb ha’ak ‘say’ (§14.5.3). Other complements usually occur without any introductory word (§13.3.2).

13.3.4.2 hosí ‘about’

For speaking verbs such as hakés ‘talk, speak’ or husu ‘ask, request’, and thinking verbs such as hanoin ‘think’, the topic of speech or thought may be introduced by hosí ‘about’. The topic may be expressed by either a clause (13.16, 13.17; 31 examples) or an NP (13.18; 10 examples). The complement introduced by hosí sometimes has the character of a direct quote, being preceded by a pause, started on a new pitch, and allowing left-dislocation and topicalised objects (13.19).

13.16 Oras e’e, lbu, ó m-usu m-osi [mane sia halo] time this mother 2S 2S-request 2S-about man PL make to’os].
garden
Now, Ibu, you asked about men working the gardens. (P3.1)

13.17 Ha’u k-akés hosí [fulan mosu].
1S 1S-talk about moon appear
I am talking about the moon appearing (i.e. new moon). (F4.2)

8 An exception is a word from the noble register, hamán ‘accompany (noble)’, from the base kmán ‘light (weight)’; however, it has undergone a major semantic shift from its more literal meaning ‘lighten’.

9 These differences are in accord with crosslinguistic tendencies (Comrie 1989:172; Givón 1980:336), in that analytic causatives correlate with less direct causation, while the syntactically more tightly-bound morphological and nuclear serial verb causatives involve more direct causation.
There is disagreement amongst consultants as to what subject-marking *hosi* can and should take. Some inflect *hosi* with the same subject marker as the preceding verb of speaking or thinking (e.g. k-akés k-osi ‘1S-talk 1S-about’; 13.16), and insist that only this type of agreement is correct. Others prefer to leave *hosi* uninflected (13.17). Yet others inflect *hosi* for 3S even if the subject of the preceding verb is not 3S (e.g. k-akés n-osi ‘1S-talk 3S-about’; 13.19). For two consultants the choice between 3S marking and agreement with the subject of the preceding verb depended on what the topic was, although we could find no clear conditioning factor.10

It is possible that the 3S inflection of *hosi* (which some speakers reject out of hand) comes about by analogy with constructions in which a speaking verb takes an object, such as ai.kanoik ‘story’, following which *hosi* takes 3S inflection to agree with that object.

Elsewhere in the grammar of Tetun *hosi* is a transitive verb meaning ‘(originate) from, go via, be at’, and can govern only NP objects. The fact that *hosi* can only take a complement clause when it itself follows a complement-taking verb indicates that *hosi* is in this construction a complementiser (and not a verb in serialisation with the complement-taking verb). That *hosi* ‘about’ is a preposition/conjunction is further shown by the fact that its complement appears to be required, and unable to be fronted (in contrast to complements of most transitive verbs). Another difference between prepositions and verbs is that verbal modifiers of an NP are usually introduced by the relative clause marker *mak*, while prepositional phrase modifiers are not. In this, too, *hosi* ‘about’ behaves as a preposition (13.20). The one verbal feature of *hosi* ‘about’ is the fact that it takes subject marking, albeit with uncertain rules.11 The ability to take subject marking is shared by a small group of other words which are synchronically not verbs (§9.3.3).

---

10 One consultant hypothesised that 3S marking was used for topics remote from the speaker (e.g. Radio Australia), while agreement with the subject was used for topics close to the speaker (e.g. me singing, my father working).

11 Note that there is similar variation in subject marking for *hosi* when it introduces the source location for transfer verbs (e.g. ‘remove X from Y’; §12.5.2.3).
13.3.4.3 Purpose conjunction

The purpose conjunction *batu* (or *bate, bat*) 'so that' (§14.7.6) can introduce clauses identifying the goal of verbs of speaking which signify either ordering (e.g. *haruka* 'order', *bola* 'call', *husu* 'request'; 13.21; 11 examples), or promising that something will be done (e.g. *hameno* 'arrange, promise'; 6 examples). It can also introduce the goal of verbs of liking (e.g. *hakara* 'like, want'; 13.22; 2 examples).

13.21 ...*n-aruka* ha'u [bate ha k-mai k-ola ó...].
3S-order 1S so.that 1S 1S-come 1S-fetch 2S
[He] ordered me to come and fetch you [to...]. (R4.65)

13.22 Na'i Maromak *n-akara* [bat ita hata'uk bá nia...].
noble God 3S-like so.that 1PI fear go 3S
Lord God wants that we fear him, [believe in him...]. (C3.7)

It appears that the purpose conjunction *ne'ebè* 'so that' can also be used in these contexts (13.23, 13.24). However, since all 18 examples in the corpus are from speakers heavily influenced by Indonesian, it is possible that this use of *ne'ebè* reflects linguistic interference (e.g. from Indonesian *sehingga* 'so that').

13.23 *Ami* halón bá matabian [ne'ebè n-alo oan
1PI plead to spirit.of.the.dead so.that 3S-make child

*isin* manas *di'ak*.
body hot good
We pray to the spirit of the dead person that (it) will make the fevered child well. (D0.44)

13.24 *Ita* hakara [ne'ebè ita fa'e malu].
1PI like so.that 1PI divide each.other
We desire that we separate from each other (e.g. by divorce). (T0.59)

Two analyses are possible for this construction. Either the clauses introduced by the conjunctions are complement clauses, or they are (like other clauses introduced by these conjunctions) non-argument, adverbial clauses. Although the evidence is not conclusive, two semantic arguments favour a complement analysis. Firstly, the semantic distinction between these clauses introduced by purpose conjunctions and complement clauses without conjunctions is not apparent. And secondly, it is not apparent that wanting or promising can be done ‘in order to’ achieve the result, as an adverbial analysis would suggest. For verbs of ordering, analysis as an adverbial clause is possible from a semantic point of view, since ordering is done ‘in order to’ achieve the result specified in the complement clause. This construction thus parallels the use of purpose conjunctions with other, non-complement-taking, predicates, such as in 13.25. Nevertheless, I propose by Occam’s razor that purpose clauses following verbs of ordering be, at least provisionally, analysed as complement clauses.

13.25 *Emi* ne’e mai [bat kesi nia].
2P this come so.that tie 3S
You come to tie him up. (Z5.103)
13.3.5 Clause-initial complements

13.3.5.1 Introduction

There are two constructions in which apparent sentential complements occur clause-initially, thus preceding the subject (if any). While the first clearly involves complementation, the analysis of the other is less certain.

13.3.5.2 Question + 'know'

The first construction involves complement-taking verbs that signify knowing (16 examples), usually hatene 'know'. The initial clause is interrogative, while the second clause indicates that the speaker does not know the answer. This ignorance may be indicated by negating a verb of knowing (13.26; 14 examples), using the verb halu'a 'forget' (1 example), or using a rhetorical question (13.27; 1 example). The two clauses are generally separated by comma intonation.

13.26 "... "[Nia n-ó oan hira hira mós] ami
3S 3S-have child how-many how-many also 1PE
la hatene." not know
["How many children does he have?""] "However many children he has, we don’t know."
(L0.31)

13.27 [Mane ne’e at sé ruma], sé n-atene dei?
man this 1RR who unknown who 3S-know only
Who this man would be, who knows?
(Z2.108)

13.28 "... "[Nia n-akara at mate ká nú sá], ita la
3S 3S-like IRR die or like what 1PI not
hatene Ibu." know mother
["Did the woman who drank pesticide want to die?""] "Whether she wanted to
die or what, we don’t know, Ibu."
(V0.98)

The initial clause in this construction is analysed as a fronted object complement of the verb of knowing. Evidence for this analysis is that the construction is found only with complement-taking predicates, and that the clause functions semantically as a complement. Fronting of the complement is not syntactically required in irrealis contexts, as shown by the final position of the sentential complement of ‘I don’t know’ in example 13.1.

The subordinate status of the initial clause is indicated both by its non-final intonation and by the fact that the adverb mós ‘also’ can occur finally within the question (13.26; 3 examples), just as it can on subjects and on initial subordinate clauses such as concessives.

---

12 These verbs of knowing appear to be a subclass of the verbs of thinking, but the extent of the subclass is not known.

Limited evidence suggests that polar questions are expressed as alternatives in this construction, by expressing the alternative either as ká lale ‘or not’ or by an expression such as ká sá ‘or what’ (13.28).
13.3.5.3 **Statement + 'want'**

A second construction in which an apparent complement occurs initially involves verbs signifying wanting (e.g. *hakara* 'like, want', *ho'i* 'do not want', *haho'uk* 'agree to'). The initial clause presents a statement, while the second indicates whether the condition specified by that statement is desired. Most such second clauses are irrealis, being either negated (13.29; 15 examples) or questioned (13.30; 4 examples). Only a minority are positive statements (2 unclear examples). The initial clause carries non-final comma intonation.

13.29  
*[Ha'u k-o'la sakaer kale'uk át él], ha'u la k-o'i.*  
1S 1S-take tamarind crooked bad this 1S not k-o'i.  
1S-do.not.want  
[I came to catch fish and instead caught a tamarind.] Me catching this bad, crooked, tamarind—I don't want to.  
(U4.16)

13.30  
*[O-'.oan sia mai labu né], emi haho'uk ká la ho'i?  
do.not.want  
The children coming here to court (your daughters)—do you agree (to it) or disapprove?*  
(O5.45)

The analysis of this construction is uncertain. There are two differences between these initial clauses and clause-final complements of verbs of wanting. Both differences are evidence (though not conclusive) against analysis of the initial clause as fronted complement of the second. Firstly, although final sentential complements were accepted during elicitation for verbs of wanting, the corpus contains no naturally occurring examples of such a construction outside poetry (fn. 5). Instead, final complements of verbs of wanting usually consist of reduced clauses. This makes analysis of initial sentential clauses as complements possible but unlikely. Secondly, although irrealis *atu* was not accepted in postverbal complements for *ho'i* 'do not want' (*la k-o'i atu bá 'not 1S-do.not.want IRR go'), it is found in clauses that precede *ho'i* (13.31; 5 examples).

13.31  
*[Ha'u at k-bá iha nebá], ha'u la k-o'i.*  
1S IRR 1S-go LOC there 1S not 1S-do.not.want  
[If we marry, we must live here.] I don't want to go over there. (or ‘As for me going over there, I don't want to.’)  
(F3.26)

An alternative analysis is that the initial clause is a subordinate non-complement clause, which is semantically interpretable as either a condition or a topic. Some examples are semantically close to example 13.32, in which the initial clause contains the explicitly conditional conjunction *kalo* ‘if’. In others, such as 13.29, the initial clause reiterates a statement and so is more like a topic.

---

13 This appears to be a subclass of the verbs of liking/wanting; however, the extent of the subclass is not known.
Nevertheless, it is unclear why irrealis contexts should favour a subordinate clause construction, when a correlation between irrealis and initial (non-complement) subordinate clauses is not observed elsewhere in the grammar. Rather, irrealis contexts were shown in the previous section to optionally trigger fronting of clausal complements of verbs of knowing. This is an argument in favour of a complement analysis of the construction currently under discussion also.

### 13.3.6 Sentential subjects

The corpus contains no unambiguous instances of complements functioning as subjects.\(^\text{14}\) Although examples such as 13.33 can readily be interpreted as containing a sentential subject, an alternative analysis as a conditional clause followed by a protasis is also possible ('If one steals money, it is not taboo').

13.32 \[Kalo fa'e malu liu] gereja la n-o'uk.\]

If divide each other further church [Mly] not 3S-agree to

If (married couples) separate totally (i.e. divorce), the church does not approve. (NO.8)

13.33 \[Hana'o loit] lulik ha'i. [Buat hotu-hotu hana'o] lulik ha'i.

steal money taboo not thing RDP-all steal taboo not

Stealing money is not taboo. Stealing any (ordinary) things is not taboo. (I0.172)

A conditional interpretation is favoured if there is comma intonation before the predicate.

13.34 \[Ai kadoek kti tán ó], lulik.

wood fragment hit on top 2S taboo

(If) the wood chips hit you, (it) is taboo. (B1.27)

### 13.4 Reduced complements

#### 13.4.1 Description

The complements discussed so far have all been sentential ones. The second type of complement clause is that of reduced complements. These have a number of identifying characteristics. They lack a subject. The implied subject is coreferential with either the subject or the object (stated or implied) of the matrix clause, depending on the semantic type of the complement-taking predicate. The verb of the complement clause takes subject marking accordingly.

Reduced complements have 'dependent time reference' (Noonan 1985:92). That is, the complement necessarily has a certain time reference (usually future) relative to the time of the complement-taking predicate. Hence such complements frequently include irrealis *atu*, no other tense-aspect marking being acceptable.

These complements allow no complementiser, and are usually not separated from the matrix clause by pause. They cannot be interrogative. They allow no clause-initial NPs (such

\(^{14}\) There are no relevant examples with /h/-initial verbs in the predicate, by which subject-marking possibilities can be determined.
Complementation as topics or fronted objects). They do, however, allow postverbal peripheral constituents, and can be independently negated (13.35; 1 elicited example only).\(^{15}\)

\[
\begin{align*}
\text{13.35} & \quad \text{Ô } m\text{-akara [at bā] bele, m\text{-akara [at la bā] mós bele.} \\
& \quad \text{2S 2S-like IRR go can 2S-like IRR not go also can} \\
& \quad \text{If you want to go, you may; if you want to not go, that is also fine. (T0.65 elicited)}
\end{align*}
\]

Complement-taking verbs for which the subject of the complement is interpreted as identical with the object of the matrix clause are classed as ‘object control’. This class of verbs includes verbs of ordering. Other verbs have ‘subject control’, whereby the subject of the complement is interpreted as identical to the subject of the matrix clause.\(^{16}\)

Syntactically, constructions in which reduced complements follow the complement-taking predicate have much in common with serial verb constructions. Both involve a shared argument, dependent time reference, and lack of intervening complementiser or pause.\(^{17}\) They differ in that:

1. Reduced complements fill the same syntactic slot as sentential complements (for those complement-taking predicates which allow both types of complement).
2. Complements and complement-taking verbs can be independently negated, while verbs in serialisation can not.
3. Certain types of reduced complements can precede the complement-taking verb, particularly if the matrix clause is irrealis. In contrast, the order of serial verbs is, in most cases, fixed, with exceptions being totally unrelated to the distinction between realis and irrealis.
4. Verbs of reduced complements can be introduced by the irrealis auxiliary \text{atu}, while the second verb in serialisation can not.
5. In positive statements, the truth value of complements depends on the nature of the complement-taking predicate (e.g. ‘remember’ versus ‘forget’), while the truth of the second verb in serial verb constructions is at least implied, even if not quite stated.

\subsection*{13.4.2 Object control: reported orders}

Verbs of ordering, such as \text{bolu} ‘call’, \text{haruka} ‘order’ and \text{katak} ‘tell’, refer to attempts to change the behaviour of the addressee. When used to report such attempts, they can take a following object NP referring to the addressee, and then a reduced complement clause, stating what the addressee is requested to do (100 examples). The complement clause is interpreted as having a subject coreferential with the object of the matrix verb.

\[
\text{Clause: Object Control} \rightarrow (S) \text{ V.Order (O.Addresssee) Reduced.Complement}
\]

\(^{15}\) I have no evidence to date of negative raising, where a negator which logically belongs in the complement is syntactically located in the matrix clause.

\(^{16}\) In English it is possible for a verb to be used with object control if used directly (‘John asked Bill to sing something’) or with subject control otherwise (‘John asked Bill what to sing’). To date I have found no such ambivalent verbs in Tetun.

\(^{17}\) The similarity leads Li and Thompson (1981:594ff.) to conclude for Mandarin Chinese that complementation constructions are a subclass of serial verb constructions, distinguished only by their semantics.
The addressee can be omitted (13.38; 9 examples), or, rarely, specified by a phrase introduced by bá ‘to’ (13.39; 1 textual plus 2 elicited examples). Fronting of the addressee NP was allowed in elicitation, although it never occurred spontaneously in the texts. This ability to be separated from the complement clause supports analysis of the addressee NP as being part of the matrix clause, and not syntactically subject of the complement.

<table>
<thead>
<tr>
<th>Verb</th>
<th>Gloss with sentential complement</th>
<th>Gloss with reduced complement</th>
</tr>
</thead>
<tbody>
<tr>
<td>hakara</td>
<td>want that</td>
<td>want to</td>
</tr>
<tr>
<td>halu’a</td>
<td>forget whether</td>
<td>forget to</td>
</tr>
<tr>
<td>haneo</td>
<td>ponder whether</td>
<td>ponder whether to</td>
</tr>
<tr>
<td>hanoin</td>
<td>think that</td>
<td>remember to, consider to</td>
</tr>
<tr>
<td>hatene</td>
<td>know that</td>
<td>know how to</td>
</tr>
</tbody>
</table>
Subject control complements are found in two positions, either clause-finally, or immediately preceding the matrix verb. I was unable to elicit any difference in meaning between the two orders. Syntactic differences are discussed below.

\[
\text{Clause Subject Control} \rightarrow \begin{cases} 
(S) \text{ V Reduced Complement} \\
(S) \text{ Reduced Complement V}
\end{cases}
\]

There are no examples in the corpus with both an object NP and a reduced subject control complement (comparable to English 'promise me to come').

13.4.3.2 Clause-final complements

Clause-final reduced complements are illustrated below for verbs of thinking (13.40; 17 examples) and liking (13.41, 13.42; 58 examples).

13.40 ...hanoin ha'i [toba].
think not lie down
[The busily playing child] doesn’t think to sleep. (I14/7)

13.41 Ha'u la k-o'i [k-ola ó].
1S not 1S-disnot.want 1S-take 2S
I don’t want to take (i.e. marry) you. (U4.5)

13.42 Sia n-akara [at tür iha rai Luka Wakeke].
3P 3-like IRR sit LOC earth Luka Wakeke
They wanted to live in the district Luka Wakeke. (F2.52)

The attention verb horan ‘feel, sense’ can take a subject control complement, but in the corpus its complements are restricted to predicates headed by adjectives and non-active intransitive verbs (23 examples).

13.43 ...ka kole ti'an mós la horan [kole].
1P be.tired already also not sense be.tired
[We were determined to win.] Even if we were tired, we didn’t feel tired. (I0.86)

The phasal verbs hahú ‘start’ and hanawa ‘finish’ also occur with following same-subject complements, which are usually but not necessarily single verbs (e.g. Sia r-ahú r-afaho ‘3P 3P-start 3P-weed’ = ‘They started weeding’; Emi hanawa harís lai ‘2P finish bathe first’ = ‘You finish bathing now’). However, of the 19 such examples for hahú ‘start’, all but one are either from direct elicitation of phasal verbs or from Indonesian-influenced teachers. Similarly, all but one of the five examples for hanawa ‘stop’ are directly elicited. It is thus possible that the use of complements with phasal verbs is marginal or Indonesian-influenced.

Same-subject complements are also found for dadi ‘happen, end up’. This verb indicates that the event specified by the subject plus complement clause ends up happening as expected (5 examples).\(^\text{18}\)

\(^{18}\) The form dadi is also a copula meaning ‘become’ (§9.4.5.4). Evidence that dadi ‘happen’ is a complement-taking verb and not an auxiliary is that it can be negated by postmodifying ha'i (13.45). For discussion of the similarities and differences between auxiliaries and complement-taking verbs in general see §10.2.
13.44 Bei Lafaek la \textit{dadi} n-á Bei Kancil...
Mr crocodile not happen 3S-eat Mr mouse.deer\[Mly\]
Mr Crocodile didn't get to eat Mr Mouse Deer [because Mr Mouse Deer was clever.]  
(F0.166)

13.45 \textit{Dadi} ha'i tukar.
(They) didn't end up swapping (rings).  
(S0.37)

13.4.3.3 \textbf{Preverbal small clause complements}

The second position for reduced subject control complements is immediately preceding the complement-taking predicate, thus following the subject, if any. In the corpus all 35 examples are for the thinking verb \textit{hatene} 'know (how to)'.

13.46 \textit{Karawa ne’e sia [r-akmatek] la r-atene!}
These monkeys don't know how to be still! (Said of children who kept giggling and moving about.)  
(X0.74)

There are two restrictions on this construction. Firstly, the complement clause must be a 'small clause' (§9.7.3), consisting of a single verb (13.46) or a verb plus a single-word object (13.48).

Secondly, the matrix clause is in this construction usually irrealis, being negative (13.46; 19 examples), or a purpose clause (13.47; 6 examples). Of the remaining examples, some are emphatically positive, contrasting with negative clauses (13.48), while the rest lack sufficient context to determine conclusively whether they are contrastive (13.49). Note that these same conditions optionally trigger the preverbal placement of nominal objects also, in what is analysed as object incorporation (§9.4.3.4).

13.47 \textit{...bate [r-a-he’e] r-atene, [r-asoru] r-atene.}
so that 3P-make-respond.greet 3P-know 3P-greet 3P-know  
[She was sent to civilise the uncivilised,] so that (they) would know how to respond 'He’e' to greetings, (and) know how to greet people.  
(F6.36)

13.48 \textit{Tán kabau [n-á hae] n-atene, ita [há hae] la hatene.}
because buffalo 3S-eat grass 3S-know 1P I eat grass not know  
Because buffalo know how to eat grass, (and) we don't know how to eat grass.  
(B1.19)

3S weave 3S-know tie-dye 3S-know spin also know  
She knows how to weave, knows how to tie-dye thread, (and) also knows how to spin.  
(N0.103)

These complements give the names of activities, rather than referring to particular instances of those activities. As such, they would be prime candidates for nominalisation if Tetun had had appropriate nominalisation strategies (Noonan 1985:129). The fact that these

---

\(^{19}\) Attempts to use the thinking verb \textit{hanoin} 'remember, think' in this construction were rejected during elicitation (*\textit{karian} la n-anoin 'work not 3S-think'; contrast: la n-anoin karian 'not remember to work').
Complementation

complements name activities, and their exact parallel with incorporated object nominals, leads me to analyse them as incorporated complements.\(^{20}\)

The fronting of reduced complements in irrealis contexts is syntactically optional, as complements can readily follow the complement-taking predicate in these contexts (13.41). It is, however, not uncommon crosslinguistically for a realis/irrealis distinction to be reflected in different complementation strategies (Noonan 1985:96).

Note that the auxiliary bele 'can' is similarly able to follow its 'complement' clause in irrealis contexts (13.50; §10.4.2), although the complement is not restricted to small clauses. The fact that bele 'can' and hatene 'know' participate in very similar constructions reflects a common pattern crosslinguistically, whereby words like 'can' (which express 'secondary concepts') use the same grammatical means as complement-taking predicates (Dixon 1995:177). It also reflects a close semantic tie, since hatene in this construction has a 'know how to' sense which overlaps with bele 'can'.\(^{21}\)

13.50 \(\ldots ha'\u2019u\ k-a-tun\ la\ bele.\)
1S 1S-make-descend not can
[Come and help me with the luggage, as] I can't get (it) down. (X0.35)

\(^{20}\) I found no mention of incorporation of complement clauses in Baker's (1988) wide-ranging discussion of incorporation, nor in Mithun's (1984) discussion of noun incorporation. Nevertheless, in the absence of any evidence of nominalisation of the complement, the Tetun evidence favours analysis of the complement clause itself as being incorporated into the complement-taking predicate.

\(^{21}\) The semantic closeness is indicated by the fact that verbs meaning 'know' are the most commonly documented diachronic source for words denoting ability (Bybee, Perkins & Pagliuca 1994:190).
14 **Beyond the clause**

### 14.1 Introduction

This final chapter is concerned with structures beyond the clause. The first section deals with constructions which contain elements which are not part of any clause, namely pragmatic topics, right-dislocated constituents, vocatives and tags. This is followed by discussions of various means of linking clauses. Such clause linkage falls into four categories.

Firstly, clauses can be embedded as arguments within other clauses, as discussed in the previous chapter on complementation.

Secondly, two or more clauses can be coordinated. Such coordination is of two types. One is applicable to a wide range of constituent types, can coordinate more than two constituents, and allows each conjunct to have the same structure as if it were not coordinated. The second type of coordination ('coordinate-dependent constructions') is only for clauses and places syntactic restrictions on the second conjunct.

Thirdly, a clause can function as an adverbial modifier for another clause. The determiner *ne’e ‘this* marks one type of adverbial clause, while others are marked as adverbial by conjunctions. Clausal parataxis can express a diversity of relationships, ranging from coordination to relationships in which the initial clause is adverbial relative to the following one; clausal parataxis is therefore discussed separately from clauses which are explicitly coordinated or explicitly adverbial.

Finally, a clause can be embedded as a relative clause modifier within an NP.

### 14.2 Extra-clausal constituents

#### 14.2.1 Overview

The structure of a sentence is summarised in the formula below, where clause* stands for one or more clauses, and represents the only obligatory constituent.

$\text{Sentence} \rightarrow \text{(Connective adverb)} \ (\text{Time}) \ (\text{Prag.Topic}) \ \text{Clause*} \ (\text{Tag}) \ (\text{Right-dislocation})$

Connective adverbs are words or phrases used to link the sentence with what went before (e.g. *nia ti’a ‘after that, then’, niama ‘then’, dadi ‘so*). Time expressions can occur either before or after pragmatic topics ('Prag.Topic'), resulting in their inclusion both amongst extra-clausal constituents and in the periphery of the clause (§9.8.2).

In addition to the extra-clausal constituents mentioned here, there are interjections (including exclamations and hesitation markers) which are not subject to syntactic rule, and which are discussed in §3.12. Vocatives too are not fully subject to syntactic rule; they are discussed in §14.2.4.

The various extra-clausal constituents differ from clause-internal constituents (such as arguments and periphery) in the following ways:
1. Constituents within a clause can be the focus of a question, while those outside the clause can not.

2. Pause is much more likely at the boundary between a clause and neighbouring extra-clausal constituents than it is between constituents within a clause.

3. In sentences consisting of more than one clause, extra-clausal constituents occur only once, either initially or finally.

4. Unlike clause-internal constituents, extra-clausal constituents cannot be part of a subordinate clause.

The final two statements raise the thorny problem of what a sentence actually is. A syntactic definition of a sentence in terms of having only a single set of extra-clausal constituents correlates strongly with intonational and intuitive clues as to where sentence boundaries lie. Nevertheless, it is possible (though unusual) for speakers to utter a tag (which is normally sentence-final) but then to continue the sentence by adding another clause anyway. This illustrates the fact that phonological and grammatical sentence boundaries are not always congruent, as well as the general difficulty of determining exactly what a sentence is, particularly in spoken discourse (Matthews 1981:29ff.; Pawley & Syder 1983:573).

Another problem related to sentence boundaries is the fact that there is a fuzzy boundary between clauses related loosely by subordination and clauses which constitute separate sentences. In particular, the reason conjunctions *tan* and *te* ‘because’ are only loosely subordinating. That is, while the constituent they introduce can be phonologically and syntactically subordinate, it can alternatively be a separate sentence both phonologically (shown by sentence-final intonation on the preceding clause) and syntactically (e.g. by having a left-dislocated subject). Such indeterminate connectivity is probably characteristic of spoken vernaculars, being found also for English connectives such as *and*, *but* and *because* (Chafe 1988:13, 18f.).

The following sections deal with the various extra-clausal constituents that have not been discussed elsewhere in this grammar, namely pragmatic topics (including left-dislocation), right-dislocation, vocatives and tags.

14.2.2 Pragmatic topics

14.2.2.1 Introduction

A clause can be preceded by an NP which identifies some sort of ‘topic’ for that clause. There is nothing in the sentence to indicate the semantic role of the NP in the following clause or clauses. The topic NP is usually coreferential with an NP within the clause, resulting in what is known as a left-dislocation construction. However, it need not be.

14.2.2.2 Left-dislocation: topic

Where the extra-clausal topic NP is coreferential with an NP within the clause (the latter being indicated by double-underlining in the examples), the clause-internal NP usually consists of a pronoun. Coreferentiality with the subject is by far the most common (137 examples), accounting for about 70% of the 190 examples with pragmatic topics.
Chapter 14

14.1 Ai\,kanoik n-osi liurai Suri Nurak ne'e nü nemai:
story 3S-about ruler Suri young this like here
The story about Ruler Suri Junior is like this:

Liurai Suri Nurak ne'e, nia mak taruka ohin liurai
ruler Suri young this 3S REL replace just.now ruler

Suri Tuan...
Suri old
This Ruler Suri Junior, he was the one who replaced the aforementioned Ruler
Suri Senior. [But he was still single...]
(F2.2)

Less common is coreference between topic and subject possessor (14.2; 22 examples) or
object (14.3; 7 examples).

14.2 Mesak Oan, kadeli nia-k á mutin; Feto lkun.
Mesak small ring 3S-POS DEF white woman tail

kadeli nia-k á mean.
ring 3S-POS DEF gold
Mesak Oan, his ring was silver; Youngest Sister, her ring was gold. (Z6.28)

14.3 Nia ne'e, ha'u k-adomi nia liu.
3S this 1S 1S-love 3S further
This one, I really pity her. (I5/7)

Only one clear example of left-dislocation on the possessor of the object was found in the
corpus.

14.4 Manu aman mean ne'e ha'u fokit k-ola ni liras.
bird male red this 1S jerk 1S-take 3S wing
This red rooster, I pluck its wing (i.e. pluck a feather from its wing). (K5.14)

14.2.2.3 Left-dislocation: 'premature subject'

Apart from discourse-motivated topics, there seems to be a significant proportion of left-
dislocation examples which are motivated by processing constraints. These are what Chafe
(1976:51ff.) calls 'premature subjects', in which the speaker chooses a subject and utters it
before formulating the clause as a whole.\(^{1}\)

Although one cannot tell for any particular example whether a left-dislocation was
motivated by discourse factors or by processing ones, there are two circumstances for which
left-dislocations are over-represented, and which can be interpreted as involving processing
constraints. Both involve constituents which put distance between the head of the would-be
subject and the head of the predicate. This distance is shortened by placing a pronominal
subject after the intervening constituent.

The first condition is where the would-be subject is postmodified by a relative clause. This
accounts for 17% of left-dislocated subject examples.\(^{2}\) Conversely, of the 59 NPs in the
corpus which are modified by relative clauses and which function as either subject or left-

\(^{1}\) I suspect that 'premature objects' are similarly possible.

\(^{2}\) Keenan and Schieffelin (1976:247) similarly report that in adolescent English, left-dislocated NPs are often
followed by interruptions to provide background information. However, their examples do not use relative
clauses.
Beyond the clause

297

dislocated subject, 23 (39%) are left-dislocated. Note that three of the examples are from written texts, indicating that this is not just a strategy of unplanned oral discourse.

14.5 Buat mak at moris. nia n-a-kiki rai.
thing REL IRR live 3S 3S-make-tremble earth
A thing (plant) that is about to shoot, it makes the ground wobble. (G0.57)

The second condition suggesting interpretation as ‘premature subject’ is where a constituent of more than one word (such as a time phrase or subordinate clause) intervenes between the would-be subject NP and the predicate (24 examples).

14.6 Ibu né. kalo nú ema r-a-horis nia musti
mother[Mly] this if like person 3P-make-live 3S must
iha oin.
LOC face
This (foreign) woman—if, for example, people were giving birth—she had to
be in the front. (J3.58)

These two conditions together account for 11 of the 18 examples in which the topic NP is headed by a pronoun (14.7). Such a topic would be expected to be relatively uncommon on discourse grounds due to the low topicality of pronouns. A further four pronominal examples have a single-word intervening constituent (e.g. nambé ‘I suppose’, loron ‘daily’). (The remaining three have a 1S pronoun ha‘u followed by a 1S subject clitic ha and subject marker k-.)

14.7 Ha‘u mak k-o‘i (hak) k-akés ne‘e. ha‘u naran N...
1S REL 1S-currently HES 1S-talk this 1S name N...
I who am currently speaking—my name is N...
(F1.1)

Final evidence that processing constraints are involved is that such conditions account for the two examples in which there is ‘double’ left-dislocation; that is, in which there is a topic NP, followed by a coreferential pronoun which is itself left-dislocated.

14.8 Iha nia fulan á. nia. su ra kalan nia tu‘an bokar.
LOC 3S moon DEF 3S every night 3S grow wide
Then the moon—it—every night it grows wider. (F4.11)

14.2.2.4 Topics with no clause-internal role

Topic NPs for which there is no clause-internal semantic role “limit the applicability of the main predication to a certain restricted domain”, and set a “spatial, temporal or individual framework within which the main predication holds” (as proposed by Chafe (1976:50) for ‘Chinese-style topics’). A single NP can set the topic for several following (intonationally linked) clauses (14.9). Short of using a separate clause, the only other means of specifying this type of topic NP in Tetun is by introducing it with a preposition (§8.6) or the Malay borrowing kalo ‘if’ (§14.7.3.1).

Of the 23 examples of such topics in the corpus, 15 (65%) refer to something which was mentioned in the immediately preceding clause (14.9). Three examples (13%) refer to something that was referred to three or four clauses previously, and five (22%) mark a change of topic (to a previous or related topic; 14.10). These topics are not contrastive. In the examples below the first line gives context (with antecedent for the topic indicated by double underlining), while the topic NP begins the second line of the example.
14.9 *Kalo fulan ida, ita kawen, ne’e n-a’ak ‘fuik-bua’*.  
if moon one 1PI marry this 3S-say betel.pepper-betel.nut Á.  
PAUSE  
One month, when we marry, this (the gift exchanged) is called ‘betel pepper and betel nut’. Hmm.  

**Fuik-bua ne’e osan bé wa’in onan...**  
betel.pepper-betel.nut this money also much IMM  
(For) this ‘betel pepper and betel nut’, there is much money (lit. ‘money is plentiful’), [many plates, much glass. From the woman there are many cloths, much rice and much meat.] (F1.22)  

14.10 *Uluk nú nia. Ikus né ... Á.*  
former.times like 3S final this PAUSE  
In former times (courtship) was like that. Nowadays ... (9 sentences from 3 speakers about how courting couples now get intimate very quickly, with the young woman often falling pregnant.)] Hmm.  

**Ema uluk, mane iha du’uk labis, feto iha uma laran.**  
person former.times man LOC self verandah woman LOC house interior  
People in former times—the man was separate on the verandah, (and) the woman was inside the house. (O5.82)

### 14.2.3 Right-dislocation: elucidation of a referent

In right-dislocation constructions, a clause is followed by an NP which further elucidates the referent of one of the clause participants. In the corpus these participants are the subject (14.11, 14.12; 18 examples), object (14.13; 9 examples) or subject complement (14.14; 2 examples). The referent may have already been specified explicitly in the clause by a pronoun (14.11) or non-pronominal NP (14.14), or it may have been left unspecified (14.12). In none of the examples in the corpus is there confusion as to which of the participants is being elucidated, since the clauses either are intransitive or have only one participant which is semantically compatible with the right-dislocated NP.

14.11 *Kalan nia toba la dukur, Sidarta ne’e.*  
night 3S lie.down not sleep Sidarta this  
At night he—this Sidarta—lay sleeplessly. (F0.178)  

14.12 *Iha nabé, wé ā?*  
LOC where water DEF  
Where is (it)—the water? (U1.43)  

14.13 *Né ita há bosu to’o, batar ai naruk.*  
this 1PI eat full(food) very maize leg long  
We eat it—sorghum—and feel very full. (Q0.164)
this 3S-POS k.o.poem house six 3S-POS k.o.poem  
This is its poem—Uma Nén's poem. (G1.1)

The right-dislocated NP generally provides 'added insurance' that the referent can be correctly identified (Givón 1990:484). Consistent with this, the right-dislocated NP sometimes has rising 'you know what I mean?' intonation. Some, at least, appear to be afterthoughts, in that the preceding clause may itself carry sentence-final intonation. Right-dislocation is considered acceptable for spoken but not for written Tetun.

14.2.4 Vocatives

Vocatives tend to form a separate intonation unit from the neighbouring clause. Most either precede the sentence with which they are intonationally linked, following any utterance-initial exclamation (14.15), or follow the sentence, including any tags (14.22). Only seldom does a vocative occur within a sentence (14.16).

14.15 ...“Ai, Bita Nahak, la'o mai...”  
EXCL Bita Nahak walk come  
[Then Buku Lasak said] “Hey, Bita Nahak, come here...” (AA1.40)

14.16 Oa ma-hó ukur ne’e, Ibu, r-aré ita moe,  
child who-have rite this mother 3P-see 2S.HON shy  
r-ata’uk.  
3P-fear  
Children who underwent this rite, Ibu, (when they) see you, they are shy (and) afraid. (P4.98)

14.2.5 Tags

Tags follow the final clause in a sentence. Each has a characteristic intonation pattern.3

The tag é (140 examples) characteristically has rising intonation, but does not form an information question and does not elicit a response. It is found in a wide range of situations where the addressee is anticipated to agree with the speaker, including explicit agreement with the addressee (Hou é ‘Yes’), suggestions (14.17), and narrative and explanatory statements (especially when the point is considered obvious; 14.18). In addition, it may be used in rhetorical questions, in which case the intended meaning has the opposite polarity to the clause (7 examples; §9.9.5.2).

14.17 Tu’an etu bá, é.  
add cooked.grain go TAG  
Take some more boiled rice, won’t you. (H0.135)

14.18 ...Sira hatene ha’i onan, é.  
3P know not IMM TAG  
[We teach girls to weave. They can’t just hang around. (Otherwise)] they won’t know (how to weave), will they? (J2.56)

3 The word lalé ‘no’ fills the same post-clausal position in polar interrogatives (e.g. Ó bá lalé? ‘2S go no’ = ‘Are you going (or) not?’). However, it is analysed as a negative pro-clause paratactically coordinated with the preceding clause.
The tag *ka* ‘maybe, or not’, usually with rising intonation, indicates uncertainty (22 examples). It usually marks polar interrogatives (14.19), but the corpus also contains an elicited example where *ka* helps mark a declarative clause as uncertain (14.20). It is also a coordinator meaning ‘or’ (§14.3.3). Because there is a close semantic relationship between the particle and the coordinator, *ka* is glossed as ‘or’ throughout.

14.19  
*A m-are kā?*  
2S 2S-see or  
Did you see (it)?  
(R4.157)

14.20  
*Keta sia de’an runa kā.*  
perhaps 3P speak. angrily maybe or  
Perhaps they would speak angrily.  
(T0.65 elicited)

The tag *tó* (27 examples), with rising intonation, occurs on both suggestions and (positive or negative) statements with which the speaker is expected to agree (14.21). It is a borrowing via Malay from Dutch *toch*. Although in many cases no reply is expected, the tag may also be used to form a biased question, to elicit what is anticipated to be a confirming answer (e.g. *Nū nia tó?* ‘like 3S TAG’ = ‘That’s how it is, isn’t it?’).

14.21  
*Leki Āt foi.wa’i ti’an n-ō neo ti’an tā?*  
Leki bad teenager already 3S-have understanding already  
TAG[Mly]  
Leki At was already a teenager, (so) already had understanding, didn’t he?  
(Z3.244)

The tag *lá* (10 examples), with falling intonation, is emphatic, and adds certainty to a statement.

14.22  
*N-a’ak “Ha’u, lá, ina!”.*  
3S-say 1S certainly mother  
(She) said “(It is) definitely I, mother!”. (Response to mother’s disbelief that the speaker was really her daughter.)  
(Z4.130)

The tag *sá* ‘just’ (10 examples) on explanations suggests that the explanation is obvious (14.23); on requests it means ‘Come on, don’t argue, just do it!’ (14.24). It has falling intonation.

14.23  
*Ha’u iha to’os dei sá. La’o bá sá?!*  
1S LOC garden only just walk go/for what  
I am just in the gardens. Where (else) would I go?! (Or: ‘What would I go (elsewhere) for?!’)  
(P4.121)

14.24  
*Fó mai sá! Fó it buat sia mai sá!*  
give come just give 1PI thing PL come just  
Just give (it) here! Just give our things here!  
(Z2.153)

---

4 Wortelboer (1955:177) notes frequent use of *lá* as a feature distinguishing the Fehan dialect from the Foho dialect of northern Belu.

5 This usage seems to overlap with that of *sá* in Kupang Malay (Steinhauer 1983:57).
The tag ő can be used following the greetings that are normally given when arriving, departing or beginning to eat or drink (5 examples). It has rise-fall intonation.

14.25  
\[ Ha'\ u\ k-liu-n\ ő. \]
1S 1S-further-IMM TAG  
I’m going on. (Greeting when leaving people one has stopped to talk with on the road.)  

(H0.36)

14.3 Phrasal and clausal coordination

14.3.1 Introduction

Coordination of two or more phrases or clauses can be made explicit by the general coordinating conjunctions nó ‘and’ and ká ‘or’. Occasionally é ‘or’, the disjunctive coordinator used in parts of East Timor (Morris 1984b), is heard in poems and songs. In addition, coordination may be expressed by parataxis, without use of an explicit coordinator.

It appears that the connective adverb lale ‘else’ can also coordinate NPs, introducing either an NP with an alternative referent (e.g. buka feto, lale mane ‘seek woman else man’) or an NP which is a paraphrase of the first (e.g. feto.ra, lale ata ‘maidservant else slave/servant’). Most of the 14 examples for this construction are from Indonesian-influenced speakers, raising the possibility that it is modelled on the Indonesian use of atau ‘or’.

The following sections will survey only the three most common forms of coordination, namely by nó ‘and’, ká ‘or’ and parataxis.

14.3.2 nó ‘and’

The conjunction nó ‘and’ coordinates a wide range of constituent types. By far the most common is coordination of NPs or NP heads (approximately 400 examples).

14.26  
\[ Sia\ r-ola\ samén\ nó\ rahenek. \]
3P 3P-fetch cement[Mly] and sand  
They fetch cement and sand.  

(X0.143)

Conjunction of more than two NPs can have nó ‘and’ between each NP (14.27), or only before the last NP (14.28). Where a non-final conjunct is not followed by nó (and sometimes when it is), it is marked by non-final intonation.

14.27  
\[ \ldots ami\ hakés\ hosí\ ama\ ida\ nó\ nia-kan\ fé-n\ nó\ ]
1PE talk about father one and 3S-POS wife-GEN and novo  
nia-kan\ oa-n.  
3S-POS child-GEN  
\ldots we have been talking about a father and his wife and his child.  

(K7.22)

---

6 Kupang Malay also has this use of ő (C. Grimes, pers. comm.).

7 The word nó is also a clause-final adverb meaning ‘also’, a verb meaning ‘exist’, and the 3S inflection of hó ‘accompany’. 
Coordination of subject NPs with human referents tends to use the comitative verb hó 'accompany' rather than nó 'and'.8 Coordination of two NPs with human referents can alternatively use the 'dual pronoun construction' described in §7.6.1.2.

Apart from NPs, nó can coordinate predicates (or parts thereof; 14.29; 36 examples), clauses (14.30; 5 examples), and NP modifiers (14.31; 11 examples). There are pragmatic restrictions on such coordination, as shown by the fact that almost 90% of the examples involve coordination of either near-synonymous pairs (as is common in ritual or poetic language; 14.29), or coordination of other elements from the same syntactic domain (e.g. two colours or two sizes; 14.30, 14.31). In contrast to hodi (§14.5.2), nó is not readily used for coordination of non-NPs unless the conjuncts are related in one of these ways. It was, for instance, deemed unacceptable (or at least very odd) to coordinate ‘come’ and ‘angry’, or ‘raise dogs’ and ‘look after the house’ with nó, but acceptable with hodi.

14.29 ...amí husu nó hakro’an bá íta bót. 1PE request and plead to 2S.HON big ...we request and plead to You (God). (C1.29)

14.30 ...hodi mai sia r-emu, nó sia r-á. bring come 3P 3P-drink and 3P 3P-eat [Then I prepared supper and] brought (it and) they drank, and they ate. (F3.16)

14.31 Nia titu iha sorin karuk nó kwana la n-aré ema. 3S look LOC side left and right not 3S-seep person He looked to the side—left and right—(but) didn’t see anyone. (K12.62)

Phonologically nó forms a constituent with the following conjunct, with pause, if any, preceding the nó (14.30).

14.3.3 ká ‘or’

Disjunction is expressed by ká ‘or, maybe’ (160 examples). It can coordinate a wide range of constituent types, including NPs (14.32), numeral phrases, verbs, prepositional phrases, predicates (14.33) and clauses (14.34).

---

8 It is possible that the coordinator nó was historically derived from the singular third person subject marking of hó ‘accompany’, namely n-ó. This is supported by the fact that ‘and’ in East Timorese Tetun is listed as hó by both Hull (1996b:321) and Morris (1984b). However, analysis of Fehan nó as a 3S inflection of hó is not viable synchronically, since nó can follow even first or second person conjuncts (e.g. emi nó ha’u 2P and 1S’ = ‘you and I’).

Comitative verbs have been posited as diachronic sources for conjunctions in some other languages also (Lord 1993:62–64).
14.32 ...dadi ne’e lahós ai abut ká sá ida, lale. 
so this indeed not plant root or what one no 
...so this wasn’t plant roots or anything (like that)—no! (K10.93)

The conjunction ká brackets with the preceding constituent. Phonologically this is shown by the fact that where there is a phonological break between the two constituents, it normally comes after the ká, which then tends to have rising intonation (14.33). Syntactically it is shown by the fact that when the disjunction is explicitly open (i.e. when the list of alternatives is not exhaustive, but is only presented as a list of possible examples), each disjunct is followed by ká (14.33).

14.33 Mane ... bá to’os. Bá hafaho ká, bá hare we ká, 
man go garden go weed or go rice water or 
nú nia.
like 3S
Men [after eating breakfast] go to the garden. (They) go weeding, or go to the rice paddy, or (whatever); it’s like that. (J2.41)

Disjuncts form the basis of alternative questions, simply by the addition of question intonation.

14.34 Belu, ó mai iha rai né, ó mane ká ó feto? 
friend 2S come LOC earth this 2S man or 2S woman
Friend, you came into this district—are you man or are you woman? (G4.1)

Polar interrogatives are formed by either conjoining a positive clause with (ká) lale ‘(or) not’ (14.35), or by appending ká to the clause without explicitly specifying the negative alternative (14.19). In the latter case, unless one accepts abstract analyses in which the negative alternative is ‘implied’, ká must be analysed not as a coordinator but as a question particle tag (§14.2.5).

14.35 Ó rona manu né ká lale? 
2S hear bird this or no
Did you hear this bird or not? (G1.62)

14.3.4 Paratactic coordination

Parataxis (or more accurately ‘asyndetic parataxis’) is here used to mean constructions in which two units are juxtaposed and linked by intonation, but are not linked explicitly by a conjunction. Parataxis can be used to coordinate a range of constituent types, including NPs (14.36), predicates (hodi... hodi... in 14.36) and clauses (14.38). Whether the relation between two intonationally linked units is one of conjunction (‘and’; 14.36), or disjunction (‘or’) is left implicit. There is a wider range of possible semantic relations between paratactic clauses than between paratactic phrases; these are discussed in the next section.

14.36 ...hodi glas, bikan, knuru, tasu, sasanan, hodi mai...

bring glass[Mly] plate spoon wok pot bring come
[The man] brings glass, plates, spoons, wok, pots, (and) brings [them to the girl’s house]. (E3.14)
14.4 Clause parataxis

When two clauses are intonationally linked, but have no explicit conjunction linking them, the two may be semantically related in a range of ways. These include conjunction ('and'), disjunction ('or'), and condition and consequent ('if'; 14.37). Example 14.38 illustrates the range and indeterminacy of the semantic relations between such clauses.

14.37 [Ó fota rahu kfui né], ita rua fa’e malu onan.
2S hit to.pieces flute this 1PI two divide each.other IMM
(If) you smash this flute to pieces, we two will separate. (F2.35)

14.38 Kanek oan ida kona ita, lalar mai n-á, ne’ek mai
wound small one touch 1PI fly come 3-eat ant come
n-á, susuk mai n-á mós habusik, la bele basa.
3-eat mosquito come 3-eat also leave.free not can slap
Even (if) we get a small wound, (and) flies come and bite, (and/or) ants come
and bite, (and/or) mosquitoes come and bite, (we) must just leave them; (we)
cannot slap (them). (F2.61)

Often two intonationally linked clauses depict simultaneous or subsequent events. In this case the two may either both have a similar level of foregrounding, or the first may provide the background time setting for the second. Backgrounding is typically achieved by aspectual marking and/or repetition (tail-head linkage).

Where the temporal relationship between the clauses is left totally implicit (e.g. without clues from connective adverbs or aspect markers), the order of the clauses is iconic. That is, the event expressed by the first clause is either simultaneous with (14.39) or subsequent to (14.40) that expressed in the first. This indeterminacy with regard to temporal relationships is the same as that found in serial verb constructions.

14.39 Dadi ferik fukun né sia mós moe, n-oran
so mature.woman elder this PL also embarrassed 3S-sense
n-ák “...”
3S-say
So the female elders were embarrassed, feeling “...” (F2.78)

14.40 Bá buka n-ola dók ida mai, sa’u, la di’ak.
go seek 3S-take shaman one come rub.medicine not good
(They) went and sought and fetched back a shaman, (he) applied medicine,
(but it) was no good. (F2.9)

Time relationships between clauses can be made explicit by the use of connective adverbs (§11.11) or by means of aspectual markers in one or both clauses. It is possible (though unusual) for aspectual marking to override the normally iconic order of clauses, as in 14.41.

14.41 [Nia uluk isin sei dauk la moras], nia n-akés
de’ak liu.
3S former.times body still not not sick 3S 3S-talk
good further
Formerly when he was not yet sick, he used to speak very well. (S0.19)

Often the initial clause is backgrounded. A particular instance of this is found in tail-head linkage, which is common in both narrative and procedural texts. Tail-head linkage involves
asserting information in one clause, and then recapitulating (part of) it as background information in the first clause of the next sentence (Thompson & Longacre 1985:209). The backgrounded clause normally has some indication that the event is now in the past, such as mós ‘finished’, hotu ‘finish, complete’, and/or ti’a (sometimes ti’an) ‘already’. Abbreviation of tail-head linkage is presumably the source for the very common sentence-initial connective nia ti’a (sometimes nia ti’an) ‘3S already’ = ‘after that; then’. In example 14.42, from a narrative text, the relevant portions of the initial assertions have double underlining, while the recapitulations have single underlining. For examples of tail-head linkage in a procedural text see Text 1 in Appendix A.

14.42 Nia ti’a ha’u k-mai. Mai ti’a.
3S already 1S 1S-come come already
Then I came. When (I) had come,

n-ó Bapa sia n-akés, ha’u k-bá nono wé á,
3S-accompany father[Mly] PL 3S-talk 1S 1S-go heat water DEF (my suitor) talked with Father and the others, (and) I went and boiled water,
kahur kopi á, fó sira r-emu.
mix coffee[Mly] DEF give 3P 3P-drink
mixed coffee, (and) gave (it) to them to drink.

Fó sia r-emu ti’a, nia n-ák á “…”
give 3P 3P-drink already 3S 3S-say HES
After (I) had given to them (and they) had drunk, he (my suitor) said “…”

Alternatives to tail-head linking are to background not a repetition or paraphrase of the previous statement, but some obvious, expected, result of it (Thompson & Longacre 1985:210). For instance, hearing is an expected result of being spoken to (14.43), and following an order is an expected result of being ordered to do something.

14.43 Na’i feto simu, n-a’ak “…” [Rona ti’a lia ne’e].
noble woman reply 3S-say hear already word this

na’i mane túr n-odi fudi la rona.
noble man sit 3S-COORD purpose not hear
The noblewoman replied “…” Having heard this message, the nobleman sat purposely not listening.

14.5 Coordinate-dependent constructions

14.5.1 Introduction

The following sections deal with two constructions in which two linked clauses are semantically coordinate, but in which the second clause is dependent on the first for its subject. Such constructions are what Foley (1986:177) calls ‘coordinate-dependent’. They differ from shared-subject serial verb constructions in that there is often a ‘comma intonation’ phonological boundary before the second clause, the two clauses can have separate peripheries, and the two can make separate assertions.

The first construction uses a verb to link the two clauses. The second uses no explicit relator, but is identified by the fact that the verb of the second clause is ha’ak ‘say’.
14.5.2 hodi: coordinate dependent

The first means of linking coordinate-dependent clauses is by the linking verb *hodi*. The subject of the second clause is obligatorily elided, being interpreted as coreferential with that of the preceding clause; *hodi* takes subject marking for that shared subject. Note that, in addition to being a linking verb, *hodi* is a transitive verb meaning ‘bring, take, use’.

There are four basic semantic relations that can hold between clauses connected by *hodi*. Although these are distinct, *hodi* is given the general gloss of ‘COORD’ in all cases. This is because the distinction between the first two relations (which are by far the most common) is largely pragmatically inferred, with the semantic relation for some examples being indeterminate.

Firstly, the two clauses can represent simultaneous actions (14.44, 14.51; 87 examples). In this case normally only the subject is shared between the two clauses.

14.44 *Ita soru hodi dakar sira.*  
1P weave COORD look after 3P  
We weave while looking after them (i.e. the children). (J2.62)

Secondly, there can be a relation of temporal succession, in which the action specified by the first clause is followed by that of the second (125 examples). This normally also involves a purposive relation, with the first done in order to then do the second. While the only required shared argument is the subject (14.45), many clauses representing successive actions share another argument as well. In particular, the object of the first clause is often interpretable as instrument (14.46; 64 possible examples), object (14 examples) or location (introduced by serial verb *bá*; 14.47; 2 examples) of the second clause. The prevalence of an instrument interpretation is not surprising given that *hodi* is also a transitive verb meaning ‘bring, take, use’, in which capacity it introduces instrument NPs in verb serialisations (§12.5.5). In fact it is sometimes not clear whether *hodi* is used as a serial verb (with an unspecified instrument NP), or as a clause linker. Both possibilities are reflected in the alternate translations of 14.46.

14.45 *Nia karian n*-odi* n-a-to'o naha bá uma laran.*  
3S work 3S-COOD 3S-make-enough baggage to house interior  
He works to supply things for inside the house. (T0.103)

---

9 Verbs functioning as clause linkers have of course been observed elsewhere. For instance, Noonan (1985:132) observes that “A few languages use verbs to translate English conjunctions like and and and then. Semantically, such conjunctions can be viewed as two-place predicates.”

Hull (1996b:4, 14) notes that *hodi* is often used in East Timor outside of Dili to mean ‘and’, while in Dili *hodi* or *atu hodi* (lit. ‘IRR COORD’) are normally used with the sense ‘in order to’.

10 Note that *hodi* is not used to connect semantically parallel verbs or predicates, which commonly occur in poetic and ritual language. These are normally juxtaposed, but can be coordinated by *nó* ‘and’. Nor does *hodi* intervene between repeated verbs that indicate repeated or continued action (e.g. *buka buka buka* ‘seek seek seek’ = ‘seek and keep seeking’).
14.46  
...ha’u  k-mama  ai-kakaluk  k-odi  taka  nia-kan  ain
1S  1S-chew  wood-power  1S-COORD  cover  3S-POS  leg
tohar  ne’e.
broken  this
...I  chew  medicine,  and  cover  his  broken  leg  (with  it)/(and)  use  (it)  to  cover  his
broken  leg.  
(K5.7)

14.47  
Ita  hada  labis  uan  ne’e,  hodi  rai  hék
1PI  build.floor  verandah  small  this  COORD  lay.down  just
hare  ne’e  bá.
rice  this  go
We  build  a  raised  floor  (in  the  rice  paddy)  to  temporarily  put  the  rice  on.
(P0.135)

Thirdly,  and  less  commonly,  hodi  may  coordinate  states  (19  examples,  mostly  elicited).

14.48  
Feto  Ikun  ksortir  di’ak  n-odi  matenk.  Ami  ksortir
woman  tail  fortune  good  3S-COORD  clever  1PE  fortune
lalek,  hodi  beik.
lack  COORD  stupid
Youngest  Sister  was  fortunate  and  clever.  We  are  unfortunate  and  stupid.
(V0.83)

Finally,  limited  evidence  suggests  that  the  initial  clause  may  specify  an  action,  while  the
second,  stative,  predicate  indicates  manner  (6  examples;  e.g.  la’o  n-odi  di’ak  ‘walk/go
3S-COORD  good’  =  ‘go  well’).

For  coordination  of  actions,  all  hodi-linked  clauses  in  the  corpus  have  positive  polarity,
and  all  have  human  or  animate  subjects.  It  is  not  clear  whether  these  are  strong  tendencies  or
syntactic  requirements.\(^\text{11}\)  For  coordination  of  states,  however,  inanimate  subjects  readily
occur,  and  negation  of  either  or  both  clauses  was  accepted  during  elicitation.

14.49  
Ema  ne’e  la  naruk  n-odi  la  kbokur  tán
person  this  not  long  3S-COORD  not  fat  on.top
ta  bá.
already  go
This  person  is  not  tall  and  on  top  of  that  is  not  fat.  
(V0.95  elicited)

Action  clauses  linked  by  hodi  can  have  separate  peripheries  (iha  rate  sia  ‘at  the  graves’  in
14.50),  include  their  own  serial  verb  constructions,  and  can  themselves  consist  of
semantically  coordinate  clauses  (e.g.  toba  la  dukur  ‘lie  down  (and/but)  not  sleep’  in  14.51).
The  two  clauses  can  be  separated  by  a  phonological  clause  boundary.  In  fact  about  35%  of
temporal  succession  sequences  and  20%  of  simultaneous  action  sequences  were  transcribed
with  a  comma  (14.51)  or  stop  (14.50)  before  hodi.

\(^{11}\)  While  reflecting  the  tendency  for  subjects  generally  to  be  animate,  and  particularly  subjects  of  active  verbs,
the  ubiquitousness  of  animate  subjects  may  also  be  an  instance  of  ‘persistence’  (Hopper  1991a:22),  by
which  aspects  of  the  history  of  a  word  tend  to  be  retained  by  it.  In  the  case  of  hodi,  the  strong  preference  for
human  or  higher  animate  subjects  for  ‘take,  use’  is  retained  when  hodi  is  used  as  a  coordinator.
...sia bá r-aló misa iha rate sia. **Hodi** hanoín
3P go 3P-make mass[Port] LOC grave PL COORD remember
ema mak mate ti’a né sia...
person REL die already this PL

[In November, priests all over the world,] they go and conduct mass at the
graves. To remember the people who have died...

(K1 4.20)

14.51 *Suri Nurak né toba la dukur, n-odi n-ú*
Suri young this lie.down not sleep 3S-COORD 3S-blow.on

**kfui nia-k ida, n-odi n-aló laran né moras tán.bá...**
flute 3S-POS one 3S-COORD 3S-make interior this sick because
This Suri Junior lay not sleeping, playing his flute, making himself sad (lit.
‘interior sick’) on account of [wanting to marry another woman].

(F2.29)

Not surprisingly on semantic grounds, none of the examples of coordination of states have separate peripheries.

When clauses are related by temporal succession, **hodi** can be preceded by the irrealis auxiliary *at(u)* (14.52; 12 examples) or by the purpose conjunction *bat(u) ‘so that’* (14.53; 10 examples) or *ne’ebé ‘so that’* (2 examples). These all force a purposive interpretation, without requiring that the purpose given in the second predicate necessarily be fulfilled. Alternatively, **hodi** can be preceded by the connective adverb *foin ‘then’* (5 examples). These co-occurrence possibilities emphasise that **hodi** is fully verbal, since only verbs may follow the auxiliary *atu* when it means ‘in order to’. Conversely, if **hodi** were a conjunction, one would not expect that it could co-occur with the purpose conjunctions or follow a connective adverb.

14.52 *Futu manu, taru, at n-odi n-ola n-ika*
fight(.cocks) bird gamble IRR 3S-COORD 3S-fetch 3S-back

**nia-kan nán né sia.**
3S-POS brother this PL

(Shewas goingto) fightcocks, (and) gamble, in order to fetch back her
brothers (who had been taken as slaves).

(G1.36)

14.53 ...

**man six this 3S cut wine chop garden DEF so.that**

**hodi** hahán ina-ama é.
COORD feed mother-father TAG

...these six men, they (lit. ‘he’) make wine, (and) work the garden, in order to
feed mother and father, eh.

(P2.95)

It is unclear why **hodi** would co-occur with other clause-linking devices, and how such doubly-linked constructions should be analysed. In many of the examples in which **hodi** co-occurs with other linking words, it is possible to interpret **hodi** as a serial verb meaning ‘use’. This does not, however, appear to be a ready way out for remaining examples, such as 14.52 and 14.53.

**14.5.3 ha’ak ‘say’**

A second type of coordinate-dependent structure involves a clause headed by **ha’ak ‘say, mean’**, in construction with a preceding clause headed by a verb of speaking. A superficially
similar construction, in which the preceding clause is not headed by a verb of speaking, will be analysed in terms of coordination rather than as a coordinate-dependent construction.

The verb ha'ak (with variants ha'ak-é, hák-é, hák, rarely k) introduces direct quotations (14.54; about 1,000 examples) or, far less commonly, indirect quotes (14.55).

14.54 ...nia bolu. N-ák "Iku mai lat". K-ák "Sá ida?..."
3S call 3S-say tail come first 1S-say what one
...she called, saying "Iku, come here". (I) said, "What is it?..." (F3.39)

14.55 Mai, n-á'ak [at n-ó nía kawen].
come 3S-say IRR 3S-accompany 3S marry
(He) came, saying he would like to marry her. (D1.7)

Frequently ha'ak is used after another verb of speaking (14.54, 14.56). In this case, unless the clause containing the verb of speaking has sentence-final intonation (and so constitutes a separate sentence; 14.54), the subject of ha'ak is obligatorily omitted, and the verb ha'ak is inflected for the subject of the preceding verb of speaking.

14.56 Bei Fahi n-atá nía n-á'ak "Ná sá?".
Mr pig 3S-answer 3S 3S-say like what
Mr Pig answered him saying "What's up?". (K8.25)

Note that other verbs of speaking are able to introduce quotes directly (§ 13.3.2). Nevertheless, there is a reasonably strong preference for using a coordinate-dependent construction with ha'ak to do so. In a sample of 250 sentences in which a verb of speaking (bolu 'call', hakés 'talk', haruka 'order', husu 'request', katak 'tell') was followed by a direct quote, 75% used a mediating ha'ak to introduce the quote, while only 25% introduced the quote directly.

In this coordinate-dependent construction the two verbs of speaking head two separate clauses. The clauses share the subject, polarity, time reference and aspect, all of which are specified on the first clause (if at all). The sharing is a consequence of ha'ak repeating part of the semantic content of the preceding verb of speaking. With regards to sharing, this construction resembles verb serialisation. However, unlike verb serialisation, it is not uncommon for ha'ak to start a new intonation contour; in fact 10% of a subsample of 100 direct quote examples were transcribed with a comma preceding the ha'ak. Another characteristic not shared with serial verb constructions is that the initial verb can be followed by peripheral constituents (14.57; 2 examples). It is not clear whether a verb of speaking followed by ha'ak should be interpreted as one assertion as in serial verb constructions ('He spoke "..."'), or as two ('He spoke. He said "..."').

14.57 Nía ti'a Bapa bá n-usu ha'u iha ke'an
3S already father[Mly] go 3S-request 1S LOC room
laran á n-ák "...
interior DEF 3S-say
Then Father went and asked me in the room, saying "..." (F3.58)

The verb ha'ak is unusual in that it allows only a limited range of modifiers. Apart from the expression ha'ak fali 'thought incorrectly' (lit. 'say in turn'), the texts contain only four

---

12 The é looks like a hesitation marker, but is particularly common following ha'ak, and is considered to 'sound good' there.
examples in which ha’ak is modified (e.g. la bele ha’ak ‘not can say’ = ‘cannot say’, ha’ak ha’i ‘say not’), all of which have ha’ak as the sole verb in the sentence. Consultants did not accept attempts to modify ha’ak with tenik ‘again’ or perfective ti’an ‘already’. It is possible, though unusual, for adverbs modifying the preceding verb to follow the complement of ha’ak (14.58; 4 examples from a church service, plus one from elicitation).

14.58 Ha’u la bele bolu emi ha’ak ‘atan’ tenik.
1S not can call 2P say slave again
I cannot call you ‘slaves’ again. (C2.6)

The same restrictions (of shared subject and modifiers) usually apply when ha’ak follows verbs which are not complement-taking predicates, but which describe some action prior to or simultaneous with the act of speaking (14.55, 14.59).

14.59 ...d6k né bá kaka n-ák “Lae, ha’u k-ó
shaman this go point.out 3S-say no 1S 1S-accompany
ida dei”.
one only
...the shaman went and pointed out (one of the girls) saying “No, I’ll take only
one”.
(F2.16)

Here there are clearly two separate assertions (of which the first can be independently negated). Subject sharing is in this context a strong pragmatic tendency rather than a syntactic rule, as shown by examples in which the implied subject of ha’ak is different to that of the preceding verb. For instance, when used following rona ‘hear’, ha’ak takes subject marking for the speaker of that which is heard, rather than for the hearer; this indicates that ha’ak means ‘say’, and is not simply a complementiser for quotes. Another example of clauses with different implied subjects is 14.60.

14.60 Tanát, n-are ha’i, n-ák “Maré ré ne’e, titu tuir ne’e”.
look.up 3S-see not 3S-say 2S-see at this look follow this
(He) looked up, (but) didn’t see (her); (she) said “Look here. Look this way”.
(O4.44)

Since argument sharing is not obligatory, such constructions in which ha’ak follows verbs other than verbs of speaking are not coordinate-dependent constructions. Rather, they are clauses coordinated by parataxis, with the subject of ha’ak conventionally (but not necessarily) omitted and interpreted as coreferential with that of the preceding clause.

14.6 Nominalisation with ne’e

When a clause describes an action which has been referred to or is inferable from the immediately preceding context or from the extralinguistic context, the determiner ne’e ‘this’ may immediately follow the clause, so nominalising it (about 50 examples). Such clauses are always subordinate. They usually occur sentence-initially, expressing a topic with respect to which the following clause is relevant (14.61, 14.62), or reiterating an earlier statement in tail-head linkage (14.63). Sometimes it is unclear whether ne’e modifies the entire clause, or merely a clause-final NP (14.63). As illustrated in 14.62, the verbs in such clauses retain normal subject marking.
14.61  Hakés hosí fila rai. [Fila rai ne’e] uluk fo hon ita... talk about turn earth turn earth this firstly 1PI (I'm) talking about turning the soil (i.e. hand-ploughing). (With regard to) this turning of the soil: firstly we [ram digging sticks into the ground.] (K11.4)

14.62  Ai! [Ha’u k-akerek e’e] hanawa hék lai. EXCL 1S 1S-write this stop just first Hey! This writing—I'll just stop for a while now. (L0.53)

14.63  Nia n-alai bá bá, sa’e liu ró. [Sa’e ró né], si 3S 3S-run go go ascend further boat ascend boat this 3P la’o.
walk/go
She ran there, (and) got up into the boat. When (she) had got up into the boat, they left. (G4.123)

Note that clauses marked by ne’e can also function as subject (14.64; few examples) or (rarely) predicative complement.

14.64  Iha nebá ami la’o, [la’o ne’e] sa’e. LOC there 1PE walk walk this ascend There (at the hill) we walked, (and) this walking was ascending. (Z1.13)

14.7  Adverbial clauses introduced by conjunctions

14.7.1  Overview

An adverbial clause is a clause which functions as modifier to another clause. Adverbial clauses which tend to precede the main clause are clauses of time, conditionals, concessives and topic clauses. Those which tend to follow the main clause are clauses specifying purpose, reason or manner.

Parataxis as a means of linking an initial backgrounded clause and a following main clause has already been discussed above, as has the use of ne’e ‘this’ to mark a topic clause. In the following sections we instead consider the various conjunctions which can be used to introduce adverbial clauses. Manner clauses will not be dealt with here, since they are introduced by the conjunction nudar ‘like, be as if’, which was discussed in §8.5.2.

14.7.2  Time

Tetun has few temporal conjunctions to make time relationships between clauses explicit, and even these are seldom used. The temporal conjunctions are hori ‘since’, natón.bá and bá.natón ‘at (time)’ and to’o ‘until’. Since all are also prepositions they are discussed further in §8.4.

Borrowing of time conjunctions from Indonesian is uncommon. However, some Indonesian-influenced speakers use the prepositional phrase bá oras ‘at time’ followed by a

---

13 Adverbial clauses are traditionally classed as 'subordinate'. Various authors have noted problems with the traditional notion of subordination, and have presented detailed alternatives (Haiman & Thompson 1984; Lehmann 1988); these, however, will not be delved into in what is only a brief overview of adverbial clauses in Tetun.
relative clause specifying a concurrent event to allow simultaneity (‘while’) to be made explicit (§14.8.3.2).

14.7.3 Condition

14.7.3.1 kalo ‘if’

Although conditionals are frequently expressed by simply juxtaposing the condition clause and the consequent, leaving the conditionality implicit, use of the Indonesian conjunction kalau ‘if, when, as regards’ is also common (200 examples). Fehan speakers recognise its Indonesian source, but accept the word as now being Tetun, distinguished from the Indonesian by its pronunciation, as kalo or, less commonly, kala.14

The word kalo is used for conditions regardless of whether they are past, present, future (14.67) or habitual (14.65), and regardless of whether they are almost certain to become true (thus translatable by English ‘when’) or whether they are counterfactual (14.66).

In 90% of examples the condition precedes the main clause, usually being separated from it by comma intonation.

14.65  
[Kalo la mama, la môn], matan dukur.
if not chew(.usu.betel) not chew.tobacco eye sleep
If (we) don't chew betel (and) don't chew tobacco, (we) get sleepy. (B0.166)

14.66  
[Kalo ha’u feto] ha’u la bele k-akur tasi wé-n
if 1S woman 1S not can 1S-cross sea water-GEN
mai buka ó.
come seek 2S
If I were a woman I wouldn't be able to cross the sea to come and look for you. (G4.3)

Alternatively, the condition follows an initial NP (14.67; 18 examples), such as a left-dislocated subject (of either clause) or a fronted object (of the conditional clause).

14.67  
Fanderen né [kalo ó m-aló nó kbit], ó mate...
foundation[Mly] this if 2S 2S-make with strong 2S die
These foundations — if you make (them) quickly, you’ll die [, or else fire will devour this house.] (R1.17)

Finally, the kalo clause can follow the main clause (6 examples). In this case it tends to introduce a reason rather than a condition, with the main clause asserting something that is true rather than something that is conditionally true.

14.68  
Baliun ne’e tá lais liu, [kalo kro’at].
axe this chop quick further as sharp.edged
This axe chops very quickly, as (it) is sharp. (H0.10)

The word kalo can also introduce a topic NP (4 examples).

---

14 Such borrowing of conjunctions is of course common crosslinguistically (Ross 1985:545; Thompson & Longacre 1985:204f.).
14.69  *Kalo* fukun, nia mate n-odi.
as.for elder 3S die 3S-bring
As for an elder, he is an elder for life (lit. 'dies taking it'). (This is in contrast
to a forestry officer, who is appointed temporarily.)

14.7.3.2 sô ‘unless’

All instances of sô (rarely sôk) ‘unless, except’ in the corpus occur following a negative
statement, and present a condition for the corresponding positive statement being true (16 examples). The positive statement optionally follows the sô clause, and is usually marked by
some adverb meaning ‘then’ (e.g. *foin*). Although the clause introduced by sô usually
constitutes a separate phonological sentence, sô is mentioned here because of its semantic
similarity to conditional conjunctions.

14.70  Ha’u k-bá ha’i k-á. [Sô emi haho’uk buat e’e.] Nia
1S 1S-go not 1S-eat unless 2P agree.to thing this then
k-tone k-á.
1S-go(to.you) 1S-eat
I won't go and eat. Unless you agree to this matter. Only then will (I) go and
eat.

The conjunction *tosako* (4 examples) appears to have a very similar meaning, but data are
limited.

14.7.3.3 surak ‘on condition that’

The conjunction *surak* (or *surak.án*) means ‘provided that, on condition that, so long as’
(19 examples). Unlike other conditional clauses, it always follows the clause with which it is
semantically tied, with the two separated by either comma or full-stop intonation.

14.71  Ô la m-á, so’in! [Surak ha’u k-ela bá ó.] 2S not 2S-eat OK so.long.as 1S 1S-leave go 2S
(If) you don't eat, so what! So long as I left (some) for you. (This was an
angry reply to the accusation that the speaker had left his brother no food.)

14.72  Ô bele bá nó, [sura]k keta m-ata’uk.
2S can go also so.long.as do.not 2S-fear
You can go too, so long as (you're) not afraid.

14.7.4 Concession: masik ‘although’

The conjunction *masik* ‘although’ introduces a concessive clause, which precedes the main
clause (13 examples).

14.73  Ina, [masik r-aboba bê], a bá dei.
mother although 3P-hit also 2S go only
Mother, even if (they) hit (you), you just go.

---

15 In their dictionaries of East Timorese Tetun, Hull (1996b) and Morris (1984b:139) list this conjunction as *maski*, which according to Morris corresponds to Portuguese *mas que*. 
Concession may alternatively be expressed by one of the adverbs mós ‘also’ or bé ‘also’ in final position in the first of two clauses. These adverbs can be supplemented by the conjunctions masik ‘although’ (14.73) or kalo ‘if’.

14.74 [Ta’u mós] bá dei.
mud also go only
Even if it’s muddy (we) will go anyway. (I26/6)

14.7.5 Reason

14.7.5.1 Introduction

There are four reason conjunctions, all of which translate as ‘because’. These are tán, tán bá, té and wain. In addition, the conjunction kalo ‘if’ can take on the meaning ‘because’ when it introduces postposed clauses (§14.7.3.1).

All the reason conjunctions can introduce postposed reason clauses. In addition tán and tán bá can introduce reason for a following clause; in this case the reason is usually given by an NP rather than a clause.

14.7.5.2 tán, tán bá ‘because’

The most ubiquitous of the reason conjunctions is tán (over 300 examples). It can introduce sentences, complete with multi-clause constructions, topicalisation and left-dislocation. Of a subsample of 100 examples with postposed reason clauses, 40% were transcribed with a full-stop before tán, and 40% with a preceding comma.

Reason clauses usually follow the main clause (14.75), but can also precede it (14.76; 12 examples) or occur in post-subject position (14.77; 3 examples).

14.75 Hetak ha’u, uluk Ina dadobe, [tán ha’u mesa
as.for 1S former.times mother pamper because 1S alone
dei].
only
As for me, in former times Mother pampered me, because I was alone (i.e. an only daughter). (Q0.174)

14.76 [Tán nta uluk sekola], dadi bolu ha’ak ‘Ama
because 3S former.times school[Mly] so call say father
Sekola’.
school
Because he formerly went to school, so (people) call (him) ‘Father School’.
(Q0.139)

---

16 This use of a focus particle meaning ‘also’ is common crosslinguistically in concessives (König 1993:980).
Mais feto na’in tolu né [tán bin
but woman CLS:human three this because elder.sister
alin], la‘o n-ela malu la di‘ak.
younger.sibling walk 3-leave each other not good
But these three women, because (they) were sisters, didn’t like (one to) go leaving the others behind.  

(F2.17)

Reasons introduced by tán can also be specified as NPs. Usually these consist of interrogative sá ‘what’ (tán sá ‘because what’ = ‘why’; 14.78; 15 examples), or are anaphoric, referring back to an earlier statement (e.g. tán lia nia ‘because word that’ = ‘for that reason’; 30 examples); however, most of the anaphoric examples are from speakers from outside the Fehan dialect area. Unlike reason clauses, most reason NPs precede the main clause.

14.78  Tán sá emi la bá fore?
because what 2P not go untwine
Why didn’t you go and untangle (them)?  

(X0.8)

The conjunction tán bá (lit. ‘because go/to’) can also be used to introduce reason (44 examples), with most examples presenting reason as an NP (e.g. tán bá ha‘u ‘because of me’; often tán bá nia ‘because of that’). Nearly all examples are from speakers from outside the Fehan dialect area.

14.79  Ami susar [tán bá kma‘un turu kona ami].
1PE be.in.difficulty because dew drip touch 1PE
We had it hard because dew dripped on us.  

(K9.75)

14.7.5.3 té ‘because’

The conjunction té always introduces reason clauses which follow the main clause (103 examples).17 It differs from tán in that the reason is nearly always (with only three exceptions) given for something that either the speaker is doing (14.80), or that the speaker and/or addressee is expected to do (14.81).18

14.80  Ha‘u k-o‘i bá [té ema r-aboba ha‘u].
1S 1S-do.not.want go because person 3P-hit 1S
I don’t want to go because people hit me (there).  

(R4.55)

14.81  Iku, Iku m-adér lai [té oan mane mai ti’an].
Iku Iku 2S-arise first because child man come already Iku, Iku get up now, because the young man (you are waiting for) has come.  

(F3.53)

In 17% of examples the té follows a full stop in the transcription, while in 25% it follows a comma; 4% have a comma following té.

---

17 Clause-initially something akin to reason can be specified by the connective phrase nia té ‘3S because’ = ‘in that case’ (18 examples). Here té is syntactically and phonologically linked to the preceding anaphoric pronoun and not to the following clause. There is limited evidence from the corpus that the form té can also be a tag on final reason clauses.

18 This strong tendency to be restricted to certain person-time combinations is identical to that found for the aspectual clitic -n ‘IMM’ (§11.5.5).
14.7.5.4 wain 'because'

There are only four examples in the corpus for wain 'because'. This conjunction is, I suspect, geographically restricted to a very small area within the Fehan dialect region, including the village of Bakateu. Its omission from the dictionaries of Hull (1996b), Mathijsen (1906) and Morris (1984b) suggests that wain is not used outside the Fehan dialect.

14.82 Bé (é) feto á kabuk ha‘i. [Wain haré ha‘i.]
but HES female DEF pregnant not because see not
But the woman wasn't pregnant, because (she) didn't see (her fiancé). (O5.67)

14.7.6 Purpose

A purpose clause expresses a motivating event which is unrealised at the time of the main clause event. In this it contrasts with reason, which generally occurs before the main clause event, and which does not require the participation of a conscious being.

In all purpose constructions the clauses occur in iconic order, with the purpose clause occurring finally. Tetun has six means of introducing purpose clauses, namely the conjunctions batu, ne‘ebé, bé and bodik, the verb hodi, and the irrealis auxiliary atu. While data on bé and bodik are scant, the other four can be shown to differ significantly in the measure of control attributed to the one who has the purpose. These differences are summed up in the following table. Note that 'pause before purpose clause' gives a very rough measure only, in that it represents the proportion of examples which were transcribed with a punctuation mark (usually comma or full stop) before the introductory word of the purpose clause.

<table>
<thead>
<tr>
<th></th>
<th>hodi</th>
<th>atu</th>
<th>bat(u)</th>
<th>ne‘ebé</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of examples</td>
<td>100</td>
<td>62</td>
<td>260</td>
<td>119</td>
</tr>
<tr>
<td>Different subject in purpose clause</td>
<td>0%</td>
<td>6%</td>
<td>53%</td>
<td>72%</td>
</tr>
<tr>
<td>Pause before purpose clause</td>
<td>36%</td>
<td>6%</td>
<td>30%</td>
<td>53%</td>
</tr>
<tr>
<td>Need purpose happen?</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

The linking verb hodi 'COORD' indicates the highest degree of control. This is indicated by the fact that it can only be used if the purpose actually happened, unless it is explicitly marked as irrealis. In addition, hodi is only used where both clauses share the same subject. One would expect that an actor has greater control over what he or she will do than over what someone else will do, with the result that having the same subject in the two clauses implies a greater degree of control (in general) than if the two clauses have different subjects. This construction has been discussed in §14.5.2.

Irrealis atu involves slightly less control, and no guarantee that the purpose happened. As an auxiliary it introduces purpose clauses only if these have no subject specified, and is discussed in §10.3.6. Note that atu can also occur within purpose clauses introduced by the conjunctions batu (14.84; 7 examples) and ne‘ebé (14.86; 13 examples).
The conjunction batu ‘so that’ (or bate, usually abbreviated to bat) involves less control by the subject of the main clause, as can be inferred from the fact that the subject of the main and purpose clauses are different in 53% of a subsample of 120 examples.  

so mother LOC face TAG so.that 3S 3S-know  
So, Ibu was in front. So that she would learn.  

(J3.61)

14.84 La'o mai bat ita atu hó malu hi'iit walk come so.that 1PI IRR accompany each other guess  
ai.kanoik. story  
Come here so that we’ll tell riddles together.  

(M0.34)

The conjunction ne’ebé ‘so that’ appears to involve still less control, as indicated by the fact that subjects of the main and purpose clauses have different referents in 72% of examples.

14.85 Nia n-akés ma-mamam [ne’ebé ha’u bele kohi].  
3S 3S-talk RDP-slow so.that 1S can catch  
He speaks slowly so that I can catch (it; i.e. understand).  

(D0.38)

14.86 Nia-kan kadun emi ko’a hodi, ne’ebé emi atu filak bá  
3S-POS udder 2P cut bring so.that 2P IRR substitute go  
bukae iha dalan.  
trailfood LOC road  
Cut off and bring its udders, so that you can use them for food on the way.  

(K10.85)

Nevertheless, batu and ne’ebé are very close in meaning. Note that both batu and ne’ebé can also introduce complements of verbs of ordering (§13.3.4.3).

The word bé ‘so that’ (11 examples, mostly from poetry and ritual language) may be a variant of ne’ebé.

14.87 M-anák [bé ha’u katak fó tone].  
2S-be.quiet so.that 1S tell give go (.to. you)  
Be quiet so I will tell you.  

(Y1.28)

As a conjunction bodik ‘for’ (6 examples) is in the corpus found mostly in elicited examples and ‘tortuous’ sentences. It is also a preposition, and is for convenience discussed and exemplified in the chapter on prepositions (§8.6.1).

---

19 The word batu apparently has restricted geographical distribution. It is listed as be atu (lit. ‘so.that IRR’) in Mathijsen’s dictionary for the Foho dialect (1906:9), and is omitted from the East Timorese dictionaries of Hull (1996b) and Morris (1984b).
14.88 ...soru tais wa'in, halo balu tais feto marobo,
weave cloth much make some cloth woman red dyed
[She ...] wove many cloths, making some red dyed women's cloths,
halo balu tais mane; (á: Nia ti'a nia)
make some cloth male PAUSE 3S already 3S
and making some men's cloths; (Stumble: Ah. Then she)—
bodik nia at hameti — at hatais.
for 3S IRR secure IRR wear cloth
for her to secure—to wear. (AA1.57)

14.8 Relative clauses

14.8.1 Introduction

Relative clauses are clauses which modify NPs. In Tetun these can either precede or
follow the head of the NP. Premodifying relative clauses have very restrictive syntax, and are
discussed in §7.4.4. This section therefore deals with postmodifying relative clauses only, in
particular with their internal structure, and with their relationship to the preceding head noun.
Most relative clauses are marked by an initial invariant subordinator mak (or ma'ak).20 Since clauses with and without mak have somewhat different possibilities, they are discussed separately.

14.8.2 Relative clause with mak

14.8.2.1 Clause structure

Subjects (14.89; 210 examples) and objects (14.90; 64 examples) are readily relativised;
in these cases a 'gap' strategy is used, whereby the relativised NP is missing from the
clause.21 Objects of serial verb constructions may be relativised just like other objects (14.97;
3 examples).

14.89 bei [mak mate kleur ti'an]
ancestor REL die long time already
ancestor(s) who died long ago (V0.161)

14.90 sá-sá [ma'ak nia fola-folan] ne'e sia
RDP-what REL 3S RDP-swallow whole this PL
the various things that she had swallowed (R6.49)

Where the non-relativised core NP in a transitive relative clause is left unspecified (leaving
a verb with no arguments), it appears to be only contextual knowledge that determines
whether the head noun is interpreted as subject or as object. In the following example, for
instance, the interest of the speaker is in who was appointed to head the committee, rather

---

20 The fact that it can co-occur with a resumptive pronoun in relativised locatives (14.92) indicates that mak is
a subordinator and not a relative pronoun. Note that this subordinator is closely related to the actor-describing prefix ma(k)- (§4.4).

21 This is consistent with Givón’s (1990:659) assertion that the ‘gapping’ strategy used in Tetun is “most
commonly found in languages with rigid word-order, where inferring the missing argument’s case-role is
presumably easier”.
Beyond the clause

than in who did the appointing, although either could in principle be described by this relative clause.

14.91  Ema [mak fote bá ulun] ne’e, nia-kan naran Ama N...
       person REL lift go head this 3S-POS name father N...
The person who was appointed to head, his name was Mr N....  (K1.7)

Relativisation of constituents other than subjects and objects is uncommon, and in the corpus is restricted to locatives (14.92; 4 examples). Here, instead of gapping, the complement of the locative preposition iha is a resumptive pronoun (ne’e ‘this’ or nia ‘3S, that’).22

14.92  ...sia boi n-ola tasi ida [mak sia atu n-akdiuk iha ne’e].
       3P choose 3-take sea one REL 3P IRR 3-play LOC this
       [Then] they chose a (part of the) sea in which they would play.  (K9.15)

Word order in relative clauses is restricted to SVO; that is, there can be no object fronting. Apart from this restriction on basic word order, and the need to be declarative, relative clauses have no other restrictions on their internal structure. They allow, for instance, the full range of aspectual modifiers (e.g. ti’an ‘already’ in 14.89, irrealis atu in 14.92), subject marking (14.92, 14.99), subordinate clauses, clause parataxis (14.94) and peripheral phrases. They also allow the full range of predicate types, including transitive (14.90) and intransitive (14.89), nominal (14.93), prepositional phrase, and clausal (i.e. body-good expressions; 14.94).

14.93  nia oan ohin [mak rán ne’e]
       3S child-GEN just.now REL blood this
       his child who used to be blood (Note: the child grew from a single clot of
       blood.)  (S2.119)

14.94  ema [mak ain át, matan át]
       person REL leg bad eye bad
       people who are crippled (or) blind  (I8/7)

14.8.2.2 Non-restrictive relative clause

The relative clauses illustrated above are restrictive relative clauses, which assist in identifying the referent of the head noun, and which account for the vast majority of relative clauses in the corpus. In contrast, non-restrictive relative clauses make a parenthetical assertion about the head noun (14.95, 14.97). They tend to supply information that is a habitually known aspect of the normal characterisation of the individual in question (14.96), and are non-contrastive. These non-restrictive relative clauses are sometimes preceded by a pause or the normally NP-final determiners ne’e ‘this’ or sia ‘PL’. They always start with mak. There appears to be no internal structural difference between non-restrictive relative clauses and restrictive relative clauses introduced by mak, and the pragmatic distinction is not always clear.

---

22 The Tetun pattern supports Keenan and Comrie’s (1977:66f.) generalisation that the lower the head noun is on the Accessibility Hierarchy, the more common it is to find it expressed within the relative clause by a personal pronoun. The hierarchy as presented in Keenan (1985:147f.) is as follows: Subject > Direct Object > Indirect Object > Object of adposition > Possessor. I have no data for relativisation on possessors.
14.95 ...ha\'u k-anoin ha\'u-kan ali-n, [mak dadi ba
1S 1S-think 1S-POS younger.sibling-GEN REL become go
manu ti\'an].
bird already
...I pity (\'think of\') my younger brother, who has turned into a bird. (K12.71)

14.96 B\'a futu manu no Lakuleik, Liurai Lakuleik [mak
go fight(.cocks) bird with Lakuleik ruler Lakuleik REL
futu manu, maka laka manu].
fight(.cocks) bird fight(.cocks) bird
(They) went cockfighting with Lakuleik, Ruler Lakuleik who fought cocks.
(G1.8)

14.97 ...su rahenek [mak tama ba].
dig.up sand REL enter go
[They slept in the sand. They] dug up the sand, which (they) went into.
(AA4.89)

14.8.2.3 Headless relative clause

Headless relative clauses are relative clauses with no preceding NP head (14 examples, in
addition to those in cleft constructions (§9.5.5)). They have the same structure as modifying
relative clauses introduced by mak.

14.98 Dadi, hela [mak adat kafir] e\'e iha rai
so leave REL tradition heathen[Mly] this LOC earth
ita-k e\'e.
1PI-POS this
So, that left those (people) who were heathen in this our land. (P1.123)

14.8.3 Relative clause without mak

14.8.3.1 Relativisation on subject and object

Relative clauses need not have the subordinator mak. Clauses without mak allow
relativisation on at least subject (14.99; 21 examples), object (14.100; 8 examples) and
object of prepositional verbs (14.101; 4 examples), and have the same word-order constraints
as relative clauses introduced by mak. Although they do allow modification (e.g. negation,
locatives), the corpus shows a strong preference textually for such clauses to be limited to the
verb, the argument NPs, and an initial ohin \‘just now\’.

14.99 Meo [ohin n-o\'o a mau-n Dele Malae] ne\'e...
hero just.now 3S-kill 2S elder.brother-GEN Dele non.native this
This hero (i.e. powerful warrior) who recently killed your elder brother Dele
Malae [has three red \‘tassels\’ on his forehead.] (AA4.138)
Beyond the clause

14.100 Fahi meak [ohin asu n-abu'a] ne'e, nia la n-aré
pig reddish just.now dog 3S-pursue this 3S not 3S-see
ti'an.
already
This reddish pig that the dog had just been pursuing, he (i.e. the dog's owner)
no longer saw. (L1.11)

14.101 ke'an [te'in bá]
room cook at
room to cook in (i.e. kitchen) (N0.132)

14.8.3.2 Time

In addition to relativising on a constituent, modifying clauses without mak allow for the head noun to be oras 'time', and the NP to refer to the time when that which is specified in the relative clause took place (34 examples). Within these modifying clauses a strict SVO word order is followed. Modifiers, such as locatives, are allowed. This construction appears to be used exclusively by those speakers (such as teachers) who are fluent in Indonesian, suggesting that it is modelled on Indonesian pada waktu 'at the time'. It thus makes up for the paucity of temporal conjunctions (particularly 'while') in Tetun.23

14.102 Diu ne'e nia n-ola bá oras [nia bá tama
saltwater.crab this 3S 3S-take at(time) time 3S go enter
raï iha alas].
earth LOC forest
This crab he obtained at a time he went hunting (lit. 'entering the land') in the forest. (L1.7)

14.8.3.3 Associated entity

There is one other class of modifying clause for which the head noun is not interpretable as coreferential with an argument in the clause. These are clauses which describe activities in which the denotatum of the head noun is typically used as instrument (14.103, 14.104), or is otherwise associated (14.105, 14.106, 14.107). Such noun-clause constructions function as conventionalised names. Relative clauses in this construction appear to be restricted to 'small clauses', whose internal structure is described in §9.7.3.

14.103 fatu [ledu]
stone mill
rollers to remove ('mill') seeds from raw cotton

14.104 nesun [fai akar]
mortar pound sago
mortar used to pound sago

---

23 According to Sasse (1993:680) such a noun plus relative clause construction is a common source for innovating substitutes for temporal conjunctions.
**Chapter 14**

14.105  *we*  [te’in]
         water  cook
         boiled water/water to use for cooking

14.106  *aikanoik*  [laran  moras]
         story  interior  sick
         sad story

14.107  *luli*  [hahoris]
         taboo  give.birth
         taboos associated with giving birth

**14.8.3.4 Comparison with attributive adjective phrases**

Given that relative clauses need not be explicitly marked as such by the presence of *mak*, there is a small class of complex stative attributes which can be analysed either as adjective phrases or as relative clauses without a subordinator. In practice the analytic difference is merely one of labelling of the modifier node. Such examples, which are uncommon, are illustrated below.

14.108  *Ema*  [wa’in  bót]  ne’e  la  r-aré.
         person  many  big  this  not  3P-see
         These very many people (or: ‘these people who were very many’) didn’t see
         (it).

14.109  *...ho’o fahi*  [mesa  bót  nú  ne’e].
         kill  pig  solely  big  like  this
         [For weddings and funerals] (they) killed pigs (which were) all big like this.

The last word in this grammar goes to *Bei* Tema, who knows the polite ways to finish a tale.

```
Lia  ha’u-k  n-akotu.  Ita  ha-badak  onan.
word  1S-POS  3S-finish  1PI  make-short  IMM
My story is concluded. We’ll cut (this discussion) short now.
```
Appendix A: Texts

1 Explanation: ‘How to make coconut oil’

The following explanation about how to make coconut oil (Text E2) was recorded after I came upon the family of Ama Bo’uk making some over a fire. The speaker is Ama Bo’uk, a man of about 60, who resides in the hamlet of Uma Katahan. Apart from replacing false starts by ‘(...)’, no corrections have been made to the text.

A1.1 Ha’u naran Am-Bo’uk, hosí Uma Katahan.  
1S name father-chubby from house Katahan  
My name is Am Bo’uk; (I am) from Uma Katahan.

A1.2 Ha’u atu hakés hosí te’in nú wén.  
1S IRR talk about cook coconut juice  
I am going to talk about cooking coconut oil (lit. ‘coconut juice’).

A1.3 Hahú, sa’e nú lai.  
start ascend coconut first  
To start, first pick (lit. ‘ascend, climb’) coconuts.

A1.4 Sa’e nú ti’a, hodi nú mai, baka.  
ascend coconut already bring coconut come split.coconut  
Having picked coconuts, bring the coconuts here, (and) split (them) in two.

A1.5 Baka ti’a, sukit.  
split.coconut already burrow  
Having split (them), prise out the meat.

A1.6 Sukit ti’a, hodi bá ledu iha motor.  
burrow already bring go mill LOC machine[Mly]  
Having prised out the meat, take (it) to mill (it) in the machine.

A1.7 Hodi hika mai, kumu.  
bring back come massage  
Bring (it) back here, and squeeze (it, to squeeze out the milk).

1 The noun wén is presumably derived from wé ‘water, liquid’ and the genitive marker -n; that is, it is liquid which derives from or is associated with something. It is glossed as ‘juice’ because the -n is retained even if the possessor is not stated, as in A1.11.
A1.8  *Kumu, kumu hotu, ta'es bá tasu.*

massage massage finish filter go wok

Squeeze, and having finished squeezing (them) all, filter (the milk through a

cloth) into a wok.

A1.9  *Ta'es bá tasu ti'a, (s,) tau ha'i bá. Te'in.*

filter go wok already HES put fire go cook

Having filtered (it) into a wok, add fire (under the wok). Cook (it).

A1.10  *Te'in, te'in dauk to'o n-akali, (á) loin ha'i.*

cook cook continue until 3S-bubble PAUSE remove heat fire

Cook (it), keep on cooking until (it) boils, (then) remove heat from the fire.

A1.11  *Loin ha'i ti'a, nia malirin, foin ti'i.*

remove heat fire already 3S cold then scoop liquid

Having reduced the heat of the fire, (when) it is cool, then scoop out the liquid

(i.e. the watery liquid which separates from the oil).

A1.12  *Ti'i ha-sai wén á,*

scoop liquid make exit juice DEF

Scoop out the liquid;

A1.13  *ti'i ha-sai ta wén ti'a,*

scoop liquid make exit already juice already

after scooping out the liquid,

A1.14  *foin tau ha'i hika bá, hodi te'in:*

then put fire back go COORD

then stoke up the fire again, and cook:

A1.15  *te'in hika nú wén ne'e.*

cook back coconut juice this

Cook the coconut oil again.

A1.16  *Te'in hika nú wén to'o tasa.*

cook back coconut juice until cooked

Cook the coconut oil again until (it) is cooked.

A1.17  *Tasa, titin mai, tau bá karon á.*

cooked put on ground come put go cloth filter DEF

(Once it is) cooked, place (the wok) on the ground, (and) put (the coconut oil

mixture) into a cloth filter.

A1.18  *(Á.) Tau bá karon ti'a, taka ti'a, habit.*

PAUSE put go cloth filter already cover already squeeze

Once it has been put in the cloth, and wrapped up, squeeze (it). (Note: The

cloth is twisted, and then squeezed between two pieces of wood which are tied

together at one end.)

A1.19  *Habit nú wén á,*

squeeze coconut juice DEF

Squeeze the coconut oil,
A1.20 nú wén (ə) turu mós, sai mós ti’a,
coconut juice HES drip finished exit finished already
(and when) the oil has finished dripping out, has finished coming out,
A1.21 ha-sai hika nú laran á iha karon á.
make-exit back coconut interior DEF LOC cloth.filter DEF
take out the solid dross (lit. ‘coconut inside’) from the cloth.
A1.22 (... Nú wén sai mós ti’an.
coconut juice exit finished already
The coconut oil has finished coming out.
A1.23 Dadi nú laran á ha-sai,
so coconut interior DEF make-exit
So the dross (you) remove,
A1.24 tau iha baskom á,
put LOC large.bowl[Mly] DEF
(and) put in a bowl.
A1.25 (... Nú wén á tau iha baskom á.
coconut juice DEF put LOC large.bowl DEF
The coconut oil (you) put in a (different) bowl.
A1.26 Ikus, foin (ə) sukat. Sukat tau bá botir á.
final then HES measure measure put go bottle[Mly] DEF
Finally, measure (it). Measure (it by) putting (it) in bottles.
A1.27 Atu te’in a.’isin kona botir hira?
IRR cook once touch bottle[Mly] how.many
How many bottles did cooking once produce?
A1.28 Nia ti’a balu te’in (...) hodi há,
3S already some cook COORD eat
After that some is cooked for eating,
A1.29 balu hodi bá fa’en, hola loit.
some bring go sell fetch money
some is taken to sell, (and) get money.

2 Folk tale: ‘Mr Pointed Head and Mr Pointed Bottom’

This popular story (Text K8) was told by Mr Hendrikus Dini, then aged almost 40, and
the transcription was subsequently checked with him on two separate occasions. The story
was carefully told, and is somewhat more explicit than is usual, in particular making little use
of zero anaphora. The text incorporates some corrections, noted in footnotes. The polite title
beì ‘grandparent, ancestor’ is glossed here with the shorter ‘Mr’.

A2.1 Bei Ulu Kmeik nó Bei Kidu Kmeik.
Mr head pointed and Mr bottom pointed
(Title:) Mr Pointed Head and Mr Pointed Bottom.
A2.2 Bá loron ida, Bei Ulu Kmeik nó Bei Kidu Kmeik, at(time) day one Mr head pointed and Mr bottom pointed One day, Mr Pointed Head and Mr Pointed Bottom,

A2.3 sia r-afaho iha to’os.
3P 3P-weed LOC garden they weeded in the garden.

A2.4 Sia r-afaho daudaun to’o loro manas, sia bá sa’e nú.
3P 3P-weed continue until sun hot 3P go ascend coconut They kept weeding until the sun was hot (i.e. until midday), (when) they went to pick (lit. ‘ascend’) coconut(s).

A2.5 Sia to’o nú hún bá, sia r-akduduk malu.
3P reach coconut trunk go 3P 3P-urge-ahead each other (When) they reached the coconut tree, they urged each other (to go up).

A2.6 Bei Ulu Kmeik na’ak, “Bei Kidu Kmeik, ó mak sa’e”.
Mr head pointed 3S-say Mr bottom pointed 2S REL ascend Mr Pointed Head said, “Mr Pointed Bottom, you go up”.

A2.7 Bei Kidu Kmeik katak bá Bei Ulu Kmeik n-a’ak, “Ó Mr bottom pointed tell to Mr head pointed 3S-say 2S mak sa’e”.
REL ascend Mr Pointed Bottom told Mr Pointed Head, “You are the one to go up”.

A2.8 Sia r-akduduk malu nú ne’e,
3P 3P-urge-ahead each other like this They urged each other like this,

A2.9 hotu ti’an, Bei Ulu Kmeik mak sa’e.
finish already Mr head pointed REL ascend and eventually it was Mr Pointed Head who went up.

A2.10 Bei Ulu Kmeik sa’e to’o nú leten bá; Mr head pointed ascend reach coconut top go Mr Pointed Head climbed up to the top of the coconut tree;

A2.11 nia sei dauk n-atetu, nia to’o nú wár,
3S still not 3S-perfect 3S reach coconut neck he had not yet reached the very top, he was up to the ‘neck’ (just below the fruit),

The speaker had said Ó sa’e uluk ‘you ascend first’, but later changed this because uluk implied that both would eventually climb the tree, whereas only one participant was to do so. Similarly in A2.7, the speaker had originally said Ó sa’e uluk. Ó mak sa’e, but requested that the initial clause be removed. I suspect the original text was fine, since uluk gave the implication that the speaker would go up later, which was strictly speaking false, but perhaps useful, nevertheless, in encouraging the friend to go up ‘first’.
(when) his head went into the coconut stalk (otherwise (called) the 'coconut thigh'). (Note: His head got stuck because it was pointy on top.)

So, he couldn't go up further,

(and) he couldn't descend again (lit. 'descended again not well').

He just dangled in the coconut spathe.

(and) he asked assistance of Mr Pig.

He called the pig, saying "Hey! Friend Mr Pig.

Do you love me or not?"
Mr Pig answered him saying “What’s up?”.

“Hey. You come and root up (i.e. dig up with your snout) and release my bottom, which has gone into the ground.”

Then Mr Pig answered him, “What will you give to me?”.

He said, “I’ll give you a garden full of cassava”.

Then Mr Pointed Bottom’s bottom was suddenly freed (i.e. he came back out of the ground).

He was very pleased.

3 During the second checking the speaker asked that N-a’ak ‘3S-say’ (which attracted no comment on the first checking) be prefaced by Hotu. Bei Kidu Kmeik simu ‘then Mr Pointed Bottom replied’. The text as it stands, however, reflects a common strategy in spoken discourse, in which the identity of the next speaker in a dialogue is often not explicitly stated. The preference for greater explicitness in a written text is not unexpected.

4 Originally the speaker used uninflected hikar ‘back’. He considers either hikar or inflected n-ikar ‘3S-back’ to be acceptable. This illustrates the fact that subject marking is generally considered to be preferred but optional (§9.3.3.3). The initial lale is a correction of Indonesian atau.
A2.34  
Nia ti’an, Bei Kidu Kmeik n-ó  Bei Fahi  
3S already Mr bottom pointed 3S-accompany Mr pig  
After this, Mr Pointed Bottom took Mr Pig

A2.35  
bá  n-atudu bá  ema ida  nia-kan to’os, mak nakonu  
go  3S-show go person one 3S-POS garden REL full  
nó  fehuk.  
with tuber  
and went and showed him someone’s garden, that was full of cassava.

A2.36  
(Á.)  Nia katak bá  Bei Fahi n-a’ak,  
PAUSE 3S tell to Mr pig 3S-say  
Hmm. He told Mr Pig,

A2.37  
“Ó m-á bá, fehuk ne’e,5  tán to’os ne’e ha’u-k”.  
2S 2S-eat IMP tuber this because garden this 1S-POS  
“You eat this cassava, because this garden is mine”.

A2.38  
Fahi ne’e, loron, kalan, bá  n-á  fehuk iha  ema  
pig this day night go 3S-eat tuber LOC person  
nia-kan to’os.  
3S-POS garden  
Day and night the pig went and ate cassava in the person’s garden.

A2.39  
(Á.)  Hotu ema to’os na’in mai n-afula,  
PAUSE finish person garden owner come 3S-spy.on  
Hmm. Then the garden owner came and spied on (him),

A2.40  
n-odi  diman;  
3S-bring spear  
bringing a spear;

A2.41  
bá  loron ida  ema to’os na’in  sona n-o’o  fahi ne’e.  
at(time) day one person garden owner pierce 3S-kill pig this  
one day the garden owner speared the pig to death.

A2.42  
Hotu, fahi mate.  
finish pig die  
So, the pig died (or ‘was dead’).

A2.43  
Hotu, Bei Ulu Kmeik, nia  at  n-anu  sá?  
finish Mr head pointed 3S IRR 3S-do.like what  
Meanwhile, Mr Pointed Head—what was he to do?

A2.44  
Haré  fali  Bei Kakae.6  
see in.turn Mr cockatoo  
(He) in turn saw (or ‘suddenly there was’) Mr Cockatoo.

---

5 The speaker had difficulty reading this transcription, in which fehuk ne’e ‘this cassava’ follows clause-final imperative bà and a comma. However, he concluded that it was acceptable. The difficulty presumably reflects the fact that right dislocation is a feature of spoken rather than written language (§14.2.3).

6 The speaker accepted this line without comment on one checking, but rejected it (particularly haré fali) the next time, preferring to substitute tekiteki ‘suddenly’. The meaning of haré fali is unclear, but it is used in several texts at a point where there is a sudden change in the situation.
A2.45 *Manu kakae ida mai rani iha nú leten.*

bird cockatoo one come perch LOC coconut top
A cockatoo came and perched in the top of the coconut tree.

A2.46 *Hotu Bei Ulu Kmeik katak bá Bei Kakae n-a’ak?*

finish Mr head pointed tell to Mr cockatoo 3S-say
Then Mr Pointed Head said to Mr Cockatoo

A2.47 “Ou, belu Bei Kakae, ó m-adomi ha’u lai.

EXCL friend Mr cockatoo 2S 2S-love 1S first
“Hey! friend Mr Cockatoo, have mercy on me now.

A2.48 Ó tulun ha’u lai”.

2S help 1S first
Help me now”.

A2.49 “Tán sá?”

because what
“Why?”

A2.50 “Ai! Ó mai, m-á sít nú kain ne’e lai.

EXCL 2S come 2S-eat be.cut.off coconut stalk this first
“Hey! You come, and chew through this coconut stalk now.

A2.51 Tán ha’u-kan ulu-n kmeik tama iha nú kain.

because 1S-POS head-GEN pointed enter LOC coconut stalk
Because my pointed head is stuck in (lit. ‘entered’) the coconut stalk.

A2.52 Ha’u sa’e la di’ak, tún la di’ak.”

1S ascend not good descend not good
I can’t go up, (and) can’t go down (lit. ‘go down not well’).”

A2.53 *Hotu, Bei Kakae katak bá nia n-a’ak,*

finish Mr cockatoo tell to 3S 3S-say
Then Mr Cockatoo said to him,

A2.54 “Hotu ó fó sá bá ha’u?”. finish 2S give what to 1S
“In the end what will you give to me?”.

A2.55 *N-a’ak, “Ha’u fó batar to’os ida bá ó”*. 3S-say 1S give maize garden one to 2S
(He) said, “I’ll give you a garden full of maize (corn)”.

A2.56 *Nia, Bei Kakae mai, rani bá nú kain; nia n-á.* then Mr cockatoo come perch go coconut stalk 3S 3S-eat
Then Mr Cockatoo came, and landed on the coconut stalk; he chewed (or ‘ate’).

---

7 This sentence originally lacked the introductory *hotu* ‘finish’, the title *Bei* before *Kakae* ‘cockatoo’, and *n-a’ak* ‘3S-say’. While none are necessary, the speaker requested during the second checking that they be added. Note that the connective adverb *hotu* ‘after that; finish’ is in this text used rather more frequently than is usual; it seems that different speakers major on different connective expressions, with some, for instance, repeatedly using *nia ti’a* ‘after that’, others *nia ma* ‘then’, and others *hotu* ‘after that; finish’. 
A2.57  Nia n-á, n-á, n-á,  
3S 3S-eat 3S-eat 3S-eat  
He chewed and chewed and chewed,

A2.58  n-á daudaun to'o nú kain sít.  
3S-eat continue until coconut stalk be.cut.off  
and kept chewing until the coconut stalk was broken through.

A2.59  Bei Ulu Kmeik nia-kan ulu-n sai n-ikar.  
Mr head pointed 3S-POS head-GEN exit 3S-back  
Mr Pointed Head's head was released again.

A2.60  Hotu, nia tún dí'a-di'ak,  
finish 3S descend RDP-good  
Then he came down well (from the top of the tree);

A2.61  mai n-ikar n-o'i rai mai.  
come 3S-back 3S-go.to earth come  
(he) came back down to the ground.

A2.62  Tún to'o rai mai, nia bolu Bei Kakae  
descend reach earth come 3S call Mr cockatoo  
Having got down to the ground, he called Mr Cockatoo

A2.63  “Mai. Ita la'o onan, belu.  
come 1PI walk IMM friend  
“Come. We'll go now, friend.

A2.64  Ita bá haré batar iha to'os”.  
1PI go see maize LOC garden  
We'll go and see the corn in the garden”.

A2.65  Bei Ulu Kmeik n-ó Bei Kakae,  
Mr head pointed 3S-accompany Mr cockatoo  
Mr Pointed Head took Mr Cockatoo,

A2.66  bá n-atudu Bei Kakae bá ema ida nia-kan to'os,  
go 3S-show Mr cockatoo go person one 3S-POS garden  
REL full with maize

and went and showed Mr Cockatoo someone’s garden, that was full of maize.

A2.67  (Á.)  “Ó m-á batar ne'e. Tán to'os ne'e ha'u-k.”  
PAUSE 2S 2S-eat maize this because garden this 1S-POS  
Hmm. “You eat this maize. Because this garden is mine.”

A2.68  Bei Kakae loron bá n-á batar, loron bá n-á batar.  
Mr cockatoo day go 3S-eat maize day go 3S-eat maize  
Daily Mr Cockatoo went and ate maize; daily (he) went and ate maize.

A2.69  To'o loron ida, ema to'os mai n-afula Bei Kakae  
until day one person garden come 3S-spy.on Mr cockatoo  
Until one day the person from the garden (i.e. the owner) came and spied on  
Mr Cockatoo,
A2.70 n-odi kahúk.
3S-bring blowgun
bringing a blowpipe.

A2.71 Loran ida Bei Kakae n-á n-odi nó lian;
day one Mr cockatoo 3S-eat 3S-bring also voice
One day Mr Cockatoo was eating noisily;

A2.72 tekitekis kahúk botu,
suddenly blowgun bang
suddenly a blowpipe made a banging noise,

A2.73 Bei Kakae nó kahúk isin (oil) toba nebá.
Mr cockatoo and blowgun contents EXCL lie down there
(and) Mr Cockatoo and the blowpipe ammunition (wow!) (came to) lie down
way over there.

A2.74 Mate kedan.
die immediately
(He) died straight away.

A2.75 Mós ti’an.
finished already
The end.
Appendix B: List of texts

The following table lists the 93 Tetun texts on which much of this study is based. Many other texts were recorded and transcribed, or were generously given to me in written form, but were not entered into the computer as whole texts. These contributed much to my language learning; however, since they were not in a form to be readily searched by computer, and did not contribute to example counts in this study, they are not listed here.

Where a speaker's full name is listed but I knew him or her only by a nickname, the latter is included in brackets. The third column gives the speaker's sex, while the fourth column (headed 'So') lists the source of the text. Where blank, it indicates that the text was recorded and later transcribed; 'W' shows that the text was presented to me in written form, while 'D' marks a text which was dictated to me.

The fifth column (headed 'Dial') lists the dialect area for the speaker. Where two are listed, it indicates that the speaker grew up in the first area, but now lives in the second. The abbreviations for the two variants of the Fehan dialect are 'W' for Wehali (the main variant on which this grammar focuses), and 'B' for Besikama (whose speech differs very little from that of Wehali). 'S' stands for the sub-dialect of Suai, while 'F' stands for the separate dialect of Foho.

<table>
<thead>
<tr>
<th>File</th>
<th>Speaker</th>
<th>Sex</th>
<th>So</th>
<th>Dial</th>
<th>Topic</th>
<th>Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA1</td>
<td>Ama Bo'uk</td>
<td>M</td>
<td>W</td>
<td>W</td>
<td>Folktales: Buku Lasak and Bita Nahak</td>
<td>1,231</td>
</tr>
<tr>
<td>AA2</td>
<td>Ama Bo'uk</td>
<td>M</td>
<td>W</td>
<td>F</td>
<td>Recent court case regarding sex before marriage</td>
<td>616</td>
</tr>
<tr>
<td>AA3</td>
<td>Hendrikus Dini</td>
<td>M</td>
<td>W</td>
<td>W</td>
<td>How to make cement bricks</td>
<td>490</td>
</tr>
<tr>
<td>AA4</td>
<td>Stanis Leto</td>
<td>M</td>
<td>S</td>
<td>F</td>
<td>Myth: Origin of fish nets</td>
<td>1,766</td>
</tr>
<tr>
<td>B1</td>
<td>Benedictus Bere Seran</td>
<td>M</td>
<td>F</td>
<td>F</td>
<td>Myth: Origin of Kateri</td>
<td>488</td>
</tr>
<tr>
<td>B2</td>
<td>Ama Bo'uk</td>
<td>M</td>
<td>W</td>
<td>W</td>
<td>Hatama Batar Mana'ik ceremony</td>
<td>474</td>
</tr>
<tr>
<td>C1</td>
<td>Kmi Laran Protestant church</td>
<td>M</td>
<td>B</td>
<td>B</td>
<td>Sermon: Fear not. God is with you</td>
<td>2,970</td>
</tr>
<tr>
<td>C2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td>Ama Bo'uk</td>
<td>M</td>
<td>W</td>
<td>W</td>
<td>Hatama Batar Mana'ik ceremony</td>
<td>355</td>
</tr>
<tr>
<td>D1</td>
<td>Emerensi Ana Telik Tefa (Mama Funan)</td>
<td>F</td>
<td>B/W</td>
<td>B/W</td>
<td>Folktales: Sikori</td>
<td>447</td>
</tr>
<tr>
<td>E1</td>
<td>Silvester Klau (Ama Klau Ki'ik)</td>
<td>M</td>
<td>W</td>
<td>W</td>
<td>Venus and stars</td>
<td>338</td>
</tr>
<tr>
<td>E2</td>
<td>Ama Bo'uk</td>
<td>M</td>
<td>W</td>
<td>W</td>
<td>How to make coconut oil</td>
<td>197</td>
</tr>
<tr>
<td>E3</td>
<td>Liman Bata</td>
<td>M</td>
<td>W</td>
<td>W</td>
<td>Mediators in courtship</td>
<td>300</td>
</tr>
<tr>
<td>F1</td>
<td>Alfonso Lala</td>
<td>M</td>
<td>F/W</td>
<td>F/W</td>
<td>Engagement traditions</td>
<td>414</td>
</tr>
<tr>
<td>F2</td>
<td>Mikhail Bere</td>
<td>M</td>
<td>W/S</td>
<td>W/S</td>
<td>Myth: Ruler Suri Junior</td>
<td>1,390</td>
</tr>
<tr>
<td>F3</td>
<td>Maria Senensis Telik (Bi Ilku)</td>
<td>F</td>
<td>W</td>
<td>W</td>
<td>My two suitors</td>
<td>519</td>
</tr>
<tr>
<td>File</td>
<td>Speaker</td>
<td>Sex</td>
<td>So</td>
<td>Dial</td>
<td>Topic</td>
<td>Words</td>
</tr>
<tr>
<td>------</td>
<td>-------------------------------</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>-----------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>F4</td>
<td><em>Ama Bo'uk</em></td>
<td>M</td>
<td>W</td>
<td>W</td>
<td>Phases of the moon</td>
<td>372</td>
</tr>
<tr>
<td>F5</td>
<td>Juliana Se'uk Seran</td>
<td>F</td>
<td>W</td>
<td>W</td>
<td>Folktale: The candle and the dirt clod</td>
<td>120</td>
</tr>
<tr>
<td>F6</td>
<td>Taek Tetik</td>
<td>M</td>
<td>B</td>
<td></td>
<td>Myth: Ruler Malaka</td>
<td>468</td>
</tr>
<tr>
<td>F7</td>
<td><em>Ama Bo'uk</em></td>
<td>M</td>
<td>W</td>
<td></td>
<td>Directions to Fahiluka</td>
<td>90</td>
</tr>
<tr>
<td>F8</td>
<td>Alfonsius Klau</td>
<td>M</td>
<td>B/W</td>
<td></td>
<td>Why Fehan people are poor</td>
<td>519</td>
</tr>
<tr>
<td>G1</td>
<td>Yohanes Seran Kehik</td>
<td>M</td>
<td>W</td>
<td></td>
<td>Myth: Origins of the noble house of Uma Nén</td>
<td>2,381</td>
</tr>
<tr>
<td>G2</td>
<td>Mikhael Bere</td>
<td>M</td>
<td>W</td>
<td>W/S</td>
<td>Myth: Suri Junior of Wehali</td>
<td>1,301</td>
</tr>
<tr>
<td>G3</td>
<td></td>
<td>W</td>
<td>W/S</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G4</td>
<td>(Ama Bo'uk)</td>
<td>M</td>
<td>W</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G5</td>
<td>Mikhael Bere</td>
<td>M</td>
<td>W</td>
<td>W/S</td>
<td>Myth: Suri Senior of Wehali</td>
<td>160</td>
</tr>
<tr>
<td>H1</td>
<td>Alfonsius Klau</td>
<td>M</td>
<td>B/W</td>
<td></td>
<td>Recent theft of a buffalo</td>
<td>363</td>
</tr>
<tr>
<td>J1</td>
<td>Juliana Se'uk Seran</td>
<td>F</td>
<td>W</td>
<td></td>
<td>Folktale: The goat and the rat</td>
<td>332</td>
</tr>
<tr>
<td>J2</td>
<td>Bui Mea</td>
<td>F</td>
<td>W</td>
<td></td>
<td>Women's work</td>
<td>411</td>
</tr>
<tr>
<td>J3</td>
<td>Bau Bete Ulun</td>
<td>F</td>
<td>W</td>
<td></td>
<td>Courtship singing, birth, baby feeding</td>
<td>878</td>
</tr>
<tr>
<td>J4</td>
<td>Alfonsius Klau</td>
<td>M</td>
<td>W</td>
<td>B/W</td>
<td>A recent dispute</td>
<td>602</td>
</tr>
<tr>
<td>J5</td>
<td><em>Na'i</em> Marianus Man</td>
<td>M</td>
<td>W</td>
<td></td>
<td>How to make steamed cassava</td>
<td>145</td>
</tr>
<tr>
<td>J6</td>
<td><em>Na'i</em> Marianus Man</td>
<td>M</td>
<td>W</td>
<td></td>
<td>The role of Uma Tabuta house</td>
<td>123</td>
</tr>
<tr>
<td>K1</td>
<td>Alfonsius Klau</td>
<td>M</td>
<td>W</td>
<td>B/W</td>
<td>Recent ordination of priests</td>
<td>503</td>
</tr>
<tr>
<td>K11</td>
<td>Hendrikus Dini</td>
<td>M</td>
<td>W</td>
<td></td>
<td>How to hoe</td>
<td>100</td>
</tr>
<tr>
<td>K12</td>
<td>Hendrikus Dini</td>
<td>M</td>
<td>W</td>
<td></td>
<td>Folktale: The crow</td>
<td>855</td>
</tr>
<tr>
<td>K13</td>
<td>Rika Lon</td>
<td>F</td>
<td>W</td>
<td>W/S</td>
<td>How to play the game <em>ai kalár</em></td>
<td>109</td>
</tr>
<tr>
<td>K14</td>
<td>Alfonsius Klau</td>
<td>M</td>
<td>B/W</td>
<td></td>
<td>Christianity and tradition</td>
<td>497</td>
</tr>
<tr>
<td>K2</td>
<td><em>Ama Bo'uk</em></td>
<td>M</td>
<td>W</td>
<td></td>
<td><em>Kôk</em> fever</td>
<td>210</td>
</tr>
<tr>
<td>K3</td>
<td><em>Ama Bo'uk</em></td>
<td>M</td>
<td>W</td>
<td></td>
<td>Healing injuries in the sacred house</td>
<td>249</td>
</tr>
<tr>
<td>K4</td>
<td>Alfonsius Klau</td>
<td>M</td>
<td>W</td>
<td>B/W</td>
<td>How to make sago pancakes</td>
<td>259</td>
</tr>
<tr>
<td>K5</td>
<td><em>Bei</em> Lalak</td>
<td>M</td>
<td>F/W</td>
<td></td>
<td>Treating illness in the sacred house</td>
<td>193</td>
</tr>
<tr>
<td>K6</td>
<td>Maria Senensis Telik</td>
<td>F</td>
<td>W</td>
<td></td>
<td>Description of a photo</td>
<td>180</td>
</tr>
<tr>
<td>K7</td>
<td>Hendrikus Dini</td>
<td>M</td>
<td>W</td>
<td></td>
<td>Description of a photo</td>
<td>190</td>
</tr>
<tr>
<td>K8</td>
<td>Hendrikus Dini</td>
<td>M</td>
<td>W</td>
<td></td>
<td>Folktale: Mr Pointed Head and Mr Pointed Bottom</td>
<td>540</td>
</tr>
<tr>
<td>K9</td>
<td>Mikhael Bere</td>
<td>M</td>
<td>W/S</td>
<td></td>
<td>Myth: Ruler Suri Senior</td>
<td>2,774</td>
</tr>
<tr>
<td>K10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L1</td>
<td>Alfonsius Klau</td>
<td>M</td>
<td>D</td>
<td>B/W</td>
<td>How <em>Bei Ulu</em> got supernatural power</td>
<td>154</td>
</tr>
<tr>
<td>L2</td>
<td>Alfonsius Klau</td>
<td>M</td>
<td>D</td>
<td>B/W</td>
<td>My mother's sickness and death</td>
<td>47</td>
</tr>
<tr>
<td>L3</td>
<td>Alfonsius Klau</td>
<td>M</td>
<td>D</td>
<td>B/W</td>
<td>Dividing a fish catch</td>
<td>78</td>
</tr>
<tr>
<td>O1</td>
<td>Emerensi Ana Telik</td>
<td>F</td>
<td>B/W</td>
<td></td>
<td>Folktale: Bui Hrik</td>
<td>2,085</td>
</tr>
<tr>
<td>O2</td>
<td><em>Tefa</em> (Mama Funan)</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O3</td>
<td>Maria Senensis Telik</td>
<td>F</td>
<td>W</td>
<td></td>
<td>Sunrise and moonrise</td>
<td>100</td>
</tr>
<tr>
<td>O4</td>
<td>Silvester Klau</td>
<td>M</td>
<td>W</td>
<td></td>
<td>Folktale: The orphan boy</td>
<td>602</td>
</tr>
<tr>
<td>File</td>
<td>Speaker</td>
<td>Sex</td>
<td>So</td>
<td>Dial</td>
<td>Topic</td>
<td>Words</td>
</tr>
<tr>
<td>------</td>
<td>---------</td>
<td>-----</td>
<td>----</td>
<td>------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>O5</td>
<td>Ina Bui Ledo, Ina Bete Paska, Ina Wa'ik</td>
<td>F</td>
<td>W</td>
<td>Courtship and marriage; girls and schooling</td>
<td>1,849</td>
<td></td>
</tr>
<tr>
<td>O6</td>
<td>Emerensi Ana Telik Tefa (Mama Funan)</td>
<td>F</td>
<td>B/W</td>
<td>Folktale: Bui Feto Fatara</td>
<td>1,207</td>
<td></td>
</tr>
<tr>
<td>O7</td>
<td>Emerensi Ana Telik Tefa (Mama Funan)</td>
<td>F</td>
<td>B/W</td>
<td>Folktale: The tamarind man</td>
<td>847</td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>Yohanes Seran Kehik (Ama Bo'uk)</td>
<td>M</td>
<td>W</td>
<td>Myth: Origins of the earth</td>
<td>1,337</td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>Yohanes Seran Kehik (Ama Bo'uk)</td>
<td>M</td>
<td>W</td>
<td>Answering questions on traditions</td>
<td>1,917</td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td>Ina Bete Ki'ik</td>
<td>F</td>
<td>W</td>
<td>Folktale: A man weeds</td>
<td>636</td>
<td></td>
</tr>
<tr>
<td>P4</td>
<td>Ina Ho'ar Bete</td>
<td>F</td>
<td>W</td>
<td>The sacred house of Uma Makta'en</td>
<td>1,211</td>
<td></td>
</tr>
<tr>
<td>R1</td>
<td>Ina Bete Ki'ik</td>
<td>F</td>
<td>W</td>
<td>Recent building of our sacred house</td>
<td>872</td>
<td></td>
</tr>
<tr>
<td>R2</td>
<td>Kornelis Taek Leon</td>
<td>M</td>
<td>W</td>
<td>Felling sago</td>
<td>241</td>
<td></td>
</tr>
<tr>
<td>R3</td>
<td>Ama Bo'uk</td>
<td>M</td>
<td>W</td>
<td>How to play the game tila kmí</td>
<td>189</td>
<td></td>
</tr>
<tr>
<td>R4</td>
<td>Ina Bete Ki'ik</td>
<td>F</td>
<td>W</td>
<td>Folktale: The orphan girl</td>
<td>3,956</td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td>Josep Seran Manek</td>
<td>M</td>
<td>W</td>
<td>Hunting</td>
<td>257</td>
<td></td>
</tr>
<tr>
<td>S2</td>
<td>Yohanes Seran Kehik (Ama Bo'uk)</td>
<td>M</td>
<td>W</td>
<td>Myth: Origins of the noble house of Uma Nén (continued from G4)</td>
<td>2,340</td>
<td></td>
</tr>
<tr>
<td>TT1</td>
<td>Piet Tahu Nahak</td>
<td>M</td>
<td>W</td>
<td>Myth: Origins of the earth</td>
<td>333</td>
<td></td>
</tr>
<tr>
<td>TT6</td>
<td>P. Seran Luan</td>
<td>M</td>
<td>W</td>
<td>Myth: Origins of the earth</td>
<td>932</td>
<td></td>
</tr>
<tr>
<td>U1</td>
<td>Oliva Ho'ar Lotu</td>
<td>F</td>
<td>W</td>
<td>Folktale: Sun’s daughter Ae Lorok</td>
<td>828</td>
<td></td>
</tr>
<tr>
<td>U10</td>
<td>Josep Seran Manek</td>
<td>M</td>
<td>W</td>
<td>Folktale: Lük Morin</td>
<td>1,115</td>
<td></td>
</tr>
<tr>
<td>U11</td>
<td>Oliva Ho'ar Lotu</td>
<td>F</td>
<td>W</td>
<td>Folktale: The sack man</td>
<td>634</td>
<td></td>
</tr>
<tr>
<td>U2</td>
<td>Maria Senensis Telik (Bi Iku)</td>
<td>F</td>
<td>W</td>
<td>Folktale: Mr Turtle and Mr Monkey</td>
<td>464</td>
<td></td>
</tr>
<tr>
<td>U3</td>
<td>Maria Senensis Telik (Bi Iku)</td>
<td>F</td>
<td>W</td>
<td>Folktale: Mr Trickster</td>
<td>164</td>
<td></td>
</tr>
<tr>
<td>U4</td>
<td>Oliva Ho'ar Lotu</td>
<td>F</td>
<td>W</td>
<td>Folktale: The crooked tamarind</td>
<td>1,090</td>
<td></td>
</tr>
<tr>
<td>U5</td>
<td>Josep Seran Manek</td>
<td>M</td>
<td>W</td>
<td>Folktale: Ruler Half-Side</td>
<td>1,066</td>
<td></td>
</tr>
<tr>
<td>U6</td>
<td>Josep Seran Manek</td>
<td>M</td>
<td>W</td>
<td>Folktale: The witch</td>
<td>490</td>
<td></td>
</tr>
<tr>
<td>U7</td>
<td>Ina Bete Ki'ik</td>
<td>F</td>
<td>W</td>
<td>Working the gardens</td>
<td>209</td>
<td></td>
</tr>
<tr>
<td>U8</td>
<td>Fransisca Abuk Nahak (Ina Fouk)</td>
<td>F</td>
<td>W</td>
<td>Growing and processing mung beans</td>
<td>688</td>
<td></td>
</tr>
<tr>
<td>U9</td>
<td>Juliana Abuk Hale</td>
<td>F</td>
<td>S/W</td>
<td>Folktale: Mr Pointed Head and Mr Pointed Bottom</td>
<td>161</td>
<td></td>
</tr>
<tr>
<td>Y1</td>
<td>Emerensi Ana Telik Tefa (Mama Funan)</td>
<td>F</td>
<td>B/W</td>
<td>The death of Naha Bauk’s father</td>
<td>411</td>
<td></td>
</tr>
<tr>
<td>Z1</td>
<td>Ama Bo'uk</td>
<td>M</td>
<td>W</td>
<td>Recent outing to Bo'as sacred house</td>
<td>457</td>
<td></td>
</tr>
<tr>
<td>Z2</td>
<td>Teresia Abuk Lo'u</td>
<td>F</td>
<td>W</td>
<td>Folktale: The kitten</td>
<td>1,607</td>
<td></td>
</tr>
<tr>
<td>Z3</td>
<td>Teresia Abuk Lo'u</td>
<td>F</td>
<td>W</td>
<td>Folktale: Leki Àt</td>
<td>3,230</td>
<td></td>
</tr>
</tbody>
</table>
Appendix B

<table>
<thead>
<tr>
<th>File</th>
<th>Speaker</th>
<th>Sex</th>
<th>So</th>
<th>Dial</th>
<th>Topic</th>
<th>Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z4</td>
<td>Teresia Abuk Lo’u</td>
<td>F</td>
<td>W</td>
<td></td>
<td>Folktale: The wild and the domesticated aubergine</td>
<td>1,507</td>
</tr>
<tr>
<td>Z5</td>
<td>Oliva Ho’ar Lotu</td>
<td>F</td>
<td>W</td>
<td></td>
<td>Folktale: Sarata Lakuk</td>
<td>1,184</td>
</tr>
<tr>
<td>Z6</td>
<td>Oliva Ho’ar Lotu</td>
<td>F</td>
<td>W</td>
<td></td>
<td>Folktale: Mesak Oan</td>
<td>926</td>
</tr>
<tr>
<td></td>
<td>TOTAL TEXTS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>64,358</td>
</tr>
</tbody>
</table>

In addition to the above texts, the computer database includes 35,400 words of examples which were entered into the computer without context. These include examples noted ‘on the run’, selected portions of texts which were otherwise not entered into the computer, and results of elicitation. The ‘text’ names for these all end in ‘0’ (e.g. M0.1 = example from page 1 of book M), or in a date (e.g. C1/8 = example recorded in booklet C on 1 August).
References


Bartkowiak, Tadeusz, 1979, Gems from the island of Timor. Ende, Flores: Nusa Indah.


da Silva, Manuel Maria Alves, 1900, Noções de Gramática Galoli, dialecto de Timor [Fundamentals of the grammar of Galoli, a dialect of Timor]. Macao: Typographia do Seminário de S. José.

das Dores, Raphael, 1907, Dicionario Teto-Portugues [Tetun-Portuguese dictionary]. Lisbon: Imprensa Nacional.

de Araújo e Corte-Real, Benjamin, 1990, A contrastive analysis between Tetun and English consonants—a preliminary study of some phonological features of both languages. Sarjana Pendidikan Bahasa Inggris, Keguruan Ilmu Pendidikan, Universitas Kristen Satya Wacana.


References


References

Foley, William A. and Mike Olson, 1985, Clausehood and verb serialization. In Johanna Nichols and Anthony C. Woodbury, eds Grammar inside and outside the clause: some approaches to theory from the field, 17–60. Cambridge: Cambridge University Press.


1982a, The Great Lord rests at the centre: the paradox of powerlessness in European-Timorese relations. Canberra Anthropology 5/2:22–33.


Handbook of Malay and Tetum, 1945. HQ NT Force.


1991b, Some remarks on the grammaticalization of classifiers in Western Austronesian, especially Malay. *VI conference on Austronesian linguistics, Honolulu*.


Keenan, Edward L. and Robert D. Hull, 1973, The logical syntax of direct and indirect questions. In Claudia Corum, T. Cedric Smith-Stark and Ann Weiser, eds You take the high node and I’ll take the low node: papers from the Comparative Syntax Festival:
the differences between main and subordinate clauses, 348–371. Chicago: Chicago Linguistic Society.


1915, *Eenige fabels en volkslegenden van de onderafdeeling Beloe op het eiland Timor [Some fables and legends from Belu on the island of Timor]*. Batavia: Albrecht.


McCarthy, John J. and Alan S. Prince, 1986, Prosodic morphology. MS, University of Massachusetts and Brandeis.


References


Ross, Malcolm D., 1985, Current use and expansion of Tok Pisin: effects of Tok Pisin on some vernacular languages. In S.A. Wurm and P. Mühlhäuser, eds Handbook of Tok Pisin (New Guinea Pidgin), 539–556. PL, C-70.

1988, Proto Oceanic and the Austronesian languages of western Melanesia. PL, C-98.


1976b, Pelajaran Bahasa Indonesia-Tetun untuk rakyat Timor-Timur [Indonesian language lessons for the population of East Timor]. Bandung: Ganaco.

Sherlock, Kevin, 1980, A bibliography of Timor including East (formerly Portuguese) Timor, West (formerly Dutch) Timor and the island of Roti. Canberra: Research School of Pacific Studies, The Australian National University.


350 References


Index

'body-good' expressions, 196–204
á definite article, 121
actor-describing verb, 70–78
addressee, 268
adjective, 52–53
  as 'possessor', 152
  as head of adjective phrase, 156
derived, 85–92, 97
full reduplication, 94
numeric, 101, 157
subject marking, 176
adjective phrase, 156, 322
  as complement, 183
  as manner modifier, 242
  as predicate, 179
distribution in NP, 132, 136
adverb, 55, 227–49
  as 'possessor', 152
derived by full reduplication, 92
adverbal clause, 246, 311–18
alienability, 145–48, 150–52
án reflexive, 232
anaphora, 106, 126
animal calls, 57, 69
apposition, 158
arithmetic, 102
aspect, 218–23, 235–40
  in serial verb construction, 251
atu irrealis, 220, 316
auxiliary, 55, 214–26, 263
bdá, 271
  'at (location)', 269
  'at (time)', 271
  'for (beneficiary, purpose)', 270
  'go', 262
  'so be it', 245
  'to (addressee)', 268
  'to (factive result)', 270
  'to (recipient)', 268
déctic particle, 243
imperative, 244
ba- an alleged prefix, 97
basuk 'very', 231
batt(u) 'so that', 285, 316
bè 'so that', 316
bele 'can', 224
beneficiary, 166, 270
Besikama variety of Tetun, 3
bodik 'for', 166, 316
bot 'very, big', 231
causative, 59–66, 257–58, 270, 281
clause, 169
dependent, 204–6
nominalised by ne'e, 310
non-declarative, 208–12
periphery, 206–8
cleft construction, 183, 194–96
clitic, 18, 19
  morphophonemics, 43
close relation, 146
coda consonant, 18
comparative, 232
complementation, 277–93
complementiser, 283
complement-taking verb, 214–16, 277
  kanoin 'I think' as adverb, 249
compound, 45–47, 84
concessive clause, 313
conditional clause, 304, 312–13
conjunction, coordinating, 56, 156, 301–3
conjunction, subordinating, 55, 160,
  163–64, 165–67, 311–18
connective adverb, 245
consonant
  /h/-insertion rule, 43
  adaptation of loans, 27
  clusters
    /k/-initial, 31–40, 44, 97
    in contractions and derivations, 31,
      44, 46, 48
  frequencies, 41
  glottal stop insertion rule, 47
  phonotactics, 18, 29
  variation, 41
  coordinate-dependent construction,
    305–10
coordination, 156, 301–3
dual pronoun construction, 157
copula, 54, 182–84
dadi 'become', 183
dau daun 'continue', 247

definite article, 121
definiteness, 107, 152, 170, 180, 187, 193
degree modifier, 231
dectic particle, 243
demonstrative, 115–20
deontic modal auxiliary, 223–26
determiner, 55, 110–31
distribution in NP, 131, 132, 136–38
dialects, 3
Dili Tetun, 3, 4, 15, 18, 24, 25, 27, 29, 36, 37, 41, 124, 147, 173, 268, 306
direction verb, 258, 265
ditransitive verb, 181
dual pronoun construction, 157
é tag, 299
East Tetun, 3, 172, 301, 306
East Timorese Tetun, 3, 4, 18, 29, 36, 100, 122, 145, 173, 190, 230, 238, 248, 302, 313, 317
ellipse of subject and object, 170
ema 'person', 135
epistemic modality, 247
equative clause, 193
exclamations, 56
existential clause, 186–92
f-a, sa- derives finger and toe names, 96
filler, 56
Foho dialect, 3, 4, 6, 36, 122, 172, 175, 176, 230, 236, 248, 300, 317
foin
‘only just’, 219
‘then’, 177, 240, 246
foot, metrical, 14
full reduplication, 93–95
morphophonemics, 45–47
generic NP, 134, 135, 224
goal (location), 161, 266
grammaticisation, 123–25, 159–60, 163, 191, 202, 214, 243, 246, 249, 250, 263–75, 284
greeting, 213
ha- 'make' derives verbs, 59–66
ha 'ak 'say', 308
ha 'i 'not', 228
hak- (-k) derives verbs, 66–70
hak- a developing subject marker?, 177
halo 'make', 242, 257, 261, 270
ham- East Timorese prefix, 62
hamatu, hamatuk 'together', 275
hanesan 'like, resemble', 166
hesitation marker, 56
hetak 'as for', 167
hika(r) 'back', 176, 241
hira 'how many', 101
hó
'accompany', 272
'have', 187, 188
ho 'i progressive, 218
hodi
'bring, use', 273
'by (time)', 275
'COORD', 306, 316
hola 'take', 260
hori
'from (location)', 162
'since (time)', 163
hosi
‘about’, 283
‘from’, 267
hotu
‘all’, 120
‘finish(ed)’, 176
hotu-hotu 'all', 120
iconicity, 251, 264, 273, 304, 316
ida 'one', 100, 123, 129
distribution in NP, 136
ida- 'idak 'each', 121
iha
'be present', 162, 186, 189
locative preposition, 160
imperative, 208, 226, 244
incorporation, 180, 293
instrument, undergoer and result nouns, derived, 79–83
intensifier, 231
interjection, 56
interrogative
as fronted complement clause, 286
disjunctive and polar, 211
information, 209
intransitive verb, 52, 54
as predicate head, 179
derived, 62, 66–70
irrealis, 220
with fronted complement, 286–88
with preverbal complement, 180, 292
ita 'IP, polite 2S', 112–13
-k
derives adjectives and nouns, 82, 88–90
possessive clitic, 145
k(a)- -k derives adjectives and nouns, 81, 86–88
ká
‘or’, 302
uncertainty tag, 300
kaliuk 'above all', 232
kalo 'if', 312
-kan possessive clitic, 145
kedan 'immediately', 237
keta
‘don’t’, 226
‘perhaps’, 248
kin relation, 146
kona ‘touch, undergo’, 184
la ‘not’, 228
tá certainty tag, 300
ladin ‘not very’, 231
lahōs ‘indeed not’, 230
tai ‘first, beforehand’, 240
tale
‘else’, 301
‘no’, 212, 228
talek ‘lack’, 190
talika(n) ‘need not’, 225
left-dislocation, 295–97
liu
‘go further’, 259
‘very’, 231
loanwords, 27
location, 207
prepositions, 160–62
ma- (k)
derives actor-describing verbs, 70–78
derives adjectives and time nouns, 83, 90
mai
‘come’, 262
deictic particle, 243
mak- -n derives actor-describing verbs, 70–78
Malay/Indonesian, 6, 103, 117, 139, 185, 226, 245, 246, 271, 285, 291, 300, 301, 311, 312, 321
ma lu reciprocal, 233
manner phrase, 164–66, 242
masik ‘although’, 313
mate ‘very, die’, 231
mensural classifier, 54, 107–9
mesa(n) ‘alone’, 276
morphophonemics, 43–48
motion-action serial verb construction, 262
motion-intransitive direction serial verb construction, 258
motion-transitive direction serial verb construction, 264
musti ‘must’, 226
mutu, mutuk ‘together’, 275
-n
‘imminent’, 239
derives adjectives and nouns, 82, 91
genitive clitic, 149–51
on location and kin nouns, 150
with full reduplication of verbs, 93
na’i
‘do heedlessly’, 93
term of address, 140
nabé ‘which, where’, 128, 129–30
natón ba ‘at (time)’, 164
ne’e ‘this’, 48, 115, 310
ne’ebé ‘so that’, 285, 316
ne’et ‘that’, 119
nebá ‘there, that’, 119
negation, 228–31
in serial verb construction, 251, 254
nemai ‘here, this’, 119
ní ‘be’, 182
nia ‘3S, that’, 114, 117
né ‘3-exist’, 186, 188
‘and’, 301
noble register, 7
nominal, 50
nominalisation
by ne’e ‘this’, 310
by partial reduplication, 78
noun phrase, 50, 132–42
appendition, 158
as modifier in NP, 155
as predicate, 193
coordination, 156
presentational, 191
noun, common, 51
as head of NP, 132
derived, 79–83
repetition, 95
noun, deverbal
as head of NP, 138
derived by partial reduplication, 78
noun, proper, 51
as head of NP, 140
names and other terms of address, 139
nickname derived by truncation, 97
nú, nu’u ‘like’, 164, 166
nuclear causative serial verb construction, 257–58
nudar ‘like, be as if’, 165–66
numeral, 54, 100–101
full reduplication, 93
in numeral phrase, 99
numeral classifier, 54, 103–9
numeral phrase, 99, 109
as predicate, 192
distribution in NP, 132, 136
ó ‘2S’, 113
ó tag, 301
object, 170
follows long recipient phrase, 268
in serial verb construction, 251
incorporation, 180
Index

topicalisation, 179
ohin 'aforementioned', 153
onan 'imminent', 238
ordinal expression, 103
orthography, 2, 4, 5, 21, 47
parataxis, 252, 303–5
partial reduplication
  derives nouns, 78, 80
  morphophonemics, 44, 47
  of verbs, 95
stress, 19
part-whole relation, 145, 196
periphery of clause, 206–8
personal pronoun, 111–15
plurality in NP, 134
poetry, 7, 18
Portuguese, 3, 15, 27, 245, 313
possession, 142–52, 186–92, 193
possessions, 146
possessor NP, 145–49
  as predicate, 193
possibility sentence adverb, 247
predicate, 178–204
  'body-good' expression, 196–204
copular, 182–84
existential, possessive, 186–92
noun phrase, 193
numeral phrase, 192
possessor, 193
prepositional phrase, 167, 192
relative clause, 194
  verbal, 178–92
prefix
  morphophonemics, 43–44, 47
  template, 18
  unstressed, 19
preposition, 55, 159–67, 263–75
prepositional phrase, 159, 167
  as complement, 183
  as predicate, 167, 192
  distribution in NP, 132, 136
prepositional verb, 56, 263–75
pro-clause, 56, 211
pronoun, 51, 110–31
  as complement, 182
  as head of NP, 139
purpose, 270, 316
quantifier, 120–21
question, 209–13
quote, 280–81, 308
-r genitive clitic, 149
re 'at (location)', 162
reason clause, 314
recipient, 268
reciprocal, 68, 233
reduced complement, 205, 288–93
reflexive, 232
register, 7
relative clause, 318–22
  as postmodifier in NP, 77, 132, 136, 156
  as predicate, 194–96
  as premodifier in NP, 132, 154
relator, 55
repetition
  of epistemes, 130
  of nouns, 95
  of verbs, 96
resi(k) 'too, overly', 231
resin 'extra', 101
result, 270
result noun, derived, 79–83
right-dislocation, 298
ruas 'two', 86, 98, 101, 157
-s derives adjectives, 97
sá
  'what', 129–30
tag, 300
sa-, fa- derives finger and toe names, 96
sé 'who, which person', 128, 129–30
sea taboo register, 8, 76
sei
  'still', 219
  with negation 'not yet', 219
sentence, 294
sentential complement, 279–83
serial verb construction, 250
sia
  plural marker, 124
  pronoun '3P', 114
small clause, 205
  distribution, 154, 184, 292, 321
só 'unless', 313
sortal classifier, 54, 104–7
source (location), 161, 266
stress, 18
Suai sub-dialect, 3, 7, 36, 38, 106, 153, 173, 219
subject, 170
  in serial verb construction, 251
  no sentential subjects, 288
  of 'body-good' expression, 199
  omitted in reduced complement, 205
subject marking, 172–78
  in serial verb construction, 251, 256
  of adverb hikar 'back', 241
  of auxiliaries, 217
  of complementiser hosí 'about', 284
suffix
  morphophonemics, 43
superlative, 232
sura 'every', 153
surak 'on condition that', 313
syllable, 14
ta 'already', 235
tag, 56, 299
tail-head linkage, 236, 255, 262, 304, 310
tán (bá) 'because', 314
té 'because', 315	
tense (none), 235
ti'a, ti'an 'already', 235
time, 206, 235–40
  prepositions and conjunctions, 163–64, 311
  relativisation of, 321
time noun, 51
  as head of peripheral NP, 206
  derived by ma- (-k), 83
  ohin as NP modifier, 153
title, 51, 139–40
tó confirmation tag, 300
to'o
  'reach', 259
  'until', 163
  'very', 231
tone
  'to (you)', 262
  deictic particle, 243
topic, 167, 297
topicalisation, 179
topic-comment, 188, 204
tosakó 'unless', 313
transitive verb, 52, 54, 263–75
  as predicate head, 179
  derived by ha- 'make', 59–66
  derived by hak- (-k), 68–70
  in serialisation, 260
tu'an 'grow', 183
verb, 51–54
  as degree modifier, 231
  as manner modifier, 242
  as predicate head, 178–92
  full reduplication, 93
  partial reduplication, 95
  repetition, 95
vocative, 299
voice (none), 13
vowel, 24–26
  frequencies, 40
  length, 19–21
  sequences, 30
  variation, 41
wain 'because', 316
word classes, 49–57
word game, 22
word templates, 15–18
zero derivation, 83