A short grammar of Tetun Dili

Catharina Williams-van Klinken,
John Hajek & Rachel Nordlinger

Table of contents

Conventions

1PE	1st person plural exclusive (*ami*)	k.o.	kind of
1PI	1st person plural inclusive (*ita*)	lit.	literally
1s	1st person singular (*hau*)	LOC	general locative (*iha*)
2P, 3P	2nd and 3rd person plural (*imi, sira*)	N	noun
2s, 3s	2nd and 3rd person singular (*o, nia*)	PL	plural (*sira, -s*)
2S.HON	2nd person singular formal (*ita*)	POS	possessive (*nia*)
Adj	adjective	Prep	preposition
Adv	adverb	PRF	perfective (*tiha*)
ANT	anterior (*ona*)	PROG	progressive (*daudauk*)
C	consonant	PROHIBITIVE	prohibitive (*keta*)
CLS:human	numeral classifier for humans (*nain*)	RDP	reduplication
CONT	continuous (*hela*)	RECIP	reciprocal (*malu*)
EXIST	there exists/ is/are, be present (*iha*)	REL	relative clause marker
FIRST	first, before doing other things (*lai*)	TAG	tag
FOCUS	focus (*mak(a)*)	TD	Tetun Dili
FUTURE	future (*sei*)	TT	Tetun Terik
GEN	genitive marker (*-n*)	V	vowel
INTR	intransitiviser (*nak-, nam-*)	Vi	intransitive verb
IRR	irrealis (*atu*)	Vt	transitive verb
JUST	only just, very recently (*foin*)		

Full stops separate words within multi-word English glosses (e.g. *laós* 'indeed.not').

The examples in this book are taken from a wide range of spoken and written texts. Personal names have in many cases been changed to guard anonymity.

In translations, gender, tense and aspect (which are often not inferable from the Tetun) are given according to the context. Brackets surround comments, as well as the translations or paraphrases of those parts of the example that are omitted by ellipsis.

1. Introduction

1.1. The language and its speakers

Tetun Dili is, alongside Portuguese, one of the two co-official languages of newly independent East Timor. It is spoken as a first language in the capital Dili, which is situated on the north coast. It is also widely and increasingly spoken throughout the country as its preferred lingua franca. It is the best known of all East Timorese languages, with dictionaries and other materials dating back at least a century.

Tetun Dili is an Austronesian language, closely related to a number of languages in the region, including Kemak, Galolen, and Mambae. Estimates vary widely, but it is likely that some 60–70% of the population of East Timor (which totals approximately 800,000) presently speak Tetun Dili as their first or second language. Perhaps 20% are first language speakers. There are two parts of East Timor in which Tetun Dili is not traditionally spoken:

the easternmost province of Lautem and the enclave of Oecussi (see Thomaz 1981). However, Tetun Dili is now spreading rapidly in both areas, particularly amongst younger speakers.[1]
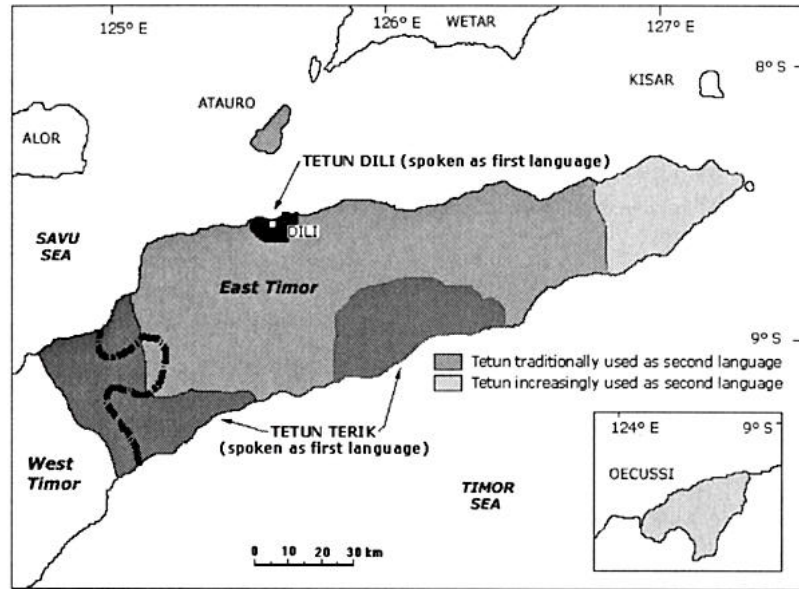
Tetun Dili has grown out of Tetun Terik, even though the areas in which Tetun Dili and Tetun Terik are spoken are a significant distance apart. Tetun Terik has approximately 300,000 speakers, most of whom live in the Indonesian province of West Timor. There are however also many speakers of Tetun Terik resident in rural areas of East Timor, particularly near the border between East and West Timor, and in the Soibada–Viqueque district on the southern coast. It is not clear to which extent these two varieties of Tetun should be viewed as dialects of the same language, since mutual comprehension can be very difficult. Reasons for this include a large number of Portuguese loans which are present in Tetun Dili but absent from Tetun Terik, as well as numerous other differences (see chapter 13 for further details).

Although speakers are very aware of the differences between Tetun Dili and Tetun Terik, it is common practice for either variety to be referred to simply as Tetun (or Tetum, according to Portuguese orthographic convention). An older name for Tetun Dili is Tetun Prasa (lit. 'marketplace Tetun').

1.2. History and contact

Tetun Dili is today native to Dili, although it was not originally indigenous to it, with the area instead being Mambae-speaking. Some form of Tetun was apparently spoken as a lingua

---

[1]   This map, based on Thomaz (1981:56), is courtesy of John Bowden, of the Australian National University.

franca within East Timor even when the Portuguese first arrived in the 1500s. Tetun subsequently became established in Dili when the Portuguese made the city their colonial capital in the late 1700s. From that time until the Portuguese departed abruptly in 1975, contact between Tetun Dili and Portuguese was intense, with Portuguese strongly influencing Tetun Dili in almost all areas of grammar, especially the phonology and lexicon. During the Indonesian period (1975-1999) direct Portuguese influence on Tetun Dili largely ceased, as Indonesian authorities endeavoured to replace Portuguese with Indonesian in all the domains where the former once dominated, such as education, the media and administration. However, since the end of Indonesian control in 1999, Portuguese has returned in a significant way to East Timor, and there are signs, especially in the media, of a strong resurgence of Portuguese influence on Tetun Dili.

Malay influence on Tetun Dili is also evident. This can be divided into two historically separate periods. Some form of Malay was used as a coastal trade language in East Timor for centuries, until its disappearance in the early 19th century. Loans from this period, which speakers of Tetun Dili consider to be fully native, include *surat* 'letter'. During the Indonesian period, Malay/Indonesian borrowings, especially of a technical or political nature (e.g. *kolusi* 'collusion'), were common. There is, however, a strong tendency, especially in written and higher level spoken contexts, to avoid Indonesianisms.

1.3. Variation

One notable aspect of Tetun Dili is the high degree of interspeaker and register variation, which makes description of the language somewhat difficult. There is no doubt that the language is presently undergoing great change, seen in the tendency, especially evident in the media, to replace Indonesianisms with new Portuguese loans. There is also a marked difference between informal spoken register, media and technical language, and religious language. Media and technical language is marked by a very high use of Portuguese loans, while religious language shows the influence of Tetun Terik, alongside a tendency to reduce somewhat the use of Portuguese loans (see also Williams-van Klinken et al. (2002) and Williams-van Klinken (2002a)).

1.4. Current status

In the original planning for independence, the East Timorese National Resistance Council (CNRT) declared that Tetun would become the national language of East Timor, while Portuguese would be made the official language. However, in formulating East Timor's new constitution, which came into effect in 2002, it was agreed that Tetun should share full co-official status with Portuguese.

Since 1981 Tetun has been the official liturgical language of the Roman Catholic Church in East Timor. Under intense pressure from Indonesian authorities after 1975 it was forced to eliminate use of Portuguese (Hajek 2000, Hull 1994). The Church's decision to adopt Tetun in this important role has been of great benefit to Tetun, in terms of raising its status and encouraging its spread across East Timor, and in the development of written materials in the language.

Portuguese is now being reintroduced as the preferred medium of instruction in primary schools, whilst Indonesian is slowly being phased out of education, as well as other sectors. Official policy is in favour of developing Tetun so that it can function fully in its official capacity in all domains. The task of development, under way since at least 1980, has now been assigned to the recently established Instituto Nacional de Lingúistica located in the

National University of East Timor in Dili. Tetun Dili's future seems very secure: it is widely used as a spoken medium, it enjoys relatively high status and, as already noted, its use as a lingua franca is rapidly spreading throughout East Timor.

The constitution gives no precise definition of Tetun, but all signs are that its core will be Tetun Dili, drawing on Portuguese for many technical loans, whilst at the same time showing some Tetun Terik influence. This is already evident in official documents, such as the Tetun translation of the nation's constitution, in which Portuguese elements are strongly evident but where Tetun Terik loans such as *wainhira* 'when' and *no* 'and' are preferred over Tetun Dili *bainhira* and *ho*.

1.5. Previous research on Tetun Dili

There is no doubt that Tetun Dili is the best described of all East Timorese languages. For useful literature reviews of all previous linguistic and anthropological work, including dictionaries and pedagogical materials in English, Portuguese, Dutch and Indonesian, readers are referred to Hull (1998) and Van Klinken (1999).

In recent years there has been substantial activity with regard to Tetun Dili, including a descriptive grammar (Williams-van Klinken *et al.* 2002), a reference grammar (Hull and Eccles 2001), a pedagogical grammar (Hull 1996), and other materials (e.g. Hull 1999, Eccles 1998).

For Tetun Terik, Van Klinken (1999) provides a detailed grammar of the Fehan variety, as spoken in West Timor. A condensed version of this research is presented in Van Engelenhoven and Williams-van Klinken (forthcoming). Hull's (1996) pedagogical manual of Tetun focuses on Tetun Dili but with substantial additional information on Tetun Terik. His dictionary (Hull 1999) also incorporates terms from both varieties.

1.6. Data for this study

The description of Tetun Dili in this sketch grammar is primarily based on a large corpus of oral language data, mainly collected in Dili between 1999 and 2001. Both the corpus and elicitation came from a wide range of speakers of different ages and backgrounds. These data were supplemented by written materials.[2]

2. Phonology

2.1. Words and syllables

The maximal number of syllables in a native Tetun word is four; e.g. *bibiliku* 'kind of drum', with the most common number being two; e.g. *moris* 'live'. Unstressed single-syllable words are restricted to grammatical terms; e.g. *la* 'not'. In Portuguese loans the number of syllables may be higher; e.g. *independénsia* 'independence'.

The syllable template in Tetun Dili is (C)(C)V(V)(C); e.g. *a.but* 'root', *kma.nek* 'wonderful', *plan.ta* 'plant', *al.dei.a* 'village', *seis* 'six', *trei.nu* 'exercise', *ou.ru* 'diamonds (cards)'.

In native Tetun words, word-initial consonant clusters always begin with /k/. These sequences tend to be split by optional vowel insertion; e.g. /'kmanek/ ['kmanekˀ] or [ka'manekˀ] 'wonderful'. Word-internal consonant sequences in underived words are restricted to /kC/ and /mC/; e.g. *hakmatek* 'quiet'; *hamnasa* 'laugh'; in compounds and derived words, a wider range is possible; e.g. *man-tolun* 'bird-egg', *halimardór* 'playful person'. A large number of new clusters have been introduced via Portuguese loans, including (s)C+liquid sequences; e.g. /'prontu/ 'ready', /es'trada/ 'road'. As in Portuguese, /sC/ clusters never appear word-initially, at least phonemically.

2.2. Stress and vowel length

Stress usually falls on the penultimate syllable; e.g. /ha'manas/ 'heat', /'fatin/ 'place'. It is however fully contrastive; e.g. /'hare/ 'unpicked rice' versus /ha're/ 'see'. Both native Tetun words and loans allow final stress; e.g. /ha'ri/ 'build', /portu'gal/ 'Portugal'. In addition, a small number of Portuguese loans have antepenultimate stress; e.g. /'lamina/ 'razor blade'.

Stressed vowels in the final syllable of a word are always long; eg. /'be/ ['be:] 'water', /ha'ri/ [ha'ri:] 'erect'.

Other vowels (namely unstressed vowels and non-final stressed vowels) are always short; e.g. /ha'tene/ [ha'tene] 'know'.

There are, however, some exceptions to the rule, leading to the possibility that vowel length may not be totally dependent on stress. Some speakers, for instance, appear to have a short final stressed /a/ in /ani'mal/ [ani'maɬ] instead of the expected [ani'ma:ɬ] 'animal', while showing the expected vowel length in /na'nal/ [na'na:ɬ] 'tongue'.

2.3. Consonants

Tetun has a maximal phoneme inventory of 22 consonants, presented below. Unless otherwise indicated in the text, the symbols have their usual IPA values in terms of voicing, place and manner of articulation.

(1)

	Bilabial	Labio-dental	Coronal	Alveo-palatal	Palatal	Velar	Glottal
Voiceless stop	p		t			k	
Voiced stop	b		d			g	
Voiceless fricative		f	s	ʃ			h
Voiced fricative		v	z	ʒ			
Nasal	m		n		ɲ		
Liquid			l		ʎ		
Trill			r				
Tap			ɾ				
Approximant	w				j		

Portuguese has had an extensive influence on the phonology of Tetun Dili. Portuguese loans are responsible for introducing the phonemes /p g v z ʃ ʒ ɲ ʎ r/, and at least partially so in the case of /j, w/. These same phonemes are also the source for numerous allophonic processes observed in the speech of many speakers. At the same time, phonetic and phonological variability is great, as a result of which many speakers do not have the full set of consonant phonemes listed above. In particular, /ɾ/ and /r/ merge for most speakers. In addition, especially for those who are not native speakers of Tetun Dili, there is the possibility of merger for: /v/–/b/, /ʃ/–/s/, /ʒ/–/z/, /ɲ/–/n/, and /ʎ/–/l/.

All voiceless stops are unaspirated, and are usually unreleased word-finally.

/t/ is lamino-dental; e.g. *tata* [tata] 'bite'.

/k/ is normally a voiceless velar stop; however in word-initial consonant clusters of the form /kC/, it has a range of pronunciations depending on the speaker; e.g. *kbit* [ᵏbiːtˀ], [kiˈbiːtˀ] or [ʔbiːtˀ] 'power'.

/b/ optionally weakens to a fricative in intervocalic position; e.g. *bibi* [ˈbibi] or [ˈbiβi] 'goat'.

/d/ is usually a post-alveolar stop; e.g. *dada* [ˈdada] 'pull'.

/v/ tends to be unstable and optionally alternates with /b/; e.g. *servisu* [serˈvisu] or *serbisu* [serˈbisu] 'work'.

/s/ is fronted alveolar; e.g. *sia* [ˈsia] 'nine'. As a result of Portuguese influence, there is optional retraction to [s̠] or [ʃ] word-finally and before voiceless stops; e.g. *tanis* [ˈtanis] or [ˈtanis̠] or [ˈtaniʃ] 'cry', *pasta* [ˈpasta] or [ˈpas̠ta] or [ˈpaʃta] 'briefcase'.

/z/ is alveolar; e.g. *zero* [ˈzero] 'zero'. Before sonorants, some speakers optionally retract /z/ to [z̠] or [ʒ], once again due to Portuguese influence; e.g. *dezmaia* [dezˈmaia] or [deʒˈmaia] 'faint'.

/ʃ/ is for many speakers, especially those with less experience of Portuguese, merged with the /s/ phoneme; e.g. *xave* [ˈʃave] or [ˈsave] 'key'.

/ʒ/ too is highly unstable across speakers and is often replaced by [z]; e.g. *joga* [ˈʒoga] or [ˈzoga] 'play'.

/h/ is optionally deleted in intervocalic position; e.g. *ahi* [ˈahi] or [ˈai] 'fire'.

/n/ is an apico-alveolar nasal. Word-finally, for some speakers it is always [n], for others it is velar [ŋ], and for others still there is only nasalisation on the preceding vowel; e.g. *jardin* [ʒarˈdĩːn] or [ʒarˈdĩːŋ] or [ʒarˈdĩː] 'garden'.

/ɲ/ is, for speakers not native to Dili, often reduced to [jn] or [n]; e.g. *linha* [ˈliɲa] or [ˈlijna] or [ˈlina] 'line'.

/l/ is lamino-alveolar; e.g. *lolon* [ˈlolon] 'trunk'. It can be velarised in syllable-final position, especially when word-final; e.g. *finál* [fiˈnaːl] or [fiˈnaːɫ] 'final'.

/ʎ/ is often reduced to [jl], [l] and more rarely [j] amongst speakers not native to Dili, e.g. *pilha* [ˈpiʎa] or [ˈpijla] or [ˈpila] 'battery'.

/ɾ/ is a voiced alveolar tap, e.g. *hirus* [ˈhiɾus] 'anger'.

/r/ is a voiced alveolar trill. For most speakers the trill [r] and tap [ɾ] are in free variation, although the trill seems the preferred rhotic in syllable-initial position, at least in careful speech; e.g. *rán* [ˈraːn] 'blood', *para* [ˈpara] or [ˈpaɾa] 'stop'. The possibility of contrast between /r/ and /ɾ/ is extremely marginal, and possible only when one contrasts a small number of Portuguese loans with original trill in intervocalic position, e.g. /kareta/ 'car', with loans and native words with intervocalic tap, e.g. /padeiɾu/ 'baker' and /hiɾus/ 'anger'. However, most speakers appear not to make a distinction, merging intervocalic rhotics into /ɾ/. As a result, we do not mark them separately in our transcription.

/w/ is found in a small number of words, most of which are loans; e.g. /ˈkwartu/ 'room' (from Portuguese *quarto*) and Tetun Terik /wainˈhira/ 'when', but some of which are native; e.g. /ˈwalu/ 'eight'.

/j/ is found only in loans; e.g. /ˈjodu/ 'iodine'.

2.4. Vowels

Tetun has five vowels: /i, u, a, e, o/.

/i/ is a high front unrounded vowel; e.g. *bibi* [ˈbibi] 'goat'.

/u/ is a high back rounded vowel. A slight fronting to [ʉ] is not uncommon in unstressed final open syllables; e.g. *trigu* [ˈtrigu] or [ˈtrigʉ] 'wheat'. Laxing to [ʊ] is possible in unstressed closed syllables; e.g. *rihun* [ˈrihun] or [ˈrihʊn] 'thousand'.

/a/ is a low unrounded central vowel. In unstressed syllables it can be raised to [ɐ], and further reduced to [ə]; e.g. *oras* [ˈoras] or [ˈorɐs] or [ˈorəs] 'time', *sasate* [saˈsate] or [səˈsate] 'satay'.

/e/ is a front unrounded mid vowel. Its height appears to vary greatly according to speaker and context. All speakers, however, tend to open the vowel when it is next to /h/; e.g. *nehek* [ˈnɛhɛk] 'ant'.

/o/ is a back rounded mid vowel. Here too speakers show considerable variation. For many, the basic allophone is open [ɔ], while for others it tends to true-mid; e.g. *mota* open [ˈmɔta] or true-mid [ˈmota] 'river'. /o/ is usually raised further to high-mid position by vowel

harmony when the following syllable contains a high vowel (/i, u/); e.g. *mohu* ['mohu] 'extinguished'. It is also usually high-mid before another vowel; e.g. *foer* ['foer] 'dirty'.

Unstressed /e/ and /o/ in Portuguese loans are frequently raised (to [i] and [u]), especially word-finally; e.g. *prezidenti* [prezi'dente] or [prezi'denti] from Portuguese *presidente* 'president'. When the unstressed vowel /e/ occurs word-initially before /sC/ clusters, it is often absent altogether, as in European Portuguese; e.g. *estrondu* [es'trondu] or [is'trondu] or [i'ʃtrondu] or ['ʃtrondu] from Portuguese *estrondo* 'din'.

When native Tetun words are pronounced carefully, sequences of vowels are easily distinguished as separate vowels. In normal speech, however, they tend to coalesce into diphthongs; e.g. *lae* ['la.e] or ['lae] 'no'. Diphthongs are common in Portuguese loanwords. However in normal speech speakers tend to reduce many of these to single vowels; e.g. *padeiru* [pa'deiru] or [pa'deru] 'baker', *tezoura* [te'zoura] or [te'zora] 'scissors'.

Although not normally marked here, when vowels immediately precede or follow nasals, they are usually nasalised; e.g. *mean* ['mẽãn] 'red', *manu* ['mãnũ] 'bird', *santu* ['sãntu] 'holy'.

2.5. Reduction of compounds and reduplications
In compounds and reduplications, the initial word is frequently truncated. In particular, its coda is often deleted, and in normal speech the vowel of the final syllable is typically weakly articulated or omitted altogether. In the latter case, the onset of the final syllable is left to function as the coda of the initial word; e.g. ['manᵘ 'tolun] or [man 'tolun] 'egg' from *manu-tolun* 'bird-egg', ['tebe'tebes] or [teb'tebes] 'really' from *tebes* 'real'.

2.6. Transcription
We have adopted a simple transcription system for Tetun Dili in this grammar. Whilst generally consistent with the IPA, we differ in using some Portuguese orthographic conventions for Portuguese loans. In particular,

a) the following (alveo)-palatal phonemes, all borrowed from Portuguese, are represented as follows:

/ʃ/ = x; e.g. *xave* 'key'   /ʒ/ = j; e.g. *janela* 'window'
/ɲ/ = nh; e.g. *kampanha* 'campaign'   /ʎ/ = lh; e.g. *barulhu* 'noise'

b) we use an accent placed on the vowel to mark non-penultimate stress, even on monosyllables, in both native and non-native terms; e.g. *mós* 'clean, also', *médiku* 'doctor'.

3. Major grammatical patterns
The basic order for Tetun clauses is subject–predicate. Verbal clauses follow the order subject–verb–object (2). The object can be topicalised by placing it in the pre-subject position (3). There is no system of voice (i.e. no passives).

(2)   *Sira hasoru nia iha nebá.*
       3P   meet 3S LOC there
       'They met her there.'

(3)   *Barak mak   ami la kompriende.*
       much FOCUS 1PE not understand
       'There is much that we don't understand.'

Non-verbal predicates directly follow the subject, without use of a mediating copula. These include noun phrases, prepositional phrases, numeral phrases (4), and the so-called 'body-good' expressions described in section 4.4.

(4)   *Sira nain   rua deit.*
       3P   CLS:human two just
       'There were only two of them.' (lit. 'They were only two.')

Tetun Dili has no grammatical inflection. There is, for instance, no agreement between the verb and subject or object, nor is their any tense–aspect–mood inflection. Instead, as illustrated in example (5), tense–aspect–mood distinctions are optionally marked by a set of words which occur in either pre-verbal or post-verbal position. When such marking is absent, context determines the interpretation. For instance, example (2) could be interpreted as 'They met her there', 'They meet her there', or 'They will meet her there'. Expressions such as *aban* 'tomorrow' or *sempre* 'always' can of course be added to force a particular interpretation.

(5)   *Hau sei   hán.*   *Hau hán hela.*   *Hau hán tiha ona.*
       1S   FUTURE eat   1S eat CONT   1S eat PRF ANT
       'I will eat.'   'I am eating.'   'I have eaten.'

Verbs and adjectives are typically negated by a preceding *la* 'not' (3). A much wider range of constituents can be negated by a preceding *laós* 'indeed not' (6); this is frequently contrastive, particularly when it negates verbs and adjectives.

(6)   *Ami iha né simu   ema   barak. Laós   ita   bót deit!*
       1PE LOC this receive person many indeed.not 2S.HON big just
       'We here receive many people—not just you!'

Non-core arguments, such as recipients and directions, are introduced by prepositions, and follow the object slot.

(7)   *Hau bele haruka surat ba nia.*
       1S   can send   letter to 3S
       'I can send him a letter.'

Serial verb constructions are used for a range of functions, including building up causative predicates (e.g. *halo moris* (make live) 'bring to life'), introducing direction or instrument, and marking consecutive actions (e.g. *mai tein* 'come (and) cook').

Within noun phrases, almost all modifiers follow the head noun (8). It is very uncommon to have more than two open-class modifiers within the one noun phrase.

(8)   *uma   bót ida*
       house big one
       'a big house'

Possessors can either precede (9) or follow (10) the head noun. When enumerating humans and some other entities, a numeral classifier is used before the numeral (9); for humans this is *nain* (also a noun meaning 'master, owner'). Relativisation occurs mainly on subjects, objects

and temporal phrases. Most relative clauses are introduced by a relative clause marker, the most common being *nebé* (also an interrogative meaning 'where, which') (11).

(9) *hau nia oan nain   tolu né*
     1S POS child CLS:human three this
     'these three children of mine'

(10) *futuru Timor nian*
     future Timor POS
     'Timor's future'

(11) *maluk sira nebé agora iha né*
     friend PL REL now LOC this
     'the friends who are now here'

Deictics mark only a two-way distance distinction, between the default *né* 'this' and the distal *nebá* 'that, there'. Plural marking is optional, and uses the definite plural marker *sira* (also the third person plural pronoun), as in (11). Noun phrases can be marked as singular indefinite by *ida*, which is also the numeral 'one' (8). The combination *né ida* (this one) can be used to mark definite singular noun phrases (e.g. *uma ida né* (house one this) 'this house'). Personal pronouns distinguish first, second and third persons, as well as singular and plural; first person plural pronouns distinguish between *ita*, which includes the addressee, and *ami*, which excludes the addressee.

There is little productive morphology in Tetun Dili. Amongst the more productive are three means of deriving agents: the Portuguese loan suffix *-dor* (e.g. *dukurdór* 'sleepy-head' from *dukur* 'sleep'), the bound root *-tén* (e.g. *baruk-tén* 'lazybones' from *baruk* 'lazy'), and compounds formed with the noun *nain* (e.g. *pinta nain* 'painter' from *pinta* 'paint'). There are many nominal compounds (e.g. *ai tahan* (plant leaf) 'leaf'), and also many so-called 'body-good' expressions. The latter, described in section 4.4, usually denote human attributes, and typically consist of a body-part noun followed by an adjective or intransitive verb (e.g. *isin diak* (body good) 'healthy, well'). They behave in some ways like compounds, and in others as two separate constituents. The prefix *ha-* derives causative verbs from adjectives and intransitive verbs (e.g. *hasai* 'remove' from *sai* 'exit'). Full reduplication derives adverbs from temporal nouns (e.g. *dadér-dadér* 'every morning' from *dadér* 'morning'), amongst other things.

For explicitly linking clauses into complex sentences, there are a range of native and borrowed (Portuguese) coordinators and conjunctions (12). Complement clauses are of two major types: sentential complements are full clauses, and are often introduced by *katak* 'that' (13); reduced complements consist of verb phrases only, and have restricted tense–aspect marking. The latter usually have no complementiser (14).

(12) *Ami mós kontenti tanba   Timor ukun-an ona.*
     1PE also happy   because Timor rule-self ANT
     'We were happy because Timor now had independence.'

(13) *Sira hatete katak hau tenki servisu iha Kupang.*
     3P say   that 1S must work LOC Kupang
     'They said that I had to work in Kupang.'

(14) *Nia la biban   atu lori labarik ba doutór.*
     3S not manage IRR take child   to doctor.
     'She didn't manage to take the child to the doctor.'

4. Words and word classes

4.1. Introduction

Tetun Dili has a relatively clear distinction between nouns and verbs, with some word forms occurring in both classes. Adjectives have much in common with verbs. There are in addition a range of smaller word classes.

In this chapter we discuss selected word classes, and also derivational possibilities.

One compound type which crosses word class boundaries, and so is mentioned here, is that of semantically coordinate compounds, in which two terms from a single semantic set are juxtaposed. Such compounding is found for nouns (e.g. *inan-aman* 'parents' from 'mother-father'), adjectives (e.g. *midar-sín* 'sweet and sour' from 'sweet-sour'), and verbs (e.g. *túnsae* 'up and down, all over the place' from 'ascend-descend').

4.2. Nouns

Syntactically, nouns and pronouns are distinguished by the fact that they can head noun phrases (which indeed is their typical function), and can be modified by numerals (e.g. *surat rua* (letter two) 'two letters').

Morphologically, some nouns form the base for verbs (see section 4.3.2), and temporal nouns can be reduplicated to derive temporal adverbs (see section 4.5).

There are three primary means of deriving nouns in Tetun Dili: by the suffix *-dór*, by nominalising a verb or adjective, or by compounding (described in sections 4.4 and 5.5).

The Portuguese loan suffix *-dór* attached to a verb X derives a noun denoting a person who habitually does X. The suffix readily occurs with Tetun roots. Although there are borrowed Portuguese nouns ending in the feminine form *-dora* (e.g. *organizadora* '(female) organiser'), and borrowed nouns denoting non-humans (e.g. *komputadór* 'computer'), derivations based on Tetun roots always use the masculine *–dór* (regardless of the referent's gender), and always refer to humans (Hajek and Williams-van Klinken in press).

(15) *husu*   'ask'   *husudór*   'person who keeps asking for things'
     tanis   'cry'   *tanisdór*   'cry-baby'
     hamnasa 'laugh' *hamnasadór*   'giggler'

Some native Tetun verbs and adjectives can be nominalised without change of form (e.g. *moris* Vi 'live', N 'life'; *koalia* Vi 'speak', N 'speaking'), as can some Portuguese loans, as illustrated by *fiar* 'believe; faith' in (16). For many Portuguese loan verbs, however, the nominalised form is borrowed separately (e.g. *investiga* Vt 'investigate', *investigasaun* N 'investigation').

(16) *Ha-metin ita nia fiar*.
     make-firm 1PI POS believe/faith
     'Strengthen our faith.'

4.3. Verbs and adjectives

4.3.1. Overview

Syntactically, verbs and adjectives have many features in common. Predicates headed by either of these word classes are typically negated by *la* 'not' (17, 18), while those headed by nouns can only be negated by the usually contrastive negator *laós* 'indeed not' (19). Adjectives and verbs accept a wider range of tense–aspect marking (described in section 8.4) than do nominal predicates.

(17) *Nia la   mai.*
      3S not come
      'She didn't come.'

(18) *Sira nia maneiras la   diak.*
      3P POS manners not good
      'Their manners aren't good.'

(19) *Nia laós mestri.*
      3S not teacher
      'He is not a teacher.'

Morphologically, both verbs and intransitive adjectives can be the root for causative verbs using the prefix *ha-*, as outlined in section 10.2.

A syntactic difference between verbs and adjectives is that when modifiers of nouns are headed by verbs, they are usually introduced by the relative clause marker (20a), whereas adjectival modifiers usually have no relative clause marker (20b).

(20) *ema   nebé rona*   *ema   moras*
      person REL hear   person sick
      '(the) person who hears'   'sick person'

Some adjectives (but not verbs) can be reduplicated to derive an adverb (e.g. *di-diak* 'carefully, thoroughly' from *diak* 'good'), or can function without reduplication as the head of a manner modifier (21).

(21) *Sira servisu diak teb-tebes.*
      3P work   good RDP-true
      'They work really well.'

Some adjectives can be reduplicated when they describe plural referents (e.g. *bo-bót* 'plural big' from *bót* 'big'); for some, this is particularly so when the referents are diverse (e.g. *selu-seluk* 'various other' from *seluk* 'other').

4.3.2. Derivation of verbs

A few transitive verbs can be detransitivised by the prefix *nak-* (e.g. *nakdulas* Vi 'spin' from *dulas* Vt 'spin') or, less commonly, *nam-* (e.g. *namkari* Vi 'scatter' from *kari* Vt 'scatter'). The derived verb does not imply that an actor was involved, and is equally applicable for events that happened without human intervention and for events that were (usually unintentionally) caused by someone.

(22) *Nia kore butaun.*   *O nia butaun nam-kore tiha ona.*
      3S undo button   2S POS button INTR-undo PRF ANT
      'She undoes a button.'   'Your button is undone (by itself).'

Some verbs are derived from nouns (including Portuguese loan nouns) with no change in form. There is considerable inter-speaker variation in such derivation. It is used for deriving some verbs denoting the use of tools from the noun for the tool, as well as for deriving verbs from abstract nouns. In the latter instance, some speakers prefer to use a construction in which the noun is made the object of the verb *halo* 'make, do'.

(23) *bahat*   'chisel'   Vt 'chisel, carve'
      bomba   'pump'   Vt 'pump up'   (Portuguese verb is *bombear*)
      esforsu 'effort'   Vt 'strive'   (Or Portuguese verb *esforsa*)
      serbisu 'work'   Vi 'work'   (Or *halo serbisu* 'do work')
      xoke   'collision' Vt 'collide with' (Portuguese verb is *chocar*)
      xave   'key'   Vt 'lock up'   (Portuguese verb is *trancar*)

4.3.3. Derivation of adjectives

Derogatory adjectives describing people who habitually do the activity of the root verb, or who display a characteristic of the root verb, adjective or noun, can be derived by attaching the bound root *-tén* (also a noun meaning 'faeces'). Although such derivations are adjectives (e.g. in that they can be intensified, and negated by *la*), they are most easily translated by nouns in English.

(24) *beik* 'stupid' *beik-tén* 'habitually stupid, never learns'
      kán   'crave'   *kán-tén*   'glutton'
      naok 'steal'   *naok-tén* 'habitual thief'
      tauk 'afraid' *tauk-tén* 'scaredy-cat, afraid of everything'

The derivation of compound adjectives is discussed in the next section.

4.4. 'Body-good' expressions

Tetun has a large number of compound-like expressions for denoting character, emotions and physical attributes of people. As noted in more detail below, most behave in some contexts as compound adjectives or nouns, and in others as multi-word predicates.

Most such expressions consist of a noun, usually denoting a body part, followed by a single-word predicate, which is typically an adjective. We have named this construction after a common exponent of it: *isin diak* (body good) 'healthy'. Other examples include:

(25) *ain badak* (leg short)   'short (of person)'
      isin molik (body bare)   'naked'
      laran ru-rua (inside RDP-two) 'in two minds, uncertain'
      laran todan (inside heavy)   'heavy-hearted, depressed'

Although many of these expressions are standardised, there are also a considerable number which are accepted by only some native speakers, or which are interpreted differently by different speakers.

Most body-good expressions behave like single-word compound adjectives in some contexts, but as two separate words in other contexts. For instance, in (26) the body-good

expression as a whole functions as the predicate of the main clause, whereas in (27) its constituent parts have different functions within the clause, with the body part noun functioning as part of the subject, and the adjective as the predicate.

(26)   *Ema   né   laran moras tanba   hau sosa kareta foun.*
      person this inside sick   because 1S buy car   new
      'This person is envious because I have bought a new car.'

(27)   *Tanba né   mak   kalan né hau nia laran né susar.*
      because this FOCUS night this 1S   POS inside this difficult
      'It was because of this, that on that night I was distressed.'

Emotion and health predicates can be introduced by the Portuguese loan *senti* 'feel'. The complement of *senti* is usually the entire body-good expression (so treating it as a compound), although it can be the adjective on its own.

(28)   *Hau senti laran beik.*   *Hau nia laran senti beik.*
      1S   feel inside stupid   1S   POS inside feel stupid
      'I feel nauseous.'

To negate such predicates, some speakers allow the negator *la* (which elsewhere negates verbs and adjectives only) to precede the adjective but not the noun. Some, however, accept it before the noun, thus analysing the noun–adjective pair as a complex adjective. Another alternative is to avoid the issue by using the contrastive negator *laós* (which can negate any constituent type) before the noun. Example (29) illustrates these three alternatives.

(29)   *Sira isin la diak.*   *Sira la isin diak.*   *Sira laós   isin diak.*
      3P   body not good   3P not body good   3P   indeed.not body good.
      'They were unwell.'

Although they are typically used adjectivally, many body-good expressions are also abstract nouns (e.g. *isin manas* (body hot) 'fevered; fever'). In this case, the expression is a compound, and its component words cannot be separated.

4.5. Adverbs

Adverbs modify constituents other than nouns. They cannot function as predicates, and cannot take arguments. With the exception of manner adverbs (all of which are also classified as adjectives), they cannot be negated or intensified.

Some adverbs are derived from adjectives by reduplication (e.g. *lai-lais* 'quickly' from *lais* 'quick', *fou-foun* 'initially' from *foun* 'new').

Temporal adverbs meaning 'every unit-of-time' can be derived from temporal nouns denoting units of time (e.g. *tin-tinan* 'annually' from *tinan* 'year', *Domingu-Domingu* 'every Sunday' from *Domingu* 'Sunday').

4.6. Closed word classes

4.6.1. Personal pronouns

The personal pronouns are listed below. Like nouns, they typically head noun phrases; unlike nouns, they cannot function as modifiers within a noun phrase.

(30)

		Singular		Plural	
1	*hau*	1S 'I, me'		*ita*	1PI 'we, us (including addressee)'
				ami	1PE 'we, us (excluding addressee)'
2	*o*	2S 'you (singular)'		*imi*	2P 'you (plural)'
	ita	2S.HON 'you (singular respectful)'			
3	*nia*	3S 'he, she, him, her, it'		*sira*	3P 'they, them'

There are a number of options for addressing people. The pronoun *o* is typically used for children. A more respectful pronoun is *ita* (also the inclusive first person plural pronoun). *Ita bót* (lit. 'you big') is used for even greater respect and formality, such as in interviews and meetings; by adding the plural marker *sira* (*ita bót sira*) it can also be used to address more than one person. Titles (e.g. *tia* 'aunt') or names (primarily for children, or people significantly younger than oneself) are frequently used instead of pronouns when addressing people.

The third person pronouns *nia* '3S' and *sira* '3P' are primarily used for persons. Although some speakers accept their use for inanimate entities, and particularly for animals, it is far more common to refer to non-humans by a noun, or to omit the noun phrase altogether. The exception is that *nia* is commonly used for non-humans when making generic statements (e.g. *Nia iha isin* (3S have content) 'They (dahlias) have bulbs').

4.6.2. Determiners and non-personal pronouns

Determiners indicate definiteness and/or number within noun phrases. Apart from personal pronouns (discussed in the preceding section), all such words can either head noun phrases (in which case we classify them as pronouns) or function as modifiers within noun phrases (in which case we classify them as determiners). Reciprocal and reflexive constructions, discussed in section 8.1.2 and 8.1.3, do not make use of pronouns.

Many noun phrases have no determiner. This is the normal structure in generic statements (31), and is usual when referring to entities which are unique within the context (e.g. *loro-matan* '(the) sun'). A determiner may also be lacking if definiteness and number don't need to be specified (32).

(31)   *Matan hodi haré.*
      eye   for see
      'Eyes are for seeing.'

(32)   *Ami lori   kuda.*
      1PE bring horse
      'We brought a horse / the horse / horses.'

The demonstratives are *né* 'this' and *nebá* 'there' (from *né bá* 'this go'). Both can either head or modify noun phrases. *Nebá* is the less common, being used only deictically, primarily to indicate that referents are relatively far from the speaker (e.g. *iha nebá* (LOC there) 'there'; *iha Suai nebá* (LOC Suai there) 'over in Suai'). Some speakers additionally use it for distant times (e.g. *momentu nebá* (moment there) 'at that time'). In contrast, *né* is used not only to indicate that items are relatively close to the speaker (e.g. *iha né* (LOC this) 'here', *uma né* (house this) 'this house'), but also for present times (e.g. *agora né* (now this) 'nowadays'), and anaphorically, to refer back to earlier-mentioned entities, actions or propositions, as in (33). Anaphoric *né* readily co-occurs with pronouns (e.g. *sira né* (3P this) 'they whom we are talking about') and proper nouns.

(33)   *Amlulik haruka lori ankor bót ida ... Ankor <u>né</u> kesi ...*
        priest   order   bring   anchor   big   one   anchor this tie
        'The priest ordered that a large anchor be brought (and put into the spring). The
        anchor was tied (to the priest's ship.)'

The definite plural marker *sira* is also the third person plural pronoun. Unlike the pronoun, the plural marker can readily be used of objects (e.g. *livru sira* 'the books') as well as humans. When used with a noun which identifies one person, *sira* indicates 'and associated people' (e.g. *Simão sira* 'Simão and his family/friends/colleagues/...'). Some speakers use Portuguese plural forms of Portuguese nouns (e.g. *livru-s* 'books'), particularly in formal contexts such as news broadcasts. At times both the Portuguese and Tetun plurals are used together (e.g. *livru-s sira* 'the books'); this is however disapproved of by many speakers. Unlike *sira*, the Portuguese plural does not indicate definiteness.

The interrogative pronouns are *sé* 'who', *nebé* 'where, which', *saida* 'what' and *sá* 'what'. All but *sé* also readily function as noun phrase modifiers (e.g. *naran saida* (name what) 'what name'). *Sá* is unique in that when it modifies a noun phrase, it must precede the noun, with the resulting construction requesting a selection from a limited set of options (e.g. *sá loron* 'which day'). However not all speakers use this construction, with some preferring post-modifying *nebé* (e.g. *loron nebé* (day which) 'which day') or *saida*. See section 11.8.1 for examples of interrogatives.

Ida 'one, a' marks a noun phrase as being singular. While it is readily used with definite noun phrases (e.g. *eskola ida né* (school one this) 'this school'), it also commonly marks noun phrases with 'new' information, in which case it is better translated into English as 'a', as in (33). Unlike other numerals, *ida* rarely co-occurs with a numeral classifier, frequently heads a noun phrase (e.g. *ida né* (one this) 'this one'), and can either follow adjectival modifiers within a noun phrase (e.g. *lingua ofisiál ida* (language official one) 'an official language') or precede the adjectival modifier (e.g. *lian ida difisil* (language one difficult) 'a difficult language' or 'a language which is difficult'), particularly if the adjective phrase consists of more than one word.

The quantifiers are *hotu* 'all', *hotu-hotu* 'all, truly all' and *balu* 'some'.[3] *Hotu* 'all' (additionally an adverb 'also' and verb 'be finished') is either the final element in the noun phrase (34), or optionally floats to the position immediately following the verb (35). Its reduplicated form *hotu-hotu* (or *hot-hotu*) 'all, truly all' cannot float to post-verbal position, and either modifies, or (less commonly) heads noun phrases (35).

(34)   *Buat sira né hotu sinál, ...*
        thing PL   this   all   sign
        'All these things are signs (that...)'

---

[3]   There are some other quantifying terms which are not determiners. Quantifying adjectives (and adverbs) include *barak* 'many, much', *uitoan* 'a few, a little', and *tomak* 'whole, entire, all'. *Ida-rua* 'a few' (lit. 'one-two') is a compounded numeral. Some speakers also use the indefinite numeral *hira* 'how many, how much' as a general quantifying term 'several'.

(35)   *Ami hotu halai.*   *Ami halai hotu.*   *Ami hotu-hotu halai.*
        1PE   all   run   1PE   run   all   1PE   RDP-all   run
        'We all ran away.'

Balu (or *balun* for some speakers) functions as a noun phrase modifier meaning 'some (of)' (e.g. *sasán balu* (goods some) 'some of the goods'), or as a noun phrase head meaning 'some' (e.g. *Balu dehan ...* 'Some (people) say ...'). It is usually used of count nouns rather than mass nouns. Its interpretation defaults to 'half' in certain contexts (36).

(36)   *Hau tama tuku rua ho balu.*
        1S   enter hour two and some
        'I start (work) at (approximately) 2.30.'

Ida-idak 'each' is primarily used in the context of individualised possession, with each referent having their own possession. It can either head or modify a noun phrase.

(37)   *Hili   tuir   imi <u>ida-idak</u> ni-nia konsiénsia.*
        choose follow 2P   each   3S.POS conscience
        'Each choose according to your own conscience.' (lit. 'Choose according to each one's conscience.')

Hirak 'certain' is always followed by a definite determiner (*né* 'this', *nebá* 'there') or a relative clause. *Hirak* is 'specific', indicating that the referents form a group, whose members are not enumerated, but could in principle be known (e.g. *ema hirak né* 'these people (whom I was talking about)'). It is primarily used in writing rather than speaking, and more common with non-human referents (such as time periods, words, or situations), than human ones.

Ruma 'some or other' indicates an unknown or unspecified quantity or type. It functions as a noun phrase modifier only, and is commonly found in the standard phrases *dala ruma* (occasion some) 'sometimes', *ema ruma* (person some) 'someone, somebody', *fatin ruma* (place some) 'somewhere', and *buat ruma* (thing some) 'something'.

(38)   *Iha   pergunta-s ruma?*
        EXIST question-PL some
        'Are there any questions?'

Seluk 'other, another' is indefinite, and neutral with respect to number (e.g. *ema seluk* (person other) 'other people, someone else'). The sequence *sira seluk* 'definite.PL other' marks definite plural 'the others' (e.g. *labarik sira seluk* (child PL other) 'the other children').

4.6.3. Numerals and numeral classifiers

Tetun numerals (including the interrogative numeral *hira*) form a unique word class in that they can combine to form complex numerals, and in that they are typically preceded by a numeral classifier when certain entities are enumerated. The numerals follow the decimal system.

Tetun Dili speakers are familiar with numerals in Tetun, Portuguese and Indonesian. These are presented below. Tetun numerals tend to be used mainly for small numbers, such as the number of children in a family. Dates, prices and arithmetic are more commonly presented in Portuguese and Indonesian, while time is given in any of the three languages. When used in phrases, Portuguese numerals are used with Portuguese nouns (e.g. *dia novi* (day nine) 'the ninth day (of the month)', in contrast to native Tetun *loron sia* 'day nine').

(39)

	Tetun	Portuguese	Indonesian
1	*ida*	*um*	*satu*
2	*rua*	*dois*	*dua*
3	*tolu*	*três*	*tiga*
4	*hát*	*quatro*	*empat*
5	*lima*	*cinco*	*lima*
6	*nén*	*seis*	*enam*
7	*hitu*	*sete*	*tujuh*
8	*walu*	*oito*	*delapan*
9	*sia*	*nove*	*sembilan*
10	*sanulu*	*dez*	*sepuluh*
100	*atus ida*	*cem*	*seratus*
1,000	*rihun ida*	*mil*	*seribu*
1,000,000	—[4]	*um milhaõ*	*satu juta*
'how much?, how many?'	*hira*		

Numerals from ten to just under a million follow the following pattern, where at least one of the first three bracketings must be specified, and 'Digit' covers the numerals 1 to 9. The only exception to the formula is that 'ten' is expressed by *sanulu*, not **ida-nulu*.

(40) (*rihun* Numeral) (*atus* Digit) (Digit-*nulu*) (*resin* Digit)

		sanulu	*resin hitu*	'17'
rihun rua	*atus tolu*	*rua-nulu*	*resin walu*	'2,328'
rihun hát-nulu				'40,000'

Tetun numerals and the quantifying adjective *uitoan* 'a little' can be reduplicated to indicate that the specified number of items is being considered as a group.

(41) *Ita   tau   musan   ha-hát*.
       1PI   put   seed   RDP-four
       'We put in four seeds at a time.'

When certain entities are enumerated, the Tetun numeral is typically preceded by a member of the small class of numeral classifiers. Classifiers are rarely used with *ida* 'one', and are seldom used with large numbers. The most common classifier is *nain* (also a noun meaning 'owner, master'), which is preferably but not obligatorily used when enumerating humans (e.g. *feto nain rua* (woman CLS:human two) 'two women').

Other classifiers classify the counted item by shape. These include *lolon* (lit. 'trunk, stem') for long cylindrical items such as cigarettes, *tahan* (lit. 'leaf') for thin flat objects such as cards, *fuan* (lit. 'fruit; heart') for whole roundish objects such as bread rolls, and *musan* (lit. 'seed') for tiny round objects such as tablets. For most objects (and all animals) no classifier is used.

---

[4]   There is no widely accepted term for 'million' in Tetun. The best-known term is the Indonesian *juta*, less known is the Portuguese *milhaõ* (*milhaun* in the Tetun transcription used here), while some writers use *tokon* (to most people known only in the expression *tokon ba tokon* 'numerous').

4.6.4. Other parts of speech

Prepositions can introduce peripheral prepositional phrases. Some can also introduce prepositional phrases functioning as oblique arguments, noun phrase modifiers, or predicates. Prepositions require a noun phrase complement, which cannot be fronted or omitted. Some prepositions are also classed as conjunctions, in that they can introduce clauses. Prepositions are listed in chapter 7.

Conjunctions connect words or clauses. They are of several sub-types. Coordinating conjunctions (section 11.2) coordinate a range of constituent types, including noun phrases and clauses. Subordinating conjunctions (section 11.3) introduce clauses which function as adverbial modifiers within a superordinate clause. Complementisers (section 11.4) introduce a clausal complement of a verb or noun. Relative clause markers (section 5.4) link a relative clause to the noun phrase that it modifies.

Tense–aspect–mood markers, considered in section 8.4, are grammatical words which mark the tense or aspect of a clause, or encode permission or obligation. Most cannot head a predicate. However a few 'auxiliaries' which occur in the pre-verbal slot can occur without a following verb if the context is understood; these are *bele* 'can', *lalika(n)* 'need not', *keta* 'PROHIBITIVE', and *seidauk* 'not yet'.

Tags usually follow the final clause in a sentence, being linked to it by intonation. The tag *ka* forms yes-no questions.

Interjections, too, are not part of a clause. Many form utterances on their own. Interjections include responses to preceding utterances (e.g. *lae* 'no', *sín* 'yes'), greetings and other politeness formulae (e.g. *kolisensa* 'excuse me'), exclamations (e.g. *eipá!* to express frustration, *koitadu!* 'poor thing!' to express sympathy), and hesitation markers (e.g. *bé* 'umm').

5. Noun Phrases

5.1. Basic noun phrase structure

All modifiers in Tetun Dili follow the head noun, with the exception of certain possessors (see chapter 6) and in some types of compounds (section 5.5). Those lexical modifiers that are typically short immediately follow the noun, followed by determiners, and followed in turn by longer phrasal modifiers.

Many noun phrases consist of a single noun without modifiers (see section 4.6.2). Not counting possessors and determiners, most noun phrases have at most one modifier; two is possible, while more than two is very rare.

The formula for most noun phrases headed by common nouns is as follows, where N represents the head noun, and Determiner* indicates one or more determiners.

(42) (Possessor) N (Noun.Phrase)   (Adjective.Phrase) (Determiner*) (Prepositional.Phrase)
                     (Numeral.Phrase)   (Relative.Clause)
                     (Possessor)   (Complement)

(43) *ami   nia   oan   sira   né*
       1PE   POS   child   PL   this
       'these our children'

(44) *ema   tasi-balu*
      person sea-side
      'person from overseas'

Note that noun phrases carry no case marking. Nor is there usually marking of gender. Exceptions include the Portuguese loan *bonitu* (MASCULINE) / *bonita* (FEMININE) 'nice, attractive' which always agrees with the gender of the human referent, and when whole noun phrases are borrowed from Portuguese. These tend to retain their Portuguese grammatical gender and number agreement (e.g. *primeira klase* 'first.FEMININE class.FEMININE').

Coordination of noun phrases is discussed along with other aspects of coordination in section 11.2.

5.2. Adjective phrases

Most adjective phrases consist of a single adjective. The adjective can however be preceded by a negator (e.g. *la diak* 'not good'), or followed by an intensifier (e.g. *diak liu* (good more) 'better').

(45) *ai-manas barak demais*
      plant-hot much too
      'too much chilli'

5.3. Numeral phrases

A numeral phrase consists of an optional numeral classifier followed by a numeral. Both word classes are discussed further in section 4.6.3.

(46) *feto   nain   tolu né*
      woman CLS:human three this
      'these three women'

5.4. Relative clauses

When the head noun is notionally the subject, object or time phrase for a relative clause, the relativised constituent is omitted from the relative clause, as in (47) and (48). Relativisation on other constituents, such as location, is rare, and speakers disagree as to how it should be done.

(47) *problema nebé mak   mosu*
      problem REL FOCUS appear
      'those problems which arise'

(48) *buat ida nebé ami presiza*
      thing one REL 1PE need
      'something which we need'

Relative clauses are an area of considerable variation across speakers and registers. Points of variation include the choice of relativiser, to what extent the relativiser is felt to be omissible, and how common (or acceptable) it is to follow the relative clause with *né* 'this', as in (49).

In everyday speech, the most common relativiser is *nebé* (47), which is also an interrogative pronoun meaning 'where, which'. *Bé* is common in formal writing (49), while

Portuguese *ke* is used by some people in speaking. The Tetun Terik relativiser *mak* is found in some fixed formulae (especially *mak ami respeita* 'whom we respect', used in formal listings of guests of honour during introductions to speeches). *Mak* (which elsewhere is a focus marker; described in section 11.6) also occurs in about a quarter of relative clauses introduced by *nebé*; in this context, it presents the relativised constituent as known information (47).

(49) *grupu hát be tuir   mai né*
      group four REL follow come this
      'the following four groups'

5.5. Nominal compounds

There are many nominal compounds, with the vast majority consisting of two terms. While some follow normal noun phrase structure, others (especially the activity-place compounds) do not. The major types of nominal compounds are as follows.

1. Possessor–head constructions, in which the second noun represents a part, product or other inalienable possession of the first, are common means of deriving terms for body parts (e.g. *ain kabun* 'calf of leg' from 'leg stomach'), plant parts (e.g. *ai tahan* 'leaf' from 'plant leaf'), animal and human bodily products (e.g. *inus bén* 'snot' from 'nose juice'), people from a country (e.g. *Timor oan* 'Timorese person' from 'Timor child'), and young of a species (e.g. *asu oan* 'puppy' from 'dog child'). Note that, unlike non-compounded possessive constructions, there is no possessor marker *nia* between the two nouns.

2. Activity–place compounds, consisting of a verb plus *fatin* 'place', provide conventionalised means of denoting places where the activity specified by the verb is usually carried out (e.g. *hela fatin* 'residence' from 'reside place'). Verb–object pairs may also be used as the first term in the compound (e.g. *fasi hena fatin* 'washing area' from 'wash cloth place').

3. Head–modifier compounds consist of a head noun followed by a modifier. The modifier is typically a noun (e.g. *banin feto* 'mother-in-law' from 'parent-in-law woman', *ema Dili* 'person of Dili' from 'person Dili', *fitun tasi* 'starfish' from 'star sea') or adjective (e.g. *ai manas* 'chilli' from 'plant hot'), but occasionally a verb (e.g. *ai habit* 'peg' from 'wood squeeze').

4. Generic–specific compounds consist of a generic classifier noun (usually *ai* 'plant', *manu* 'bird' or *ikan* 'fish') followed by a second term with more specific meaning (e.g. *ai kameli* 'sandalwood', *manu pombu* 'pigeon', *ikan lele* 'catfish').

5. Possession–*nain* compounds. The person who owns or controls a major entity can be designated by a noun denoting the entity, followed by *nain* 'master, owner' (e.g. *loja nain* 'shop keeper', *tós nain* '(agricultural) garden owner, farmer'). Compounds denoting someone who is recognised as being or doing a particular thing can similarly be formed by compounding a verb or adjective with a following *nain* (e.g. *kaben nain* 'married person' from 'marry owner', *matenek nain* 'expert' from 'clever owner').

6. Possessive and existential constructions

6.1. Introduction

Possession in Tetun can be expressed either within a noun phrase or within a clause. One type of clausal possession is closely related to the existential construction, while the other is closely related to nominal possession.

The same grammatical constructions are used to express a wide range of semantic relationships, including true possession ('my house'), part-whole relationships ('the car's tyres'), kin ('her father'), and more general relationships of association ('his work', 'the nation's official language').

6.2. Verbal possessive and existential constructions

Both existential and verbal possessive clauses use *iha* (which is also a general locative preposition meaning 'in, at, on...', discussed in section 7.2). Possessive clauses usually follow the standard subject–verb–object order, although some speakers also accept the far less common subject–object–verb order.

(50) *Hau la iha aimoruk.*
      1s not have medicine
      'I don't have (any) medicine.'

In existential clauses, the subject usually precedes the verb if the subject is definite, or if the clause is negative; in this order the clause indicates the presence rather than absolute existence of the subject referent (51). The reverse order, in which the argument follows the verb, occurs if the argument is indefinite (52), or if it is modified by a relative clause. This verb-initial order typically introduces new, albeit usually minor, participants into the discourse. Often there is ambiguity between this verb-initial construction and a possessive construction in which the subject is elided; in the latter case the clause specifies that something is possessed by, or available to, an unspecified person (53).

(51) *Labarik sira iha ka?*
      child PL EXIST or
      'Are the children here?'

(52) *Iha aimoruk oi-oin, ho ida-idak nia doze.*
      EXIST medicine RDP-type with each 3S.POS dose
      'There are various types of medicine, each with its own dosage.'

(53) *Halo oinsá mak sira bele mai? La iha meius ida.*
      do how FOCUS 3P can come not EXIST means one
      'How could they come? There is no way.' (or: 'They have no way.')

6.3. Nominal possessive and associative constructions

In nominal possession, there are two noun phrases: a possessed noun phrase, and a possessor. The latter is usually followed by a possessive marker.

Where a possessor functions as predicate (54), or as a headless noun phrase in any other position (55), it is followed by the possessive marker *nian* (or *ninian*).

(54) *Rai né sé nian?*
      land this who POS
      'Whose land is this?' (lit. 'This land is whose?')

(55) *Hau nian mak né.*
      1S POS FOCUS this
      'Mine is this one.'

Within a noun phrase, the possessor may either precede or follow the head noun. When a possessor follows the head noun it is marked by *nian* (or *ninian*), while a preceding possessor is marked by *nia* (or *ninia*).

(56) *Maria nia uma* = *uma Maria nian*
      Maria POS house house Maria POS
      'Maria's house'

The short form of the possessive marker (i.e. *nia* and *nian*) is required after pronouns; the exception is the 3s pronoun *nia*, whose possessive form is either *nia* or (less commonly) *ninia* (and not *nia nia*). After noun possessors, the short form of the possessive marker is by far the more common, but the long form (i.e. *ninia* and *ninian*) is used too, mainly in formal registers.

The possessive marker is not used in possessive compounds (e.g. *hudi tahan* 'banana leaf'; see section 5.5), and is usually omitted when a possessed noun specifies a location relative to the possessor noun (e.g. *iha uma laran* (LOC house inside) 'inside the house'), as described in section 7.2. Otherwise the possessive marker is normally obligatory in Tetun Dili. (This contrasts with Tetun Terik and liturgical Tetun, in which it is omissible after pronouns.)

Within noun phrases, it is more common for the possessor to precede the head than to follow it. The possessor–possessum order is obligatory when the relationship is one of inalienable possession. This includes part–whole relationships (e.g. *hau nia ulun* (1S POS head) 'my head', *ai tahan* (plant leaf) 'leaf'. It also includes location relative to an entity (e.g. *iha kareta laran* (LOC car inside) 'inside the car'), kin relationships (e.g. *Maria nia maun* (Maria POS older.brother) 'Maria's older brother'; the exception is *familia* 'family', which allows either order), and *naran* 'name' (e.g. *Senhora nia naran* (Mrs POS name) 'the lady's name'). This possessor–possessum order is also required when the possessum is a deverbal Tetun noun (e.g. *sira nia hakarak* (3P POS desire) 'their desire').

Pronominal possessors nearly always precede the possessum noun in everyday speech. The reverse order is found in the press, but is often judged to be odd by native speakers (e.g. *hau nia uma* (1S POS house) 'my house'; ?*uma hau nian*). The possessor–possessum order is also strongly preferred when a possessor noun refers to particular individuals (e.g. *Abete nia uma* (Abete POS house) 'Abete's house').

The reverse order of possessum–possessor is possible with personal possessions (e.g. *kareta UNAMET nian* (vehicle UNAMET POS) 'UNAMET vehicle'), but usually expresses a more general relationship of association between the possessor and possessum (e.g. *kultura Timor nian* (culture Timor POS) 'culture(s) of Timor', *organizasaun feto nian* (organisation woman/female POS) 'women's organisation', *aimoruk kabun moras nian* (medicine stomach sick POS) 'medicine for upset stomachs'). The term filling the possessor slot is often descriptive, and can be an adjective (e.g. *modelu uluk nian* (model past POS) 'an old model').

This possessum–possessor order is common in formal and fixed phrases, and in formal registers.

A single noun phrase can have both a preceding and a following possessor. In this case the preceding one is closer to a true possessor, while the following one indicates a more general associative relation.

(57)   *sira nia diak aban-bainrua   nian*
       3P POS good tomorrow-two.days.time POS
       'their future good'

7. Prepositions

7.1. Introduction

Prepositions can all introduce prepositional phrases functioning as peripheral constituents within a clause. Some can, in addition, introduce prepositional phrases functioning as oblique arguments of verbs, modifiers of noun phrases, or predicates. The complements of prepositions must be noun phrases, and (unlike the objects of transitive verbs) cannot be fronted or omitted.

The class of prepositions overlaps with the class of conjunctions (i.e. some terms can introduce either noun phrase complements or clauses), and with the class of transitive verbs. See section 9.3 for some terms which we have analysed as serial verbs but which are translated by prepositions in English.

7.2. Location and direction

The very common preposition *iha* introduces either location or time, and can be translated according to context as 'in, at, from, to'.

(58)   *Jacinto moris iha tinan 1957, iha distritu Manatuto.*
       Jacinto born LOC year 1957 LOC district Manatuto
       'Jacinto was born in the year 1957, in the district of Manatuto.'

Ba (the unstressed equivalent of the verb *bá* 'go') and *mai* (the prepositional equivalent of the verb *mai* 'come') can introduce goal location (59, 60), recipient (61), addressee, and beneficiary noun phrases. When any of the latter denote someone other than the speaker, *ba* must be used. When it denotes the speaker (or a group including the speaker), most Tetun speakers strongly prefer *mai* (61), although some use *ba* as in Tetun Terik.

(59)   *Halai ba foho,   hau ba Suai.*
       run to mountain 1s to Suai
       '(When we) ran away to the countryside, I went to Suai.' (said in Dili)

(60)   *Ami sai hosi Gleno, tún   mai Dili.*
       1s exit from Gleno descend come Dili.
       'We left Gleno, and came down to Dili.' (said in Dili)

(61)   *Nia fó paun mai hau.*   *Hau fó paun ba nia.*
       3s give bread come 1s   1s give bread to 3s
       'He gave me bread.'   'I gave him bread.'

Hosi 'from, originate from' expresses source location (60, 62). It also expresses the 'source' for a range of verbs, including the source material for verbs of making, and the giver for verbs of receiving (63).

(62)   *Hau hosi Ermera.*
       1s from Ermera.
       'I am from Ermera.'

(63)   *Hau simu lisaun hosi mestra   ho mestri   sira.*
       1s receive lesson from teacher.female and teacher.male PL
       'I received lessons from the teachers.'

Tó 'until, reach' is one means of introducing goal location, although *ba* and *mai* (discussed above) are more common for this function. (*Tó* is also a transitive verb 'reach', an intransitive verb 'arrive, suffice', and a temporal preposition and conjunction; see section 7.4.)

(64)   *Ami sae   bá tó   foho   leten.*
       1PE ascend go reach mountain top
       'We climbed up to the top of the mountain.'

Besik 'close to, near (location)' can introduce location for a preceding verb phrase.

(65)   *Ami hela besik kapela.*
       1PE reside near chapel
       'We stayed near the chapel.'

When the location relative to an entity is to be expressed more exactly (e.g. 'in' versus 'on'), a possessed location noun is used after the noun referring to the entity. Many of these location nouns, when used outside of locative expressions, refer to body parts. Some (not listed below) are restricted to certain semantic contexts (e.g. *iha ... tutun* 'at the summit, point, end').

The following list includes common general-purpose expressions resulting from using the preposition *iha* 'LOC' in conjunction with a location noun. Note that the same nouns are also used with other location and direction prepositions (e.g. *hosi* 'from', as in (68)) and with verbs of motion (e.g. *sae* 'ascend').

(66)

Location phrase	Meaning	Noun meaning
iha...oin	'in front of'	'face'
iha...kotuk	'behind'	'back'
iha...laran	'in, within, inside; amongst; during (time)'	'the seat of emotions, thoughts, and character'
iha...klaran	'in the middle of; between'	
iha...lét	'between; among, amongst'	
iha...sorin	'beside'	'side'
iha...sorin bá	'on the other side of'	'side go'
iha...sorin mai/né	'on this side of'	'side come/this'
iha...sorin/liman lós	'on the right side of'	'side/hand right'
iha...sorin/liman karuk	'on the left side of'	'side/hand left'
iha...leten	'on, on top of, above'	'upper'
iha...kraik	'at the base of; under, below'	'lower'
iha...okos	'underneath, under, below'	
iha...ninin	'at the edge of'	'edge'

(67)   *Sira túr iha ita nia oin.*
      3P   sit LOC 1PI POS face
      'They sit in front of us.'

(68)   *Sira sai hosi kareta laran.*
      3P   exit from car   inside
      'They got out of the vehicle.'

7.3. Accompaniment (with)

The preposition *ho* 'with' introduces a wide range of semantic roles, including accompaniment (69), instrument (e.g. *ho tudik* 'with (a) knife'), manner (e.g. *ho páz* (with peace) 'peacefully'), and means of transport (e.g. *ho kareta* 'by car'). *Ho* is also a coordinator meaning 'and' (see section 11.2.1). A Portuguese loan *sein* 'without' is used as an antonym of *ho* by some speakers.

(69)   *Nia halimar hamutuk ho labarik sira seluk.*
      3S play   together with child   PL   other
      'She plays together with the other children.'

7.4. Time

The temporal prepositions discussed here are all also conjunctions; that is, they can introduce either noun phrases or clauses. The other temporal preposition is *iha* 'LOC, at (time)', which, along with the expression *iha ... laran* (LOC ... inside) 'during', is discussed in section 7.2. There are also some temporal conjunctions which are not prepositions (see section 11.3). In addition, temporal relationships may be indicated by tense–aspect–mood markers (discussed in section 8.4), and temporal connective adverbs (e.g. *depois, hafoin* 'and then').

The concept 'before (time)' may be expressed by the Portuguese loan *antes* or *antes de* (e.g. *antes setenta* (before seventy) 'before (19)70'), or by the primarily liturgical and formal preposition *molok*.

(70)   *Molok sira tama, sira tenki koalia lai.*
      before 3P enter 3P must speak FIRST
      'Before they go in, they must talk.'

'During', especially in formal and written contexts, may be expressed by *duranti* (or *durante*) 'for (period of time), during, when'. An alternative is to use an expression ending in *nia laran* 'POS inside', or the combination *duranti ... nia laran* 'during ... POS inside'.

(71)   *Ami hela iha ai-laran durante tinan tolu (nia laran).*
      1PE reside LOC jungle   during year three POS inside
      'We lived in the jungle for three years.'

While there are several alternatives for expressing 'after', it is much more common to put clauses in chronological order, and optionally link them with a connective adverb meaning 'then' (e.g. *depois, hafoin*). One alternative is the preposition *depois*, or the more Portuguese-influenced *depois de* (e.g. *depois de referendum* 'after the referendum'). Another is the somewhat verbal *liu* 'after, pass' (e.g. *liu semana tolu* (pass week three) 'after three weeks').

The concept 'from, since' can be expressed by the Portuguese loan *dezde* (e.g. *dezde dominggu* 'since Sunday') or by two prepositions and conjunctions which derive from verbs meaning 'start': the Portuguese loan *komesa* and the more liturgical or purist *hahú* (which appears to be gaining currency in everyday speech).

(72)   *Hau servisu komesa tuku walu, tó tuku hát.*
      1S work start   o'clock eight until o'clock four
      'I work from eight o'clock to four o'clock.'

'Until' can be expressed by *tó* 'until, up to', illustrated in (72).

7.5. Other

Kona-ba 'about, concerning (a topic)' (lit. 'touch to') introduces a topic of speech or thought.

(73)   *Hau mós halo intervista kona-ba fronteira.*
      1S also do   interview about   border
      'I also conducted an interview about the (East Timor–Indonesia) border.'

Hanesan 'like' indicates likeness in some respect. As a preposition, it follows a verb phrase. It is also a verb 'be like', and an expression meaning 'for example, for instance, like'.

Nudar 'in the capacity of, as' is primarily liturgical and formal. It often introduces that part of a referent's identity that is particularly relevant to the current discussion.

(74)   *Nudar ema nasionalista, hau senti laran susar tebes, ...*
      as   person nationalist 1S feel inside distress truly
      'As a nationalist, I feel truly distressed, (because this current policy is not good).'

Entre 'between, among' primarily introduces co-participants for an abstract noun (e.g. *dialogu entre sira* 'dialogue amongst them'). It is primarily used by more Portuguese-influenced speakers.

Tuir 'according to (one's opinion/desire, an information source, criterion, ability, person...)' frequently specifies the justification for a statement or reason for an action. It is also a transitive verb meaning 'follow, obey'.

(75) *Nia haruka ita hamulak, tuir   ita ida-idak nia fiar.*
      3S order 1PI pray   follow 1PI each   POS belief
      'He orders us to pray, each according to his own beliefs.'

Konformi (or *konforme*) 'according to, depending on (a person, a source of data, one's desires or abilities...)' is used in Tetun as both a preposition and conjunction. It is very similar to *tuir*.

(76) *Diak liu   fó   susu ba bebé konformi   nia hakarak.*
      good more give milk to baby according.to 3S want
      'It's better to give milk to the baby according to the baby's wishes.'

Tanba and *tan* 'because' are primarily conjunctions (see section 11.3), but can also introduce noun phrases (e.g. *tanba né* 'because.of this').

8. Clauses

8.1. Verbal clauses

8.1.1. Basic word order

Tetun verbal clauses have subject–verb(–object) word order (although objects may be fronted, see below). Unlike Tetun Terik, Tetun Dili has no subject marking on verbs.

(77) *Nia dada hau nia liman.*
      3S pull 1S POS arm/hand
      'He pulled my arm.'

Oblique arguments are usually introduced by prepositions (see chapter 7), and typically follow the object slot. Such arguments include goal (78), source, recipients (79), addressees, and instruments. Instruments and some direction phrases can alternatively be introduced by serialised verbs, discussed in chapter 9. Certain complement clauses also occur in this position (see section 11.4).

(78) *Ami seidauk haruka nia ba eskola.*
      1PE not.yet send   3S to school.
      'We haven't yet sent her to school.'

Recipients are usually introduced by the prepositions *ba* 'to (away from speaker, not towards speaker)' or *mai* 'come (towards speaker)', with the prepositional phrase following the object, as in (79). The object can however follow the recipient, particularly if the object noun phrase is long.

(79) *UNICEF fó   kadernu   mai ami.*
      UNICEF give exercise.book come 1PE
      'UNICEF gave us exercise books.'

Subjects and objects can readily be omitted if the referent can be inferred from the context, or if the referent is irrelevant. Determining the intended referents of omitted arguments seems to be based on pragmatics rather than on grammatical rules. An exception is that in serial verb constructions and reduced complements, the subject is obligatorily omitted, and necessarily co-referential with either the subject or the object (depending on the construction) of the preceding verb.

The object can be fronted to before the subject slot, in order to mark the object as contrastive or topical. In the former case it is often marked with the focus marker *mak(a)* (80) (see section 11.6); in the latter, with anaphoric *né* 'this' (81).

(80) *... So   Maria nia pozisaun deit maka ita la hatene.*
      only Maria POS position just FOCUS 1PI not know
      '(We know that Simão supports Portuguese, and João supports English.) It is only Maria's position that we don't know.'

(81) *Ami halai, ami lori kuda ida. ... kuda né ami bá troka   fali   hahán.*
      1PE run   1PE bring horse one   horse this 1PE go exchange in.turn food
      'When we ran away, we brought a horse. (When our food ran out), this horse, we went and exchanged for food.'

Unlike the clauses discussed above, the single argument in existential clauses can either precede or follow the verb. Verbal existential and possessive clauses are discussed in section 6.2. Under the influence of Portuguese, subjects of a number of other intransitive verbs can follow the verb, if the subject is not definite and presents a new participant into the discourse. Such verb-initial constructions occur in writing and formal speech, but are rare in informal speech. Verbs allowing following subjects include *akontese* 'happen', *moris* 'live' and *mosu* 'appear'.

(82) *Maibé se ita uza lian   né, sei   mosu problema barak.*
      but   if 1PI use language this FUTURE appear problem many
      'But if we use this language, many problems will crop up.'

8.1.2. Reciprocals

A reciprocal construction is formed by placing *malu* 'RECIP, each other' in a clause's object, addressee or recipient slot. The regular reciprocal construction has a (semantically) plural subject, and indicates that at least one of the referents of the subject is acting towards one or more of the others in the way described by the verb. There is no overt indication as to whether the referents are all acting simultaneously (e.g. *halibur malu* 'gather.together RECIP'), or not (as one would expect of *husu ba malu* 'ask to RECIP'). To use this construction, the referents must have comparable status with regard to the specified activity. For example, *Sira hanorin malu* '3P teach RECIP' could be used of teaching amongst adults, but would not be appropriate for talking about adults teaching little children.

In a second type of reciprocal construction, one of the participants is presented as the instigator of a shared action. In this case, the subject refers to the instigator, and secondary participants are introduced by *ho* 'with' following the reciprocal marker *malu*. Example (83) illustrates the contrast between the two reciprocal constructions.

(83) *Adino baku malu ho Pedro.*
      Adino beat RECIP with Pedro
      'Adino and Pedro beat each other (Adino started it).'

Adino ho Pedro baku malu.
Adino and Pedro beat RECIP
'Adino and Pedro beat each other. (There is no indication as to who started it).'

Some fixed phrases using the reciprocal construction include: *diak malu* (good RECIP) 'be on good terms with one other'; *hadau malu* (snatch RECIP) 'compete for resources'; *diskuti malu* (discuss/argue RECIP) 'discuss, argue with each other'; *fahe malu* (divide/separate RECIP) and *haketak malu* (separate RECIP) 'separate from each other'. The expression *hán malu* (eat RECIP), when used with a human subject means 'fight (by any means), argue'; when the subject refers to what has been said, it means the expressed views do not agree with one another.

8.1.3. Reflexives

Tetun Dili reflexives are formed in two related ways. In the first, the noun *án* 'self' is possessed by a pronoun which has the same person and number as the subject. Most such noun phrases function as direct object; however *án* can also occur in addressee or beneficiary prepositional phrases. In the second type of reflexive, *-an* (an unstressed variant of *án*) is a clitic attached to the verb; this type is restricted to reflexivisation on the object. The synonymous examples below illustrate these alternatives. Note that *rasik* 'own', in the example below, is not uncommon in reflexives; it further stresses that it was the subject referent himself or herself who performed the action denoted by the verb.

(84) *Nia oho rasik nia   án.* = *Nia oho-an rasik.*
      3S kill own 3S.POS self   3S kill-self own
      'He himself killed himself.'

The two variants are not always interchangeable. Expressions with *-an* which do not have a variant with a possessed *án* include *foti-an* (lift self) 'boast', *bók-an* (move self) 'move', *haraik-an* (lower self) 'be humble, humble oneself', and translations of Portuguese reflexives such as *esforsa-an* 'strive' (from Portuguese *esforçar-se*).

8.2. Non-verbal clauses

In non-verbal clauses, the predicate is juxtaposed to the subject, without a copula. Non-verbal predicates include noun phrases (85), numeral phrases, prepositional phrases (87), and 'body-good' expressions (see section 4.4). In clauses with nominal predicates, the old information is usually given in the first noun phrase, and the new information in the second (85); however the reverse is also possible, in which case the fronted 'new' information is generally focused by *mak* (86).

(85) *Hau doutór.*
      1S   doctor
      'I am (a) doctor).'

(86) *Hau mak   doutór.*
      1S   FOCUS doctor
      'It is me who is the doctor.'

(87) *Nia laós   iha Dare.*
      3S indeed.not LOC Dare.
      'He's not in Dare.'

8.3. Negation

The two basic negators are *la* and *laós*, both of which immediately precede the negated constituent. *Laós* may be used to negate all types of constituents, including noun phrases (88), adjectives, verbs, time phrases, prepositional phrases (87) and dependent clauses. It is strongly contrastive, especially before verbs and adjectives, where *la* is the norm.

(88) *O hanoin deit ema   nia hakarak, laós   Nai Maromak nia hakarak!*
      2S consider just person POS desire indeed.not Lord God   POS desire
      'You consider only people's wishes, not God's wishes.'

La, on the other hand, is an unstressable proclitic, and is used only to negate verbs and adjectives (e.g. *la tama* 'not enter', *la bót* 'not big').

Ladún 'not very' similarly precedes the adjective or verb which it modifies (e.g. *ladún todan* 'not.very heavy').

Lae 'no' is a negative pro-clause. It is often used as the second alternative in a yes–no question. *Lae*, or more commonly *lae ida*, is also used as a negative answer to a yes–no question, to contradict a previous positive statement, or to reinforce a preceding negative statement.

(89) *"Todan ka lae?"*   *"Lae ida."*
      heavy or no   no one
      "(Is it) heavy?"   "No."

Nunka 'never' (from Portuguese *nunca*) negates a following verbal or adjectival constituent (e.g. *nunka servisu* 'never work').

In speaking (but never in writing) a negated term is often followed by *ida* (lit. 'one'). This is most common with *la* (90), but also occurs with other negative terms such as *lakohi* 'not want', *nunka* 'never', *seidauk* 'not yet' and *lae* 'no' (89). Usually the negated term is a single-word adjective or verb, but *ida* can also follow the object in a negative clause. This *ida* does not co-occur with tense–aspect marking (e.g. *ona* 'ANT', *atu* 'IRR').

(90) *Bosok! Ami la fiar   ida!*
      lie   1PE not believe one
      '(You are) lying! We don't believe (it)!'

8.4. Tense, aspect and mood

Verbs in Tetun are not marked for tense, aspect or mood. Such distinctions can instead be revealed through the range of optional tense–aspect markers listed below. Some of these appear before the verb (marked 'pre-verbal' below) and some after ('post-verbal'). Frequently, however, it is context that determines the tense–aspect interpretation. The sentence *Hau bá* 'I go', for instance, can be interpreted as 'I went', 'I am going', or 'I will go'. The basic set of tense–aspect–mood markers are discussed below. Others which are not discussed include: *lalika(n)* 'need not, should not, don't' (pre-verbal), *keta* 'PROHIBITIVE', *kedas* 'immediately, straightaway; in advance, beforehand' (post-verbal), and *nafatin* 'continue, keep on' (post-verbal).

Sei 'FUTURE' (pre-verbal): *Sei* marks something as definitely happening in the future. The negation of future *sei* is *sei la* 'will not'.

(91) *Tuir   mai hau <u>sei</u>   haktuir terus   no susar   ...*
      follow come 1S FUTURE relate   suffering and difficulty
      'Next I will recount the sufferings and difficulties (that we experienced in August 1999).'

Sei 'still' (pre-verbal): *Sei* can also mean 'still'. In this case, its negation is *seidauk* 'not yet'.

(92) *"Modo   sei iha?"*   *"Modo   iha   hela."*
      vegetable still EXIST   vegetable EXIST CONT
      "Are there still any vegetables?" "There are."

Atu 'IRRealis, want to, intend to, be going to, about to' (pre-verbal): *Atu* usually indicates a desire, intention or imagination of a future event. While it occurs mostly with non-stative verbs (93), it can occur with states as well (94). Unlike *sei* 'FUTURE', *atu* can be used if the (once) future event is known not to have come about, and so is a common way of marking counter-factuality. *Atu* can also occur with complements which are not necessarily desired or intended, but which are nevertheless about to happen; in this case it marks immediate future (95). It is frequently found in purpose clauses, and in complements of a wide range of verbs and nouns, including those of wanting (e.g. *gosta* 'like', *lakohi* 'not want, refuse'), ability (e.g. *biban* 'manage'), and fear (e.g. *barani* 'brave, dare').

(93) *Hau bá, tinan walu, <u>atu</u> halo sia.*
      1S go year eight IRR make nine
      'When I went, I was eight years old, about to turn nine.'

(94) *... sira sei kík,   antes sira <u>atu</u> bót.*
      ... 3P   still small before 3P IRR big
      '(We want to discipline our children while) they are still small, before (they) are older.'

(95) *Nia <u>atu</u> mate.*
      3S IRR die
      'He is about to die.'

Tiha (frequently pronounced [tia]) 'PeRFective' (post-verbal): *Tiha* is usually used with non-stative verbs, indicating that the activity has been completed. With stative verbs and adjectives, *tiha* indicates that the state has been entered. For some speakers, *tiha* can also be used in commands (e.g. *Túr tiha!* (sit PRF) 'Please sit!').

(96) *Hau lés <u>tiha</u> hau nia ropa.*
      1S tear PRF 1S POS clothes
      'I tore my clothes.'

Ona 'ANTerior, already' (post-verbal): *Ona* usually indicates that the event/state is underway or has already happened, and that its happening is still relevant. So, example (97) does not indicate when the person came, but does imply that he is still here or still on the way. *Ona* is also used in statements of what must be done (e.g. *Mai ona!* (come ANT) 'Come here!') Some speakers also use *ona* in informal speech to mark past time in general.

(97) *Nia mai <u>ona</u>.*
      3S come ANT
      'He has come.' (or 'He is on his way here.')

Tiha ona 'PERFECT' (post-verbal): This common sequence indicates perfect aspect. It is mainly used in intransitive non-stative clauses, indicating that the activity has finished and still has effect (98). Its function thus overlaps with *ona* 'ANT', although there are differences between them. For example, *ona* readily occurs with states, whereas *tiha ona* seldom does. With the same verb, *ona* will suggest a stative interpretation and *tiha ona* an activity, as illustrated by the difference between examples (99) and (100). In other examples, *tiha ona* has a past perfective function, while *ona* expresses simple past.

(98) *Molok o mai, hau hán <u>tiha</u> <u>ona</u>.*
      before 2S come 1S eat ANT PRF
      'Before you came, I had already eaten.'

(99) *Nia kaben <u>ona</u>.*
      3S marry ANT
      'She is married (is in the state of marriage).'

(100) *Nia kaben <u>tiha</u> <u>ona</u>.*
      3S marry PRF ANT
      'She has already married' (e.g. appropriate if one is asking whether the anticipated wedding has happened yet)

Hotu 'be finished' (post-verbal): When this intransitive verb meaning 'be finished' immediately follows a non-stative intransitive verb, it indicates that the action denoted by the verb has finished. *Hotu* is also a quantifier meaning 'all' and an adverb 'also'.

(101) *Sira tiru <u>hotu</u>, mak ita foin bele fila mai uma.*
      3P   shoot finish FOCUS 1PI just can return come house
      'Only after they had finished shooting, would we be able to return home.'

Foin 'JUST, only just, very recently' (pre-verbal): When used with punctual verbs, *foin* indicates that the activity has been very recently completed (102). With stative verbs, it indicates that the state has recently been entered. With non-stative durative verbs like *hán* 'eat', *foin* may be interpreted as indicating either that the activity has just finished, or that it is only just in progress, as illustrated by the alternative translations for (103). The sequence *foin atu* 'JUST IRR' means 'just about to'.

(102) *Sira <u>foin</u> tó   iha Dili.*
      3P   JUST arrive LOC Dili
      'They have only just arrived in Dili.'

(103) *Ami <u>foin</u> hán.*
      1PE JUST eat
      'We have just eaten.' or 'We are just in the process of eating.'

Daudauk (or *dadauk*, *dadaun*, *daudaun*) 'PROGressive' (post-verbal): *Daudauk* occurs with non-stative durative verbs which involve steady progress towards an end-point (e.g. *halai* 'run', *hemu* 'drink'), and indicates that the activity is in progress at the reference time. It can

be used in bi-clausal constructions to indicate that the events of one clause are progressing at the same time as the events of the other clause.

(104) *Hau la bele sai tanba   hau sei hanorin daudauk Senhora tó   tuku   rua.*
      1S not can exit because 1S still teach   PROG   Mrs   until o'clock two
      'I can't go out because I'm teaching this lady until two o'clock.'

Hela 'CONTinuous' (post-verbal): Like *daudauk*, *hela* indicates that a condition is ongoing at the reference time, and is frequently used in bi-clausal constructions to mark the fact that the two clauses are simultaneous. *Hela* is however more general in that it can be used with stative predicates as well as dynamic ones. *Hela* is also an intransitive verb meaning 'reside, stay'.

(105) *Ohin   hau hetan nia, mais la koalia ho nia, tanba   nia animadu hela.*
      just.now 1S get   3S but not speak with 3S because 3S busy   CONT
      'I met him just now, but didn't talk with him, because he was busy.'

Bele 'can, may' (pre-verbal): *Bele* indicates ability or permission. In purpose clauses, it frequently marks future expectation, as in the final clause of (107).

(106) *Nia bele kura ema   nia moras.*
      3S can cure person POS sickness
      'He could cure people's illnesses.'

Tenki 'must' (pre-verbal): This loan (from Portuguese *tem que* lit. 'it.has that') usually indicates obligation, but can also be used in the sense of 'be inevitable, be necessary'.

(107) *Sira dehan katak ami tenki koalia Portugés,   para bele aprende lai-lais.*
      3P   say   that 1PE must speak Portuguese so.that can learn   RDP-quick
      'They (the teachers at boarding school) said we had to speak Portuguese, so that (we) would learn it quickly.'

Lai 'FIRST, before doing something else' (post-verbal): *Lai* indicates that something must be done before some other (explicit or implicit) event occurs (108). It is also commonly used in polite invitations, requests and commands. A negated clause with *lai* means 'not yet', and is appropriate for a temporary prohibition.

(108) *Ita deskansa lai.*
      1PI rest   FIRST
      'We'll rest for a while (before...)'

8.5. Intensifiers, comparatives, and superlatives

Most intensifiers follow the verb or adjective which they modify (e.g. *malirin demais* (cold too) 'too cold', *tauk atu-mate* (afraid IRR-die) 'scared to death'). The exception is the informal *para* 'so', which precedes the modified term.

(109) *Malae sira ain para ás!*
      foreign PL leg so   tall
      'The foreigners are so tall!'

Speakers vary with respect to which intensifiers they prefer to use under which circumstances. The common intensifiers are:

(110) *demais*   'too (much), excessive' (Portuguese loan)
      liu   'more, extremely, completely, most' (also Vt 'pass')
      tebes, tebe-tebes 'truly, indeed' (primarily formal)
      lós   'very, truly' (also Adj 'straight, true, right')
      atu-mate   'very' (lit. 'IRR-die') (not formal; especially for negative qualities)
      la-halimar   'very, really' (lit. 'not-play/joke') (informal)
      liu tán   'even more' (lit. 'more in.addition')

The intensifier *liu* 'more, extremely, most' is also used to form comparatives and superlatives. The object of comparison may be introduced directly by *liu*, especially when it is the subject referent that is compared with something else (111). Alternatively it may be introduced by *duké* (from Portuguese *do que*), particularly in more complex comparisons, such as (112). Although *liu* can be translated as either 'more' or 'most', a superlative interpretation can be forced by ensuring that the description would fit only one referent (e.g. *ikus liu* (last more) 'very last', since 'last' is not a gradient concept) or focusing the entity or the description with *mak* (113).

(111) *Kanek bót át liu kanek kík.*
      wound big bad more wound small
      'A big wound is worse than a small one.'

(112) *Diak liu   ita hán lai, mak bá, duké orsida hamlaha iha dalan.*
      good more 1PI eat FIRST FOCUS go   than soon   hungry LOC road
      'It is better to eat before going, than to later be hungry on the way.'

(113) *Nia mak   bót liu.*
      3S FOCUS big more
      'It is he who is most senior.'

8.6. Manner modifiers

Manner modifiers follow either the verb they modify or the object. Most manner adverbs are also classified as adjectives (e.g. *diak* 'well, successfully; good', *makás* 'strongly, loudly, fast...; strong, loud, fast...'). These can be modified by an intensifier or negated by *la* 'not'.

(114) *Nia tesi justisa la lós.*
      3S cut justice not true
      'He judges unjustly.'

A small number of manner adverbs are derived by reduplicating an adjective (e.g. *di-diak* 'carefully, thoroughly' from *diak* 'good; well, successfully').

(115) *Lao di-diak!*
      walk RDP-good
      'Walk carefully!'

Other ways of specifying manner include a manner adverb introduced by *halo* 'do, make' (e.g. *fasi halo di-diak* (wash do RDP-good) 'wash well'), or by a prepositional phrase introduced by *ho* 'with' (116).

(116) *Ita lao   ho ansi deit.*
      1PE walk/go with hurry just
      'We go in a hurry (without preparation).'

9. Serial Verbs

9.1. Introduction

Tetun Dili serial verb constructions consist of two verbs which co-occur in a single clause, neither of which is subordinate to, nor modifies, the other. The verbs share tense, aspect, negation, illocutionary force, and at least one argument. In particular, the subject of the second corresponds to the subject, object or recipient of the first verb, depending on the construction. In most of the constructions, one of the two verb slots is restricted to a limited set of verbs, while the other is more open.[5] Note that causative serial verb constructions are discussed along with other causative constructions in chapter 10.

9.2. Motion–action

In motion–action serialisation, the two verbs specify two separate, sequential, actions, of which the first is *bá* 'go (away from, or at least not towards, the speaker)' or *mai* 'come (towards the speaker)'. This construction usually implies that the motion of going or coming was done in order that the activity of the second verb can be carried out. Adverbs can follow either verb in this construction.

(117) *Hau bá harís lai-lais   bebé.*
       1s   go bathe RDP-quick baby
       'I'll go and quickly bathe the baby.'

(118) *Malae   UNAMET sira haruka ami mai fali sae   kareta.*
       foreigner UNAMET PL   order 1PE come again ascend vehicle
       'The UNAMET (UN agency running the 1999 referendum) foreigners ordered us to come back and get into the vehicles.'

9.3. Motion–direction

There are several types of motion–direction serialisation. In each, the first verb is typically one of motion (e.g. *monu* 'fall') or causation of motion (e.g. *haruka* 'send'), while the second specifies the direction of motion.

In one sub-type, an intransitive verb of motion is followed immediately by one of a small class of intrinsic direction verbs (*tún* 'descend', *sae* 'ascend', *tama* 'enter', *sai* 'exit').

(119) *... kafé né bele monu tún   tiha ba rai   deit.*
       coffee this can fall descend PRF to earth just
       '(If the coffee beans hang too long on the tree) the coffee can just fall down to the ground.'

In the second, direction relative to the speaker is specified by *bá* 'go' or *mai* 'come'. These deictics are only loosely tied to the preceding verb, and follow any adverbs.

---

[5]   These Tetun constructions thus have features consistent with those discussed for serial verbs in many publications (e.g. Crowley in press; Durie 1997).

(120) *Ami tún   fali mai.*
       1PE descend again come
       'We came back down.'

Bá and *mai* can also co-occur with verbs of bringing and giving, to indicate which direction the object is moving in (121), and with verbs of speaking, to specify the direction of the speech (e.g. *hatete mai* (tell come) 'tell me/us').

(121) *Fó livru mai lai!*
       give book come FIRST
       'Please give me/us the book!'

Finally, there is a range of other transitive and intransitive verbs which occur in constructions with a preceding verb of motion or posture, and which in some way further specify its direction, path or position. These include *haleu* 'surround, encircle, around' (122), *liu* 'pass, go past', *liu hosi* (pass from) 'via' (123), *tuir* 'follow, along' (124), and *hakat* 'cross'. This class also includes *hasoru* 'meet, oppose' (125) and *kontra* 'oppose' (from the Portuguese preposition *contra*), which both also occur in constructions with non-motion verbs (126).

(122) *Sira túr haleu   nia.*
       3P   sit surround 3s
       'They sat around him.'

(123) *Favór ida, esplika hela mai hau liu hosi email.*
       favour one explain leave come 1s pass from email
       'Please explain (it) to me via email.'

(124) *Sira lao tuir   tasi ibun.*
       3P   walk follow sea mouth
       'They walked along the seashore.'

(125) *Nia lao hasoru tráfiku.*
       3s walk meet   traffic
       'He walks against the direction of traffic.'

(126) *Brazil joga kontra IMPETTU.*
       Brazil play oppose IMPETTU
       'Brazil played (soccer) against IMPETTU (a Timorese youth organisation).'

9.4. Instrument–action

In instrument–action serialisation, *lori* or *hodi*, both verbs meaning 'carry, bring, take', introduce an instrument for a following verb (127, 130). The initial verb here has some verbal characteristics (e.g. in that its object can be fronted).

(127) *Nia lori hena kesi kanek.*
       3s take cloth tie wound
       'She used cloth to bind up the wound.'

The instrument phrase can alternatively follow the other verb, as in (128). In this construction, however, *lori* and *hodi* are prepositions (e.g. in that the complement cannot be fronted or omitted).

(128) *Sira selu hahán <u>hodi</u> dolar.*
       3P   pay food take dollar
       'They pay for food with (American) dollars.'

9.5. Cause–effect

In cause–effect (Durie 1997:331) or resultative serialisation, the second verb specifies the result of the activity described by the first verb. This second slot appears to be lexically restricted, and includes the directionals, such as *sai* 'exit' in (129), and *mate* 'die, dead' (130).

(129) *La bele ajuda <u>dada</u> <u>sai</u> labarik.*
       not can help pull exit child
       'You cannot help pull out the baby (when assisting in childbirth).'

(130) *Ita hodi aimoruk né <u>rega</u> <u>mate</u> tiha dút   sira né.*
       1PI use medicine this spray dead PRF grass/weed PL   this
       'We used this pesticide to spray these weeds so they died.'

Intransitive *hela* 'stay, reside' follows transitive transfer verbs (e.g. of giving, putting, leaving, or throwing), to indicate that the object remains in the new location, at least temporarily.

(131) *La bele <u>rai</u> <u>hela</u> labarik nia hahán mak   hán.*
       not can put stay child   POS food FOCUS eat
       'You cannot set aside the baby's food for a while before (the baby) eats.'

10. Causation

10.1. Introduction

Causative constructions in Tetun Dili can be formed in various ways: by using the prefix *ha-*; by using one of a number of serial verb constructions; and through periphrastic causative constructions formed with *halo* 'make, do'.

10.2. Prefixation with *ha-*

Causative transitive verbs can be derived from adjectives and instransitive verbs by the prefix *ha-* (e.g. *halakon* 'get rid of' from *lakon* 'disappear', *hamaus* 'quieten, subdue' from *maus* 'quiet, tame'). Typically, the resulting derivation has a human agent, and a non-human patient, or a human patient who has no control over the situation (such as a baby).

There is considerable inter-speaker variation as to the number of roots which accept *ha-*, and in whether *ha-* also derives verbs for involuntary activities. This variation is largely due to the varying degrees of influence of Tetun Terik, in which *ha-* is far more productive. Some speakers for whom *ha-* is productive apply it to Portuguese roots as well (e.g. *haforsa* 'strengthen' from *forsa* 'strong'), while other speakers reject such derivations.

10.3. Serialisation with *halo* 'make, do'

Causation may also be expressed by a serial verb construction consisting of *halo* 'make, do' immediately followed by a verb or adjective indicating the caused event. The object follows the second verb.

(132) *Ita tenki <u>halo</u> <u>moris</u> fali ita nia unidade.*
       1PI must make live   again 1PI POS unity
       'We must re-establish our unity.'

There are some roots for which speakers alternate freely between the *ha-* and the *halo* constructions, with no apparent difference in meaning (e.g. *habadak*, *halo badak* 'shorten', from *badak* 'short'). There are other roots (e.g. *foer* 'dirty'), for which some speakers accept a derivation with *ha-*, while others use only *halo*. For yet other roots, causation in Tetun Dili is only ever expressed by *halo* (e.g. *halo kole* 'make tired').

10.4. Periphrastic causative with *halo* 'make, do'

There is another, periphrastic, construction with *halo*, which has more syntactic and semantic freedom than do the previous two constructions. In this construction *halo* 'make, do' has an object noun phrase, following which is a predicate expressing what the object referent has been caused to do or be. This periphrastic construction is readily used for indirect causation, and for inanimate causers (e.g. 'the poison made him die').

(133) *Buat sira né <u>halo</u> hau moe   tún   ba sidade.*
       thing PL   this make 1S ashamed descend to city
       'These things make me ashamed to go down to the city.'

10.5. Serialisation with *fó* 'give'

There are also two causative constructions based on the verb *fó* 'give'. Some verbs can be serialised to an immediately preceding verb *fó* 'give' to create standard causative expressions (e.g. *fó empresta* (give borrow) 'lend'; *fó harís* (give bathe) 'bathe'; *fó hatene* (give know) 'inform').

(134) *Hau seidauk <u>fó</u>   <u>hán</u> bebé.*
       1S not.yet give eat baby
       'I haven't fed the baby yet.'

(135) *Ami <u>fó</u>   <u>aluga</u> ami nia uma ba malae   Tailándia.*
       1PE give rent 1PE POS house to foreigner Thailand
       'We are renting out our house to Thai people.'

The second *fó* construction specifies what is given (by the object of *fó* 'give'), to whom (by a beneficiary phrase), and what the recipient is supposed to do with the item given (by the second verb in the construction). There is a strong implication that the recipient did indeed carry out the purpose for which he or she was given the item.

(136) *Nia <u>fó</u> fali nia   uma né ba ema   seluk <u>aluga</u>.*
       3S give again 3S.POS house this to person other rent
       'He instead gave his house to someone else to rent (and that person rents it).'

11. Sentences

11.1. Introduction

This chapter deals with structures beyond the basic clause, including various means of linking and conjoining clauses, topics, focus, commands and questions. Note that relative clauses are discussed in section 5.4.

11.2. Coordination

11.2.1. Conjunction

Ho (also a preposition meaning 'with', discussed in chapter 7.3) primarily coordinates noun phrases (e.g. *Tetun ho Portugés* 'Tetun and Portuguese') and adjectives (e.g. *mean, mutin ho azúl* 'red, white and blue'). A second conjunction *i* (from Portuguese *e* 'and') is mainly used to coordinate clauses, sentences, and non-stative verb phrases (137). Tetun Terik *no* is widely used (often to the total exclusion of the other coordinators) to coordinate all types of constituents in liturgical and other formal settings, but is uncommon in everyday speech in Tetun Dili. Clauses and verb phrases can also be coordinated simply by juxtaposing them, particularly in speech (138).

(137)   *Ami komesa buka fali   malu, i   komesa hanoin   atu hari.*
        1PE begin   seek again RECIP and start   consider IRR erect
        (When we returned from being evacuated to Darwin:) 'We started looking for each other again, and started thinking about setting up (this NGO).'

(138)   *..., iha disiplina, iha edukasaun, morál, buat sira né hotu.*
           have discipline, have training,   morals, thing PL this all
        '(We discipline our children, so that when they reach school, they) have discipline, have good upbringing, morals, all these things.'

Contrast may be indicated by *maibé* 'but' (used in all registers) or *mais* 'but' (based on Portuguese *mas*).

(139)   *Ami atu bá Ainaro, maibé la bele.*
        1PE IRR go Ainaro but   not can
        'We wanted to go to Ainaro, but couldn't.'

11.2.2. Disjunction

The usual marker of disjunction in spoken Tetun Dili and in liturgical Tetun is *ka* 'or'. The Portuguese loan *ou* is more common in non-liturgical writing. Both can coordinate a wide range of constituents, including noun phrases, adjectives, predicates and clauses. Yes–no questions use a final *ka lae* 'or not' or simply a final *ka* (see section 11.8.2).

(140)   *Baku ka, hakilar ka, tolok. Né hau lakoi.*
        beat or scream or insult this 1S refuse
        (I'm unwilling to discipline my child by rough means.) '(Like) by hitting, or shouting, or saying rude things. I am unwilling to do that.'

(141)   *Ninia rezultadu hanesan ou besik malu.*
        3S.POS result   same   or near RECIP
        'His results are the same or similar.'

11.3. Adverbial clauses

Adverbial clauses specifying time, condition, and concession usually precede the main clause, while those encoding reason or purpose tend to follow it.

Although adverbial clauses may be simply juxtaposed to the main clause, with no conjunction (142), it is more common for such clauses to be introduced by conjunctions.

(142)   *Azeiti   la iha, mina Bemoli   mós diak.*
        olive.oil not exist oil   k.o.cooking.oil also good.
        'If (you) don't have olive oil, *Bemoli* oil is fine too.'

Many temporal conjunctions meaning 'before', 'during', 'after', 'since' and 'until' are also prepositions, and are discussed in section 7.4. Others temporal conjunctions include *bain-hira* 'when, whenever', *kuandu* 'when, whenever, if' (from Portuguese *quando* 'when') and *momentu* 'while' (from the Portuguese noun *momento* 'moment').

(143)   *Bain-hira nia tama, ami hán hela.*
        when   3S enter 1PE eat CONT
        'When he came in, we were eating.'

The conditional conjunction is *se* 'if' (from Portuguese). An alternative (common in the liturgical register) is to use the clause-final adverb *karik* 'perhaps'. Another alternative is the combination of *se ... karik*.

(144)   *Se imi hakarak, imi sae   deit mai iha Ailoklaran.*
        if 2P want   2P ascent just come LOC Ailoklaran.
        'If you want, just come up here to Ailoklaran.'

Concessive clauses can be introduced by the conjunctions *maski* (or *maske*) 'although, even though' (from Portuguese, mainly liturgical and formal), *biar* (from Indonesian, considered by only some speakers to be acceptable in Tetun), or *mezmu, mezmuké* and *embora* (all from Portuguese and used only occasionally by more strongly Portuguese-influenced speakers). Alternatives to the conjunctions are to mark an initial concessive clause with *bele* 'can' or (in informal speech) with clause-final *mós* 'also'.

(145)   *Nia sai   hanesan ema bót maski sei klosan.*
        3S become like   person big although still youth
        'He has become like an important person, even though he is still young.'

Reason conjunctions include *tanba* 'because' (from *tán-ba* 'because-to/for'), *tán* 'because' (mainly liturgical), *purké* 'because' (from Portuguese *porquê*; mainly used by more Portuguese-influenced speakers), and *basá* (literally '*ba-sá* 'for what'; liturgical).

(146)   *Ropa   né la bele uza lai, tanba   sei bokon hela.*
        clothes this not can use FIRST because still wet   CONT
        '(You) can't use these clothes yet, because they are still wet.'

Purpose clauses may be introduced by *hodi* 'so that, in order to; and' if the purpose is indeed achieved (147), or by *para* 'so that, in order that' (from Portuguese *para que*) if the purpose is not necessarily achieved. Purpose verb phrases are alternatively often introduced by the irrealis marker *atu* (148).

(147)   *Sira hú   bibi dikur hodi   bolu ema   tomak atu mai hasoru marinheiru.*
        3P blow goat horn in.order.to call person whole IRR come meet   sailor.
        'They blew goat horns to call all the people to come and oppose the (foreign) sailors.'

(148)   *Hau bá merkadu atu sosa sasán, maibé osan la iha; hau fíla   fali.*
        1S go market IRR buy goods but   money not have 1S return again
        'I went to the market to buy things, but didn't have money, (so) I came back again.'

11.4. Complement clauses

Tetun Dili has two types of complement clauses, namely sentential and reduced.

Sentential complements allow most of the possibilities found in main clauses, including having independent tense, aspect, mood, negation, and subject. For many verbs, such complements are commonly introduced by the complementiser *katak* 'that' (which in Tetun Terik is a verb meaning 'say'). These include verbs of speaking (e.g. *dehan* 'say', *fó hatene* (give know) 'inform'), thinking and knowing (e.g. *hatene* 'know', *fiar* 'believe'), and attention (e.g. *haré* 'see', *rona* 'hear').

(149)   *Nia husu ba ami dehan <u>katak</u> "Nusá? Imi sai ka lae?"*
        3s ask to 1PE say   that   how   1PI exit or not
        He asked us, "What's the situation? Are you leaving or not?"

Sentential complements of some verbs, including verbs of wanting and trying, may be introduced by the purposive conjunction *para* 'so that' (150), while complements for verbs of thinking and knowing, as well as direct quotes, may occur without a complementiser (151).

(150)   *Ita mós buka <u>para</u> partisipante kongresu né reflekta ema hotu.*
        1PI also seek so.that participant congress this reflect person all
        'We also seek that the participants in this congress reflect the entire population.'

(151)   *... hodi   hatene labarik nia isin tún   ka sae.*
        in.order.to know child POS body descend or ascend
        '(The midwife will weigh the child) in order to know whether the child's weight is going up or down.'

A reduced complement lacks a subject, and has restricted tense–aspect possibilities, but can be independently negated. Temporally, the events specified by these complements must either co-occur with or follow the event of the main clause. Reduced complements occur for a wide range of verbs, including verbs of thinking and knowing (e.g. *hatene* 'know (whether to/how to)'), wanting and liking (e.g. *gosta* 'like', *lakohi* 'not want'), trying (e.g. *buka* 'seek (to)'), beginning (e.g. *hahú* 'start'), and *la serve* 'unworthy'.

(152)   *Ita hakarak harí dame.*
        1PI want   erect peace
        'We want to establish peace.'

Verbs of ordering (e.g. *haruka* 'order, send') and of causing (e.g. *husik* 'allow') take an object noun phrase followed by a reduced complement; the object is interpreted as the semantic subject of the complement verb phrase.

(153)   *Sira haruka ami sae   kareta.*
        3P order 1PE ascend vehicle
        'They told us to get into the car.'

Reduced complements of many verbs can optionally be introduced by the irrealis marker *atu* or the purposive conjunction *para* 'so that'.

(154)   *Husik ami atu ukun-an rasik.*
        leave 1PE IRR rule-self own
        'Let us rule ourselves.'

11.5. Topics

A clause may be preceded by an intonationally-linked noun phrase functioning as topic. Some topics are marked by final *né* 'this' (155), but many are recognised as topics only by their position (156). Many topics are co-referential with a noun phrase within the clause (in a left-dislocation construction, as in (156)), but some are not (155). Initial topics may alternatively be introduced by the prepositions *ba* 'as for' or *kona-ba* 'concerning'.

(155)   *<u>Ami ema sanulu né</u>, mane hitu, feto ami nain tolu deit.*
        1PE person ten this male seven female 1PE CLS:human three just
        '(Of) the ten of us, there were seven men, and only three of us women.' (lit. '... the men were seven, the women were only us three.')

(156)   *<u>Hau ho hau nia oan feto ida</u>, ami nain   rua mai iha Delta né.*
        1s and 1s POS child female one 1PE CLS:human two come LOC Delta this
        'And then I and my daughter—the two of us came here to Delta.'

11.6. Focus

Constituents which are fronted within a clause are frequently followed by the focus marker *mak* (or *maka*, especially in writing). This is quite common (but not obligatory) for fronted question constituents (157) and relativised constituents (see section 5.4). Outside of questions and relative clauses, *mak* is usually contrastive, especially for subjects and fronted objects. Frequently the referent of the focused constituent is explicitly contrasted with other entities with which it forms a set, as in (158). Even when the contrast is not explicit, the fact that *mak* is used means that the focused fronted constituent is interpreted as exclusive (157). The part of the clause which follows *mak* presents known or uncontroversial information.

(157)   *"Se <u>mak</u> fó hán imi?"*   *"Ami rasik <u>mak</u>   buka hahán."*
        who FOCUS give eat 2P   1PE own   FOCUS seek food
        "Who fed you?"   "We ourselves sought (our own) food."
                                    (i.e. Nobody else provided for us.)

(158)   *Foho   la iha. So Dili deit <u>mak</u>   iha.*
        mountain not have only Dili just FOCUS have
        'The mountains/countryside don't have it. Only Dili has it.'

Mak is also used to introduce the second of two clauses; here it can be translated 'and then, only then'.

(159)   *Sira tún   mai iha Dili <u>mak</u>   hatene Tetun.*
        3P descend come LOC Dili FOCUS know Tetun
        'Only after they came down to Dili did they know Tetun.'

11.7. Commands and invitations

Commands usually have no explicit marker of imperative mood. The subject is often, but not necessarily, omitted. Common aspectual terms in commands are *ona* 'ANTERIOR, already' and *lai* 'FIRST, before doing something else', the latter being less insistent.

(160)   *Selu osan mai!*
        pay money come
        'Pay me!' (said in a children's game)

Commands or invitations for the addressee to do something without the speaker may end in *bá* (lit. 'go') as in (161), while invitations to do something with the speaker can be prefaced by *mai ita* 'come 1PI' (162).

(161) *Tulun hau bá!*
       help 1s go
       'Help me!'

(162) *Mai ita hotu hán.*
       come 1PI all   eat
       'Let's all eat.'

11.8. Questions and answers

11.8.1. Information questions and their answers

In information questions, questioned constituents usually occupy the same place in the sentence as the answer would. Questioned subjects are usually followed by the focus marker *mak* (157), as are questioned objects when they are optionally placed clause-initially (163). The interrogative pronouns are *sé* 'who', *saida* 'what', *sá* 'what', and *nebé* 'where, which'. As already noted at 4.6.2, *sá* is idiosyncratic in that, when it modifies a noun phrase, it precedes the head noun; in this case it means 'which', requesting a selection of one item from a restricted set (e.g. *sá kór* 'which colour').

(163) *"Kareta nebé mak   ita atu sae?"*   *"Maria sira nian."*
       vehicle which FOCUS 1PI IRR ascent   Maria PL POS
       "Which car are we going to get into?"   "Maria (and associates...)'s one."

(164) *"O nia uma iha nebé?"*   *"(Hau nia uma) iha Balide."*
       2S POS house LOC where   1S   POS house LOC Balide
       "Where is your house?"   "(My house is) in Balide."

Other interrogative forms include:

(165) *hira*   'how many, how much'
       tansá   'why'
       tanba sá   'why'
       bain-hira   'when (future)'
       hori-bain-hira 'when (past)'
       oinsá   'how, by what means/method'
       nusá   'what's happening' (when clause-final),
                         'why' (when clause-initial)

(166) *"Hakerek ita   nia naran oinsá?"*   *"A B E L."*
       write   2S.HON POS name how   A B E L
       "How do you write/spell your name?"   "A B E L."

Of these, adverbs questioning reason (e.g. *tansá* 'why'), along with *bain-hira* 'when (future)', are usually fronted (167).

(167) *"Tanba sá   mak   nia la iha tempu?"*   *"Tanba nia iha enkontru ida."*
       because what FOCUS 3S not have time   because 3S have meeting one
       "Why doesn't he have time (to see me)?"   "Because he has a meeting."

Indirect questions have the same structure as main-clause interrogatives, as illustrated by the final clause in (169).

11.8.2. Polar and alternative questions and their answers

There are various means of forming polar ('yes–no') questions. One can simply add question intonation to a declarative clause, add a final question tag *ka* (lit. 'or') (168) or *ka lae* 'or not' (169), or specify both the positive and the negative options in full. Answers to positive-polarity questions can be *sín* 'yes' or *lae* 'no', but it is more common to echo part of the question that agrees with the answer. Answers to negative-polarity questions usually consist of an echo (168). The negative answer to a question as to whether something has already happened is *seidauk* 'not yet'.

(168) *"Uluk joven sira la fán sasán ka?"*   *"Uluk la fán."*
       past youth PL   not sell goods or   past not sell
       "In the past, didn't the youth sell goods?" "In the past (they) didn't sell."

(169) *"Nusá? Imi sai ka lae?"*
       whats.up 2P exit or not
       "How is it? (Will) you leave or not?"

       "Ami atu sai, maibé ami la hatene lós ami atu sai ba nebé."
       1PE IRR exit but   1PE not know   true 1PE IRR exit to where
       "We want to leave, but we don't rightly know where to go."

Other alternative questions are formed by specifying the alternatives, separated by *ka* 'or'. When the final alternative is also followed by *ka*, it indicates that the list of alternatives is open, being representative of the options, rather than a complete list.

(170) *"Ema mai iha né barak ka, oitoan?"*   *"Ema nain   sanulu karik."*
       person come LOC this many or few   person perhaps ten   perhaps
       "Did many people come here, or only a few?" "The people numbered perhaps ten."

12. Word list

The following word list is presented by semantic fields. Confirmed Portuguese loans are marked by '(P)'. Note that native Tetun vocabulary makes up the majority of the basic vocabulary, including body part and plant terms. A perusal of the grammar will show that all pronouns and determiners are of native stock, while prepositions and conjunctions include Portuguese loans.

The abbreviations used for kin term glosses follow normal anthropological usage: C child, D daughter, e elder, F father, H husband, M mother, P parent, S son, Si sibling, Sp spouse, W wife, y younger, Z sister. For instance, PSiD means 'parent's sibling's daughter'.

Kin terms reflect an interesting mixture of native Tetun and borrowed Portuguese words, imposed upon two markedly different kin systems. In particular, there are two contrasting means of classifying uncles, aunts and cousins in Tetun Dili.

Traditionally, cousins who are related through same-sex parents (MZC, FBC) are referred to as one's brothers and sisters, and their parents (MZ, FB) as one's parents. Cross-cousins (MBC, FZC) are distinguished from these in the terms used, and by the fact that in much of Timor cross-cousins (or in some cases, certain types of cross-cousins) can traditionally marry one another.

Onto this system is superimposed the Portuguese kin system, by which all cousins (PSiC) are terminologically distinguished from one's brothers and sisters (PC), and one's parents' siblings are distinguished from one's parents. Some Tetun Dili speakers draw the traditional distinctions, while others use the same terms with the Portuguese meanings.

As may be seen in the list of plant terms, many plant parts are obligatorily possessed. For instance, the thorn of an orange tree is *sabraka tarak* 'orange thorn', while a thorn of unidentified extraction is *ai tarak* 'plant thorn', rather than just *tarak* 'thorn'.

Body Parts

1. head	*ulun*
2. hair	*fúk*
3. eye	*matan*
4. ear	*tilun*
5. nose	*inus*
6. mouth	*ibun*
7. lip	*ibun kulit* 'mouth skin'
8. tooth	*nehan*
9. tongue	*nanál*
10. arm, hand	*liman*
11. elbow	*liman sikun* 'arm bend'
12. finger	*liman fuan* 'hand round.item'
13. fingernail	*liman kukun* 'hand nail'
14. breast	*susun*
15. back	*kotuk*
16. stomach, belly	*kabun*
17. liver	*aten*
18. leg, foot	*ain*
19. knee	*ain túr* 'leg sit'
20. body hair	*fulun*
21. skin	*kulit*
22. blood	*rán*
23. bone	*ruin*
24. flesh	*isin*
25. urine	*mí*
26. faeces	*té*
27. body	*isin*

Kin terms

1. Sp	*kaben*
2. H	*lain, katuas oan* 'old.man child', *kaben mane* 'spouse male'
3. W	*fén, ferik oan* 'old.woman child', *kaben feto* 'spouse female'
4. M (often also: MZ)	*inan, amá* (P), *maen* (P)
5. F (often also: FB)	*aman, apá* (P), *pai* (P)
6. C (often also: FBC, MZC)	*oan*
7. S (often also: FBS, MZS)	*oan mane* 'child male'
8. D (often also: FBD, MZD)	*oan feto* 'child female'
9. SiS	*sobrinhu* (P)
10. SiD	*sobrinha* (P)
11. eC	*oan bót* 'child big'
12. yC	*oan ikun* 'child tail'
13. DH	*mane foun* 'man new'
14. SW	*feto foun* 'woman new'
15. eB (often also: MZeS, FBeS)	*maun, manu* (P, term of address only)
16. eZ (often also: MZeD, FBeD)	*bin, mana* (P, term of address only)
17. ySi (often also: MZyC, FByC)	*alin*
18. yB (often also: MZyS, FByS)	*alin mane* 'ySi male'
19. yZ (often also: MZyD, FByD)	*alin feto* 'ySi female'
20. B (of female)	*nán* (term rarely used)
21. Z (of male)	*feton*
22. PSiS (often MBS, FZS)	*primu* (P)
23. PSiD (often MBD, FZD)	*prima* (P)

24. ZH, WB	*kunhadu* (P), *rian*
25. BW, HZ	*kunhada* (P)
26. MB, PZH (for some also: FB)	*tiu* (P)
27. BZ, PBW (for some also: MZ)	*tia* (P)
28. PP, PPSi	*abó* (P), *avó* (P)
29. CC	*bei-oan* 'grandparent-child'
30. SpP	*banin* (term of reference only)
31. SpF	*banin mane* 'SpP male', *sogru* (P)
32. SpM	*banin feto* 'SpP female', *sogra* (P)

Animals

1. bird	*manu*
2. wing	*liras*
3. (bird/...) egg	*(manu/...) tolun*
4. rat	*laho*
5. dog	*asu*
6. tail	*ikun*
7. pig	*fahi*
8. fish	*ikan*
9. prawn	*boek*
10. crab	*kadiuk*
11. snake	*samea*
12. lizard	*lafaek rai maran* 'crocodile land dry'
13. gecko	*teki*
14. worm, small crawling creature	*ular*
15. mosquito	*susuk*
16. louse	*utu*
17. leech	*ular susu rán* 'worm suck blood'
18. crocodile	*lafaek*
19. cuscus	*meda*
20. fly	*lalar*
21. cockatoo	*kakatua*

Plants

1. plant, tree, wood	*ai*
2. bark	*ai kulit* 'plant skin'
3. leaf	*ai tahan* 'plant leaf'
4. thorn	*ai tarak* 'plant thorn'
5. seed	*ai musan* 'plant seed'
6. seed (to plant)	*fini*
7. flower	*ai funan* 'plant flower'
8. betel nut	*bua*
9. coconut	*nú*
10. banana	*hudi*
11. cassava	*ai-farina* 'plant-flour' (P)
12. sago	*akar*
13. rice plant	*hare*
14. raw husked rice	*fós*
15. cooked rice	*etu*

Natural world

1. soil, earth	*rai*
2. stone	*fatuk*
3. cave	*fatuk kuak* 'rock hole'
4. sand	*rai henek*
5. beach	*tasi ibun* 'sea mouth'
6. water	*bé*
7. sea	*tasi*
8. river	*mota*
9. cloud	*kalohan*
10. rain	*udan*
11. thunder	*rai tarutu* 'earth bangs'
12. lightning	*rai lakan* 'earth flames'
13. sky	*lalehan*
14. wind	*anin*
15. sun	*loro-matan*
16. moon	*fulan*
17. night	*kalan*
18. star	*fitun*
19. fire	*ahi*
20. smoke	*ahi suar* 'fire smoke'
21. ashes	*ahu kresan*
22. forest	*ai laran* 'plant inside'
23. mountain	*foho*
24. mangrove	*ai tasi* 'plant sea'

Human artefacts

1. garden (agricultural)	*tós*
2. garden (flower)	*jardín* (P)
3. house	*uma*
4. roof	*uma kakuluk* 'house roof'
5. rope	*tali*
6. canoe	*bero, korakora*
7. path/way	*dalan*
8. road	*estrada* (P), *lurón,*

	dalan	
9. machete	*katana* (P), *taha*	
10. axe	*baliu(n)*	
11. spear	*dima*	
12. knife	*tudik*	
13. cloth	*hena*	
14. mosquito net	*muskiteiru* (P)	
15. fish-hook	*kail*	
16. lamp/light	*ahioan* 'fire child'	
17. torch	*lámpada* (P)	
18. wok	*taxu* (P)	

Colour
1. black — *metan*
2. white — *mutin*
3. green — *matak, verdi* (P)
4. red — *mean*
5. yellow — *kinur*

Properties
1. big — *bót*
2. small — *kik*
3. old (things) — *tuan*
4. new — *foun*
5. hot — *manas*
6. cold — *malirin*
7. good — *diak*
8. bad — *át*
9. near — *besik*
10. far — *dók*
11. wet — *bokon*
12. dry — *maran*
13. long — *naruk*
14. short — *badak*
15. short (height) — *ain badak* 'leg short'
16. tall — *ain ás* 'leg tall'
17. heavy — *todan*
18. light (weight) — *kamán*
19. sick — *moras*
20. well (healthy) — *isin diak* 'body good'

Verbs
1. see — *haré*
2. hear — *rona*
3. know (something) — *hatene*

4. know (someone) — *konhese* (P)
5. speak — *koalia*
6. afraid — *tauk*
7. sleep, lie down — *toba*
8. sit — *túr*
9. wake up (Vi) — *hadér*
10. wake up (Vt) — *fanu(n)*
11. stand — *hamrík*
12. walk — *lao*
13. run — *halai*
14. swim — *nani*
15. fly — *semo*
16. wash — *fasi*
17. scratch — *koi*
18. hold — *kaer*
19. split (wood) — *fera*
20. tie — *kesi*
21. dig — *ké*
22. stab — *sona*
23. poke — *fai*
24. fall — *monu*
25. drop (Vt) — *hamonu* 'cause-fall'
26. bathe — *harís*
27. eat — *hán*
28. drink — *hemu*
29. die — *mate*
30. kill — *oho*
31. give — *fó*
32. come — *mai*
33. go — *bá*
34. laugh — *hamnasa*
35. cry — *tanis*
36. sing — *kanta* (P), *hananu*
37. burn (Vi) — *ahi hán* 'fire eat'
38. burn (Vt) — *sunu*
39. cook — *tein*
40. dry in sun (Vt) — *habai*
41. blow — *hú*
42. hit, beat — *baku*
43. throw — *soe*
44. search — *buka*
45. shoot — *tiru*
46. bite — *tata*
47. cough — *mear*
48. vomit — *muta*

13. Comparison with Tetun Terik

13.1. Introduction

Tetun Dili, as a lingua franca much influenced by Portuguese, has its roots in the vernacular Tetun Terik, spoken along parts of the south coast of Timor, and along the East Timor–West Timor border. In this chapter (based largely on Williams-van Klinken (2002b)) we briefly outline some of the differences between these two major varieties of Tetun.

13.2. Lexicon

A major lexical difference between Tetun Dili and Tetun Terik is that the former has absorbed a far larger percentage of Portuguese loans than the latter. Even in everyday speech, limited counts suggest that the percentage of word tokens that are borrowed from Portuguese is of the order of 25% in Tetun Dili. The percentage rises markedly in the language used in the press. Portuguese loans are ubiquitous for abstract and modern vocabulary. They are considerable for all open classes (nouns, verbs, adjectives; borrowed adverbs are largely restricted to higher registers or more Portuguese-influenced speakers), and are also common for prepositions and conjunctions (especially in high registers). There is, however, no Portuguese vocabulary for pronouns, determiners or tense–aspect terms. Portuguese is frequently used for numerals. There is one borrowed Portuguese suffix, the agentive *–dór* (section 4.2.

Where loans are used for concepts which are already labelled in Tetun Terik, the loan sometimes largely replaces the Tetun Terik term, which is then not widely understood in Dili (particularly amongst younger people):

(171)

Tetun Terik			**Tetun Dili**
(ai) kalete	N	'bridge'	*ponti*
babebar	N	'butterfly'	*borboleta*
baur	N	'rainbow'	*arkuiris*
fehuk ai	N	'cassava'	*ai farina*
folin kmán	Adj	'cheap'	*baratu*
hori	Prep	'since (time)'	*dezde*
tekitekis	Adv	'suddenly'	*derepenti*

Alternatively, both terms are used in Dili, with the same meanings, but with somewhat different registers or sets of speakers. For instance, the Portuguese loan may suggest greater Portuguese influence, or the Tetun Terik term may be largely restricted to the more purist liturgical and formal registers (labelled 'formal' below) or may sound more rural.

(172)

Tetun Terik			**Use in Dili**	**Tetun Dili alternative**	
ama(n)	N	'father'	(formal, rural, addresses only God)	*pai, apá*	(common terms of address)
basar	N	'market'	(mainly rural)	*merkadu*	
dalaruma	Adv	'sometimes'		*asvéz(es)*	
hananu	Vi	'sing'	(mainly formal)	*kanta*	
molok	Prep	'before'	(mainly formal)	*antes*	
tinan	N	'year'		*anu*	(used with Portuguese numbers)

It is also possible for the meaning of the native term to have become narrower than it is in Tetun Terik, while a Portuguese loan has taken over the rest of the meaning. This is illustrated by Tetun Terik *hanorin* Vt 'teach, learn', which in Dili is restricted to 'teach', with the Portuguese loan *aprende* taking over the meaning of 'learn'.

Finally, there are also many concepts for which speakers can choose between a Portuguese loan and an Indonesian one. The use of Indonesian loans is however largely avoided in writing and in formal situations, and Indonesian-educated speakers appear to be acquiring an increasing number of high-level Portuguese loans.

(173)	**Portuguese loan**			**Indonesian loan**
	distritu	N	'government district'	*kabupaten*
	xefi suku	N	'village head'	*kepala desa*
	orijinál	Adj	'original'	*asli*
	dominasaun	N	'domination'	*dominasi*

13.3. Phonology

Tetun Dili has lost one phoneme which occurs in Tetun Terik. The glottal stop has simply disappeared (e.g. TT *to'os* = TD *tós* 'gardens'). Another Tetun phoneme /w/ has almost completely merged with /b/ in native Tetun Dili roots (e.g. TT *wé* = TD *bé* 'water'), and now appears most frequently in loans, eg *warun* 'stall' (from Indonesian *warung*). A further change in native Tetun words is the loss of word-initial consonant clusters, which in Tetun Terik all begin with /k/. In most instances the /k/ is dropped (e.g. TT *ktodan* = TD *todan* 'heavy'). However sometimes an /a/ is inserted between the consonants (e.g. TT *kmán* = TD *kamán* 'light (weight)'). Those words which retain the consonant cluster are mainly used in conservative liturgical vocabulary (e.g. *kbit* TT 'strong', TD '(spiritual) power').

As a result of numerous loans from Portuguese, and a much smaller number from Malay / Indonesian, Tetun Dili has gained new phonemes (e.g. /v/), syllable and word patterns (e.g. stop-/r/ sequences as in *treinu* 'training'), and stress patterns (e.g. ante-penultimate stress, as in *prémiu* 'prize'). For further details see chapter 2.

13.4. Inflectional morphology

A major difference between Tetun Terik and Tetun Dili is that the latter has lost most Tetun morphology, both inflectional and derivational.

In Tetun Terik, verbs are marked for the person and number of the subject, but in Tetun Dili, verbs have an invariant form. Tetun Terik subject marking is expressed by an initial consonant. Where the citation form begins with /h/, the /h/ is replaced by the subject marker, namely 1S *k-*, 2S *m-*, 3S *n-*, 3P *r-* or *n-* (depending on the dialect); 2P and 3P remain as /h/. Verbs beginning with other consonants are marked only for 1S subjects, by placing a *k-* before the verb. (Note that /kC/ sequences are the only consonant sequences occurring in Tetun Terik roots.)

(174)	TT: *Ha'u k-bá k-ó*		*sia.*		TD: *Hau bá ho sira.*
	1S	1S-go 1S-accompany	3P		1S go with 3P
	'I go with them.'				'I go with them.'

In Tetun Terik, when a noun in certain constructions is inalienably (or near-inalienably) possessed (e.g. being a kin term or body part noun), the noun is followed by a genitive *-n*

suffix (or an *-r* for plural possessors, in some varieties of Tetun). In Tetun Dili, this /n/ has become a permanent part of the noun, being present even in non-possessive constructions. (e.g. TT *Ina!* 'Mother!', but *ha'u-kan ina-n* '1S-POS mother-GEN'; TD *inan* 'mother').[6]

13.5. Derivational morphology

Most Tetun Terik derivational affixation and reduplication has been lost, being replaced by an increased use of 'zero' morphology (i.e. the use of a verb form as a noun, and vice versa, without any change of form), some new possibilities in verb serialisation and compounding, and a Portuguese loan suffix.

Derivation from noun to verb is by the prefix *ha-* in Tetun Terik; in Tetun Dili the verb has the same form as the noun (e.g. TT *ha-tolu* 'lay an egg' from *tolun* 'egg', TD *tolun* 'lay an egg; egg').

Derivation from verb to abstract noun is by partial reduplication in Tetun Terik; here too Tetun Dili uses the same form for both noun and verb (e.g. Fehan dialect of TT *fa-fiar* 'faith, trust' from *fiar* 'believe', TD *fiar* 'faith, trust; believe'). Alternatively, Tetun Dili borrows an abstract noun from Portuguese (e.g. *fé* 'faith').

Derivation from verb to instrument noun is, in Tetun Terik, by either partial reduplication or prefixation (e.g. TT *kakusan* 'button' from *kusa* Vt 'button, lock'). This is not found in Tetun Dili.

Causativisation of verb and adjectives by the prefix *ha-* is much more productive in Tetun Terik than in Tetun Dili. Conversely, the alternative strategy of preceding the root with the verb *halo* 'make, do' is much more common in Tetun Dili than in Tetun Terik (e.g. TT *ha-maran* (make-dry); TD *halo maran* (make dry) 'cause to become dry').

Actor derivations are markedly different between the two varieties of Tetun. In Tetun Terik, terms deriving actors are derived by a circumfix *mak- -n* (or *ma- -k*, depending on the phonology of the root) (e.g. *mak-leo-n* 'guardian' from *leo* 'protect'). In Tetun Dili, only a few such derivations survive in the liturgical register; otherwise, actor terms are derived using the Portuguese loan suffix *-dór* (e.g. *hamnasadór* 'giggler' from *hamnasa* 'laugh'), the bound root *-tén* (e.g. *baruk-tén* 'lazybones' from *baruk* 'lazy'), or the noun *nain* 'owner, master' (e.g. *servisu nain* 'diligent worker' from *servisu* 'work').

There is other, less productive, morphology which exists in Tetun Terik but has been lost from Tetun Dili. This includes the use of *k- -k* to derive adjectives (e.g. TT *k-silu-k* 'snapped' from *silu* 'snap'), and *hak- -k* to derive reciprocal verbs (e.g. TT *hak-tuda-k* 'throw (e.g. spears) at one another' from *tuda* 'throw (e.g. spear)').

---

[6]　　　　There are a few remnants of genitive marking in Tetun Dili. Some speakers require *-n* on the Portuguese kin loans *tia* 'aunt' and *tiu* 'uncle' when they are possessed; some also attach *-n* to *rai* 'land, country' when it is possessed with the meaning 'country' (so distinguishing *hau nia rai* 'my land (that I own)' and *hau nia rai-n* 'my country').

One construction that seems to be more widespread in Tetun Dili than in Tetun Terik is the compound-like serialisation of *fó* with a following verb (e.g. TD *fó aluga* (give rent) 'rent out'), as described in section 10.5.

13.6. Grammar

With regard to numerals, Tetun Terik uses a wider range of numeral classifiers than does Tetun Dili (e.g. the Fehan dialect uses *matan* when counting traditionally-significant domestic animals), and Tetun Terik has some derived adjectival numerals (e.g. *rua-s* 'two' from the number *rua* 'two') which are not found in Tetun Dili.

Possessive marking varies from dialect to dialect within Tetun Terik. In Tetun Dili it is restricted to *nia* after the possessor (e.g. *hau nia karau* (1S POS buffalo) 'my buffalo'). Options found in dialects of Tetun Terik include use of the bare pronoun (e.g. *ha'u kabau* '1S buffalo'), and following the possessor with *nia* (as in Tetun Dili) or *-kan* (e.g. *ha'u-kan kabau*).

Possibilities for negation also differ. Tetun Terik has a postposed *laek* 'without'. In Tetun Dili this is used in the liturgical register, but in everyday speech is replaced by a regular negated verb *la hó* 'not have' (175). In spoken Tetun Dili, negation is often expressed by a negator before the verb, and *ida* (literally 'one') after it (176). This use of *ida* appears to not be found in Tetun Terik.

(175) TT: *Nia inan   laek.*   TD: *Nia la ho   inan.*
            3s mother without   3s not have mother
            'She is motherless.'   'She has no mother.'

(176) TT/TD: *Nia la bá.*   Informal TD: *Nia la bá ida.*
            3s not go   3s not go one
            'He didn't go.'   'He didn't go.'

Tetun Dili allows some verbs to precede the subject (177). This construction does not exist in Tetun Terik (except for existential verbs), and appears to be a calque on Portuguese.

(177) TD: ..., *bele akontese   problema   barak.*
                 can occur   problem   many
                 '(If care is not taken,) many problems can occur.'

14. Texts

14.1. Introduction

The following texts illustrate a range of genres. Portuguese loans are underlined to aid the reader in recognising the frequency of such loans, particularly in press reports. These texts contain also two Indonesian loans, namely *Polda*, the acronym for the Indonesian police headquarters, and *Politeknik*, the Indonesian term by which the technical college at Hera was known.

The two written texts are presented in their original form to illustrate the type of spelling variation that is widespread in East Timor today. Our own transcription follows the original where there is a difference.

14.2. Oral narrative

A young man recounts his experiences during the destruction of Dili at the hands of the Indonesian military and their local militias in September 1999. He is talking to another young Timorese man at a refugee base in Australia, soon after his evacuation from Dili.

(178) *Ema   sunu bispu   nia uma.*
        person burn bishop POS house
        People were burning the bishop's house.

(179) *Ema   sunu Câmara hanesan né,*
        person burn chamber like   this
        While people were burning the Câmara (Eclesiástica, the Ecclesiastical Chamber)

(180) *hau sei lao   hela ho Pedro,*
        1S   still walk/go CONT with Pedro
        I was driving with Pedro,

(181) *be   ami sei lao   dobra   tún-sae   sa.*
        well 1PE still walk double ascend-descend TAG
        well, we were still going to and fro, right!

(182) *Ami bá Becora, bá Becora nebá,*
        1PE go Becora, go Becora there
        We went to Becora—over to Becora,

(183) *tún   fali mai para tiha iha Lecidere,*
        descend again come stop PRF LOC Lecidere
        came back down and stopped in Lecidere,

(184) *haré tiha bispu   nia uma,*
        see PRF bishop POS house
        looked at the bishop's house,

(185) *halai liu mai tiha iha Polda.*
        run pass come PRF LOC police.HQ
        then drove on quickly to the provincial police headquarters.

(186) *Polda,   entaun hau tama kala   dala   hát ka dala   lima ida karik.*
        police.HQ so   1s enter perhaps instance four or instance five one perhaps
        So at the police headquarters, I entered the building maybe about four or five times.

(187) *... ho madre sira mak   tama iha nebá.*
           with nun PL FOCUS enter LOC there
        ... It was with the nuns that I finally went in.

(188) *Tama bá, madre dehan "Diak ka lae?"*
        enter go nun   ask   good or not
        When we were inside, a nun said "Are you OK?"

(189) *Hau dehan "Kalma deit" sa.*
        1s   say   calm   just TAG
        I replied "Just relax", right.

(190)   *Tama bá, sira be <u>seguransa</u> iha odamatan né,*
      enter go 3P REL security   LOC door   this
      We went in, and those who were guards at the door—

(191)   *<u>polísia</u> sira né, haré bá, <u>kolega</u> fali.*
      police PL   this see   go friend instead
      the policemen—when I looked at them, they turned out to be friends (of mine).

(192)   *Sira husu dehan "Atu bá nebé?"*
      3P ask say   IRR   go where
      They asked me "Where do you intend to go?"

(193)   *"Hau tama lai   iha laran né."*
      1S   enter FIRST LOC inside this
      "I'm just going in."

(194)   *Sira dehan "<u>Entaun</u> liu   bá."*
      3P   say   so   pass go
      Then they said "Well then, go on."

14.3. Email to a friend

This email was written by a 19 year old man to one of the authors. The writer was born in Baucau, but lived for an extended period in Dili. He had worked as a translator for UNTAET, which accounts for his initial English email-style greeting of 'Hi'.

(195)   *Hi maun   boot, Diak ka lae?*
                 bót
      Hi older.brother big   good or not
      Hi, older brother, How are you?

(196)   *Oinsa kona ba maun   nia <u>visita</u> ba Timor <u>Leste</u>?*
      Oinsá kona-ba   vizita
      how   about   older.brother POS visit to Timor East
      How are things with your next visit to East Timor?

(197)   *Hau <u>espera</u> katak buat hotu lao   diak deit*
      1S   hope   that   thing all   walk/go well just
      I hope that everything is going well

(198)   *iha maun   nian <u>actividade</u> e   <u>servisu</u> iha Dili e   iha Australia.*
                   nia   aktividade i   i   Austrália
      LOC older.brother POS activity   and work   LOC Dili and LOC Australia
      with your activities or work in Dili and in Australia.

---

(199)   *<u>Novidade</u> diak ba maun*
      news   good for older.brother
      (I have) good news for you[7]

(200)   *katak hau <u>oras</u> ne'e dadun   iha <u>universidade</u> Indonesia*
                   né   daudaun   Indonézia
      that 1S time this PROG   LOC university   Indonesia
      that right now I am at the University of Indonesia,

(201)   *tamba   hau la <u>konsege</u> hetan <u>bolsu</u>   <u>studu</u> ba Australia*
                                          estudu ba Austrália
      because 1S   not manage get   scholarship study to Australia
      because I didn't manage to get a scholarship to study in Australia,

(202)   *<u>mais</u> hau sei   <u>esforsu</u>[8] an <u>estuda</u> didiak   iha ne'e*
                                       di-diak   né
      but 1S FUTURE effort   self study RDP-good LOC this
      but I will strive to study hard here

(203)   *<u>para</u>   atu <u>servi</u> povo   Timor <u>Leste</u>.*
                     povu
      so.that IRR serve the.people Timor East
      so that (I can) serve the people of East Timor.

14.4. Newspaper report

This report comes from the daily newspaper *Suara Timor Lorosae* published in Dili on 20th December 2000. It is rather typical of current press writing in East Timor, showing elevated use of Portuguese loans, especially of a technical nature.

(204)   *Loron <u>Quarta Feira</u> liu   ba (13/12) <u>Gabinete Transisaun</u> Timor Lorosae*
               Kuarta-feira   bá   Tranzisaun
      day   Thursday   pass go   cabinet   transition Timor east
      Last Thursday (13/12), the Transitional Cabinet of East Timor

(205)   *hakotu tiha ona katak fatin Politeknik   Hera nian*
      decide PRF ANT that   place polytechnic Hera POS
      decided that the Hera polytechnic (college)

(206)   *sei   <u>usa</u> hodi fatin <u>treino</u>   ba <u>primeiro batalhao</u>   <u>Falintil</u>*
               uza   treinu   primeiru batalhaun
      FUTURE use for   place training for first   battalion Falintil
      will be used as a training location for the first Falintil battalion

---

[7]   This sentence appears ellipsed with the expected initial *Hau iha* 'I have' absent.

[8]   Such use of a Portuguese noun form as a verb in Tetun is quite common (see section 4.3.2). It is however more usual to use the Portuguese verb *esforsa* in this construction.

(207) *para*   *atu* *prepara* *forca-s*   *armada-s* ba Timor Lorosae
                              forsa-s   *armada-s*
       so.that IRR prepare force-PL armed-PL for Timor east
       to prepare the armed forces for East Timor

(208) *hodi*   *hasoru* *processo* ba *independencia* Timor Lorosae nian
                            prosesu   *independénsia*
       in.order.to meet   process to independence Timor sun-rise POS
       to participate in the process towards East Timor's independence

(209) *ba tinan oin mai.*
       for year face come
       next year.

(210) *Decisaun ne'e be hasai*   *husi* *governo*   *ne'e*
       Desizaun nebé   *hasai*   *husi* *governu*   *né*
       decision REL   cause-exit from government this
       This decision that was taken by the government

(211) *bele* *considera* hanesan *decisaun* *provisoria* la'os   *definitiva*[9]
              konsidera   *desizaun provizória laós*
       can consider like   decision provisional indeed.not definitive
       can be considered as a provisional, not definitive, one,

(212) *e*   *concerteza* *to'o* loron ida wainhira ita nia *condicao* *desenvolve* diak ona
       i   *konserteza tó*   *bain-hira*   *kondisaun dezenvolve*
       and certainly until day one when   1PI POS condition develop   good ANT
       and certainly one day when our situation has improved,

(213) *fatin Politeknik* *ne'e sei*   *usa* fali ba nia   fatin *origem* hanesan *escola*
                      né   *uza*   *place orijen*   *eskola*
       place polytechnic this FUTURE use again for 3S.POS place original like   school
       the polytechnic location will be used again for its original purpose as a school.

15. References

Crowley, Terry. in press. *Serial verbs in Oceanic: a descriptive typology.* Oxford: Oxford University Press.

Durie, Mark. 1997. Grammatical structures in verb serialization. In *Complex Predicates,* ed. by Alex Alsina, Joan Bresnan, and Peter Sells, 289–354. Stanford: CSLI Publications.

Eccles, Lance. 1998. Perfective and inchoative verbal markers in Tetun. *Studies in Languages and Cultures of East Timor* 1:39–51.

---

[9]   Note that both *provizória* and *definitiva* are here borrowed in the feminine form with final *-a*, even though adjectives are more commonly borrowed in the masculine form (*provizóriu* and *definitivu*).

Hajek, John. 2000. Language planning and the sociolinguistic environment in East Timor: Colonial practice and changing language ecologies. *Current Issues in Language Planning* 1:400–414.

Hajek, John, and Catharina Williams-van Klinken. in press. Um sufixo românico numa língua austronésia: *-dor* em Tetum (A Romance suffix in an Austronesian language: *-dor* in Tetun). *Revue de linguistique romane.*

Hull, Geoffrey. 1994. A national language for East Timor. In *Language reform: History and future. La Réforme des Langues: Histoire et Avenir. Sprachreform: Geschichte und Zukunft,* ed. by István Fodor and Claude Hagège, 347–366. Hamburg: Helmut Buske Verlag.

—— 1996. *Mai kolia Tetun: a beginner's course in Tetum-Praça: the lingua franca of East Timor* (second ed.) North Sydney: Australian Catholic Relief and the Australian Catholic Social Justice Council.

—— 1998. The languages of Timor 1772–1997: a literature review. *Studies in Languages and Cultures of East Timor* 1:1–38.

—— 1999. *Standard Tetum-English dictionary.* Sydney: Allen & Unwin in association with the University of Western Sydney Macarthur.

Hull, Geoffrey, and Lance Eccles. 2001. *Tetum Reference Grammar.* Winston Hills: Sebastiao Aparício da Silva Project, in conjunction with Dili: Instituto Nacional de Linguística, Universidade Nacional de Timor Lorosa'e.

Thomaz, Luis Filipe F.R. 1981. The formation of Tetun-Praça, vehicular language of East Timor. In *Papers on Indonesian languages and literatures,* ed. by Nigel Phillips and Anwar Khaidir, 54–83. Paris: Cahier d'Archipel 13.

van Engelenhoven, Aone, and Catharina Williams-van Klinken. forthcoming. Tetun and Leti. In *The Austronesian languages of Asia and Madagascar,* ed. by Sander Adelaar and Nikolaus Himmelmann. Richmond: Curzon Press.

van Klinken, Catharina. 1999. *A grammar of the Fehan dialect of Tetun, an Austronesian language of West Timor.* Canberra: Pacific Linguistics, C-155.

Williams-van Klinken, Catharina. 2002a. High registers of Tetun Dili: Portuguese press and purist priests. In *Proceedings of the 2001 Conference of the Australian Linguistics Society,* ed. by Cynthia Allen, 8. Canberra: http://linguistics.anu.edu.au/ALS2001/proceedings.html.

—— 2002b. Tetun Dili: A portrait of language change in progress. Paper presented to *9th International Conference on Austronesian linguistics,* Canberra.

Williams-van Klinken, Catharina, John Hajek, and Rachel Nordlinger. 2002. *Tetun Dili: a grammar of an East Timorese language.* Canberra: Pacific Linguistics, 528.

Index

Languages of the World/Materials
LINCOM's Descriptive Grammar series