

# Two Tetuns: Language change before your very eyes

Catharina Williams-van Klinken

(Dili Institute of Technology and Melbourne University)

6th International East Nusantara Conference on Language & Culture

Kupang, 4-6 August 2010

## 1. One Tetun or two?

### Two named *Tetuns*

- Tetun Terik (TT): vernacular Austronesian language, spoken on the south coast of East Timor and along both sides of the Indonesian-East Timor border.
- Tetun Dili (TD), the major lingua franca throughout East Timor. Also known as Tetun Prasa (*prasa* ‘market place’) and Tetun Portu (‘Portuguese Tetun’).

### One language?

- Article 13 of the East Timorese constitution says, “1. Tetum and Portuguese shall be the official languages in the Democratic Republic of East Timor. 2. Tetum and the other national languages shall be valued and developed by the State.”
- Hull (2002): “Tetum is the most widely used vernacular in East Timor and the country's national language.” Tetum in East Timor has been influenced by Portuguese, and in West Timor by Indonesian; this differential influence “hinders mutual intelligibility between the two varieties” (east and west of the border).
- A TT speaker said to me as we were crossing the border last week, “Tetun Terik and Tetun Dili are the same, it’s only the *logat* that is different.” This is commonly expressed by East Timorese TT speakers.

### Two languages? At least two very different varieties:

- My staff when shown TT texts: total puzzlement.
- Myself on being asked to be take minutes of my first-ever TD meeting, when my TT was already reasonably good: “I can’t, as I don’t speak Portuguese.” The reply: “But we’re speaking Tetun!”
- Myself at the same conference, trying to interpret a question to me during a break:

- (1) TD: *Senyora hamlaha ona ka?*  
My TT guess: *Ibu n-ala’a onan ka*  
                  madam 3S-hungry about.to QUESTION  
TD: ‘Are you hungry yet?’  
My TT guess: ‘Are you nearly hungry?’

## 2. History

The history of Tetun Dili has been written about by Thomaz (1981, 2002), Fox (1997), Hull (1996:xiii, 2002), and Nordholt (1971), but much remains unknown.

Major dates:

- 1514: The Portuguese arrived in Timor. Tetun was already a lingua franca in the east of the island.
- 1641: They established a capital at Lifau in Oecusse.
- 1642: Fall of the kingdom of Wehali at the hands of Portuguese-authorized “Topasses” troops.
- 1769: The Portuguese moved their capital east to Dili, in a Mambae speaking region. They ruled indirectly through local rulers, using Tetun and Portuguese.
- 1898: They established a college in Soibada, in the East Tetun-speaking region. Around this time, Portuguese rule became more direct. Both Tetun and Portuguese were used in Dili.
- 1975: Indonesian invasion. Indonesian replaced Portuguese as the language of government, education and media.
- 1982: A liturgical form of Tetun replaced Portuguese as the language of the Catholic mass.
- 1999: Vote for independence. UN rule used English. Many materials translated from English to Tetun by untrained translators. Tetun suddenly used for government, administration, media, and some education. Start of phase-out of Indonesian in schools. Indonesian terms avoided in formal Tetun.
- 2002: Independence. Tetun and Portuguese chosen as official languages.

So, Tetun has been in close contact with Mambae as a result of Dili’s location, with Portuguese as a result of centuries of colonisation, with Malay as a lingua franca of the region, and, for a shorter period, with Indonesian.

Unless otherwise specified, information on Tetun Terik is from van Klinken (1999), and on Mambae is from Hull (2001, 2003) and Hajek (2006). For a description of Tetun Dili, see Williams-van Klinken *et al.* (2002).

## 3. Vocabulary

Tetun Terik	Tetun Dili	Likely source
<p><b>Large native vocabulary</b></p>	<p><b>55% of TT terms lost, 20% changed in pronunciation or meaning.</b> (Based on Fehan words, less for East Tetun, but still large.)</p> <p>TD is not used for <i>adat</i>, poetry, traditional stories, farming, hunting, traditional knowledge and beliefs.</p>	<p>Language not used in most traditional contexts.</p>
<p><b>Low percentage of loans.</b></p> <p>12% of 3700 terms in my dictionary file are from Indonesian.</p> <p>In spoken texts, percentage of Indonesian tokens range from 0.1-10%.</p>	<p>Nearly 50% of the 11300 entries in my lexical database are from Portuguese.</p> <p>In texts, the percentage of Portuguese tokens are about 30-40% in newspapers, 10-30% in formal speech, 10-25% in casual speech. (Casual youth speech also mixes 7% Indonesian).</p>	<p>Much greater exposure to Portuguese, many more educated people, language used for more non-native functions</p>

## 4. Style

Tetun Terik	Tetun Dili	Likely source
<p><b>Repetition is aesthetically pleasing:</b></p> <ul style="list-style-type: none"> <li>poetic parallelism</li> <li>head-tail constructions</li> </ul>	<p><b>Such repetition is avoided</b>, even in speaking, but especially in writing:</p> <ul style="list-style-type: none"> <li>almost no poetic parallelism.</li> <li>little head-tail construction</li> </ul>	<p>European and Indonesian writing styles. Not used for poetry and story-telling. Loss of poetic terms.</p>

## 5. Phonology

Tetun Terik	Tetun Dili	Likely source
<p>Glottal stop: <i>na'an</i> 'meat' versus <i>naan</i> 'sister of a man'</p>	<p><b>Lost glottal stop:</b> <i>naan</i> 'meat; sister of a man'.</p> <p>Retained a little in liturgical Tetun and writing (often incorrectly placed).</p>	<p>TT glottal unstable, Mambae has no glottal</p>
<p>/w/ and /b/ are separate phonemes: <i>wee</i> 'water' versus <i>bee</i> 'also'; <i>wani</i> 'bee', <i>banin</i> 'parent in-law'</p>	<p><b>/w/ merged into /b/:</b> <i>bee</i> 'water', <i>bani</i> 'bee', <i>banin</i> 'parent-in-law'.</p> <p>Only retained in <i>walu</i> 'eight', to avoid merger with <i>balu</i> 'some'.</p>	<p>Mambae has no 'w'</p>
<p>Word-initial consonant clusters beginning in /k/: <i>ktodan</i> 'heavy', <i>kmaan</i> 'light weight'</p>	<p><b>Most lost</b> by deleting /k/: <i>todan</i> 'heavy' or inserting /a/: <i>kamaan</i> 'light weight'.</p> <p>Retained some /kl, kn, kr/: <i>klaran</i> 'centre', <i>knuuk</i> 'nest', <i>kraik</i> 'down'. These clusters are used in Mambae (as is /km/, which is lost in TD).</p> <p>Retained some in conservative liturgical register: <i>kbiit</i> 'power', <i>kmanek</i> 'good'.</p>	<p>Mambae, reinforced by Portuguese and Indonesian.</p> <p>Simplify pronunciation to unmarked syllable structure (typical of creoles...)</p>
<p>13 consonants</p>	<p><b>Add 10 consonants</b> from Portuguese: /p g v z ʃ ʒ ʎ ɲ r j/</p> <p>For nearly all speakers, tap /r/ merges with trill /r/. Some speakers merge other consonants with the nearest TT phoneme: [ʃavi] / [sabi] 'key', [sejór] / [senór] 'sir'.</p>	<p>Portuguese loans</p>
<p>Penultimate stress: <i>hatene</i> 'know'</p>	<p><b>Add stress patterns from Portuguese:</b> Some Portuguese words have final stress or antepenultimate stress: <i>animál</i> 'animal', <i>armáriu</i> 'cupboard'.</p> <p>Some with antepenultimate stress optionally regularised by deleting final syllable: <i>armari</i>.</p>	<p>Portuguese loans</p>
<p>Syllable: (C)V(C) Word: up to 4 syllables</p>	<p><b>Add syllable templates and longer words from Portuguese:</b> <i>estrundu</i> 'noise', <i>autodeterminasaun</i> 'self-determination'</p>	<p>Portuguese loans</p>

## 6. Morphology

Tetun Terik	Tetun Dili	Likely source
<b>Has subject prefixes on verbs.</b>	<b>Subject prefixes lost</b> , even in conservative liturgical Tetun.	Neighbouring languages; simplification.
<b>Genitive –n on vowel-final inalienably possessed nouns</b> in certain constructions: <i>ha'u-kan ina-n</i> '1s-POS mother-GEN'	<b>Genitive –n has become part of the lexeme:</b> <i>inan</i> 'mother'. Retained on <i>rai</i> 'land, country' when it refers to one's district or country: <i>hau nia rai-n</i> 'my district/country'. Some speakers use it on Portuguese kin terms <i>tia</i> 'aunt' and <i>tiu</i> 'uncle': <i>nia tiu-n</i> 'his/her uncle', though not on <i>avoo</i> 'grandparent'.	Simplification and reanalysis; Mambae
<b>Some derivational morphology:</b>	<b>Most morphology replaced by conversion ('zero derivation') and analytic constructions.</b>	Mambae is even more analytical, and has none of these derivations. Reduced morphology is typical of creoles.
<b>Derivation of actor nouns by circumfix</b> <i>mak-</i> -n / <i>ma-</i> -k: <i>makfa'an</i> 'seller' from <i>fa'an</i> 'sell'	Use compounds with bound roots. 1. <i>nain</i> , literally 'owner, master': <i>matenek nain</i> 'expert,' from <i>matenek</i> 'smart, educated'; recent innovations <i>lee nain</i> 'reader' from Portuguese loan <i>lee</i> 'read' 2. <i>-teen</i> , literally 'faeces', for negatively valued terms: <i>naok-teen</i> 'thief' 3. Portuguese suffix <i>-dór</i> , borrowed as a bound root and applied to Tetun roots as well: <i>hamnasa-door</i> 'giggler' from <i>hamnasa</i> 'laugh'. (Described in detail in Williams-van Klinken and Hajek 2009).	Replace morphology with compounding. • <i>nain</i> fits areal pattern (Hull 2001: 107) • <i>teen</i> : Hull (2001:107) suggests reinterpretation of proto form * <i>teras</i> 'hard' • Portuguese <i>-dór</i> common, fits existing pattern, reinterpreted as root.
<b>Rare serialisation with <i>foo</i> 'give'</b> : <i>foo hatene</i> 'introduce, inform'.	<b>Increased serialisation with <i>foo</i> 'give'</b> and a following verb: <i>foo haan</i> 'give eat' = 'feed'; <i>foo empresta</i> 'give borrow' = 'lend'.	Productive in Malay. Malay calque from 19 <sup>th</sup> century? Expansion of existing low-frequency TT construction?
<b>No number marking on nouns.</b>	<b>Plural marking on Portuguese nouns</b> is not uncommon: <i>foo livru-s</i> 'give book-s'.	Portuguese categories and rules applied in Tetun.
<b>No number or gender agreement on adjectives.</b>	<b>Some agreement on Portuguese loans</b> , by Portuguese-influenced speakers: <i>kompanyias lokais</i> 'company.plur local.plur = 'local companies'	Portuguese categories and rules applied (or mis-applied) in Tetun.

## 7. Noun phrases and numerals

Tetun Terik	Tetun Dili	Likely source
<p><b>Possessive marker depends on dialect:</b> Fehan –<i>kan</i> or short form of pronoun, East Tetun <i>nia</i> or bare pronoun and genitive suffix (Morris 1984:xv)</p>	<p><b>Possessive marker is <i>nia</i>.</b> Church Tetun uses <i>nia</i> or bare pronoun.</p>	<p>Follow a prestigious dialect, reinforced by Mambae.</p>
<p><b>Relativiser in Fehan is focus marker <i>mak</i>:</b> <i>bei mak mate</i> ‘ancestors who died’</p>	<p><b>Relativiser is <i>nebee</i> ‘where’, or Portuguese <i>ke</i>:</b> <i>bei-ala nebee mate</i> ‘ancestors who died’ Mambae uses <i>bae</i> ‘where’ (Hajek 2006). Perhaps other TT dialects use <i>nebee</i>: Mathijsen’s (1967) Bible stories in Atambua dialect from early 1900s use <i>ne’ebe</i></p>	<p>Mambae calque reinforces one dialect?, Portuguese loan.</p>
<p><b>Abstract nouns are uncommon in texts</b></p>	<p><b>Abstract nouns are more commonly used:</b> <i>edukasaun</i> ‘education’, <i>urbanizasaun</i> ‘urbanisation’, <i>dezenvolvimentu</i> ‘development’. Often used to derive verbs: <i>halo planu</i> ‘make plan’ = ‘plan’, <i>halo orasaun</i> ‘make prayer’ = ‘pray’.</p>	<p>Influence of Portuguese and Indonesian education.</p>
<p><b>Nouns cannot take complement clause.</b></p>	<p><b>Nouns can take complements:</b> <i>direitu atu rona rádio</i> ‘right irr hear radio’ = ‘(the) right to listen to radio’. Rare except in writing, and appears restricted to Portuguese nouns</p>	<p>Portuguese calque in high-level Tetun</p>
<p><b>Determiners</b></p>	<p><b>TD appears to follow East dialect;</b> it is quite different to Fehan in West Timor. More detailed descriptions needed.</p>	
<p><b>Numeral classifiers: TT has a half dozen classifiers.</b></p>	<p><b>Reduced: Retain human classifier, lost animal classifier (<i>matan</i> ‘eye’), classifiers for objects are seldom used.</b></p>	<p>Simplification, Mambae has less classifiers than TD, Portuguese has no classifiers.</p>
<p><b>Numeric adjectives: ruas/ruak ‘two’ and toluk/tolus ‘three’ can be used when there is no classifier: ita ruas ‘1PI two’ = ‘we two’</b></p>	<p><b>Numeric adjectives lost.</b></p>	<p>Simplification, Mambae?</p>
<p><b>Taboo on numeral ‘3’.</b></p>	<p><b>Taboo almost lost, although some still don’t accept being photographed in groups of three.</b></p>	<p>Regional variation?, Christianisation / urbanisation.</p>

## 8. Prepositions, conjunctions, complementisers

Tetun Terik	Tetun Dili	Likely source
<p><b>Loan prepositions and conjunctions from Indonesian few and rarely used.</b></p>	<p><b>Loan prepositions and conjunctions from Portuguese common:</b></p> <p><i>i</i> ‘and’ (used for clauses, <i>ho</i> for NPs, TT <i>no</i> in church and much writing; about 7400 examples, 38% oral);</p> <p><i>para</i> ‘so that’ (&gt;3000 examples)</p> <p><i>maski</i> ‘although’ (1183 examples, 12% oral)</p>	Portuguese loans
<p><b>Use possessed noun to express ‘during’ and ‘between’:</b></p> <p><i>iha ... na laran</i> ‘LOC ... POS inside = ‘during’, <i>iha ... na leet</i> ‘LOC... POS space’ = ‘between’</p>	Retain these, but often use Portuguese loans: <i>durante</i> ‘during’ (2577 examples, 14% oral), <i>entre</i> ‘between, among’ (1274 examples, 13% oral)	Loan follows European and Indonesian structure
<p><b><i>hosi</i> ‘from, about (a topic)’</b></p>	<p><b>New preposition <i>kona ba</i> ‘about (a topic)’</b> (4638 examples, 26% oral), retain <i>hosi</i> as ‘from’</p>	Mambae calque ( <i>en lao</i> ‘touch go’)
<p><b>Complementisers rare:</b> only <i>hosi</i> ‘about’ and <i>batu</i> ‘so that’, both rare. Usually no complementiser: <i>hatene o m-oo buat di’ak ida</i> ‘know 2S 2S-have thing good one’ = ‘know you have a good thing’</p>	<p><b>Complementiser <i>katak</i> ‘that’ from TT verb <i>katak</i> ‘say’</b> common in all registers (15400 examples) but particularly common in formal and written texts: <i>hatene katak hau monu</i> ‘know that I fell’</p>	Grammaticalisation; follow European and Indonesian structure
<p><b>Time: mainly parataxis,</b> few conjunctions, and rarely used. No conjunction or preposition ‘before’</p>	<p><b>Time: Conjunctions very common,</b> especially <i>bainhira</i> ‘when’ (6300 examples most of which are conjunctions, from TT question <i>wainhira</i> ‘when? (in the future)’ (based on Portuguese <i>kuandu</i> ‘when; when?’?), and <i>kuandu</i> ‘when, if’ from Portuguese (2000 examples).</p> <p>Two conjunctions for ‘before’, though both uncommon in casual speech: <i>molok</i> (590 examples), Portuguese <i>antes</i> (similar frequency). Both easily misinterpreted.</p>	Follow European and Indonesian structure: Pressure to translate literally.
<p><b>Condition:</b> often by parataxis, can use <i>kalo</i> from Indonesian.</p>	<p><b>Condition:</b> Can use parataxis. Non-liturgical writing normally uses Portuguese <i>se</i> (over 4000 examples), some additionally use <i>karik</i> ‘maybe’; liturgical TD uses <i>karik</i> without <i>se</i>.</p>	Portuguese loan, follows European and Indonesian structure.

## 9. Clauses

Tetun Terik	Tetun Dili	Likely source
<b>Negation by preverbal <i>la</i> ‘not’ or emphatic <i>lahoos</i>.</b>	<b>Gain discontinuous negation</b> for short constituents in spoken TD: <i>Hau la baa ida</i> ‘1S not go one’ = ‘I didn’t go.’	Mambae calque
<b>Post-posed <i>lalek</i> ‘without’</b> : <i>karian lalek</i> ‘work without’ = ‘jobless’.	<b>Lost, replaced by negated verb <i>la iha</i></b> ‘not have’ (TD <i>la ho inan</i> ‘not have mother’; Fehan <i>inan lalek</i> ‘mother-less’). Variant form <i>laek</i> is still used in the liturgical register: <i>rohan-laek</i> ‘end-less’ = ‘eternal’.	Regularisation of grammar.
<b>SOV order allowed</b> in irrealis clauses, although low frequency: <i>Ha’u kopi k-emu ha’i</i> ‘1S coffee 1S-drink not’ = ‘I don’t drink coffee’.	<b>SOV order allowed, and still low frequency</b> , but not acceptable in writing. Speakers are unaware of the construction, even if they use it.	Remove structure from formal Tetun that isn’t in European languages and Indonesian.
<b>‘try’ expressed by a post-modifier</b> : <i>hemu kokon</i> ‘drink try’ = ‘have a taste’	<b>TT order used but rare.</b> <b>Most reverse the order</b> : <i>koko naok</i> ‘try (to) steal’, often with the auxiliary <i>atu</i> (60 examples): <i>koko atu hahuu</i> ‘try IRREALIS start’ = ‘try to start’.	Calque on Portuguese and Indonesian, and matches complement-taking verbs in Tetun, removing exceptions.
<b>Verb never precedes subject, except presentative <i>noo</i> ‘exist, there is/are...’</b>	<b>Verb-initial order allowed for some verbs</b> in addition to <i>iha</i> ‘exist, there is’, particularly by Portuguese-influenced speakers. Portuguese <i>akontese</i> ‘occur’, <i>falta</i> ‘be absent, be missing’ Tetun: <i>mosu</i> ‘appear’ (P. <i>aperecer</i> ), <i>hela</i> ‘remain’ (P. <i>ficar</i> ), <i>moris</i> ‘live’ (P. <i>viver</i> ).	Portuguese calque
<b>No true copula</b>	<b>New copula <i>hanesan</i></b> , lit. ‘like; alike; for example’: <i>Ida nee hanesan krimi ida</i> ‘one this like crime one’ = ‘This is a crime.’ Becoming common in newspapers and translations; not found in oral texts or liturgical Tetun.	Portuguese influence.
<b>No passive</b>	<b>Passive developing</b> in written and translated texts: <i>governu nebee lidera hosi Alkatiri</i> ‘government REL lead from Alkatiri’ = ‘the government led by Alkatiri’. (For further details see Williams-van Klinken (in press).)	Calque on Portuguese, English, Indonesian, pushed by translation demands.
<b>Tense-aspect</b>	<b>TD appears to follow East dialect</b> ; it is quite different to Fehan in West Timor. More detailed descriptions needed.	

## Conclusion

Tetun Dili shows evidence of: extensive loans, simplification, removal of exceptions, calquing on Mambae and Portuguese, grammaticalisation, reanalysis of grammatical morphemes as part of the root. Most deep-rooted differences from TT, found in the speech of all Tetun Dili speakers, are attributable to Mambae influence and simplification. Portuguese influence is much more common in writing, formal speech, and the speech of educated people than in casual uneducated speech. These higher registers have also been influenced by Portuguese, English and Indonesian styles of language.

## References

- Costa, Luís, 2000, *Dicionário de Tétum-Português*. Lisbon: Edições Colibri.
- Fox, James J., 1997, The historical position of Tetun among the languages of the Timor area. *Tetun language conference, Darwin*. [??Check whether it has been published.]
- Hajek, John, 2006, Language contact and convergence in East Timor: the case of Tetun Dili. In Aikhenvald, Alexandra Y. and Robert M. W. Dixon, eds *Grammars in contact: a cross-linguistic typology*. Oxford: Oxford University Press.
- Hull, Geoffrey, 1996, *Mai kolia Tetun: a beginner's course in Tetun-Praça: the lingua franca of East Timor* (second edition). North Sydney: Australian Catholic Relief and the Australian Catholic Social Justice Council.
- 2001, *Mambai language manual (Ainaro dialect)*. Sydney: Sebastião Aparícu da Silva Project.
- 2002, *The languages of East Timor: some basic facts*. [http://www.portphillip.vic.gov.au/default/CommunityGovernanceDocuments/The\\_Languages\\_of\\_East\\_Timor\\_Some\\_Basic\\_Facts.pdf](http://www.portphillip.vic.gov.au/default/CommunityGovernanceDocuments/The_Languages_of_East_Timor_Some_Basic_Facts.pdf) - downloaded 27/7/2010. Accessed 27/7/2010.
- 2003, *Southern Mambai (Manbae-Ainaru nor Same)*. Dili: Instituto Nacional de Linguística, Universidade Nacional Timor Lorosa'e.
- Mathijsen, A., 1906, *Tettum-Hollandsche woordenlijst met beknopte spraakkunst*. Batavia: Albrecht & Co.
- 1967, *Testamento Tuan nian dale hodi lia Tetum*. Ende, Flores: Pertjetakan Arnoldus.
- Morris, Cliff, 1984, *Tetun - English dictionary*. Canberra: Pacific Linguistics, C-83.
- Nordholt, H.G. Schulte, 1971, *The political system of the Atoni of Timor*. The Hague: Martinus Nijhoff.
- Siegel, Jeff, 1997, Mixing, leveling and pidgin/creole development. In Arthur K. Spears and Donald Winford, eds *The structure and status of pidgins and creoles*. Amsterdam/Philadelphia: John Benjamins.
- Thomaz, Luis Filipe F.R., 1981, The formation of Tetun-Praça, vehicular language of East Timor. In Nigel Phillips and Anwar Khaidir, eds *Papers on Indonesian languages and literatures*, 54-83. Paris: Cahier d'Archipel 13.
- 2002, *Babel Loro Sa'e: O Problema Linguístico de Timor-Leste*. ??: Instituto Camões.
- 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, in press, Voice and valency in Tetun Dili. In Ewing, Michael and Marian Klamer, eds *Typological and areal analyses: Contributions from East Nusantara*. Canberra: Pacific Linguistics.
- Williams-van Klinken, Catharina, John Hajek, and Rachel Nordlinger, 2002, *Tetun Dili: a grammar of an East Timorese language*. Canberra: Pacific Linguistics, 520.
- Williams-van Klinken, Catharina, and John Hajek, 2009, Double agent, double cross? Or how a suffix changes sides in an isolating language: *dór* in Tetun Dili. *Eleventh International Conference on Austronesian Linguistics*, Aussois, France.