

# OS CONTATOS LINGUÍSTICOS EM TIMOR-LESTE: MUDANÇAS E REESTRUTURAÇÃO GRAMATICAL

Davi Borges de Albuquerque\*

**Resumo:** Timor-Leste é uma pequena ilha localizada no sudeste asiático e tem como línguas oficiais o português e o Tétum-Praça. Há milhares de anos esta ilha é alvo de contato entre povos e línguas pelo fato de fazer parte de uma rota comercial asiática documentada há séculos desde um período anterior aos europeus. Através de evidências arqueológicas e linguísticas é possível mapear migrações de povos pré-austronésios, austronésios e papuásicos, em tempos pré-históricos e em tempos históricos. Este artigo procura analisar o contato de línguas em Timor-Leste, quais línguas nativas sofreram mudanças devido ao contato e quais sofreram reestruturação gramatical devido a um contato ainda mais intenso. Ainda, as línguas de origem austronésia e papuásica de Timor-Leste sofreram mudanças linguísticas distintas, como o Tétum, o Manbae e o Tokodede que tiveram contato intenso e mudanças induzidas pelo contato, enquanto outras línguas, como o Baikenu, o Makasae e o Bunak, sofreram poucas mudanças linguísticas.

**Palavras-chave:** Timor-Leste. Contato de línguas. Mudança linguística. Reestruturação gramatical.

**Abstract:** East Timor is a small island located at Southeastern Asia and present Portuguese and Tétum-Praça as its official languages. For thousand years this island has suffered contacts with several people and languages due to its place within a commercial trade route in Asia which is documented for several centuries, even before European. Through archeological and linguistic evidences, it is possible to map the migration of Pre-Austronesian, Austronesian and Papuasic people dating pre-historical and historical times. This paper intends to analyze different degrees of language contact in East Timor, contact-induced change underwent on Timorese native languages, and in some cases grammatical restructuring due to intense contact. Furthermore, Austronesian and Papuan languages of East Timor have had divergent linguistic change, some languages, such as: Tetum, Mambae, and Tokodede, have intense contact and great contact-induced change, others languages, such as: Baikenu, Makasae, and Bunak, have only slight contact-induced change.

**Keywords:** East Timor. Language Contact. Contact-induced Change. Grammatical Restructuring.

---

\* Doutor em Linguística pela Universidade de Brasília (UnB). Foi professor cooperante da Universidade Nacional Timor Lorosa'e (UNTL), em Dili, Timor-Leste, e professor substituto da Universidade Federal de Sergipe (UFS). É membro do Núcleo de Pesquisa NELIM (Núcleo de Estudos de Ecolinguística e Imaginário), na Universidade Federal de Goiás, Goiânia, Brasil. E-mail: [albuquerque00@hotmail.com](mailto:albuquerque00@hotmail.com).

## 1 – Introdução<sup>1</sup>

Timor-Leste é uma pequena ilha localizada no sudeste asiático e há milhares de anos é alvo de contato linguístico entre diversos povos. Através de evidências arqueológicas e linguísticas é possível mapear migrações de povos pré-austronésios, austronésios e papuásicos, em tempos pré-históricos e em tempos históricos. Após as primeiras migrações pré-austronésias e papuásicas, em tempos pré-históricos, ocorreu uma migração austronésia, de povos Ambônicos, originários de Molucas, em direção à ilha de Timor. Vale lembrar que Timor fez parte da rota comercial do sudeste asiático, tendo contato com os reinos indianizados, o império chinês, povos árabes e os sultanatos malaios, e, finalmente, o contato com os colonizadores europeus, portugueses e holandeses.

Após sofrer uma invasão da Indonésia em 1975 que se estendeu até 1999, a República Democrática de Timor-Leste reconheceu, na constituição de 2002, a língua portuguesa e a língua Tétum<sup>2</sup> (em sua variedade Tétum-Praça<sup>3</sup>) como línguas oficiais, e a língua inglesa e o Malaio (em sua variedade *bahasa indonesia*) como línguas de trabalho. Ainda, há 16 línguas nativas pertencentes a duas filiações genéticas diferentes, a saber: 12 línguas austronésias e 4 línguas papuásicas que se encontram distribuídas pelo território leste-timorense de acordo com o mapa:

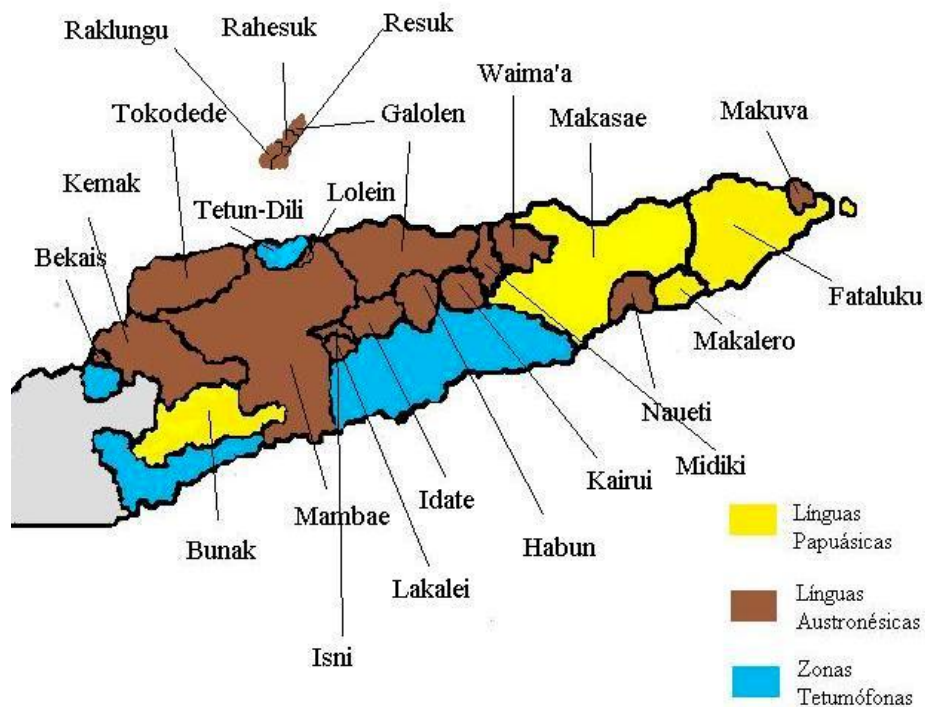
Mapa 1. Timor-Leste e a distribuição das línguas nativas pelo seu território

---

<sup>1</sup> Abreviaturas utilizadas neste trabalho: 1sg ‘1ª pessoa do singular’, 1EXC ‘1ª pessoa do plural exclusiva’, 1INC ‘1ª pessoa do plural inclusiva’, 2 ‘2ª pessoa’, 3pl ‘3ª pessoa do plural’, CAU ‘causativo’, CLF.HUM ‘classificador humano’, CNTF ‘marcador diretivo centrífugo’, COP ‘cópula’, DEF ‘definido’, DEM ‘demonstrativo’, DIST ‘distancial’, ENF ‘ênfase’, EXCL ‘exclamação’, EXI ‘existencial’, FOC ‘foco’, GEN ‘genitivo’, IN ‘inalienável’, LOC ‘locativo’, NEG ‘negação’, NOM ‘nominalizador’, PERF ‘perfectivo’, PL ‘plural’, POS ‘possessivo’, PROG ‘progressivo’, PRX ‘proximidade’, RED ‘reduplicação’, REL ‘relativizador’ e TOP ‘topicalizador’. Foram utilizadas também as abreviaturas específicas dos autores citados no decorrer deste artigo, porém elas se encontram explicadas antes dos exemplos em que são empregadas.

<sup>2</sup> Grafias alternativas *Tetum*, *Tétum* e *Tetun*. Optou-se pela grafia *Tétum*, e para seus compostos também, como *tetunófono*, *Tétum-Praça* pelo fato de ser a escrita mais adaptada à ortografia portuguesa, bem como ser essa forma a mais usada em publicações acadêmicas.

<sup>3</sup> Há duas principais variedades, a saber: Tétum-Praça e Tétum-Terik (ALBUQUERQUE, 2011b; THOMAZ, 2002). Há certa controvérsia em relação à variação dialetal da língua Tétum. Alguns autores, como Klinken (1999) e Hull e Eccles (2001), subdividem a variedade Tétum-Terik em duas, a saber: Tétum-Terik e Tétum-Fehan, acabando por considerar estas como duas variedades diferentes, aumentando o número para três variedades diferentes.



(Fonte: Albuquerque, 2011a, adaptado)

No presente trabalho apresentarei as estruturas linguísticas da língua Tétum, comparando as variedades Tétum-Praça (que sofreu diversas mudanças linguísticas devido ao contato) e Tétum-Fehan (que sofreu menos contato linguístico) por ser falada em regiões geográficas mais isoladas. Ainda, serão apresentados dados das demais línguas leste-timorenses, procurando estabelecer quais línguas sofreram mais mudanças gramaticais com o contato linguístico e quais sofreram menos e, assim, mantiveram maior número de estruturas linguísticas nativas. A comparação entre as duas variedades da língua Tétum, complementada pelos dados das demais línguas faladas em Timor-Leste, servirão para a discussão da relação entre mudança linguística e contato de línguas, bem como a classificação das línguas leste-timorenses que, pelo fato de apresentarem poucas estruturas de morfologia flexional, são apontadas como pidgins, crioulos, entre outros.

## 2 – Histórico dos contatos

A língua Tétum é usada como língua franca em uma grande parte da ilha de Timor desde tempos anteriores à chegada dos portugueses, provavelmente a partir do século XV, quando o reino tetumófono *Wehale* se espalhou a partir da região centro-leste da Ilha de Timor e conquistou grande parte do território e de outros povos (THOMAZ, 2002).

Ainda, há indícios de que outras línguas serviram como língua franca no sudeste asiático, são elas: o *Bazar Melayu* – variedade crioula do Malaio – usada inicialmente nos portos para a comunicação entre os diversos comerciantes, aproximadamente no século XV e o Crioulo Português de Malaca (CPMal) que teve um papel importante durante o século XVI. Desta maneira, há também a possibilidade de que essas duas línguas tenham influenciado em alguma medida o Tétum-Praça<sup>4</sup>.

O Tétum-Praça é a variedade que possui o status de língua oficial e é aquela que funciona como língua franca até os dias atuais. Ela é uma das línguas leste-timorenses que sofreu um intenso contato linguístico, conforme foi mencionado anteriormente. O número de contatos entre povos, migrações, rotas e atividades comerciais etc. foi grande e ocorreram em períodos históricos distintos. A seguir encontra-se um resumo dos prováveis contatos, juntamente com suas influências nas línguas e uma provável datação:

- Em tempos pré-históricos ocorreu um contato entre os povos papuásicos, que já habitavam a ilha de Timor, com o povo que realizou a primeira migração austronésia. Este longo contato linguístico entre as línguas austronésias e papuásicas em Timor foi o fator mais importante para a formação de uma área linguística (HULL, 2001);
- A região central de Timor-Leste apresenta indícios de habitação de povos pré-austronésios distintos dos papuásicos, que já habitavam a ilha, e deram origem às línguas papuásicas. Desta maneira, as línguas da região central do país – Mambae, Tokodede e Kemak – apresentam a influência de substratos distintos que não são de origem austronésia, nem papuásica;
- A segunda migração austronésia (de origem ambônica) causou uma série de empréstimos e simplificações gramaticais, além disso, outro fato interessante é que esse povo foi totalmente assimilado à população que já habitava a ilha;
- O contato com os comerciantes de diversas origens étnicas deve ter se realizado com o uso do *Bazar Melayu* (VERSTEEGH, 2008, p. 175). Este fator causou uma série de empréstimos e outro processo de simplificação de algumas estruturas gramaticais, principalmente no Tétum (HULL, 2005), mas também nas demais línguas faladas em Timor-Leste (HULL, 2001);
- No século XV o reino tetumófono de Wehale iniciou sua expansão e dominação dos demais reinos. Desta forma, a língua Tétum começou a ser aprendida como a língua do

---

<sup>4</sup> Esperança (2001) apresenta evidências de uma possível influência do Português Crioulo de Bidau (PCB) na formação do Tétum-Praça.

reino dominador. Assim, a aquisição não nativa (MCWHORTER, 2008) do Tétum causou mais uma nova série de simplificações gramaticais;

- A mudança da capital do Timor português, no ano 1769, de Lifau – hoje a região do enclave de Oecussi – para Díli, que era uma região falante de Mambae, causou mais uma série de simplificações gramaticais;
- Nos últimos anos, o Tétum-Praça, em sua variedade Tétum-Díli<sup>5</sup>, como língua urbana vem recorrendo à língua portuguesa para realizar um grande número de empréstimos relativos ao mundo moderno.

### 3 – A língua Tétum

O Tétum-Praça (doravante TP), quando comparado ao Tétum-Fehan (TF)<sup>6</sup>, apresenta um grande número de exemplos que corroboram para a hipótese da reestruturação gramatical devido aos diversos contatos mencionados anteriormente. Outra evidência de reestruturação gramatical é que diversas mudanças ocorreram em todos os níveis de análise linguística: fonologia, morfologia, sintaxe e léxico do TP.

Na fonologia, o TP possui diversos empréstimos fonológicos; na morfologia, a simplificação de estruturas morfológicas e algumas morfossintáticas; no léxico, uma grande quantidade de empréstimos linguísticos do português e do malaio.

De maneira distinta, o TF mantém um número considerável de estruturas linguísticas nativas. Somente a sua fonologia é que se apresenta mais simplificada, quando comparada com a do TP, pois o TF não apresenta os empréstimos fonológicos do português. Porém, muitas estruturas nativas foram mantidas no TF, principalmente as estruturas morfossintáticas.

Entre essas estruturas gramaticais presentes no TF<sup>7</sup>, podemos citar: uma série de prefixos marcadores de sujeito no verbo TF, ausente em TP; um conjunto de afixos operadores de valência e derivação em TF, com somente alguns resquícios em TP; o uso de reduplicação parcial e total em TF, enquanto este processo é limitado em TP; TP não apresenta a cópula TF *nii*; TF possui dois verbos existenciais: *iha* e *nó*, enquanto o TP manteve somente o *iha*; o sistema de negação complexo do TF foi reduzido em TP; TF apresenta um sistema de

---

<sup>5</sup> Hull (2002) apresenta um esboço gramatical do Tétum-Dili, já Albuquerque (2011, p. 56) fornece uma discussão detalhada a respeito da diferencial dialetal e variação linguística da língua Tétum.

<sup>6</sup> Variedade do Tétum-Terik que possui estudos linguísticos significativos, como a descrição elaborada por Klinken (1999), que será usada neste trabalho, por isso sua escolha para a presente análise.

<sup>7</sup> Os dados presentes nos exemplos e as tabelas sobre o Tetum-Fehan foram extraídos de Klinken (1999) e outros foram gentilmente cedidos pela autora.

classificação nominal marcada por classificadores numerais, o TP mantém somente três desses classificadores.

A seguir, será feita uma breve análise, com a apresentação de alguns dados, de cada uma das características apontadas anteriormente.

O TF possui uma série de prefixos marcadores de sujeito no verbo que são flexionados de acordo com a pessoa, os quais se perderam totalmente em TP:

TF	Oa	ne'e	oo	m-atene	lale?	
	criança	PRX	2sg	2-conhecer	não	
TP	Oan	ne'e, ó	hatene	ka	lae?	
	criança este	2sg	conhecer	ou	não	
	'você conhece esta criança?'					
1. TF	Sia	r-alai	r-ola	r-ikar	loro-sa'e=n	baa.
	3pl	3-correr	3-ir	3-costas	sol-subir=GEN	DIST
TP	Sira	halai	ba	lorosa'e		
	3pl	correr	CNTF	leste		
	'eles fugiram em direção a leste'.					

TF apresenta um conjunto de afixos operadores de valência e derivação que se apresentam produtivos em TF e possuem somente alguns resquícios em registros literários do TP. A tabela abaixo os apresenta em TF:

Tabela 1. Derivação de nomes não abstratos para verbos

Afixo	Raiz	Glossa	Nome	Glossa
RDP parcial	simu	'receber'	sa-simu-n	'coisa recebida'
k(a)- -k	lele	'flutuar'	k-lele-k	'barco'

A seguir encontram-se alguns exemplos destes mesmos afixos em TP:

2. *fahe* ‘dividir’ > *fafahek* ‘divisão’, *la’o* ‘andar’ > *lala’ok* ‘andamento, progresso’, *seluk* ‘outro’ > *saseluk* ‘substituo’, *mii* ‘urinar’ > *mamiik* ‘bexiga’.

A reduplicação parcial e total de vocábulos para expressar diversas noções semânticas é produtiva em TF, já em TP este processo é limitado e não produtivo em TP. Digno de nota é que tanto no exemplo em (3) e na tab. (1), há também exemplos de reduplicação, porém a função para esse tipo de reduplicação é bem específica, a derivação. A reduplicação em TP somente é usada em alguns itens lexicais da língua e que de certa forma já foram cristalizados, sendo a maioria são adjetivos e advérbios, com o sentido de serialização ou intensificação, ou para fazer o plural de alguns nomes que são empréstimos linguísticos<sup>8</sup>:

3. Reduplicação em TP (ALBUQUERQUE, 2011b, p. 116):

*bain* ‘tempo’ > *baibain* ‘todo o tempo’

*loron* ‘dia’ > *lorloron* ‘todo dia’

*ida* ‘um’ > *idaidak* ‘um por um’

*liu* ‘acima, passar’ > *liuliu* ‘especialmente, acima de tudo’

Tabela 2. Reduplicação parcial em TF

Sílaba	Base	Glossa	Reduplicação	Glossa
CVV	hoo	‘ter’	ha-hoo-k	‘poses’
CVCVC	sukat	‘medir’	sa-sukat	‘ferramenta para medir’
CCVCVC	krakat	‘ter.raiva’	k-ra-rakat	‘raiva’
Trissílaba	hakees	‘falar’	ka-ke-es	‘fala (n.)’

O TP não apresenta a cópula *nii*, existente em TF. TP não possui nenhuma cópula:

<sup>8</sup> A reduplicação em Tétum-Praça foi analisada detalhadamente por diferentes linguistas, destacando-se Esperança (2001), Avram (2008) e Albuquerque (2011b, p. 116).

4. TF    lale,    tais    oo=k            nii    nia  
           NEG   roupa   2sg=GEN    COP   3sg
- TP    lae,    o=nia            hena   mak   ida    ne'ebá.  
       NEG   2sg=POS    roupa   FOC   um    DIST  
       'Esta roupa não é tua.'

Em relação aos verbos existenciais, TF apresenta dois verbos: *iha* e *nó*, enquanto o TP manteve somente o *iha*:

5. TF    Ai kanoik    nee    hori    rai    moris   noo    kedas.  
           história    este    desde   terra    nascer   EXI    imediatamente
- TP    Istoría            ne'e    iha    hori    rai    moris.  
       história            este    EXI    desde   terra    nascer  
       'Esta história existe desde o início do mundo.'

6. TF    Ei!    Noo    feto    ida.    Noo    feto    ida    noo    oa!  
           EXCL   EXIST   mulher   um    EXI    mulher   um    e    criança
- TP    Ei!    Iha    feto    ida.    Iha    feto    ida    ho    oan!  
       EXCL   EXI    mulher   um    EXI    mulher   um    e    criança  
       'Ei! Há uma mulher! Há uma mulher e uma criança!' (Dito por um homem que acabou de vê-los em uma ilha distante)

O TF apresenta um sistema de negação complexo em que a negação é marcada por diferentes itens lexicais ou gramaticais, sendo distintos também de acordo com a classe gramatical do item que está negando. Esse sistema do TF foi reduzido para três itens em TP: a negação verbal *la=*, a negação nominal *la'ós* e a negação absoluta *lae*.

7. TF    Oo!    Buat    e'e    Bei    Beur            ha'i!  
           oh    coisa   PRX   Sr.    enganar        não



- TP Oo! Buat ne'e lae señór bosok.  
EXCL coisa este NEG senhor mentira  
Oh! Esta coisa (na realidade uma pessoa) não é o Sr. Enganador!
8. TF Ne'e lahoos ema lian. Manu lian.  
PRX realmente.não pessoa voz pássaro voz
- TP Ne'e ne'e lae ema lian manu lian.  
Isto TOP NEG pessoa voz pássaro voz  
Isto não é a voz de uma pessoa. (É a) voz de um pássaro.

O TF possui um sistema de classificação nominal marcada por classificadores numerais, o TP mantém somente três dos classificadores, porém com o uso facultativo e limitado a situações formais. Atualmente, só o classificador numeral humano [+ HUM] ainda é utilizado:

9. TF feto kawa'ik na'in neen ne'e  
mulher velho CLF.HUM seis PRX
- TP ferik ne'e (na'in) neen.  
velha este seis  
'estas seis velhas (senhoras).'

Tabela 3. Classificadores numerais

Classificador	Significado	Classificador para
na'in	nobre	pessoas
matan	olho	búfalo e porco
lolon	galho	objetos cilíndricos grandes, ex. vela, peixes.
tahan	folha	objetos chatos e finos, ex. papel, roupa.
fuan	fruta, coração	objetos arredondados grandes, ex. coco, ovo.
musan	semente	objetos arredondados pequenos.

Esses fatores fizeram com que as diferentes variedades citadas da língua Tétum fossem classificadas de diferentes maneiras: processo de crioulização (HULL, 2001); crioulo (ETHNOLOGUE, 2009); pidgin (HAGÈGE, 2002 *apud* HAJEK, 2007). Na realidade, se verificarmos os dados linguísticos do Tétum, das demais línguas leste-timorenses e da bibliografia linguística a respeito do assunto, é possível perceber que tais classificações estão incorretas.

O Tétum-Praça, quando comparado ao Tétum-Fehan, apresenta um grande número de exemplos que corrobora para a hipótese de reestruturação gramatical (HOLM, 2004; LUCCHESI & BAXTER, 2009), que ocorreu por causa da transmissão linguística irregular, ou aquisição não nativa, conceitos que serão comentados abaixo. Em relação aos aspectos linguísticos e sócio-históricos, vale a pena salientar os seguintes: os diversos contatos linguísticos sofridos pelos diferentes povos leste-timorenses; a expansão dos reinos tetumófonos de *Servião* e *Belo*, dominando e impondo a língua Tetun aos povos vizinhos, o que facilitou seu uso como língua franca. Outra evidência de reestruturação gramatical é que diversas mudanças ocorreram em todos os níveis de análise: fonologia, morfologia, sintaxe e léxico, mas somente nas línguas de Timor-Leste que se localizavam nas zonas que sofreram maior contato linguístico, enquanto as línguas em regiões mais isoladas apresentam apenas um número reduzido de empréstimos e mudanças linguísticas, como é o caso das zonas falantes de Tétum-Fehan.

Seguindo o conceito de transmissão linguística irregular (LUCCHESI, 1994), ou de aquisição não nativa (MCWHORTER, 2007), que, grosso modo, são semelhantes e consistem na interrupção e/ou alteração de alguma forma na transmissão linguística de uma geração para outra, defende-se aqui que o Tétum-Praça sofreu erosão de sua morfologia flexional, por causa de diversos processos sócio-históricos que influenciaram a mudança linguística, entre eles: o contato com os comerciantes de diversas origens étnicas mediado pelo uso do *Pazar Melayu* por volta dos séculos XIII e XIV; no século XV, o reino tetumófono de *Wehale* iniciou sua expansão e dominação dos demais reinos. Dessa forma, a língua Tetun começou a ser aprendida como a língua do reino dominador; a mudança da capital do *Timor Português*, no ano 1769, de Lifau (hoje a região do enclave de Oecussi) para Díli, que era uma região falante de Manbae.

Ainda, trabalhando com a hipótese de Ansaldo e Nordhoff (2009) de que simplicidade e complexidade gramaticais não estão ligadas com a idade da língua, mas com o meio ambiente linguístico, em outras palavras, caso o meio ambiente linguístico está cercado por línguas morfologicamente ricas, logo influenciará na formação de uma morfologia mais rica; caso o meio ambiente linguístico seja isolante, influenciará na formação de uma língua mais

isolante, argumento também que o TP vem desenvolvendo uma “nova” morfologia flexional que está em vias de gramaticalização por causa da convivência extensa com diferentes variedades da língua portuguesa do século XVI até o XIX, principalmente com o Crioulos Portugueses Asiáticos, principalmente os Crioulos Portugueses de Malaca e Macau, e no século XX até a atualidade com a língua portuguesa e a língua inglesa. A formação dessa morfologia flexional, que será descrita abaixo, consiste no outro aspecto da reestruturação gramatical que, além de incluir a erosão da morfologia flexional, consiste na transferência funcional, o uso de um lexema com uma nova função gramatical, conforme foi estudado amplamente por Siegel (2008).

#### **4 – Línguas Austronésias**

As línguas austronésias faladas em Timor-Leste sofreram mudanças devido ao contato em escalas distintas. As línguas faladas na região centro-norte do país, região mais afetada pelo contato linguístico, apresentam uma reestruturação gramatical, como o Mambae, Tokodede e o Tétum-Praça, como foi analisado anteriormente. Já as línguas que ocupavam regiões relativamente isoladas conseqüentemente sofreram menos mudanças linguísticas com o contato, apresentando um maior número de retenções da proto-língua original, foram elas o Tétum-Terik, Baikenu e os dialetos da ilha de Ataúro (ALBUQUERQUE, 2009; HULL, 2001).

A seguir serão analisados alguns dados dessas línguas leste-timorenses. Será possível perceber por meio dos dados apresentados que as mudanças linguísticas ocorrem de maneira natural, mesmo sendo condicionadas pelo contato de línguas e povos, e que o par ‘complexidade/simplicidade linguística’, além de ser uma característica um tanto relativa, não tem ligação com a idade da língua, com as mudanças e nem com os contatos linguísticos, bem como não é um parâmetro importante para a classificação de uma língua como pidgin ou crioula.

##### **4.1 – Mambae**

A língua Mambae é falada em diversos distritos por toda a região central de Timor-Leste. Em um de seus primeiros estudos linguísticos a respeito do Mambae, Hull (2001) afirma que essa língua não possui flexão e nem outras estruturas gramaticais, o que não está em concordância com os dados que serão apresentados mais adiante. Posteriormente, o próprio autor revisou esta análise (HULL, 2003). Tanto o Manbae, quanto o Tokodede, outra língua da mesma região de Timor-Leste, são apontadas como possuidoras de estruturas crioualizadas, pelo fato de

apresentarem um histórico de contato intenso, reestruturação gramatical e morfologia analítica, conforme alguns crioulistas afirmam, como McWhorter (2007).

O que ocorre, porém, é algo diferente. As línguas Mambae e Tokodede, por estarem localizadas na região central de Timor-Leste, foram aquelas que sofreram maior impacto com os sucessivos contatos e, assim, ocorreu um verdadeiro processo de reestruturação gramatical com a perda de diversas estruturas morfossintáticas, fazendo com que certos itens lexicais fossem gramaticalizados para assumir a posição e a função dos itens gramaticais perdidos. Vale lembrar que esses processos são mais bem compreendidos em Manbae pelo fato de esta ser uma das línguas mais estudadas de Timor-Leste, além do Tétum<sup>9</sup>.

Assim, a língua Manbae possui a presença de morfologia flexional, tanto de morfemas que são resquícios da protolíngua (HULL, 2001), como também os mesmos cognatos gramaticalizados do Tétum-Praça, estando um estágio mais avançado de gramaticalização em Manbae<sup>10</sup>, pois nesta língua são sufixos ou clíticos. Em Hull (2003a) e Albuquerque (2013), o Manbae apresenta os seguintes afixos e clíticos: prefixo *n-* marcador de 3ª pessoa do singular, em (11); o prefixo *ma-* ‘intransitivizador’<sup>11</sup>, em (12); os sufixos *-ga* para marcar definição e *-id* indefinição, em (13), (14) e (16); *-ser* para marcar o definido plural, em (13) e (115); o enclítico *=ni* para marcar posse, em (14) e (16); os sufixos que marcam a flexão de gênero são *man* ‘masculino’ e *hin* ‘feminino’, e *ama* ‘macho’ e *ina* ‘fêmea’<sup>12</sup>, em (15) e (17)<sup>13</sup>:

## 10. Flexão pessoal em Manbae

---

<sup>9</sup> É possível citar os seguintes estudos linguísticos sobre o Manbae nos últimos anos: Corte-Real e Hull (1998), em que há uma análise de vários textos Manbae; Hull (2003) em que o autor faz um esboço gramatical da língua; Albuquerque (2013) em que o autor realiza um estudo do contato de línguas e seu impacto em Manbae; Fogaça (2013) em que é elaborada uma descrição fonética e fonológica da língua.

<sup>10</sup> Os cognatos entre Tétum-Praça e Manbae são os seguintes: Tt. *Ida Mb. -id*, Tt. *sira Mb. -ser*, Tt. *mane Mb. -man*, Tt. *feto Mb. -hin*, Tt. *aman Mb. -ama* e Tt. *inan Mb. -ina*, sendo que em Tetun Prasa são lexemas gramaticalizados, enquanto no Manbae estes mesmos lexemas são usados como formas livres (lexemas propriamente dito) ou formas presas (sufixos).

<sup>11</sup> Esses dois primeiros prefixos, *n-* e *ma-*, são retenções do proto-austronésio, de acordo com Hull (2001), Hull (2003) e Albuquerque (2013). Há vários outros afixos, não citados aqui, apontados nas publicações anteriores, que também são retenções dessa proto-língua.

<sup>12</sup> Hull (2001) afirma que esses e vários outros traços tipológicos fazem parte da área linguística de Timor-Leste, ou área timórica oriental. Digno de nota é que Timor-Leste encontra-se inserido em uma área linguística maior, conhecida como Nusantara Oriental (KLAMER, REESINK e STADEN 2007), e compartilha esses mesmos traços com as línguas faladas nela.

<sup>13</sup> Os dados do Manbae foram extraídos de Albuquerque (2013, p. 269).

au ét 'eu vejo'  
 ó ét 'tu vês'  
 ua **n-ét** 'ele vê'  
 it/ am ét 'nós vemos'  
 im ét 'vós vedes'  
 ró ét 'eles veem'

11. salmatan **ma-laé** pod atub met  
 porta INT-abrir para pessoa tudo  
 'a porta está aberta para todas as pessoas'

12. méstri-**ga** babar tél iskolanti-**ser** man pada.  
 professor-DEF mandar PERF estudante-PL para casa  
 'O professor mandou os estudantes para casa.'

13. au=**ni** ama né ruif-**id** laó aus  
 1sg=POS pai dar osso-IND para cão  
 'Meu pai deu um osso ao cachorro.'

14. lóba-**man-ser** balikan istuda  
 criança-MS-PL não.quer estudar  
 'Os meninos não querem estudar.'

15. léb kdéi néi-**ga** it=**ni** sos  
 poder deixar LOC-este 2sg=POS compra  
 'Pode deixar aqui suas compras.'

16. ua ténki sos maun-**ama**  
 3sg tem comprar pássaro-macho  
 'Ele tem que comprar um galo.'

## 4.2 Tokodede

A língua Tokodede, língua materna do distrito de Liquiçá, apresenta morfologia isolante e sintaxe paratática e é considerada, juntamente com o Mambae e o Tétum-Praça, uma das línguas mais afetadas pelos contatos linguísticos (HULL, 2001). A reestruturação gramatical sofrida pelo Mambae foi semelhante a do Tokodede, com a simplificação de grande parte das estruturas morfossintáticas. O Tokodede se destaca, ainda, por apresentar menos resquícios das formas antigas com apenas a retenção parcial de uma série de pronomes possessivos.

17. aka'u so'a di'a kuartu telu, sala ru, diuhati.  
 1POS casa ter quarto três sala dois banheiro  
 'minha casa tem três quartos, duas salas e um banheiro'

As estruturas atuais da língua que diferenciam o Tokodede são as seguintes: uma possível formação de gênero discursivo, com o uso dos vocábulos TK. *mane* 'homem' ou *hine* 'mulher' para especificar o sexo da 1ª pessoa do discurso, como no exemplo:

18. aipil mane-kami uza kami=ni'i lapar, mane-kami di'a osa dada  
 quando homem-1EXC usar 1=POS coisa homem-1EXC ter dinheiro muito  
 'quando nós (masc.) usarmos nossos bens, então nós (masc.) teremos muito dinheiro'

Ainda, o Tokodede compartilha com o Tétum-Praça e o Mambae uma série de marcadores aspectuais e modais (17), e uma flexão, quanto ao número e à definição, do 'relativizador': TP *ne'ebé*, *ida-ne'ebé*, *sira-ne'ebé*; MB. *bae*, *id-bae*; TK. *mane* ~ *ke*, *iso-mane* ~ *iso-ke*:

19. a: peni bali her aldea besi vila  
 1sg morar PERF LOC aldeia perto vila  
 'eu morei em uma aldeia perto da vila'

20. a: pene mane iso-ke (iso-mane) doso aka'u sa heu  
 1sg ver homem REL queimar 1POS casa novo  
 'Eu vi o homem que queimou minha casa nova'

Os exemplos acima apontam possivelmente para um novo processo de reestruturação gramatical, desta vez, porém, de maneira contrária ao do ocorrido anteriormente, com a emergência de novas categorias gramaticais sendo marcadas como: número, definição, grau de animacidade e gênero natural. Isso está ocorrendo principalmente pelo fato de as comunidades

de fala dessas línguas de Timor-Leste estarem estabilizadas e por algumas gerações estão sendo adquiridas de forma natural e pela comunidade nativa (MCWHORTER, 2007). Entre essas comunidades estão: o Tokodede, que atualmente é a língua materna de um distrito inteiro, Liquiçá, e símbolo de identidade deste povo, e o Mambae falada em vários distritos e símbolo de identidade de uma etnia inteira que vive nas regiões montanhosas de Timor-Leste.

### 4.3 Kemak

A língua Kemak falada na fronteira com a Indonésia, principalmente no distrito de Bobonaro, encontra-se em intenso contato com a língua papuásica Bunak. Este contato aparenta ser antigo, pois o Kemak, de origem austronésica, apresenta diversos traços tipológicos papuásicos, o que fez com que sua classificação fosse duvidosa durante um longo período. Entre esses traços o Kemak apresenta uma divisão de classe de nomes e pronomes entre alienável e inalienável, divisão esta que faz parte da tipologia das línguas papuásicas (KLAMER, REESINK & STADEN, 2007) e, ainda, a presença de uma série de pronomes duais, que é outro traço tipológico das línguas papuásicas, ambos emprestados da língua Bunak.

21. ita=bua            taai    karaka ulala  
1INC=DUAL    NEG    gostar    gostar  
'Nós dois não gostamos de cobra.'

22. imi=bua            la        uma    baipila  
2pl=DUAL        ir        casa    quando  
'quando vocês dois irão para casa?'

23. au            enu    tsilu    bea  
1sg        beber    PROG    água  
'eu estou bebendo água'

24. au:            asu    di:a    au:    uma  
1POS    cão    LOC    1        casa  
'meu cão está em minha casa'

25. au:            nipa-ga            baraga

1POS    dente-IN        doente  
'meu dente está doendo.'

#### 4.4 Baikenu

Baikenu é falado no enclave de Oecussi e, na realidade, trata-se de um dialeto da língua Dawan, que é falada em grande parte do lado oeste da ilha de Timor, que é território indonésio. O enclave de Oecussi, chamado de Lifau na época colonial, foi a capital de Timor Português até o ano de 1769, quando foi transferida para Díli.

A influência tetumófona parece ter sido pouca no Baikenu, assim como o contato linguístico sofrido, pois a língua Dawan apresenta somente uma leve influência do Malaio. Das línguas austronésicas leste-timorenses, então, o Baikenu é a que mantém o maior número de estruturas nativas, como uma série de classificadores e diversos marcadores sintáticos, assim como a série completa de pronomes pessoais sujeito, objeto e possessivo, ainda, esta série completa dos pronomes variam entre tônicos e átonos (MCWHORTER, 2008):

26. Qau        qòt  
      1        cortar  
'eu cortei.'

27. Nuif      qi        qau      qote  
      Osso    esse    1        cortar  
'eu cortei esse osso.'

28. Hò        m-qoet        nuif?  
      2sg        2-cortar      osso  
'você cortou o osso?'

#### 5 – Línguas Papuásicas

As línguas de origem papuásica faladas em Timor-Leste são apenas quatro, a saber: Fataluku, Makalero, Makasae e Bunak. A língua Makasae que se espalhou por uma grande parte do território a leste foi a única que sofreu reestruturação gramatical, enquanto a comunidade Fataluku, apesar de apresentar um maior número de falantes em relação às demais línguas



papuásicas, manteve-se sempre como comunidade monolíngue e relativamente fechada a influências externas. O povo Bunak migrou há um tempo considerável para a região oeste de Timor-Leste e apresenta algumas influências do Tétum-Terik. Finalmente, os falantes de Makalero são em número reduzido e mantiveram-se isolados ao sul do território Makasae, mantendo muita das estruturas gramaticais nativas<sup>14</sup>.

### 5.1 Makasae

O Makasae por ter dominado uma grande parte do território leste da ilha e ter tido contato por um longo período com uma grande parte dos povos austronésicos dentro da ilha de Timor, assim como das migrações vindas do exterior, a língua Makasae é a língua papuásica que mais apresenta traços de reestruturação gramatical com a simplificação de suas estruturas morfossintáticas.

### 5.2 Fataluku

O Fataluku, língua papuásica falada no extremo leste da ilha no distrito de Lautém, manteve uma série de estruturas fonológicas e morfossintáticas complexas por ser uma comunidade tradicional de Timor-Leste, reconhecida por sua resistência ao contato com outros povos, assim como seu relativo isolamento, manutenção e apego a diversos itens tradicionais da cultura material e imaterial (GOMES, 1972). No nível linguístico pode ser percebida esta resistência pelo fato dessa comunidade ser uma das únicas monolíngues em Timor-Leste, resistindo muitas vezes ao aprendizado da língua portuguesa, indonésia e até o TP<sup>15</sup>.

Na morfossintaxe Fataluku há uma série de afixos e clíticos com diversas funções, entre elas mudança de valência e flexão derivacional (ENGELHOVEN, 2009). Abaixo há exemplos de sufixo verbalizador (VEB), clítico sequencial (SEQ), e clítico e afixo de ligação (LIG):

29. Doutor Asumoto            i            vahin naka-ne            tu            tali

---

<sup>14</sup> Como o objetivo do presente trabalho era analisar as mudanças linguísticas e reestruturações gramaticais devido ao contato, a análise dos dados das línguas papuásicas será apenas comentada brevemente, pois se tratam de casos distintos das línguas austronésias, que serão analisados separadamente em um trabalho futuro.

<sup>15</sup> Obviamente não são todos os indivíduos Fataluku que se mantêm monolíngue. Aqueles que trabalham, estudam, ou desenvolvem qualquer outra tarefa na capital, Díli, ou em outros distritos, necessitam ao menos aprender uma das línguas – preferencialmente o indonésio, ou o TP – necessárias para a comunicação com os timorenses de outras etnias, ou falantes de outras línguas.

doutor Asumoto          3sg    dente mancha-VEB    SEQ    muito  
kapar-e                hai      n-apam-e  
mal-VEB                PERF    LIG-puxar-VEB

‘O doutor tirou o dente manchado de Asumoto que estava muito ruim.’

30. I h=iti-jet-ana                                  ia      umani=t            taratori me.  
3sg LIG=profundo-cair-ANA    DEM    quem=SEQ    trator    levar  
‘A questão (aquilo que cai) era quem usaria o trator’.

### 5.3 Bunak

A língua Bunak falada no distrito de Bobonaro e em contato principalmente com o Kemak e o Tétum-Fehan apresenta fonologia, morfologia e léxico bem distintos das línguas vizinhas.

Na morfologia, o Bunak manteve o paradigma de pronomes pessoais inclusivos-exclusivos, singular-dual-plural, a diferença entre pronome sujeito e pronome objeto, e a marcação da oposição alienável x inalienável. Somado a esses paradigmas o Bunak desenvolveu uma série de inovações (BERTHE, 1963), entre elas: o afixo *ha-* causativo, o infixo *-l-* marcador de dualidade, *-o* e *-i* como sufixos marcadores de sg. e pl. nos nomes, e outros elementos adnominais que realizam marcações de número e elementos adverbiais que marcam TAM e alteram a valência verbal.

## 6 – Algumas conclusões

De acordo com as evidências apresentadas, percebeu-se que as línguas nativas de Timor-Leste sofreram mudanças gramaticais distintas conforme a natureza do contato com os outros povos. O Baikenu, o TF e os dialetos de Atauro por estarem localizados em regiões mais isoladas sofreram mudanças gramaticais leves com o contato linguístico; os povos falantes das línguas papuásicas, além de estarem localizados fora da grande área onde ocorreram os contatos linguísticos (a região central de Timor-Leste), estão em regiões mais periféricas; ainda, os povos papuásicos já haviam se fixado em Timor-Leste em um período pré-histórico; as demais línguas austronésias – Mambae, Tokodede, Kemak, entre outras – sofreram um grande impacto pelos sucessivos contatos linguísticos, o que as levou a uma reestruturação gramatical, perdendo uma grande parte de suas estruturas morfossintáticas complexas.

As estruturas morfossintáticas que se perderam no contato linguístico foram reconstruídas em Hull (2001) e entre elas encontram-se atualmente alguns resquícios nas línguas, como os que foram apresentados anteriormente: clíticos e afixos marcadores de pessoa, os diversos afixos marcadores de flexão nominal e verbal, as diferentes formas de classificação nominal, a flexão de posse, as formas complexas de negação.

Desta maneira, pode-se afirmar que a natureza das mudanças linguísticas que ocorreram por causa do contato nas línguas de Timor-Leste afetou muito mais algumas línguas do que outras, de acordo com o esquema abaixo:

– MUDANÇAS

+ MUDANÇAS

Baikenu >	Tétun-Fehan >	Línguas >	Tétum-Praça >	Mambae
		Papuásicas	Kemak	Tokodede

O esquema acima, então, consiste em um resumo do que foi apresentado, mostrando as línguas que sofreram menos mudanças e mais mudanças linguísticas resultantes dos contatos que ocorreram com diversos povos na ilha de Timor. Ainda, Hajek (2007), quando analisa as gramáticas em contato das línguas Baikenu, Tétum e Mambae, chega à mesma conclusão: o Baikenu é a língua que sofreu menos mudança devido ao contato, enquanto o TP sofreu uma mudança mediana se comparado ao Mambae, que sofreu uma reestruturação gramatical.

Porém deve-se ter o cuidado de não confundir mudanças estruturais induzidas pelo contato com crioulização e pidginização. Pidgin é definido por Holm (1988) como uma língua reduzida resultante de um extenso contato entre povos falantes de diferentes línguas geneticamente distantes, enquanto o crioulo consiste na nativização e/ou estabilização de um pidgin. Ainda, segundo o mesmo autor, essa situação ocorre geralmente quando há um povo com status social superior aos demais, e quando o contato é entre povos falantes de línguas geneticamente próximas – como é o caso de Timor-Leste – ocorre ‘coineização’. Desta maneira, com as evidências históricas e linguísticas apresentadas, percebe-se que ocorreu apenas uma ou outra das condições necessárias para o surgimento de pidgins e/ou crioulos. Por isso, não é correto afirmar que as línguas de Timor-Leste na realidade são pidgins, crioulos, ou que sofreram processos de crioulização. O que ocorreu foi somente a influência de diferentes processos sócio-históricos de mudança linguística.

Finalmente, de acordo com a hipótese formulada por McWhorter (2008), que as línguas mantêm alto nível de complexidade gramatical naturalmente, não há possibilidade nenhuma de

uma língua sofrer mudanças gramaticais aleatórias que a tornem semelhante a crioulos. O único fator que altera essa tendência natural das línguas de se manterem complexas é se por motivos sócio-históricos elas passarem a ser adquiridas em massa por não nativos e o caso extremo são os pidgins e crioulos (MCWHORTER, 2007).

## Referências

- ALBUQUERQUE, D. B. Pré-história, história e contato lingüístico em Timor Leste. *Domínios de Lingu@agem*, v.6, n.2, p.75-93, 2009.
- \_\_\_\_\_. O Português de Timor Leste: contribuição para o estudo de uma variedade emergente. *Papia*, v. 21, n.1, p.65-82, 2011a.
- \_\_\_\_\_. *Esboço gramatical do Tetun Prasa: língua oficial de Timor-Leste*. Dissertação (Mestrado em Linguística). Programa de Pós-Graduação em Linguística, Universidade de Brasília, Brasília. 2011b.
- \_\_\_\_\_. Ecologia dos contatos lingüísticos em Manbae, Timor-Leste. In: COUTO, E. K. N. N.; ALBUQUERQUE, D. B.; ARAÚJO, G. P. (Org.). *Da Fonologia à Ecolingüística*. Ensaio em homenagem a Hildo Honório do Couto. Brasília: Thesaurus, 2013. p. 251-283.
- ANSALDO, U.; NORDHOFF, S. Complexity and the age of languages. In E.O. ABOH, E. O.; SMITH, N. (eds.). *Complex processes in new languages*. Amsterdam/ Philadelphia: John Benjamins, 2009.
- AVRAM, A. A. An Overview of Reduplication and Compounding in Tetun Dili. *Revue Roumaine de Linguistique*, v.53, n.4, p.427-448, 2008.
- BERTHE, L. Morpho-syntaxe du Buna' (Timor central). *L'Homme*, v. 3, n. 1, p. 106-116, 1963.
- CORTE-REAL, B. A.; HULL, G. First texts in Mambai-Ainaro. *Studies in languages and cultures of East Timor*, v. 1, p. 69-87, 1998.
- ENGELHOFEN, A. On derivational processes in Fataluku, a non-Austronesian language in East Timor. In WETZELS, L. (ed.). *The Linguistics of Endangered Languages*. Contributions to Morphology and Morphosyntax, p. 333-362. Utrecht: LOT 2009.
- ESPERANÇA, J. P. T. *Estudos de Lingüística Timorense*. Aveiro: SUL, 2001.
- FOGAÇA, H. A. O. *Estudo fonético e fonológico do Mambae de Same*. Uma língua de Timor-Leste. Dissertação (Mestrado em Linguística). Programa de Pós-Graduação em Linguística, Universidade de Brasília, Brasília, 2013.

- GOMES, F. A. *Os Fataluku*. Tese de Doutorado. Lisboa: Universidade Técnica de Lisboa, 1972.
- HAJEK, J. Language Contact and Convergence in East Timor: The Case of Tetun Dili". In: AIKHENVALD, A.; DIXON, R. (eds.). *Grammars in Contact: A Cross-Linguistic Typology*. Nova York: Oxford University Press, 2007. p. 163-178.
- HOLM, J. *Languages in Contact. The Partial Restructuring of Vernaculars*. Nova York: Cambridge University Press, 2004.
- HULL, G. A Morphological Overview of the Timoric Sprachbund. *Studies in Language and Culture of East Timor*, v.4, p. 98-205, 2001.
- \_\_\_\_\_. *Southern Mambai (Manbae-Ainaru Nor Same)*. Dili: Instituto Nacional de Linguística/Universidade Nacional Timor Lorosa'e, 2003.
- \_\_\_\_\_. The Malay Lexical Element in Tetum. *Studies in Languages and Cultures of East Timor*, v.7, p.84-124, 2005.
- HULL, G.; ECCLES, L. *Tetum Reference Grammar*, Sydney: Sebastião Aparício da Silva Project, Dili: Instituto Nacional de Linguística/Universidade Nacional de Timor Lorosa'e, 2001.
- KLAMER, M.; REESINK, G.; STADEN, M. East Nusantara as a Linguistic Area. In: MUYSKEN, P. (ed.). *From linguistic areas to areal linguistics*. Amsterdam: Benjamins, 2007.
- KLINKEN, C. *A Grammar of the Fehan Dialect of Tetun, an Austronesian Language of West Timor*. Canberra: Pacific Linguistics, 1999.
- LEWIS, P. (ed.). *Ethnologue: Languages of the World*. 16a ed. Dallas: SIL International, 2009.
- LUCCHESI, D. Variação e norma: elementos para a caracterização sociolinguística do português do Brasil. *Revista Internacional de Língua Portuguesa*, v. 12, p. 17-28, 1994.
- LUCCHESI, D.; BAXTER, A. A transmissão linguística irregular. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (orgs.). *O português afro-brasileiro*. Salvador: EDUFBA, 2009. p. 101-124.
- HAGÈGE, C. *Morte e Rinascita delle Lingue*. Milão: Feltrinelli, 2002.
- MCWHORTER, J. *Language Interrupted: Signs of Non-Native Acquisition in Standard Language Grammars*. Nova York: Oxford University Press, 2007.
- \_\_\_\_\_. Why does a language undress? Strange cases in Indonesia. In: MIESTAMO, M.; SINNEMÄKI, K.; KARLSSON, F. (eds.). *Language Complexity: typology, contact, change*. Amsterdam: John Benjamins, 2008. p. 167-190.

THOMAZ, L. F. *Babel Loro Sa'e: O Problema Lingüístico de Timor Leste*. Lisboa: Instituto Camões, 2002.

SIEGEL, J. *The Emergence of Pidgin and Creoles Languages*. Nova York: Oxford University Press, 2008.

VERSTEEGH, K. Non-Indo-European Pidgins and Creoles. In KOUWENBERG, S.; SINGLER, J. V. (eds.) *The Handbook of Pidgin and Creoles Studies*. p. 158-186. East Sussex: Wiley-Blackwell, 2008.

WILLIAMS-VAN KLINKEN, C.; HAJEK, J; NORDLINGER, R. *Tetun Dili: A grammar of an East Timorese language*. Canberra: Pacific Linguistics, 2002.