

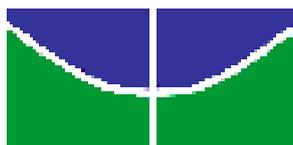
UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE LETRAS
DEPARTAMENTO DE LINGUÍSTICA, PORTUGUÊS E LÍNGUAS CLÁSSICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

**ESBOÇO GRAMATICAL DO TETUN PRASA: LÍNGUA OFICIAL DE TIMOR-
LESTE**

DAVI BORGES DE ALBUQUERQUE

Brasília – DF

2011



UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE LETRAS
DEPARTAMENTO DE LINGUÍSTICA, PORTUGUÊS E LÍNGUAS CLÁSSICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

**ESBOÇO GRAMATICAL DO TETUN PRASA: LÍNGUA OFICIAL DE TIMOR-
LESTE**

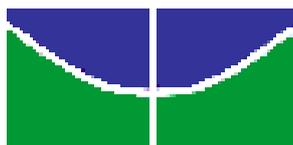
DAVI BORGES DE ALBUQUERQUE

Dissertação apresentada ao Departamento de Linguística, Português e Línguas Clássicas da Universidade de Brasília, como parte dos requisitos para a obtenção do grau de Mestre em Linguística.

Orientador: Prof. Dr. Hildo Honório do Couto

Brasília – DF

2011



UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE LETRAS
DEPARTAMENTO DE LINGUÍSTICA, PORTUGUÊS E LÍNGUAS CLÁSSICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

**ESBOÇO GRAMATICAL DO TETUN PRASA: LÍNGUA OFICIAL DE TIMOR-
LESTE**

DAVI BORGES DE ALBUQUERQUE

Banca Examinadora:

Prof. Dr. Hildo Honório do Couto (LIP – UnB) - Presidente
Prof. Dr. José Olímpio Magalhães (UFMG) - Membro externo
Prof. Dr. Antônio Augusto Souza Melo (LIP – UnB) - Membro
Profa. Dra. Enilde Leite de Jesus Faulstich (LIP – UnB) - Suplente

Brasília – DF

201

Dedico esta dissertação
À minha mãe, Elizabete,
E à minha noiva, Aurelie.

AGRADECIMENTOS

Agradeço a meu amigo e orientador, professor Hildo Honório do Couto, por ter me apoiado e guiado nas mais diversas questões tanto na vida acadêmica, quanto fora dela;

Carinhosamente, agradeço a minha família, que sempre esteve ao meu lado, não importa qual fosse a situação, especialmente minha mãe, Maria Elizabete Borges de Albuquerque, e minha noiva, Aurelie Marie Franco Nascimento;

Agradeço à Catharina Williams-van Klinken, que esteve constantemente disponível para esclarecer dúvidas e travar conversas emocionantes sobre a língua Tetun;

Agradeço também a todos os cidadãos leste-timorenses com quem tive contato, que, de alguma maneira, me ensinaram coisas valiosas para a vida, sejam essas coisas importantes para a vida acadêmica, a vida profissional ou a vida pessoal;

Um agradecimento especial ao Sr. Domingos dos Santos, sua esposa Judite Ximenes, e seus filhos, já que durante a elaboração desta dissertação se disponibilizaram constantemente a tirar dúvidas e fornecer dados sobre o Tetun;

Aos meus alunos do curso de *Português Instrumental 1* na Universidade Nacional de Timor Lorosa'e (UNTL), que me ensinaram muita coisa sobre as línguas de Timor-Leste e a cultura leste-timorense;

Aos amigos: Jessé Fogaça Soares, grande amigo e companheiro nos estudos linguísticos leste-timorenses; João Cleto do Nascimento, que me ensinou sobre a cultura leste-timorense e muitas outras lições que ficaram como ensinamentos para a vida; César Cavalcante, sempre que precisei ele esteve lá para ajudar; Nilsen Christiani Borges, grande conhecedor de história leste-timorense e amigo para a vida inteira; e Paulo Jefferson Pilar Araújo e Everaldo José Freire que me auxiliaram com conversas linguísticas instigantes e disponibilizando materiais fundamentais para minha pesquisa.

Eh... Foin sae Timor oan,
ó mak fini diak no esperansa,
ba ó nia rain.

Eh... Foin sae Timor oan,
hakat no hatudu ba
ba mundu tomak
nee ó nia diak.

(Abito Gama, cantor leste-timorense)

SUMÁRIO

ABREVIATURAS UTILIZADAS	09
LISTA DE MAPAS	11
LISTA DE TABELAS	12
LISTA DE FIGURAS	13
RESUMO	14
ABSTRACT.....	15
1 INTRODUÇÃO	16
2 BREVE HISTÓRICO DA ILHA DE TIMOR	19
2.1 Pré-história.....	20
2.2 Proto-história	24
2.3 História	28
2.4 Os dias atuais	33
3 AS LÍNGUAS DE TIMOR-LESTE	40
3.1 Filiação genética.....	40
3.2 Multilinguismo	44
3.3 A ecolinguística de Timor-Leste	50
4 A LÍNGUA TETUN	53
4.1 A documentação no período colonial e os primeiros estudos linguísticos	54
4.2 Os estudos tetumófonos na atualidade	55
4.3 A classificação da língua Tetun	56
4.4 O Tetun nas diferentes situações sociais	57
4.5 A variação da língua Tetun	59
5 O LÉXICO	67
5.1 O núcleo Austronésio.....	68
5.2 A influência do Malaio.....	70

5.3 A camada lusófona	72
6 FONOLOGIA	76
6.1 Fonologia segmental	76
6.2 A sílaba	89
6.3 O acento	91
6.4 A ortografia	96
7 MORFOLOGIA.....	99
7.1 A ausência de morfologia flexional e a gramaticalização	99
7.2 Morfologia derivacional	111
7.3 As classes de palavras	127
8 SINTAXE	149
8.1 Parataxe	149
8.2 O sintagma nominal	152
8.3 O sintagma verbal	158
8.4 Demais sintagmas	165
9 CONCLUSÕES	177
10 REFERÊNCIAS BIBLIOGRÁFICAS	179
APÊNDICE	188

ABREVIATURAS UTILIZADAS

1sg '1ª pessoa do singular'
1pl.exc '1ª pessoa do plural exclusiva'
1pl.inc '1ª pessoa do plural inclusiva'
2sg '2ª pessoa do singular'
2pl '2ª pessoa do plural'
3sg '3ª pessoa do singular'
3pl '3ª pessoa do plural'
ABL 'preposição ablativa'
AGT 'sufixo agentivo'
ANA 'pronome anafórico'
CAU 'prefixo causativo'
CL 'classificador'
CL.DIV 'classificador para divindade'
CLF.HUM 'classificador para humano'
CNTF 'marcador diretivo centrífugo'
CNTP 'marcador diretivo centrípeto'
COP 'cópula'
COM 'preposição comitativa'
CONT 'aspecto contínuo'
DAT 'preposição dativa'
DEF 'definido'
DEM 'aspecto demarcativo'
DIM 'diminutivo'
DIST 'distancial'
EGR 'aspecto egressivo'
ENF 'ênfase'
EXCL 'exclamação'
EXI 'existencial'
FEM 'feminino'
FOC 'foco'
FUT 'futuro'

GEN 'genitivo'
IN 'inalienável'
IND 'indefinido'
ING 'aspecto ingressivo'
INT 'prefixo intransitivizador'
INTG 'interrogativo'
IRR 'modo irrealis'
ITE 'aspecto iterativo'
LOC 'preposição locativa'
NEG 'negação'
MSC 'masculino'
NOM 'nominalizador'
PERF 'aspecto perfectivo'
PF 'tempo-aspecto perfeito'
PF.R 'aspectualidade perfeito-recente'
PL 'plural'
PONT 'aspecto pontual'
POS 'possessivo'
PROG 'aspecto progressivo'
PROS 'aspectualidade prospectiva'
PRX 'proximativo'
REC 'recíproco'
RED 'reduplicação'
REL 'relativizador'
SEM 'aspecto semelfactivo'
SG 'singular'
TOP 'topicalizador'

LISTA DE MAPAS

Mapa 1. Timor-Leste e regiões adjacentes	21
Mapa 2. Distrito de Manatuto	22
Mapa 3. Timor-Leste e a distribuição das línguas nativas pelo seu território	44

LISTA DE TABELAS

Tabela 1. Os fonemas da variedade Tetun Terik	62
Tabela 2. Segmentos consonantais do Tetun Prasa	77
Tabela 3. Segmentos vocálicos do Tetun Prasa	83
Tabela 4. Compostos gramaticais do TP	93
Tabela 5. O paradigma pronominal do TP	124
Tabela 6. Classificadores numerais	135
Tabela 7. Os quantificadores numerais	136

LISTA DE FIGURAS

Figura 1. A filiação do Proto-Timórico	41
Figura 2. O grupo Fabrônico de línguas Timóricas	41
Figura 3. O grupo Ramelaico das línguas Timóricas	42
Figura 4. As línguas papuásicas de Timor-Leste e suas filiações	43
Figura 5. A língua Tetun, suas variedades e subvariedades	61

RESUMO

A língua Tetun, em sua variedade Tetun Prasa, é a língua oficial da República Democrática de Timor-Leste, desde a constituição de 2002, juntamente com a língua portuguesa. Timor-Leste, em um pequeno território, possui cerca de 16 línguas pertencentes a filiações genéticas diferentes. Desta maneira, há evidências de que a variedade Tetun Prasa já funcionava como língua franca de Timor em um período anterior à chegada dos portugueses, por volta do ano 1515. A presente dissertação objetiva elaborar um esboço de gramática do Tetun Prasa. Para tanto, discorre sobre a história da ilha; analisa os contatos linguísticos, a situação atual de multilinguismo e as contribuições anteriores de estudiosos a respeito da língua Tetun; aponta aspectos notáveis do léxico tetumófono; procura fazer uma proposta de análise fonológica da língua para servir de base à descrição morfossintática e, finalmente, descreve a morfologia e a sintaxe desta variedade do Tetun.

Palavras chave: Timor-Leste; Tetun Prasa; léxico; fonologia; morfossintaxe; descrição linguística.

ABSTRACT

The variety of Tetun language, called Tetun Prasa, is the official language of Democratic Republic of East Timor together with Portuguese language since the constitution of 2002. East Timor has in a narrow territory about 16 languages with different genetic affiliations. Thus, there are evidences that Tetun Prasa had had a status of lingua-franca before the Portuguese arrived in the island, approximately 1515. The present dissertation has the objective of elaborate a Tetun Prasa grammatical sketch. Thus, it will be discussed Timor Island history, followed by analysis of previous studies on Tetun, language contact, current multilingual situation, and conspicuous aspects of Tetun lexicon. Finally, a phonological analysis will be elaborated and followed by Tetun Prasa morphological and syntax studies.

Keywords: East Timor; Tetun Prasa; lexicon; phonology; morphosyntax; language description.

1. INTRODUÇÃO

Timor-Leste conquistou sua independência recentemente, no ano de 1999, livrando-se de uma dominação estrangeira, de origem indonésia, e em 2002 obteve o reconhecimento internacional como nação, elegendo em sua constituição a língua portuguesa e o Tetun Prasa (objeto de estudo da presente dissertação) como línguas oficiais¹. É considerado por várias entidades como o país mais novo do planeta. Por esses motivos, destaca-se no cenário internacional, despertando interesse de diversos países com a intenção de atuar de alguma maneira na reconstrução dessa nação cujo território consiste na parte leste de uma pequena ilha localizada no sudeste asiático.

As atuações internacionais que merecem destaque são as do governo português e do governo brasileiro, que enviam profissionais nos diversos setores: educacional, administrativo, judiciário e militar, além da ONU e de um grande número de ONGs. O fator que aproximou os países lusófonos, principalmente Portugal e Brasil, de Timor-Leste foi a escolha da língua portuguesa como língua oficial, juntamente com o Tetun Prasa, na constituição da República Democrática de Timor-Leste, do ano de 2002.

Em um território relativamente pequeno, cerca de 14.600 km², Timor-Leste possui um número de aproximadamente 16 línguas nativas de filiações genéticas distintas (HULL, 2001b). As pesquisas linguísticas que tem Timor-Leste como objeto de estudo são poucas, mas vêm alcançando avanços significativos. Dessa maneira, pode-se afirmar com certo grau de certeza que as 16 línguas nativas de Timor-Leste pertencem a duas filiações genéticas diferentes: 12 línguas de origem austronésia e 4 línguas de origem papuásica².

As línguas austronésias são: Tetun, Manbae, Tokodede, Kemak, Bekais, Galolen, Habun, Kawaimina, Makuva, Idalaka, Wetarês e Lolein. As quatro línguas papuásicas são: Fataluku, Makasae, Makalero e Bunak. Timor-Leste, então, apresenta um quadro de multilinguismo, onde algumas dessas línguas são L1 de um distrito/povo

¹ Para o presente estudo utilizo a forma *Tetun Prasa*, já que esta é usada na ortografia da língua, assim como pelos falantes nativos da língua, as ortografias alternativas, muito utilizadas em outros trabalhos são: *Tétum-Praça* ou *Tetum Praça*.

² Adotei nesta dissertação o número de Hull (2001b) pelo fato deste ser mais acurado linguisticamente, pois o autor conduziu uma pesquisa, durante um longo período, tendo como objeto as diversas línguas nativas de Timor-Leste aqui citadas. Outras fontes fornecem números diferentes, entre elas Fox (2000) e a mais nova edição do *Ethnologue* (Lewis, 2009) apresentam um número um pouco superior ao aqui adotado, que fica entre 18 ou 19 línguas. Ainda, os autores mencionados também discordam em relação à presença ou ausência de algumas línguas em território leste-timorense.

específico e possuem um alto número de falantes, como é o caso do Manbae, Makasae, Fataluku e Tetun, enquanto outras línguas não são mais adquiridas pela geração mais nova, possuem um número reduzido de falantes e/ou são utilizadas somente em situações sociais específicas, como é o caso do Habun, Bekais, Makalero e Makuva.

Diante dessa complexa situação de multilinguismo, o Tetun Prasa (uma variedade da língua Tetun) surgiu como língua franca entre os diferentes grupos etnolinguísticos que habitavam a parte leste da ilha de Timor em um período anterior à chegada dos portugueses, provavelmente no século XV (ALBUQUERQUE, 2009). Além disso, a variedade Tetun Prasa possuiu papel importante no decorrer dos conturbados momentos da história do povo leste-timorense e, desta maneira, está associada à própria noção de identidade deste povo (ALVES, 2005).

Resumindo, Timor-Leste é uma nação lusófona que apresenta uma situação linguística pouco estudada, pode contribuir de maneira significativa à linguística, possui um histórico de séculos de contato com o colonizador português e, nos últimos anos, vem fechando diversos acordos governamentais com o Brasil, que envia mão de obra especializada todos os anos.

Assim, esta dissertação tem o objetivo de realizar o primeiro esboço gramatical da variedade Tetun Prasa em língua portuguesa, assim como no Brasil, e contribuir aos estudos linguísticos para o registro e a análise da língua mais usada no território leste-timorense. A escolha deste tema justifica-se, ainda, pelo fato de o presente autor ter atuado como professor de língua portuguesa no país, nos anos de 2008 e 2009, e ter tido contato com a língua Tetun e seus falantes; também, pelo hiato que há na bibliografia linguística a respeito das línguas nativas de Timor-Leste, somente poucas línguas possuem algum tipo de estudo linguístico, outras permanecem sem nenhum tipo de registro nem análise até a atualidade, este é o primeiro material de descrição da língua Tetun, ou de qualquer outra língua de Timor, em língua portuguesa.

Para tanto, inicialmente farei uma apresentação da história da ilha de Timor no capítulo 2. A seguir, no capítulo 3, discutirei a situação atual das línguas de Timor-Leste, analisando fatores como: filiação genética, multilinguismo e ecolinguística. No capítulo 4, encontra-se uma introdução à língua Tetun. Após esses capítulos iniciais, entrarei na análise linguística do Tetun Prasa propriamente dita, com o capítulo 5 possuindo uma breve descrição do léxico tetumófono e o capítulo seguinte, capítulo 6, uma análise da fonológica. No capítulo 7 serão analisadas as categorias gramaticais e

outros aspectos da morfologia do Tetun Prasa para, em seguida, no capítulo 8, descrever a sintaxe, diferenciando estruturas linguísticas nativas das emprestadas pelo contato linguístico.

A metodologia baseia-se principalmente nos dados coletados em campo pelo autor, e na pesquisa bibliográfica dos estudos linguísticos prévios. Os dados do Tetun Prasa foram coletados em momentos diferentes. No primeiro momento, os dados foram coletados no período que fixei residência em Timor-Leste, 2008 e 2009. Esses primeiros dados, porém, foram coletados visando um estudo comparativo das línguas de Timor-Leste. Em um segundo momento, coletei os dados para a averiguação das análises presentes nesta dissertação. Digno de nota é que os falantes de Tetun que se disponibilizaram para serem nossos informantes eram timorenses residentes no Brasil, principalmente em Brasília, São Paulo e Goiânia. Ainda, as principais referências bibliográficas consultadas de estudos linguísticos prévios sobre a língua Tetun foram: a gramática do Tetun, dialeto Fehan, de Klinken (1999); a gramática do Tetun de Hull e Eccles (2001) e, mais recente, um esboço gramatical do Tetun Dili, elaborado por Williams-van Klinken, Hajek e Nordlinger (2002).

Ainda, a presente gramática apresenta diversas contribuições em relação às anteriores, já que se aprofunda principalmente no contato de línguas, no léxico e na morfologia, inexistentes nas gramáticas dos outros autores, assim como há uma análise diferencial para a fonologia do Tetun Prasa, o multilinguismo em Timor e a variação da língua Tetun. A sintaxe, o nível de análise mais explorado nas publicações prévias, foi descrita aqui somente em suas estruturas principais, parataxe e os sintagmas, deixando certas estruturas problemáticas, que também não foram abordadas em estudos anteriores, para uma análise futura.

Finalmente, pretende-se alcançar os objetivos propostos aqui através da pesquisa bibliográfica, da análise dos dados coletados em campo, em diferentes momentos no decorrer de vários anos, e, assim, contribuir para a teoria linguística, com a descrição do Tetun Prasa, e para o povo timorense, oferecendo-lhes esta dissertação como mais um auxílio e uma ferramenta para eles conquistarem uma independência genuína. Pois com a posse de instrumentos que levam ao conhecimento, entre eles o conhecimento linguístico, que é fundamental, os novos cidadãos leste-timorenses poderão exercer sua cidadania efetivamente.

2. BREVE HISTÓRICO DA ILHA DE TIMOR

A ilha de Timor foi habitada por diversos povos desde tempos pré-históricos. Ela foi alvo de migrações de origem papuásica em tempos pré-históricos; há evidências de povoamento anterior a este período papuásico de povos conhecidos como ‘pré-austronésios’ ou ‘pré-papuásicos’; em um período mais recente, as migrações austronésias ocorreram em direção à ilha; já no período histórico houve diversos contatos com os grandes impérios e civilizações da região do sudeste asiático, a saber: a China, Índia e os reinos indianizados, a civilização árabe e os sultanatos malaios; a colonização portuguesa e as sucessivas disputas territoriais e comerciais entre a coroa portuguesa e a Holanda (representada pela Companhia das Índias Orientais); na década de 40, do século XX, Timor-Leste, na época conhecido como Timor Português, foi dominado pelo Japão no período da 2ª Guerra Mundial; na década de 70 do século XX, sofreu uma dominação da Indonésia. Finalmente, na atualidade com a presença da ONU e de muitas ONGs que oferecem assistência à República Democrática de Timor-Leste, este país encontra-se com um intenso contato com a Austrália e a língua inglesa, pois ela funciona como uma língua de comunicação entre os diversos estrangeiros que residem lá.

Segundo Thomaz (1994), para um melhor estudo do passado da ilha de Timor, pode-se dividir os processos históricos em três fases distintas: pré-história, proto-história e história. A pré-história é caracterizada principalmente pela ausência de registros escritos, ou qualquer documento dessa natureza; a proto-história consiste no período anterior à chegada dos europeus no continente asiático; há nesse período, porém, registros escritos e outros tipos de documentação que fazem algum tipo de referência (direta ou indireta) à ilha de Timor e às outras ilhas da região, mas de maneira não sistemática; o período histórico é caracterizado pela chegada dos europeus à Ásia e o início da produção de uma documentação sistemática com descrições das ilhas da região, diversos tipos de registros: de viagens, de atividades colonizadoras, comerciais, missionárias, correspondências etc.

Este capítulo adota a divisão de Thomaz (1994) em período pré-histórico (2.1), proto-histórico (2.2) e histórico (2.3) para discorrer sobre o processo histórico ocorrido na ilha de Timor, com suas várias ondas migratórias, contato de povos, e,

consequentemente, contato de línguas e a colonização europeia. Ainda, acrescentamos os dias atuais (2.4) para falar da atual situação de Timor-Leste.

2.1 A pré-história

Sobre este período a metodologia utilizada nos estudos históricos não é eficaz; desta maneira, as ciências que podem revelar alguma coisa sobre esse passado distante são a arqueologia e a linguística, principalmente com o método histórico-comparativo, conforme apresentei em Albuquerque (2009).

As pesquisas arqueológicas realizadas em Timor-Leste revelaram que esta ilha já era ocupada em um período entre 35.000 A.P.³ e 30.000 A.P. (O'CONNOR, SPRIGGS e VETH, 2002). Outras evidências baseadas na análise da tecnologia náutica, na pesca e na presença de animais conseguiram trazer datações mais específicas sobre as migrações que ocorreram no passado. Por exemplo, a datação de alguns artefatos encontrados utilizados para pesca, como anzóis, redes, entre outros, foi de 10.000 A.P. (O'CONNOR e VETH, 2005), e remete ao uso dessa tecnologia de pesca anterior ao povo austronésio (que desenvolveu essa tecnologia em um período posterior). Foram encontrados também vestígios da presença de um marsupial originário da Papua, que data 9.000 A.P. (O'CONNOR, 2006), esta é outra evidência da presença da ocupação humana em Timor-Leste de povos de origem não austronésios, assim como essas informações são evidências para a datação da migração austronésia para a ilha.

As informações que a linguística traz são diferentes das apresentadas pela arqueologia. As evidências linguísticas mostram que o povo, que deu origem às línguas nativas austronésias, chegou à ilha por volta do século X e era de origem butônica, sudeste das Ilhas Celebes.

³ A.P. sigla de 'antes do presente' e muito utilizada, geralmente, em trabalhos arqueológicos, por isso utilizo esta sigla quando me referir à arqueologia.



MAPA 2. Distrito de Manatuto (Fonte:

<http://webcarta.net/carta/mapa.php?id=10433&lg=pt>)

As línguas de Timor-Leste de origem austronésia são descendentes de uma única língua, o Proto-Timórico, e os processos que geraram as diversas línguas modernas e suas respectivas reestruturações gramaticais⁵ aconteceram em um período histórico mais recente, por volta do século XIII e, depois, no século XV (HULL, 2001b, p.100). Essas migrações e o processo de reestruturação gramatical sofrido pelas línguas de Timor-Leste nesse período serão explicados em 2.2.5.

A linguística acaba por se diferenciar da arqueologia em suas hipóteses e conclusões quanto ao povoamento pré-histórico de Timor-Leste. As evidências são obtidas com o método histórico-comparativo aplicado às línguas nativas, principalmente às línguas papuásicas (Fataluku, Makalero, Makasae e Bunak) e ao subgrupo Ramelaico (Manbae, Tokodede, Kemak e Idalaka) das línguas Timóricas.

⁵ Utilizo aqui o termo 'reestruturação gramatical' (Holm, 2004) e na pesquisa recente que venho realizando sobre os contatos linguísticos em Timor-Leste. Já Hull (2001b) e outros linguistas que estudam essas mesmas línguas insistem em chamar esse fenômeno de 'crioulização'.

Em relação às línguas papuásicas, pode-se afirmar que elas possuem uma origem comum e uma filiação genética com as línguas faladas na península Bomberaica da Papua Ocidental, e o povo falante da língua bomberaica que as originou parece ter migrado para o Timor cerca de 2000 a.C. A análise dos cognatos dessas línguas papuásicas (HULL, 2004, p.28) indica que elas possuem um ancestral comum, ou seja, somente um povo falante de uma só língua papuásica que migrou cerca de quatro mil anos atrás para o Timor.

As línguas Ramelaicas (Manbae, Tokodede, Kemak e Idalaka), porém, apresentam outro substrato que indica a presença de línguas pré-austronésias, e, conseqüentemente, a presença de um povo pré-austronésio anterior ao povo neo-bomberaico que inseriu a língua papuásica que se fragmentou e gerou as quatro línguas conhecidas na atualidade. Esse grupo de línguas merece destaque por ter uma natureza híbrida. Essas línguas são também chamadas de semi-austronésias por possuir apenas superficialmente alguns elementos austronésios, quando na realidade elas possuem três substratos distintos, o que leva a crer que na região do Monte Ramelau, antes de ser introduzido o Idalaka, ali eram faladas três línguas pré-austronésias distintas (HULL, 2001a, p.4).

Dessa forma, através da análise dos cognatos das diferentes línguas identificou-se a presença de mais de um povo pré-austronésio que habitou primeiramente a região e efetuou-se um primeiro contato com o povo papuásico recém-chegado, por volta de 4.000 A. P.; posteriormente com a migração austronésia houve diversos contatos destes com os povos papuásicos, já dispersos pelo território, e com os povos pré-austronésios de que se conhece quase nada.

Abaixo se encontram alguns exemplos de substratos distintos nas línguas Manbae, Tokodede e Kemak, acompanhados das reconstruções para o Proto-Austronésio e/ou Proto-Malaio-Polinésio e dos itens lexicais em Tetun Prasa:

1. Substrato pré-austronésio:

PAN, PMP	Tetun	Manbae	Tokodede	Kemak
*kedi,	kiik	loba	X	loba
*dikiq ‘pequeno’				
*ma-diŋdiŋ ‘frio’	malirin	bisa	tupudu	suma

*ma-raŋaw ‘seco’	maran	glia	seu	X
*ma-baseq ‘molhado’	bokon	X	tita	glia
*nabuq ‘cair’	monu	mou	blasi	manahu

2.2 Proto-história

A proto-história é o período histórico de Timor-Leste anterior à chegada dos colonizadores europeus (holandeses e portugueses). Neste período, foram elaborados os primeiros registros sobre a ilha de Timor. Ainda, os fatores históricos de maior destaque foram: o estabelecimento das rotas comerciais e as sucessivas dominações e influências das sociedades indianas e malaias, do reino islâmico e do império chinês. As pesquisas linguísticas sobre esse período são poucas, mas alcançam conclusões significativas sobre o contato entre esses povos e a presença deles em Timor, como será comentado mais adiante.

2.2.1 Os reinos indianizados

A sociedade com maior profundidade temporal que influenciou o sudeste asiático foi a sociedade indiana. Sua influência data dos primeiros séculos da nossa era. A partir desse período, algumas sociedades com base indiana começaram a se formar: em Camboja, por volta do século III; Samatra, a partir do século VII; Java, no século XIII. A influência indiana nessas sociedades foi marcada pela inserção de bens relativos à cultura material, toda uma herança tecnológica que essa sociedade possuía, e, principalmente, de bens culturais imateriais como a inserção das religiões budista e hindu, e de diversas influências na área das artes: na literatura, na arquitetura, na pintura, que podem ser vistas até a atualidade nessas sociedades.

Timor-Leste teve contato com esses reinos indianizados, pois há vários registros históricos que citam a ilha de Timor, assim como o interesse comercial desses reinos no sândalo branco dessa ilha. Tais contatos, porém, parecem ter sido irregulares, pois as influências culturais e linguísticas de origem indiana em Timor são poucas. Culturalmente, apenas a região de Suai parece ter sofrido uma influência maior com alguns traços indianos em motivos arquitetônicos, na literatura oral e nas danças (THOMAZ, 2002, p.78). Linguisticamente, a influência indiana é irrisória já que há somente em Tetun, e também no português falado em Timor-Leste, poucas palavras do sânscrito, ou de outras línguas indianas, como o Tamil e Malayalam. Estes empréstimos

de origem indiana na língua Tetun e no português falado em Timor, porém, foram introduzidos via Malaio, ou via algum dos portugueses crioulos do sudeste asiático, o Crioulo Português de Malaca (CPMal) ou Crioulo Português de Macau (CPMac). A seguir encontram-se alguns exemplos desses empréstimos (ESPERANÇA, 2001, p.41; THOMAZ, 1995, p.165):

2. *jagra* ‘um tipo de açúcar mascavo’ do malayalam *chákkara*;
3. *jaka* (português *jaca*) ‘fruto comestível’ também de origem do malayalam *chákkka*;
4. *mainato* ‘lavadeiro, criado’ também do malayalam *mainattu*;
pardau ‘padrão de valor de búfalos’ do sânscrito *pratâpa* ‘calor ardente, brilho’, por extensão semântica passou a denominar uma moeda da Índia e, posteriormente, também por extensão semântica, termo usado para designar o valor de um búfalo;
5. *sarón* ‘espécie de saia’ do sânscrito *saranga* ‘vestido’, via malaio *sarong*;
topaz ‘mestiço, nativo assimilado a cultura portuguesa’ do dravídico *tuppâsi*.

2.2.2 Os sultanatos malaios

A maior influência que aparece nas línguas e na cultura timorense é originária dos povos falantes da língua malaio. Como foi apresentado anteriormente, durante um período de aproximadamente quinhentos anos, por volta do século VIII até o século XIII, o reino indianizado de Java dominou o comércio nos mares do sudeste asiático. No final do século XIII, porém, o reino Jau de Majahapit, localizado em Java Oriental, inicia uma investida ao reino de Sriwijaya e assume o controle das rotas comerciais (THOMAZ, 1994, p.547). No entanto, um pouco depois já entra em decadência, no século XV, e, então, Malaca começa paulatinamente a assumir o controle dessa área do sudeste asiático para posteriormente, em 1511, ser dominada por Afonso de Albuquerque (LOUREIRO, 1995, p.30).

Essa hegemonia de Malaca é que merece destaque para nós, pois foi a partir de Malaca que se difundiu a religião hindu, a língua malaia e outros traços culturais nesta parte do mundo. Contudo, para Timor-Leste a influência malaia parece ter vinda também da parte indonésia da ilha, ou seja, Timor-Leste durante esse período histórico

recebeu influências da cultura malaia graças às relações comerciais com Malaca e também pela sua proximidade com a sociedade de cultura malaia que habitava a parte oeste da ilha (LOBATO, 2000a, p.356).

As influências históricas do Malaio⁶ no Tetun foram significativas na configuração atual da língua, afetando todos os níveis de análise linguística: fonologia, morfossintaxe e léxico. Na fonologia, pode-se apontar a inserção dos fonemas /p/ e /g/, que não existiam anteriormente e vieram juntamente com empréstimos lexicais como *dapur* ‘cozinha’, *surat* ‘papel, carta’, *katuas* ‘velho’ entre outros, e a posse passou a ser expressa com um marcador possessivo enclítico =*nia* que veio do verbo malaio *punya* ‘possuir’.

2.2.3 Possível influência islâmica?

A importância dos árabes no processo histórico do sudeste asiático é fundamental, mas quando falamos de uma influência árabe em Timor, e em um processo histórico timorense, pouco, ou quase nada, sobre os árabes há para ser dito, somente algumas informações pontuais. Há evidências da circulação dos árabes pelos mares do sudeste asiático já no século VIII; posteriormente, começa a aparição de pequenos reinos, e, finalmente, a adoção do islamismo pelo grande sultão de Malaca (THOMAZ, 2002, p.83). Em Timor, há registros de que no século XVII muçulmanos do reino de Macassar povoaram uma região de Manatuto, alguns timorenses converteram-se ao islamismo, e lutaram contra os portugueses ao lado dos árabes. Há pouca influência em Timor-Leste de origem árabe: são elas algumas de natureza cultural, como resquícios de narrativas árabes adaptadas ao Tetun Terik, danças e arquitetura, outras de natureza linguística, como os escassos empréstimos árabes em Tetun: *hotar* ‘amaldiçoar’ (< Árabe *xuṭbah* ‘rezar’, via Malaio *berkhotbah*), *arak* ‘bebida fermentada’ (< Árabe ‘*araq* ‘bebida forte’, via Malaio *arak*), *pateka* ‘melância’ (< Árabe *baṭekh*).

2.2.4 O império chinês

⁶ Há duas influências históricas do contato com a língua Malaio, a primeira data do século XV com o apogeu dos sultanatos malaios e o uso do *Pazar Melayu* como língua franca, a segunda é recente e data do período da invasão indonésia de 1975 a 1999, conforme será descrito nas seções seguintes. Ainda, a tarefa de identificar possíveis influências do contato linguístico entre Tetun e o Malaio é difícil pela afiliação genética próxima das duas línguas, pois as duas pertencem à família Malaio-Polinésia.

Finalmente, o último povo que nos interessa aqui nesse processo histórico do sudeste asiático que envolve Timor-Leste são os chineses. O primeiro fato digno de nota sobre o império chinês da época é que os primeiros registros escritos que fazem referência a Timor parecem que são de origem chinesa. Em Eccles (2004, p.178), o autor faz uma análise das documentações de origem chinesa que fazem referência ao Timor. As documentações são as seguintes:

- *Registro das várias Nações Estrangeiras*⁷ de Zhao Rugua, que data do século 13;
- *Breve Registro das Nações das Ilhas* de Wang Dayuan do ano de 1349;
- *Investigações dos Oceanos Orientais e Ocidentais* de Zhang Xie do ano de 1617.

A obra de Zhao Rugua é um dos primeiros documentos escritos a fazer referência ao Timor. Ele faz uma descrição das várias nações da rota comercial asiática, entre elas Timor, e as várias ilhas que fazem parte da Indonésia atualmente, com o intuito de registrar o que há de vantajoso nas relações comerciais com esses povos, como são os costumes destes e quais são os produtos que podem ser achados em cada território. O registro deixado por Wang Dayuan, que data do ano de 1349, pouco tem a dizer sobre o Timor. Ele apenas elenca algumas características relevantes de interesse ao império chinês, e ao se referir ao povo e aos costumes timorenses faz somente afirmações de natureza pejorativa e preconceituosas. O documento legado a nós por Zhang Xie em 1617, consiste em uma reunião do conhecimento dos navegantes chineses sobre o sudeste asiático, que, mesmo com o fechamento da China e a proibição de viagens pelos mares, continuavam a negociar com os vários povos nativos que habitavam a ilha de Timor.

Mesmo com o intenso contato com os chineses até a atualidade, pois há uma grande comunidade chinesa em Díli que é responsável pelos estabelecimentos comerciais na capital, não há influência de nenhum dialeto chinês, ou outra língua de origem chinesa, nas línguas nativas de Timor-Leste⁸. Isso ocorre pelo fato de a comunidade chinesa ser maleável e de fácil assimilação; desta forma, os imigrantes

⁷ As traduções dos títulos originais chineses são apenas tentativas de minha autoria, já que não há traduções desses documentos para a língua portuguesa.

⁸ Há alguns empréstimos de origem chinesa em Tetun Prasa, porém parecem terem sido inseridos na língua via Malaio e/ou CPMac.

chineses, em sua maioria, eram falantes de Crioulo Português de Macau. Na atualidade, a comunidade chinesa de Díli é falante de *bahasa indonesia*.

2.2.5 As migrações Ambônicas (austronésias)

Anteriormente, foram apresentadas as diversas e possíveis influências na formação do povo, da cultura e das línguas timorenses. A seguir, serão discutidas as ondas migratórias que culminaram no processo de reestruturação gramatical que as línguas nativas de Timor-Leste sofreram.

O principal deles é apontado por Hull (2001b, p.100) como duas ondas sucessivas de migrações em massa para o Timor-Leste em um intervalo de tempo relativamente pequeno. O autor afirma que a primeira onda de migração ocorreu provavelmente no século XIII com a introdução do *Ambonês Antigo*, um conjunto dialetal inserido na ilha, que teve intenso contato com a língua austronésia já falada em Timor, desde o século X, foi introduzida por meio das migrações das Celebes. A segunda onda de migração foi da língua Malaio que foi introduzida com os comerciantes, provavelmente no século XV, que navegavam as rotas comerciais asiáticas. Nesse período, o *Pazar Melayu* tornou-se a língua franca de grande parte do sudeste asiático, pois era a língua usada nas relações comerciais, e também foi uma língua regional de troca. Segundo Hull (2001b, p.101), essas duas ondas migratórias acabaram por fazer com que as línguas nativas do Timor-Leste sofressem um processo de criouliização intenso e em um curto tempo.

2.3 Período histórico

Os portugueses chegaram à ilha de Timor em 1515. Não se estabeleceram na ilha de maneira adequada pelos seguintes fatores: as atividades comerciais em Malaca eram mais rentáveis, e a ilha de Solor também tinha o sândalo branco da ilha de Timor e era mais eficaz para os portugueses extraí-lo de Solor. Entre os motivos pelos quais era melhor para os portugueses explorar Solor, destaca-se o fator que os portugueses já haviam montado estabelecimentos nessa ilha, principalmente a fortaleza de Ende, que posteriormente foi dominada pelos holandeses (LOBATO, 2000b, p.364).

A maioria dos reis timorenses aceitou pacificamente o batismo, a conversão ao catolicismo e o domínio português. Houve alguns, porém, que se rebelaram e outros que se aliaram aos holandeses, o que fez com que Portugal fizesse alianças com os reinos

fiéis à coroa portuguesa para acabar com esses reinos rebeldes (LOBATO, 2000b, p.93). Portugal passou a lutar mais pelo território timorense no século XVII por ter perdido a ilha de Solor e o sultanato de Malaca para os holandeses, que também tinham interesse em Timor. Os sucessivos embates e as constantes disputas entre portugueses e holandeses veio a cessar temporariamente somente com o Tratado de 1661 que fixou os limites portugueses e holandeses (OLIVEIRA, 2004, p.122). O apoio, porém, da Holanda aos reinos rebeldes, principalmente no lado oeste da ilha de Timor, e os diversos combates com Portugal continuaram durante os séculos seguintes.

2.3.1 Os crioulos asiáticos de base lexical portuguesa

Os portugueses, como se instalaram em diversos territórios pela Ásia - Goa, Macau, Malaca, e em várias ilhas da Insulíndia, além do Timor-Leste, tiveram intenso contato com os diferentes povos que habitavam essas regiões, assim como tiveram contato com as diversas culturas e as várias línguas faladas por esses povos. Desta forma, esse contato fez com que fossem gerados vários crioulos de base portuguesa, entre eles: o crioulo português de Malaca (doravante CPMal), o crioulo português de Macau (CPMac) e o crioulo português de Bidau (CPB), em Díli.

O CPMal foi usado como língua franca em um vasto território do sudeste asiático durante o século XVI e a influência da língua portuguesa nas diversas línguas nativas dessa região provavelmente foi via CPMal ou do Malaio-Português, já que o número de portugueses era pequeno, assim como o nível de fluência dos nativos em língua portuguesa também era reduzido. Alguns exemplos de itens emprestados para o Tetun Prasa via CPMal são *dotoor* ‘médico’ e *kareta* ‘carro’.

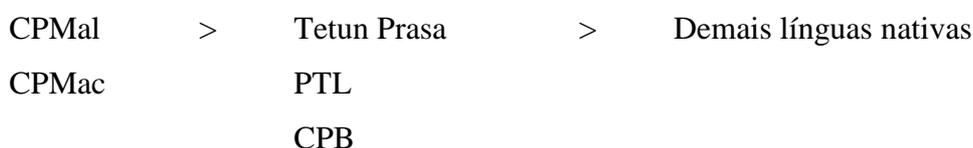
O CPMac parece ter influenciado de maneira decisiva a língua Tetun Prasa, já que a maioria dos funcionários da administração portuguesa em Lifau que foram transferidos para Díli eram provenientes de Macau e/ou falantes de CPMac. A tarefa de identificar as fontes dos diversos empréstimos linguísticos de origem portuguesa em Tetun Prasa é difícil, pois há certas evidências linguísticas e sócio-históricas de que o CPMal tenha contribuído na formação do CPMac (BATALHA, 1958) e, desta forma, os dois crioulos portugueses são muito próximos. O que se pode afirmar com maior grau de certeza é que os empréstimos de origem portuguesa não vieram diretamente da língua portuguesa, mas dos CPs. Alguns exemplos de itens emprestados para o Tetun Prasa via CPMac: *abusadoor* ‘aquele que é descortês’, *kanku* ‘vegetal amargo

comestível’ (< Mandarin *kang-kong*), kusi ‘pote para armazenamento’ (Mandarin *guànzǐ* ou *guan4zi*) e *s’á* ‘tea’ (< Mandarin *chá* ou *cha2*).

As condições sócio-históricas do surgimento do crioulo português de Bidau (CPB) são ligeiramente diferentes dos demais CPs asiáticos. O CPB formou-se em um período posterior à transferência da capital para Díli, em 1769, e em sua formação contribuíram o CPMac e o Tetun Prasa, ele apresenta, porém, diversos traços tipológicos dos CPs asiáticos, conforme foi apresentado em Baxter (1990). Ainda, Esperança (2001) apresenta uma hipótese de que houve um período em que provavelmente ocorreu uma influência mútua do CPB > Tetun Prasa e Tetun Prasa > CPB⁹.

2.3.2 O papel da língua portuguesa

A colonização portuguesa, que data do século XVI em diante, marcou de maneira significativa as línguas de Timor-Leste, tanto de origem austronésia, quanto de origem papuásica. Esta influência, porém, não ocorreu diretamente da língua portuguesa para as línguas nativas, pois os dados históricos apontam uma quantidade ínfima de europeus falantes de português. Desta maneira, segundo evidências históricas e linguísticas, a influência lusófona na realidade ocorreu via CPs, com os CPMal e CPMac como as duas línguas que mais influenciaram o Tetun Prasa e a formação de uma variedade do português falada em Timor-Leste, que chamei simplesmente de Português de Timor-Leste (PTL) em Albuquerque (2010b, 2011). Posteriormente, as línguas nativas serviram-se do Tetun Prasa para obter os empréstimos lexicais necessários. O esquema abaixo apresenta as influências do contato linguístico em Timor-Leste:



Digno de nota é que o esquema acima pretende apenas resumir a complexa situação linguística que ocorria em Timor-Leste durante os séculos XVIII e XIX. Vale a pena lembrar que em períodos distintos houve uma influência maior de uma ou outra

⁹ Uma discussão sobre o Tetun Prasa ser ou não um crioulo será feita no capítulo 4.

das línguas mencionadas sobre as demais; provavelmente, o século XVIII foi marcado por uma influência maior do CPMal, enquanto o século XIX foi marcado pela influência maior do CPMac.

Os empréstimos fonológicos são os seguintes: /v/, /ʃ/, /z/, /ɲ/, /ʎ/; os empréstimos lexicais são de diversas áreas, que cobrem campos semânticos de itens da cultura material e imaterial que não faziam parte dos povos nativos de Timor-Leste. Alguns exemplos:

- empréstimos lexicais de diversos campos semânticos: /livru/ ‘livro’, /borasa/ ‘borracha’, /burokrasia/ ‘burocracia’, /demokratiku/ ‘democrático’, /kruz/ ‘cruz’ etc.;
- o sufixo agentivo *-dór* que é bem produtivo na língua atualmente, tanto em empréstimos do português como *administradór* ‘administrador’, como em palavras nativas *husudór* ‘pessoa que pergunta constantemente’ ou ‘perguntador’ de *husu* ‘perguntar’, conforme os linguistas Hajek e Williams-van Klinken (2003) analisaram;
- diversas palavras gramaticais algumas estão em alternância de código, outras são compostas, como /kuandu/ ~ /bainhira/ ‘quando’, /ke/ ~ /nebee/ ‘que (relativo)’, /agora/ ~ /orasnee/ ‘agora’ (‘horas’+‘dêitico de proximidade’), /mais/ ~ /maibee/ ‘mas’, /i/ ~ /no/ ~ /ho/ ‘e’, entre outras.

Os itens que aparentam terem sido emprestados diretamente da língua portuguesa são as conjunções e preposições, pois, além de estarem em alternância de código, não há registros dessas formas nos dicionários antigos do Tetun (SILVA, 1889; DORES, 1907). Ainda, os itens lexicais relativos ao mundo modernizado, por motivos óbvios, também não possuem registros nos primeiros dicionários do Tetun, o que nos faz classificá-los como empréstimos recentes e via direta da língua portuguesa.

2.3.3 A mudança da capital

Os portugueses, quando chegaram ao Timor, por volta do ano de 1515, desembarcaram na parte oriental da ilha, onde atualmente é o enclave de Oecussi. A capital de Timor, então, era esta região, que na época era chamada de Lifau (OLIVEIRA, 2004, p. 100).

A administração portuguesa concentrava-se em Goa, porém Macau fornecia um grande número de funcionários para os outros territórios administrados pela coroa. A ilha de Timor não foi inicialmente alvo de grande interesse para a administração portuguesa, que preferiu as ilhas vizinhas, pertencentes aos arquipélagos de Maluco e Moluca (THOMAZ, 1994). Ainda, o sultanato de Malaca forneceu uma fonte de renda imensurável durante o tempo que foi dominado por Portugal (LOBATO, 2000b). Somente com a perda de diversas fortalezas e territórios para a Holanda foi que Portugal teve interesse maior em Timor.

Com as inúmeras perdas da coroa portuguesa para a Holanda, a capital da ilha de Timor, que se fixava em Lifau, teve que ser transferida para Díli, em 1769, após os portugueses serem expulsos do lado oeste da ilha. Outro fator digno de nota consiste nas questões sobre a língua, ou línguas, que eram usadas para se realizar a comunicação entre os diversos membros da administração portuguesa (portugueses e funcionários de Goa, Malaca e Macau) com os timorenses nativos de diferentes etnias, falantes de línguas diferentes. Algumas evidências históricas apontam para o uso do *Pazar Melayu* como uma língua franca no sudeste asiático, por volta do século XV, que posteriormente foi substituído pelo CPMal no século seguinte. Ainda, o CPMac foi muito falado em Díli durante o século XVIII pelo grande número de funcionários da administração portuguesa ser de origem macaense, além de que provavelmente se usava também o CPB.

Esta mudança de capital pode ter contribuído ainda mais para um processo de reestruturação gramatical do Tetun Prasa, já que essa região era falante de Manbae. Um estudo sistemático da influência da língua Manbae como um substrato do Tetun Prasa, porém, ainda precisa ser feito, algumas influências de ordem fonológicas são claras, mas em relação aos demais níveis de análise Tetun Prasa (morfologia, sintaxe, léxico) não é possível afirmar nada com muita certeza.

A influência maior da língua Manbae no nível fonológico foi a perda da oclusiva glotal e do glide bilabial, presentes nas demais variedades da língua Tetun, porém, não existem no inventário fonológico da língua Manbae e no Tetun Prasa:

6. [‘ha.ʔu] > [‘ha.u] ‘1sg’
7. [‘la.ʔɔs] > [‘la.ɔs] ‘não’
8. [‘we.e] > [‘be.e] ‘água’
9. [la.‘wa.rik] > [la.‘ba.rik] ‘criança’

No nível morfossintático pode-se citar a perda completa da flexão verbal, a tendência à parataxe (com a perda de preposições e conjunções), a queda dos genitivos e a perda de outras estruturas complexas, como: classificadores, cópula, sistema de negação e verbos existenciais. Digno de nota é que este processo de reestruturação gramatical do Tetun Prasa não ocorreu simplesmente pelo fato histórico da transferência da capital de Lifau para Díli, como acreditam alguns linguistas (HULL, 2001b).

10. TP. nia koalia ba hau ho hirus, (...)
 3sg falar com 1sg com raiva
 agora o halo nia mamar ona.
 agora 2sg fazer 3sg acalmar

MB. ua bogaes nor au (...) hirus, (...)
 3sg falar com 1sg raiva
 agora ó pun ua tnegá bel soi.
 agora 2sg fazer 3sg calma PERF

‘Ele falou comigo com raiva, porém agora ele foi acalmado por ti’.

11. TP.Ami sai loro-loron (...)dadeer, (...) meiodia ho kalan.
 MB.Aem sai namada bus, lelrae, (...) hoda.
 1pl.exc sair todo.dia manhã tarde noite

‘Nós saímos todo dia pela manhã, pela tarde e pela noite’.

Há evidências em outras subvariedades da língua Tetun, como a subvariedade falada no distrito de Suai (Tetun Belo), outra falada no distrito de Viqueque (Tetun Foho), e a falada na parte oeste da ilha na Indonésia, que este processo foi ocorrendo paulatinamente desde o período da segunda migração austronésia (Ambônica) por volta do século X, passando pela expansão e dominação do reino tetumófono de Wehale no século XV, e culminando no período de colonização portuguesa efetiva que ocorreu somente em meados do século XVIII.

2.4 Os dias atuais

Nos últimos anos, Timor-Leste sofreu outras influências linguísticas por motivos de natureza política. As duas grandes influências nos tempos modernos que podem ser citadas foram: a invasão indonésia que se estendeu de 1975 até a 1999, e a construção da República Democrática de Timor-Leste, que está se efetuando até o presente momento com diversos auxílios internacionais, sendo que o de maior destaque é o da ONU¹⁰.

2.4.1 A dominação indonésia

No ano de 1974, em Portugal, ocorre a revolução dos Cravos que faz com que o período de ditadura nesse país acabe e, se aproveitando disso, muitas colônias portuguesas se emancipam. Timor-Leste, desta maneira, nesse mesmo ano passa por um período conturbado com diferentes partidos políticos brigando pelo poder.

Os três principais partidos políticos de Timor-Leste foram: União Democrática Timorense (UDT), que defendia uma emancipação gradual e sob a defesa e apoio de Portugal para não romper os laços com a metrópole; Associação Social-Democrata Timorense (ASDT), composta em sua maioria por jovens, buscava o apoio popular e defendia um processo de independência em curto prazo, mas veio a radicalizar suas ações, transformando-se em Frente de Timor Leste Independente (FRETILIN); Associação Popular Democrática de Timor (APODETI) possuía um número reduzido de membros e uma ideologia conservadora, que defendia a integração a Indonésia, desde 1965 era governada por um regime militar e ditatorial liderado por Suharto.

No ano seguinte, em 1975, a UDT tentou realizar um golpe e tomou a capital, declarando de forma imediata a independência de Timor-Leste, o que facilitou a invasão indonésia, que já havia declarado a Portugal seus interesses de explorar a ilha de Timor neste mesmo ano. Indonésia, com um grande poderio militar, bombardeou a capital de Timor e enviou rapidamente um grande número de tropas, efetivando de maneira eficaz sua dominação.

O período de dominação indonésia, que se estendeu até 1999, foi marcado por uma intensa repressão, seguida de várias atrocidades, como massacres e torturas. Mas destacou-se também por ter reconstruído grande parte das estruturas destruídas na II

¹⁰ Digno de nota é que Timor também sofreu uma breve invasão japonesa (1942-1945), que deixou apenas alguns resquícios linguísticos, como um ou outro empréstimo usado somente pelos mais velhos, em línguas como o Tetun e o Manbae.

Guerra Mundial e por ter fornecido uma ampla infraestrutura aos distritos mais populosos de Timor-Leste, o que a administração portuguesa não o fez.

Na área linguística, a dominação indonésia adotou uma política linguística de *indonesização* da população timorense, assim como o planejamento linguístico indonésio mostrou-se ter extrema eficiência. Nos vinte e quatro anos em território timorense, a Indonésia conseguiu: diminuir o pouco que restava do uso da língua portuguesa pela população timorense banindo-a totalmente seu uso, sendo utilizada somente pelos revolucionários como *língua de resistência*; a língua Tetun como possui um caráter cultural e identitário para a nação timorense foi reduzido seu uso lentamente no dia-a-dia da população; o *bahasa indonesia* (língua oficial e nacional da Indonésia) foi implantado de maneira intensa por meio de um grande fluxo de professores e materiais didáticos, assim como da importação de diversos itens da cultura material indonésia como livros, revistas, músicas, programas de rádio e televisão, produtos de consumo em geral, entre outros.

Assim, no curto intervalo de tempo que a Indonésia permaneceu em Timor-Leste, esta deixou marcas linguísticas e culturais que podem ser vistas até a atualidade. Entre as marcas culturais podem ser mencionadas: a alimentação e produtos alimentícios em geral serem todos de origem indonésia; grande parte do entretenimento da população como televisão, rádio e até brincadeiras infantis; uma parcela da população timorense continua sendo fiel e simpatizante à dominação indonésia. Em relação às línguas: o *bahasa indonesia* é falado até a atualidade por uma grande parte da população, incluindo zonas rurais; os materiais didáticos e técnico-científicos são em grande parte em *bahasa indonesia*, o que faz com que a população letrada de Timor-Leste somente conheça a terminologia científica nesta língua; há uma grande influência do *bahasa indonesia* no Tetun Prasa falado pela população não escolarizada. Entre essas influências encontram-se:

- os numerais lusófonos e tetumófonos são em grande parte desconhecidos pela população não escolarizada; assim, são usados os numerais do *bahasa indonesia* no dia-a-dia, tanto números baixos, como *satu* ‘um’, *dua* ‘dois’, *tiga* ‘três’ *empat* ‘quatro’ e *lima* ‘cinco’, quanto números altos para expressar preços, horas e datas. Exemplos de preços: *dua dolar lima puluh sen* ‘dois dólares e cinquenta centavos’, *tiga dolar dua puluh sen* ‘tres dólares e vinte centavos’; horas: *jam lima* ‘5:00 h’, *jam dua tiga*

puluh ‘2:30 h’; datas: *seribu sembilan ratus tujuh puluh empat* ‘1974’, ou *tujuh puluh empat* ‘nos anos de 74’, e *tanggal tiga-puluh* ‘(a data de) hoje é dia 30’;

- certos itens culturais como *bapa* ‘senhor’ ao invés do lusismo ‘señór’, *warung* ‘restaurante indonésio’, comidas indonésias *gado-gado* ‘vegetais com molho de amendoim’, *bakso* ‘um tipo de sopa com carne’, *padang* ‘famoso método indonésio de conservar a comida pré-cozida’, o uso da palavra *kiri* ‘esquerda’ para fazer o transporte público parar, entre outros.

Há também uma influência do *bahasa indonesia* sobre a terminologia científica e em outros termos referentes ao mundo globalizado, já que o *bahasa indonesia* utilizou-se do inventário lexical da língua inglesa, é possível encontrar, assim, alguns vocábulos comuns ao português, inglês, Tetun e *bahasa indonesia*. Desta forma, há uma grande confusão sobre a pronúncia e a grafia adequada desses termos, e, a partir disto, pode-se verificar a produtividade deles e suas terminações na língua portuguesa e no *bahasa indonesia*. Os exemplos a seguir encontram-se na seguinte ordem Português, Tetun Prasa e *bahasa indonesia*:

- Pt. –aun, Tp. –aun e Bh. –i: definição – definisaun – definisi, comunicação – komunikasaun – komunikasi, constituição – konstituisaun – konstitusi, declaração – deklarasau – deklarasi, plantação – plantasaun – plantasi, população – populasau – populasi;
- Pt. –ade, Tp. –ade e Bh. –tas: comunidade – komunitas – komunitas, estabilidade – stabilitas, facilidade – fasilitas – fasilitas, faculdade – fakulda – fakultas, identidade – identitas – identitas, capacidade – kapasitas – kapasitas, nacionalidade – nasionalitas – nasionalitas;
- Pt. –mento, Tp. –mentu e Bh. –men: argumento – argumentu – argumen, departamento – departamentu – departemen, documento – dokumentu – dokumen;

- Pt. –nsia, Tp. –nsia e Bh. –nsi: conferência – konferénsia – konferensi, correspondência – korrespondénsia – korespondensi, província – provínsia propinsi¹¹.

2.4.2 A “invasão” anglófona

Após a conquista da independência no ano de 1999 e a posterior constituição da República Democrática de Timor-Leste, que data de 2002, a nação leste-timorense recebeu um grande apoio de governos de várias nações espalhadas pelo mundo, de diversas ONGs e principalmente da ONU, que instalou um governo provisório visando a criação, restabelecimento e a manutenção de diversos órgãos e instituições necessários para se estabelecer um governo democrático.

O impacto da atuação internacional pode ser visto de diversas formas; as que nos interessam para o presente trabalho são aquelas relacionadas a questões linguísticas. Na área linguística, as influências internacionais foram principalmente em questões relacionadas à política linguística, sociolinguística e contato de línguas.

A política linguística de Timor-Leste na atualidade baseia-se no que diz a constituição de 2002 a respeito das línguas que promulga o status de língua oficial ao Português e ao Tetun, e o status de língua de trabalho ao Inglês e *bahasa indonesia*. Aqui, porém, podem ser percebidas as atuações de instituições internacionais, já que há argumentos históricos, sociais, econômicos, políticos, linguísticos, entre outros, que justifiquem a presença do português, Tetun e malaio. Mas os argumentos a favor da escolha da língua inglesa são puramente políticos.

Ainda, o planejamento linguístico é quase nulo. Entendendo aqui política e planejamento linguístico de acordo com Calvet (2007), que diferencia ‘política linguística’ como o conjunto de decisões em relação a(s) língua(s) e ‘planejamento linguístico’ como a implantação, ou não, deste conjunto de decisões. Não há planejamento linguístico quase nenhum em Timor-Leste, já que as instituições sociais ceste-timorenses (escolas, universidades, hospitais, comércios) oferecem ao público seus respectivos produtos e serviços nas mais diversas línguas. As principais são inglês, chinês e malaio, sendo uma raridade encontrar tais serviços e produtos em português ou

¹¹ Exemplos retirados de Williams-van Klinken (2003).

Tetun. Ou seja, não há nenhum tipo de posicionamento em relação ao planejamento linguístico do que está previsto na constituição.

A única iniciativa de planejamento linguístico é em relação ao ensino de língua portuguesa no nível pré-secundário da rede de ensino público de Timor-Leste. Há uma preocupação de natureza política em se implantar a língua portuguesa nos primeiros anos de ensino educacional. Para tanto, houve alguns censos para verificar a situação da rede de ensino de Timor-Leste, entre eles o *Timor-Leste Census of Population and Housing* (NATIONAL BOARD OF STATISTICS, 2006) e o *Relatório do Desenvolvimento Humano de Timor Leste* (PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO, 2002). O primeiro, *Timor-Leste Census of Population and Housing* (NATIONAL BOARD OF STATISTICS, 2006), aponta uma taxa alta para o nível de falantes de português, cerca de 37%, e para os falantes de Tetun, cerca de 86%. Já o segundo, *Relatório do Desenvolvimento Humano de Timor Leste* (PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO, 2002), aponta uma taxa de 5% para os falantes de língua portuguesa e de 82% para falantes de Tetun. Ainda, o *Relatório do Desenvolvimento Humano de Timor Leste* (PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO, 2002) realizou também um censo nas escolas para verificar a situação do ensino de língua portuguesa e verificou níveis preocupantes em relação a média de professor/aluno, que era de 62 alunos para cada professor, e o domínio da língua portuguesa dos professores timorenses que era baixo, de acordo com o *Relatório do Desenvolvimento Humano de Timor Leste* (PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO, 2002, p.55), quando “3000 professores foram sujeitos a um teste realizado pela Missão Portuguesa (...) apenas 158 (5%) obtiveram aprovação, a maioria dos quais vivia em Dili ou em Baucau”, que são os distritos mais urbanizados do país. Com base nesses dados citados acima, o governo de Timor-Leste vem realizando uma política de incentivo e melhoria do ensino de língua portuguesa no nível pré-secundário. Esta foi a única medida de planejamento linguístico adotada pelo governo.

Vale acrescentar que nem mesmo a ortografia padrão, conforme foi autorizada pelo Decreto do Governo N° 1/2004 de 14 de Abril, vem sendo observada, já que os próprios documentos e instituições governamentais apresentam diversas grafias distintas

para os mesmos itens lexicais, e grafias estas que estão em desacordo com a ortografia padrão. Isso será discutido mais adiante no capítulo 6.

No campo da sociolinguística, principalmente na capital Díli, com o número cada vez maior de estrangeiros de origens diferentes atuando em território leste-timorense, a língua inglesa vem sendo utilizada como uma ‘língua franca internacional’, ou ‘língua de comunicação dos estrangeiros’. Além disso, um grande afluxo de infraestrutura acadêmica de origem australiana (livros e outras publicações, instrumentos, professores, pesquisadores) faz com que uma parte da população timorense associe a língua inglesa ao ‘sucesso’, a um status socioeconômico alto. Somado a isso, o *bahasa indonesia* retirou uma grande parte de seu inventário lexical moderno também da língua inglesa, diferenciando-se do Malaio falado na Malásia e em Singapura que mantém um número maior de itens nativos, ou de influência árabe. Dessa maneira, a população timorense simpatizante da Indonésia, assim como a juventude ‘seduzida’ pelo mundo anglófono, vem adquirindo um grande interesse em aprender a língua inglesa em detrimento de suas línguas nativas, da língua Tetun e do português.

A língua inglesa também desempenha um papel notável em relação ao contato com o Tetun e as demais línguas nativas leste-timorenses. Isso ocorre pelos fatos apresentados anteriormente: a língua inglesa vem despertando um interesse de uma parcela da população nos últimos anos e o léxico o *bahasa indonesia* possui um grande número de itens lexicais de origem inglesa. Esses dois fatores levam os aprendizes de inglês como língua estrangeira a trazer inúmeros empréstimos anglófonos de maneira indevida para a língua Tetun e para suas línguas maternas¹²:

12. Variação da fala refletida na ortografia:

armáriu ~ armari ‘armário’

fotokópia ~ fotokopi ‘xerox, fotocópia’

klásiku ~ klasik ‘clássico’

kómiku ~ komik ‘engraçado’

magnétiku ~ magnetik ‘magnético’

polísia ~ polisi ‘polícia’

¹² Os exemplos seguem a ortografia padronizada proposta pelo INL (2002).

3. AS LÍNGUAS DE TIMOR-LESTE

No capítulo anterior, foi possível observar as diversas influências linguísticas que a ilha de Timor sofreu no decorrer da história: a povoação papuásica, migrações austronésias em períodos distintos, contatos com os povos vizinhos dominantes, sendo o principal os sultanatos malaios, e a colonização europeia.

O processo histórico de Timor-Leste ajuda a entender sua situação linguística atual. Para tanto, analisaremos de maneira sucinta a filiação genética e as relações internas das línguas nativas de Timor-Leste (3.1) para podermos discutir as características do multilinguismo leste-timorense (3.2) e algumas premissas sobre a ecolinguística (3.3).

3.1 Filiação genética

As línguas de Timor-Leste pertencem a duas filiações genéticas distintas: Austronésia e Papuásica. Até a atualidade poucos são os estudos de natureza histórica que se debruçaram sobre a classificação das línguas leste-timorenses. Os únicos estudos que apresentaram uma proposta de classificação dessas línguas foram os estudos pioneiros de Capell (1943a, 1943b, 1944) e recentemente os estudos de Hull (2001b, 2004). Ainda, em Albuquerque (2010d), organizo uma lista das línguas faladas em Timor-Leste, juntamente com breves informações tipológicas e a situação sociolinguística atual de cada uma delas.

Na proposta recente de Hull (2001b) para as línguas de origem Austronésia ele lançou a hipótese de que elas descendem de um ancestral comum, chamado por ele de Proto-Timórico (fig.1). Ele incluiu neste grupo 12 línguas: Bekais, Tetun, Habun, Kawaimina, Makuva, Galolen, Wetarês, Manbae, Tokodede, Kemak, Idalaka e Lolein. Estas foram subdivididas em dois subgrupos: o Fabrônico (fig.2) e o Ramelaico (fig.3), e estes se ramificam ainda mais, de acordo com a localidade geográfica da língua:

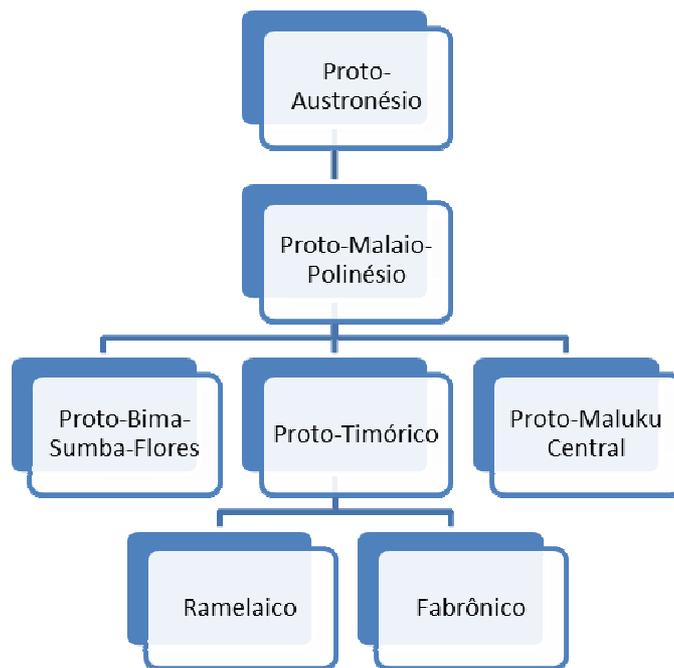


FIGURA 1. A filiação do Proto-Timórico

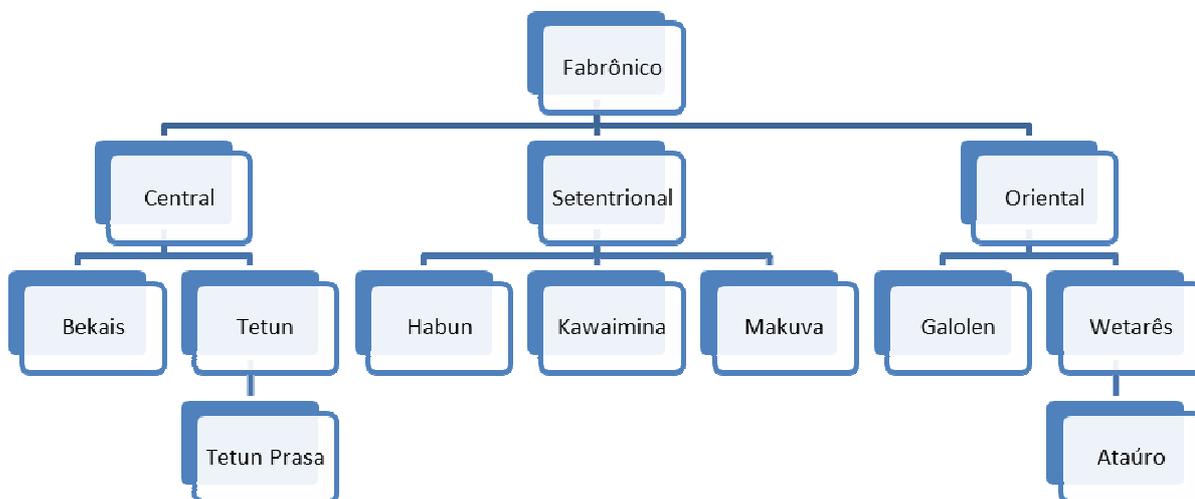


FIGURA 2. O grupo Fabrônico de línguas Timóricas

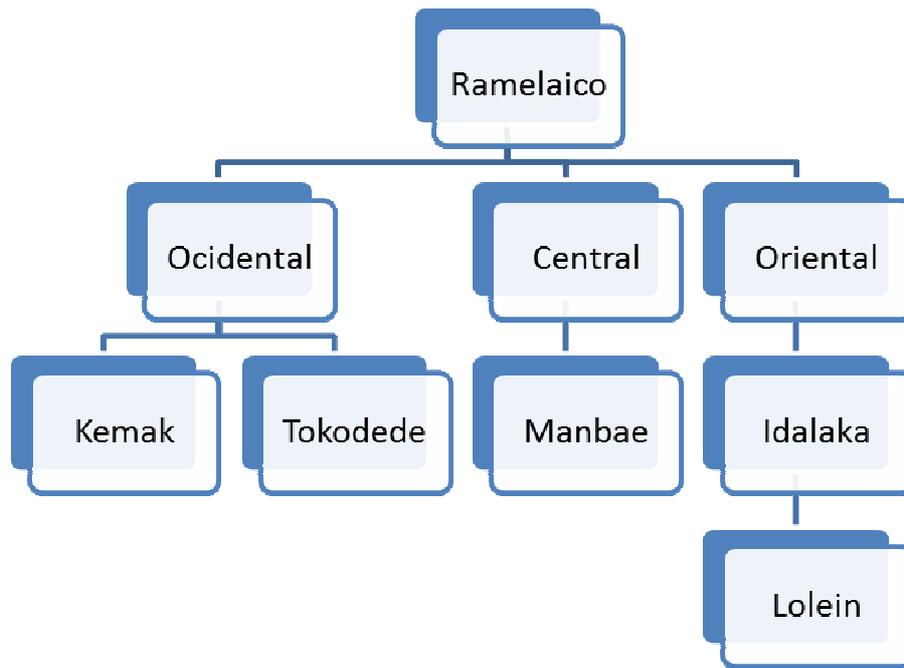


FIGURA 3. O grupo Ramelaico das línguas Timóricas

Em outro trabalho, Hull (2004) analisou as demais línguas e classificou-as como pertencentes ao agrupamento Trans-Nova-Guiné, já que essas línguas possuem um ancestral comum, que seria o proto-bomberaico, nome dado a uma suposta proto-língua que teve sua origem na península Bomberaica, localizada em Papua-Nova-Guiné. Dessa maneira, dentro do grande agrupamento Trans-Nova-Guiné, as línguas leste-timorenses de origem papuásica pertencem à família Neo-Bomberaica. Ainda, as quatro línguas¹³: Bunak, Fataluku, Makasae e Makalero, separaram-se em períodos históricos diferentes, o que faz com que sua localização geográfica seja descontínua, como é o caso da língua Bunak que se separou primeiro do ancestral comum, e a língua Makalero, que provavelmente originou-se da língua Makasae (fig.4).

¹³Segundo Engelenhoven (2009), há indícios da existência de outra língua papuásica, denominada Rusenu ou Nisa, que provavelmente foi extinta na década de 50 do século XX. Foi achado, porém, um semi-falante que forneceu alguns dados linguísticos. Estes dados apresentam indícios de que essa língua é papuásica e provavelmente desenvolveu-se a partir do Fataluku.

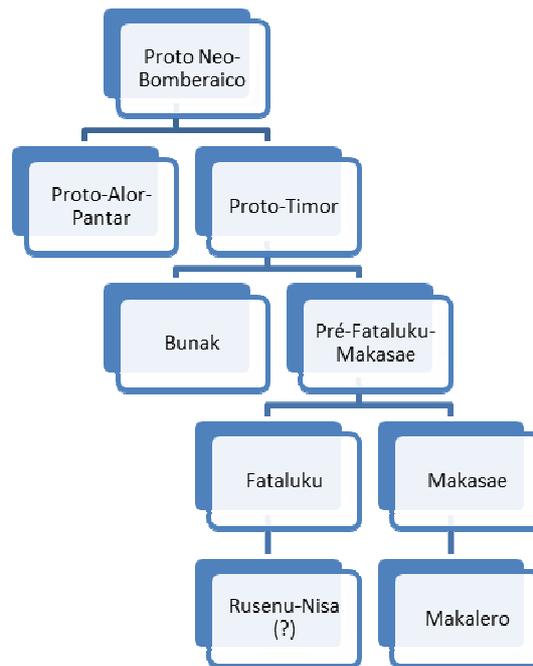
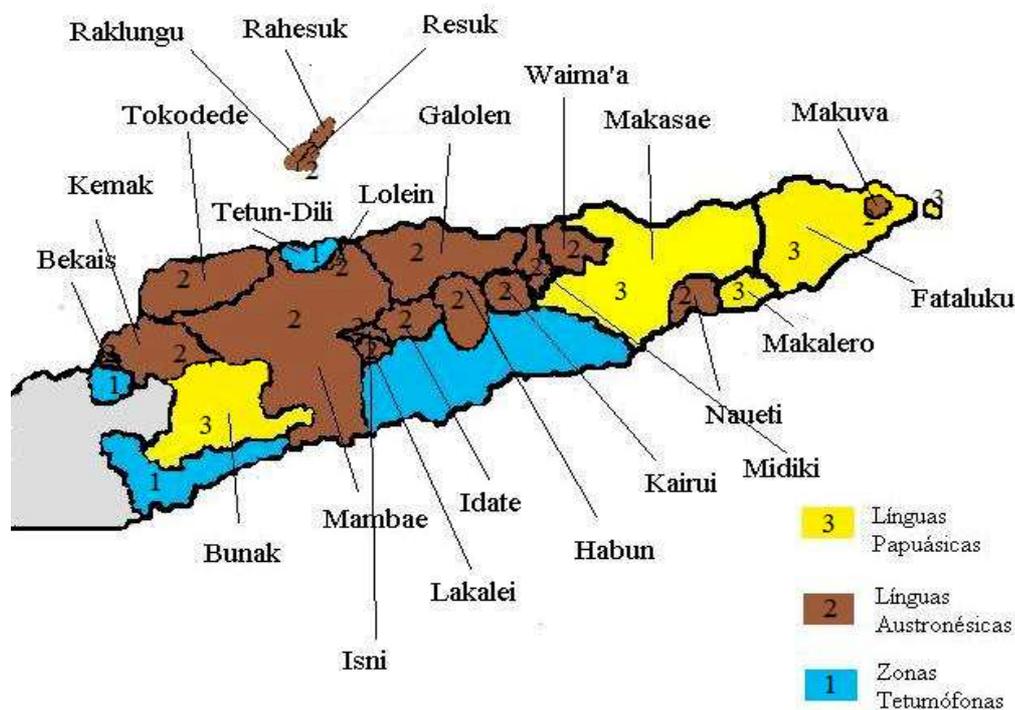


FIGURA 4. As línguas papuásicas de Timor-Leste e suas filiações

Pode-se perceber, entretanto, que há predominância das línguas papuásicas a leste do território leste-timorense (Mapa 3), com exceção somente do Bunak, enquanto os demais povos se espalharam somente por aquela região. A região central de Timor-Leste é predominantemente Manbae, e as regiões adjacentes de outros membros da família Ramelaica: Tokodede, Kemak e Idalaka. Assim, a configuração atual das línguas nativas de Timor-Leste pelo seu território é complexa, porém pode ser esquematizada de acordo com o mapa abaixo:



MAPA 3. Timor-Leste e a distribuição das línguas nativas pelo seu território

3.2 Multilinguismo

Conforme o mapa linguístico visto anteriormente, Timor-Leste possui um amplo conjunto de línguas, juntamente com suas variedades e subvariedades, sendo faladas em todo seu território em situações sociolinguísticas distintas: casamentos interétnicos, comunicação entre habitantes falantes de diferentes línguas, línguas midiáticas etc. Entende-se aqui multilinguismo como a competência do indivíduo de produzir enunciados compreensíveis pelo falante nativo em uma ou mais língua(s) alvo(s) específica(s) (HAUGEN, 1953). Apesar dessa definição, proposta por Haugen (1953), ter sido bastante discutida na atualidade, pelo fato de os linguistas acharem que ela e outras definições mais antigas de bilinguismo e multilinguismo serem mais restritas e ligadas somente à fluência em outra língua (EDWARDS, 1994, p. 56), esta funciona é válida como um ponto de partida para a análise inicial do multilinguismo em Timor-Leste, iniciando com o ambiente familiar, passando pela atitude do indivíduo em relação à(s) língua(s) e, posteriormente, discutida através dos conceitos de etnicidade, territorialismo e nacionalismo.

Partindo do ambiente familiar de aquisição de língua pela criança leste-timorense, deve-se distinguir certos conceitos culturais do que acontece realmente no dia-a-dia do cidadão de Timor-Leste, ou seja, a teoria da prática. A cultura imaterial leste-timorense está amparada em certos conceitos patriarcais e católicos, que demandam a supremacia do homem, da figura paterna, do primogênito e do filho, em detrimento da mulher, da esposa, e da filha. Assim, em relação a situações de multilinguismo no ambiente familiar, a cultura do povo leste-timorense determina que a língua que deve ser adquirida pelos filhos do casal é sempre a língua do pai. Pode-se afirmar que isto acontece na teoria, pois no dia-a-dia dos habitantes de Timor-Leste isto não ocorre por diversos fatores que serão analisados adiante.

Dessa maneira, na prática, surge o primeiro problema a ser analisado: a escolha e o uso da língua a ser falada em casa entre os pais, e entre os pais e a criança. Esse problema que surge na análise ocorre por diversos fatores, apesar da cultura leste-timorense determinar a predominância da língua do pai, isto na maioria das vezes não acontece.

Para se analisar a situação de multilinguismo no ambiente domiciliar, optou-se por se fazer uso da tipologia da escolha de línguas pela família proposta por Romaine (1995, p. 183). Segundo a autora, há seis tipos de situações em que diferentes línguas são usadas pelo pai e pela mãe para a comunicação com a criança:

- *Uma pessoa – uma língua*: os pais falam diferentes L1, possuem certo grau de fluência na língua do outro, a língua comunitária é variedade de um dos pais, e usam cada um sua respectiva língua para se comunicar com a criança;
- *Língua domiciliar não dominante/ uma língua – um ambiente*: a mesma situação anterior, porém a língua usada é a língua não comunitária;
- *Língua domiciliar não dominante sem apoio comunitário*: os pais falam a mesma língua, mas não é a língua comunitária;
- *Língua domiciliar não dominante dupla sem apoio comunitário*: os pais falam línguas diferentes e ambas não são a língua comunitária;
- *Pais não nativos*: os pais falam a mesma língua, que é a língua comunitária, porém um dos pais se comunica com a criança usando uma língua distinta;

- *Línguas mistas*: os pais e a comunidade são bilíngues, e cada um deles usam as diferentes línguas para se comunicar com a criança.

Porém, antes de se realizar uma análise da situação do multilinguismo leste-timoreense de acordo com os tipos propostos acima, deve-se levar em conta que o Tetun Prasa está presente em toda a sociedade de Timor-Leste de alguma forma, salvo raras exceções. O TP, dessa maneira, pode ser adquirido como L1, L2, em diferentes situações de diglossia, dependendo da comunidade, e em um mero conhecimento passivo em alguns casos isolados que serão comentados. Assim, esses tipos propostos por Romaine (1995), que serão descritos a seguir com exemplos, não são absolutos, mas tentativas de se descrever a realidade dos vários falantes que é, na maioria dos casos, fluida.

A maioria das famílias leste-timoreenses se enquadram no tipo de *línguas mistas*, já que a comunidade é predominantemente multilíngue, assim como os pais, conforme verifiquei em diversas visitas às famílias nas zonas urbanas de Dili e Baucau, e nas zonas rurais de Maliana e Fatumaca. Tal fato ocorre principalmente por causa dos casamentos interétnicos serem comuns, já que, de acordo com a cultura de Timor-Leste, o pretendente para efetuar o casamento deve pagar o *barlaki* ‘dote’ à família da noiva. O *barlaki* é pago com búfalos (animal de importância fundamental na cultura dos vários povos de Timor-Leste), podendo também ser pago em dinheiro, jóias e pedras preciosas, porém o preço muitas vezes é alto, chegando a alcançar cerca de 300 cabeças de búfalos, quando a família da pretendida recebe muitas propostas. Isto faz com que a parcela de homens solteiros que desejam casar partam para outros distritos distantes de onde nasceram para adquirir uma esposa. Assim, o noivo, quando consegue em outro distrito ganhar a noiva junto à família, pagando o *barlaki*, estabelece sua residência no distrito de sua futura esposa. Consequentemente, os filhos nascerão neste contexto linguístico.

Por causa dos casamentos interétnicos serem comuns em Timor-Leste, há também a ocorrência dos tipos: *língua domiciliar não-dominante/ uma língua – um ambiente, uma pessoa – uma língua* e *língua domiciliar não-dominante dupla sem apoio comunitário* ora com o pai e a mãe falantes de diferentes línguas, ora com apenas um deles falantes de Tetun Prasa. Há casos em que ambos os pais possuem um grau maior de fluência no TP; há outros em que apenas um deles possuem uma fluência maior em TP, e outros casos, estes em menor número, que ambos não possuem fluência

em TP, falando apenas algumas palavras. Nesses diferentes casos, a criança geralmente adquire o Tetun Prasa, que é a língua comunitária, como L1 pelo fato do *input* recebido das línguas dos pais, ou de um deles, diferentes do TP ser mínimo e a comunidade reforçar e incentivar o uso do TP. A aquisição do TP como L1 nesses casos somente não ocorre em comunidades extremamente isoladas, como a do *Foho Laelako*, ou em comunidades monolíngues, que são somente duas em Timor-Leste, no distrito de Lautém (falante apenas de Fataluku) e no enclave de Oecussi (falante apenas de Baikenu).

Pude verificar em uma visita a uma comunidade relativamente isolada o tipo *Língua domiciliar não-dominante sem apoio comunitário*. Esta comunidade localiza-se no Monte Laelacu, ou *Foho Laelako*¹⁴, que é uma cadeia montanhosa no interior do distrito de Maliana e fica próxima a fronteira indonésia. Nesta comunidade, conduzi algumas entrevistas e verifiquei alguns casais mais velhos, falantes apenas da língua Kemak. Seus filhos adquiriram o Kemak como L1 e somente vieram a aprender o Tetun Prasa, e demais línguas, nos anos iniciais escolares ou somente aprenderam o Tetun Prasa e o *bahasa indonesia* em idades mais avançadas, por necessidades sociais. Ainda, a maioria das vilas do distrito de Lautém, localizado no extremo leste da ilha de Timor, que é um distrito falante da língua Fataluku, são reconhecidas por serem uma das únicas comunidades leste-timorenses monolíngues; dessa forma, os pais são falantes apenas de Fataluku e os filhos adquirem somente a mesma língua dos pais como L1.

O único tipo que não foi encontrado na situação de multilinguismo domiciliar em Timor-Leste foi o *pais não-nativos*, pois na situação domiciliar em que ambos os pais são falantes do TP não há motivo no contexto sociocultural de Timor-Leste para os pais deixarem de se comunicar e transmitir o TP, língua comunitária, para a criança.

Sobre o status das línguas em Timor-Leste pode-se afirmar o seguinte: a constituição de 2002 da República Democrática de Timor-Leste considera a língua portuguesa e o Tetun Prasa como línguas oficiais, e a língua inglesa e o *bahasa indonesia* como línguas de trabalho. O Tetun Prasa é falado como L1 por 18% da população, enquanto 82% da população leste-timorense possui alguma fluência nele, pois é a língua que funciona como língua franca, ou seja, língua de comunicação entre os diferentes grupos etnolinguísticos que possuem L1 distintas.

¹⁴A etimologia deste topônimo é um pouco incerta em relação ao vocábulo *laelako*, enquanto *foho* em Tetun é ‘morro, montanha’. O vocábulo *laelako* poderia ser *lae* ‘NEG’ e *lakon* ‘derrota’ que poderia ser traduzido como ‘invencível’, ou seja, aquele que não é derrotado.

As demais línguas do estado - a língua portuguesa, a língua inglesa e o *bahasa indonesia* -, não possuem falantes nativos. A língua portuguesa e a língua inglesa são faladas por uma pequena parcela da população, 5% para aquela e 2% para esta. Já *obahasa indonesia* é falado por 42% da população, por causa dos fatores sócio-históricos discutidos no capítulo anterior¹⁵. Digno de nota, é que a parcela da população leste-timorense falante de língua portuguesa, assim como a falante de língua inglesa, vem crescendo significativamente nos últimos anos, por causa da política linguística adotada pelo estado. Porém, a língua inglesa vem destacando-se e ganhando um vasto número de simpatizantes e um espaço valioso na cultura popular.

Ainda sobre as línguas portuguesa, inglesa e o *bahasa indonesia* estarem em uma situação de diglossia na sociedade leste-timorense: há a L1 (as várias línguas nativas) que possui um status domiciliar, familiar e/ou regional; há o Tetun Prasa que possui o valor de língua nacional, interétnica e língua franca; e há a presença das línguas estrangeiras (a língua portuguesa, a língua inglesa e o *bahasa indonesia*) que possuem um valor de língua internacional, língua de cultura e são portadoras de um alto prestígio social.

O segundo fator importante ao analisar o multilinguismo é a atitude do indivíduo em relação à(s) língua(s): esta atitude, assim como o uso e a escolha da língua a serem usadas em situações multilíngues, é conhecida como ideologia linguística, ou ideologia de língua. Lanza (2007) analisa as várias definições do termo e também diversos estudos de caso. A autora afirma que a importância de se estudar a ideologia linguística do falante é fundamental, pois ela pode variar de um falante para outro, assim como varia de uma comunidade para outra; a ideologia linguística está ligada diretamente com a escolha da língua feita pelos pais nas situações domiciliares e familiares, e também influencia na realização linguística da conversa entre os pais e os filhos (LANZA, 2007, p. 53)¹⁶.

O terceiro fator, a etnia e a etnicidade, tem um papel importante na concepção das interações e atitudes do indivíduo leste-timorense em relação aos demais indivíduos, assim como na identificação de si mesmo e do grupo. Segundo Edwards (1994, p. 125), não há uma definição explícita de etnicidade e um estudioso, ao tentar fazê-lo, abre a

¹⁵Dados extraídos de *National Board of Statistics* (2006) e *Programadas Nações Unidas para o Desenvolvimento* (2002).

¹⁶Vale a pena lembrar que apesar de a ‘ideologia linguística’ afetar a realização linguística entre pais e filhos, a análise dos elementos linguísticos conversacionais que são frutos dessa comunicação é feita por uma disciplina à parte, chamada ‘análise interacional’.

caixa de Pandora: logo o autor identifica uma série de parâmetros que podem ajudar a solucionar tal problema. Utilizei-me dos seguintes traços, que se revelaram pertinentes ao presente estudo: a identificação das fronteiras grupais, e a separação entre o conceito subjetivo e o conceito objetivo de etnicidade.

A identificação das fronteiras entre os diferentes grupos etnolinguísticos de Timor-Leste é ponto fundamental para o estudo do multilinguismo, já que os principais traços usados pelo povo leste-timorense para se definir como grupo e se separar de outros grupos são: a língua, características físicas, o território e a postura em relação ao outro. Dessa forma, ao fazer uso dos traços que o povo de Timor-Leste utiliza para se identificar como grupo, também fiz uso do parâmetro apontado acima: a separação entre o conceito subjetivo e o conceito objetivo de etnicidade. A seguir será discutido cada um dos elementos, explicando o porquê e como são utilizados pelo próprio povo na sua definição como grupo, separando-se dos demais grupos etnolinguísticos de Timor-Leste e de outros.

A língua é um fator identitário muito forte entre os diferentes grupos leste-timorenses, pois, na maioria das vezes, ela está ligada a um povo somente e a um território também; as únicas línguas que não são faladas somente por um grupo específico são: Tetun, Manbae e Makasae. As características físicas são fatores de etnicidade em Timor-Leste por causa de certos traços físicos notáveis serem específicos somente de um grupo, como é o caso dos Fataluku, que são identificados pelos próprios povos vizinhos como mais altos e de pele mais claras, ou os povos falantes de outras variedades de Tetun dos distritos de Suai e Viqueque serem considerados mais baixos e de pele mais escura. Finalmente, a postura em relação ao outro consiste no comportamento do indivíduo em relação ao estrangeiro e aos povos vizinhos, os chamados *Firaku* que são indivíduos da parte mais oriental do país, considerados os mais falantes e mais violentos, enquanto os *Kaladi* são indivíduos da parte mais ocidental do país, e são considerados mais calados, calmos e taciturnos.

O conceito de territorialismo será discutido separadamente ao do nacionalismo, pois ambos são fatores culturais importantes em Timor-Leste. Considero aqui a noção de territorialismo ligada à divisão dos distritos de Timor-Leste, que em sua maioria refletem as divisões entre os pequenos reinos existentes em tempos pré-coloniais. Estes reinos eram de certa forma pequenas aldeias com uma formação social complexa, organizada em castas, e ainda havia entre os diversos pequenos reinos leste-timorenses

uma intensa rede de alianças com o objetivo de união para a constante guerra existente entre uns e outros. Estes reinos eram falantes cada um de uma língua diferente, o que faz com que esta herança cultural seja mantida atualmente, com a ideia de que cada distrito seja falante somente de uma língua e seja reflexo de um pequeno reino antigo.

Finalmente, o conceito de nacionalismo não se encontra muito presente na identidade do povo de Timor-Leste. Principalmente se for pensado de maneira mais atual e ocidentalizada, pois o povo leste-timorense possui uma concepção de identidade grupal, pensando em um grupo maior que ultrapassa as aldeias e os distritos. Porém esta identidade de um grupo maior não coincide com a concepção de estado-nação. Entre os elementos culturais que são comuns aos diversos grupos são: a variedade Tetun Prasa ser encarada como uma língua extra-grupal (ou seja, nacional); a religião católica misturada com alguns elementos animistas, como veneração aos antepassados e a animais, principalmente o crocodilo e o búfalo; um conjunto semelhante de contos e lendas sobre a ilha de Timor ter se originado de um crocodilo e o povo ter migrado da ilha de Makassar; presença de alguns rituais relacionados a organização dos clãs, *uma lulik* a casa sagrada¹⁷, e a casamentos, a entrega do *barlaki*, como foi mencionado.

3.3 A ecolinguística de Timor-Leste

A ecolinguística, também conhecida como ecologia da linguagem, é definida como o estudo das interações entre as línguas e seus respectivos meio ambiente, de acordo com Haugen (1972, p. 324). Desta forma, este ramo da linguística encara a língua como um fenômeno natural, social e psicológico. Por isso, preocupa-se em analisar o meio de comunicação, a língua (L), em uso efetivo em uma determinada sociedade, o território (T), e falado por um grupo específico, povo (P). Em outras palavras, “para que haja uma L, é necessário que exista um P, cujos membros vivam e convivam em um determinado T”, conforme Couto (2007, p.20).

Em um trabalho anterior (ALBUQUERQUE, 2010c), analisei a configuração ecolinguística de Timor-Leste, discorrendo sobre a língua (L), o território (T) e o povo (P). Nas seções (3.1) e (3.2) analisei brevemente diferentes aspectos (filiação genética, multilinguismo) sobre as línguas faladas em Timor-Leste (L), e no capítulo anterior

¹⁷A ‘casa sagrada’ é uma estrutura da cultura material timorense, grosso modo, semelhante ao totemismo, dividindo a sociedade e as famílias de maneira análoga aos clãs. Para uma análise antropológica dos diferentes povos de Timor, ver Traube (1986) para os Mambae, Clamagirand-Renard (1980) para os Ema, falantes da língua Kemak, e Gomes (1972) para os Fataluku.

(cap. 2) e na seção anterior (3.2) descrevi características fundamentais das diferentes etnias (P). Dessa forma, falta somente analisar as questões relativas ao território (T).

A ilha de Timor é uma pequena ilha localizada no Sudeste da Ásia, mas que se localiza também próximo ao norte da Austrália, cerca de 650 km. Somente a parte leste da ilha é o território da nação timorense, que constitui cerca de 14.600 km², enquanto a parte ocidental da ilha é território indonésio. Geologicamente, as ilhas da região de Nusantara Oriental¹⁸ fazem parte da grande placa tectônica australiana e são ilhas vulcânicas. O solo de Timor-Leste (FOX, 2000) é um solo predominantemente barroso, mas que possui a presença de alguns materiais rochosos, entre eles rochas metamórficas, sedimentares e vulcânicas. Ainda, o relevo é montanhoso com o seu ponto mais alto, o Monte Ramelau, com 2.963 metros de altitude; nos diversos vales entre os montes e as montanhas encontram-se muitos rios caudalosos, especialmente na época das chuvas, entre outubro e dezembro. Diferente do resto do território, ao longo do litoral há uma planície, assim como a presença de banco de corais e recifes. O clima é de monções, caracterizando-se por longos períodos secos: de março a junho e de agosto a novembro, enquanto de dezembro a fevereiro e julho são os dois períodos em que chegam as monções que trazem chuvas à região. As únicas áreas férteis são os vales que acumulam água e tornam o solo produtivo para a agricultura. As demais regiões não são próprias para agricultura, mesmo os solos considerados mais produtivos são pouco férteis, possuem somente a predominância de cálcio, enquanto apresenta ausência dos demais nutrientes necessários. Estes fatores limitam, e limitaram, a agricultura de Timor a poucos produtos a serem plantados e, conseqüentemente, a povoação do território¹⁹.

Ainda, no mesmo artigo citado acima (ALBUQUERQUE, 2010c) analisei as publicações existentes sobre ecolinguística de Timor-Leste. Em minha análise, observei que a noção de ecolinguística dos autores está intimamente relacionada com política linguística. Os trabalhos são os seguintes: Hajek (2000), Taylor-Leech (2005) e Wendel (2005).

No primeiro deles, Hajek (2000) analisa a política e o planejamento linguístico formulado para o Timor-Leste no decorrer da história. Ele analisa o chamado período

¹⁸ Grupo de pequenas ilhas que se estende desde a ilha de Flores a leste, na Indonésia, até a região da baía de Cenderawasih e da Bird's Head a oeste, na Papua Nova-Guiné.

¹⁹ Segundo Fox (2003, p. 106), os principais produtos agrícolas são: arroz (*Oryza sativa L.*), milho-miúdo (*Panicum viride L./Setaria italica L.*), gergelim (*Sesame orientale L.*), sorgo (*Andropogon sorghum Brot.*) e ervilha d'angola (*Cajanus cajan Millspaugh.*), de acordo com as reconstruções lingüísticas do Proto-Austronésio, ou do Proto-Malaio-Polinésio.

português, que se estende de 1515 até 1975; o período de dominação indonésia, de 1975 até 1999; o período de independência que se estende de 2002 até a atualidade. Hajek (2000) afirma que as autoridades não se importavam com as línguas nativas e a sobrevivência dessas línguas até os dias atuais baseou-se em uma ecologia local das línguas com a capacidade de se adaptar e sobreviver, e aponta como solução atual a valorização de uma língua nativa – a língua Tetun, que é língua oficial – considerado um marco, com uma mudança da ecologia linguística voltada, agora, para uma língua nativa. Críticas, porém, ainda são feitas, pois se deve manter uma ecologia saudável, centrada em todas as línguas nativas, e não somente concentrada em uma língua, neste caso o Tetun.

Taylor-Leech (2005), em seu artigo, argumenta da mesma maneira que Hajek (2000). A autora, porém, vai além e apresenta, segundo ela, o que seriam ‘forças’ que podem influenciar e deslocar a ecologia das línguas em Timor-Leste, a saber: a presença da língua inglesa, que foi inserida através da presença de entidades internacionais de assistência que trouxeram junto com elas seus respectivos funcionários; o legado do colonialismo; fatores relacionados à política e identidade linguística. Ainda, como proposta da autora, ela vai além de uma simples análise, pois considera uma solução para a valorização das línguas nativas o reconhecimento da sociedade timorense como uma sociedade multilíngue e a expansão do uso das línguas nativas em diferentes ‘espaços sociolinguísticos’ (TAYLOR-LEECH, 2005, p.119), como: educação, judiciário e imprensa.

Finalmente, Wendel (2005) apresenta em seu artigo uma grande diferença dos anteriores, pois analisa primeiramente diversas questões teóricas e metodológicas a respeito da ecolinguística, e analisa também o ‘meio ambiente’ leste-timorense. A proposta de Wendel (2005, p.73) é aquela que pode ser considerada como a mais ‘ecológica’ de todas, pois o autor, ao mesmo tempo em que alerta sobre a dominação das potências, como a língua portuguesa, que é a língua do colonizador, e a língua inglesa que é a língua mundial e, de certa forma, uma língua sedutora, ele também alerta sobre o equilíbrio ecológico com o ambiente, já que as línguas nativas não podem sobreviver sem o contato linguístico, que já acontecia antes da chegada dos portugueses no século XVI, e o completo isolamento, ou valorização da língua nativa através do excesso de ‘purismos’, podem levá-las a situação de línguas ameaçadas, ou até à extinção.

4. A LÍNGUA TETUN

Apesar de uma quantidade considerável de empréstimos lusófonos, como será analisado no capítulo seguinte sobre o léxico tetumófono, a língua Tetun é de origem austronésia, mais especificamente à família Malaio-Polinésia. Em sua variedade Tetun Prasa, ela é falada de alguma maneira por 82% da população, enquanto é L1 de 18% da população total de Timor-Leste, aproximadamente a população total do distrito de Dili que tem essa variedade como L1, de acordo com os diferentes censos realizados pelos diferentes programas da ONU: *National Board of Statistics* (2006) e *Programadas Nações Unidas para o Desenvolvimento* (2002). Segundo os mesmos censos, a variedade chamada Tetun Terik, língua materna do distrito de Viqueque e uma parcela da população de Suai, é falada como L1 por 6% dos leste-timorenses e as demais variedades do Tetun somam cifras inferiores a 2%.

Salvo os distritos e a parcela da população onde o Tetun em alguma de suas variedades são falados como L1, conforme citado acima, a variedade Tetun Prasa encontra-se majoritariamente em uma situação de diglossia, pois a maior parte da população leste-timorense adquire uma língua nativa, geralmente associada ao seu distrito, como língua materna e, logo em seguida, nas primeiras interações sociais fora de seu ambiente familiar próximo, há uma necessidade da criança aprender rapidamente e de maneira informal o Tetun Prasa por necessidades comunicativas. As demais línguas, como a língua portuguesa e a língua inglesa, são aprendidas como L2/LE em uma fase posterior, como foi analisado no capítulo 3, na seção sobre o multilinguismo leste-timorense (3.2).

A língua Tetun possui um status de língua franca em Timor-Leste desde aproximadamente século XV (ALBUQUERQUE, 2009; THOMAZ, 2002). Os portugueses chegaram à ilha de Timor no início do século XVI e já a encontraram sendo falada em um vasto território. Porém aproximadamente no mesmo período, século XV, a língua franca do sudeste asiático era o *Pazar Melayu*, variedade crioula do Malaio. Esse fator faz com que seja possível datarmos a ascensão do império tetumófono na ilha de Timor durante este mesmo período, meados do século XV.

Outras informações históricas que servem como argumento para uma possível datação do surgimento do Tetun Prasa como língua franca podem ser encontradas em documentos da administração portuguesa. Há uma documentação rica que aponta a

existência de dois grandes reinos. O primeiro, chamado de reino dos Belos, que se localizava na região central do que é hoje Timor-Leste e tinha em sua figura central o rei de Luca. O segundo, chamado de reino de Servião, que ficava na parte oeste da ilha, o que corresponde atualmente a território indonésio, e a figura central era o rei de Wehale. Essas informações e outras, a respeito do período histórico de Timor-Leste encontram-se na coletânea de documentos do período colonial português no oriente, em Sá (1961).

No presente capítulo serão apresentados os estudos que já foram realizados tendo como objeto a língua Tetun. Primeiramente, falaremos sobre os estudos realizados nos tempos coloniais e os primeiros estudos linguísticos (4.1); para, em seguida, nos debruçarmos nos estudos linguísticos mais recentes e sistemáticos sobre o Tétum (4.2). Depois, discorreremos sobre a classificação do Tetun Prasa (4.3), os registros tetumófonos nas diferentes situações sociais (4.4) a variação dialetal da língua (4.5).

4.1A documentação no período colonial e os primeiros estudos linguísticos

O primeiro registro sobre as línguas nativas timorenses data do século XVIII. Porém foram listas de palavras das línguas Galolen, Makasae, Tetun e Waima'a, compiladas por um viajante francês chamado F. E. de Rosily, que nunca foram publicadas e datam do ano de 1772.

O primeiro trabalho que foi dedicado ao estudo da língua Tetun foi realizado por Sebastião da Silva, um padre português, que publicou um dicionário Português-Tetun (SILVA, 1889). Anos mais tarde, Raphael das Dores continuou o trabalho do padre Sebastião da Silva e publicou um dicionário Tetun-Português, no ano de 1907 (DORES, 1907).

O início do século XX foi marcado pelo forte presença portuguesa nos estudos sobre a língua Tetun. O padre Manuel Mendes Laranjeira publicou sua *Cartilha Tétum*, em 1916 (LARANJEIRA, 1916); Júlio Garcêz de Lencastre publicou um trabalho intitulado *Algumas regras gramaticais da língua tétum e vocabulário*, em 1929 (LENCASTRE, 1929); e outro padre, chamado Abílio Fernandes, publicou outra obra didática sobre o Tetun o *Método prático para aprender o tétum*, em 1937 (FERNANDES, 1964). Ainda, no ano de 1952, o padre Artur Basílio de Sá realiza um estudo linguístico sobre o Tetun, com ênfase na fonologia e propõe um sistema

ortográfico com bases fonéticas em seu *Notas sobre linguística timorense: Sistema de representação fonética* (SÁ, 1952).

Os trabalhos, citados anteriormente, publicados sobre a língua Tetun foram obras realizadas por padres portugueses e seguem uma tradição portuguesa e missionária já conhecidas, com ênfase nos estudos lexicográficos e em obras didáticas voltada para a aprendizagem da língua nativa. As grandes contribuições desses trabalhos estão: no processo de formação da ortografia padronizada do Tetun Prasa (INL, 2002, 2004), pois, dos diversos problemas encontrados na formulação de uma ortografia do Tetun, cada um dos autores apresentaram uma solução distinta; e no registro lexicográfico do Tetun em períodos históricos distintos, o que possibilita estudar e mapear os empréstimos e as mudanças linguísticas.

O primeiro estudo de natureza classificatória sobre as línguas timorenses, incluindo assim o Tetun, foi de autoria do australiano Arthur Capell. Ele teve um contato com diversos refugiados do Timor-Leste durante o período da II Guerra Mundial, pois o Timor-Leste foi invadido pelo Japão em 1942. Seu extenso artigo intitulado *Peoples and Languages of Timor* foi dividido em três partes para a publicação, nos anos de 1943 e 1944. A primeira parte de seu artigo consiste em uma revisão bibliográfica; na segunda parte, o autor realiza uma análise fonológica e morfossintática das línguas a que ele teve acesso, São elas: Tetun, Tokodede, Manbae, Galolen, Waima'a, Baikenu, Bunak e Makasae²⁰, e a terceira parte consiste na tradução de alguns textos e uma lista comparativa de 170 itens lexicais.

Durante a ocupação indonésia, a grande parte dos estudos linguísticos tem o Tetun Belo como enfoque. Fransiskus Monteiro elaborou um dicionário Tetun-Indonésio em 1985, e contribui com outros autores em um estudo de fonologia e morfossintaxe do Tetun Belo (TROEBOES *ET AL.*, 1987), e outro de morfologia da mesma variedade (SALIWANGI *ET AL.*, 1991). O Tetun falado em Timor-Leste foi somente estudado por um conjunto de linguistas em 1992, em um trabalho intitulado *Sistem morfologi kata tugas bahasa Tetum*.

4.2 Os estudos tetumófonos na atualidade

²⁰O autor ainda teve acesso e analisou em seu artigo três línguas faladas na parte oeste da ilha de Timor. São elas: uma variedade do Tetun, Dawan e Helong. Como o estudo dessas línguas está fora do escopo do presente projeto, falar-se-á delas somente quando necessário.

Nos últimos anos destacaram-se os seguintes trabalhos sobre a língua Tetun: Esperança (2001) apresenta uma coletânea de artigos sobre os mais variados estudos linguísticos, com análise linguística do Tetun, questões de política linguística e contato de línguas; Hull (2000, 2001a, 2001b, 2004) realizou diversos trabalhos na área de linguística histórica, verificando a filiação genética, a classificação interna das línguas e reconstruindo a proto-língua; Williams-van Klinken vem realizando trabalhos na área de descrição (KLINKEN, 1999; WILLIAMS-VAN KLINKEN, HAJEK e NORDLINGER, 2002) e análise linguística sobre temas notáveis da gramática do Tetun, como os verbos seriais, a expressão de atributos, entre outros.

4.3 A classificação da língua Tetun

A variedade Tetun Prasa da língua Tetun é considerada por muitos autores como um crioulo (LEWIS, 2009), uma língua formada por processo de crioulição (HULL, 2001b), ou um pidgin (HAGÈGE, 2002 *apud* HAJEK, 2007). Nesta gramática, considero o Tetun Prasa apenas como uma variedade (ou dialeto) da língua Tetun teve as mudanças linguísticas intensificadas por processos de contato de línguas e, assim ocorreram aquisição não nativa (McWhorter, 2007) e reestruturação gramatical (Holm, 2004). Estes dois processos alteram as estruturas gramaticais, tornando-as, de certa forma, semelhantes a crioulos, o que não faz da língua um crioulo genuíno²¹.

Com os dados dos estudos de natureza comparativa, citados anteriormente (HULL 2000, 2001b), foi possível classificar a língua Tetun como pertencente ao tronco Austronésio e à família Malaio-Polinésia (PMP). A lista de cognatos PAN e Tetun Prasa, localizada no **Apêndice**, está organizada de acordo com os diferentes campos semânticos e consta de diversos itens lexicais e gramaticais reconstruídos para o Proto-Austronésio (PAN) comparados com o Tetun Prasa. Ainda, esta lista, adaptada de Blust (1999), apresenta mais de 120 itens lexicais e gramaticais reconstruídos para o PAN. Desta lista, o Tetun Prasa possui aproximadamente 84 prováveis cognatos com o PAN, o que corresponde a cerca de 70% de taxa de cogação com o PAN. Esta taxa de cogação é um forte argumento a favor da classificação da língua Tetun como pertencente ao tronco Austronésio e a família Malaio-Polinésio²².

²¹ Há necessidade de ser investigadas as diferentes subvariedades do Tetun Prasa, já que o Tetun Dili se assemelha a uma língua mista, ou língua entrelaçada. Esta investigação, porém, foge do escopo deste trabalho.

²² As demais variedades da língua Tetun apresentam uma taxa de cogação ainda maior quando comparada com o PAN, por não ter sofrido tanto com o contato linguístico como o Tetun Prasa. Logo, as

Além disso, pode-se perceber a regularidade na mudança linguística do PAN para o Tetun Prasa:

- As consoantes uvulares do PAN *C, *S, *R, *q e *N caíram em TP, como em 2, 3, 5, 8, 9, 11, 19, 20, 21, 63, 64, 65, 77, 79 e vários outros, em alguns casos *N > n e *C > t;
- *z > d como em 4, 40, 77;
- O *t do PAN mantém-se t em TP como em 22, 24 e 30;
- *b sofre espirantização, tornando-se f em Tetun, como em: 15, 33, 66, 74, 75, 89;
- Hull (2000) realiza um estudo exaustivo das correspondências fonológicas do TP com as reconstruções do PAN e do PMP.

4.4 O Tetun nas diferentes situações sociais

No único trabalho existente sobre os diferentes registros do Tetun Prasa, Williams-van Klinken (2001) enumera duas modalidades principais: o registro católico e o registro midiático. Ainda, acrescento mais um registro aos dois identificados pela autora, que chamo de *Tetun literário*. Vale mencionar que entendo aqui o termo ‘registro’ como é utilizado amplamente na sociolinguística: ‘diferenças linguísticas correlacionadas com diferentes ocasiões de uso’ (FERGUSSON, 1994, p.16). Essas ocasiões podem ser formais, informais ou rituais, e trata-se de como o falante modifica estruturas e como estas se tornam específicas de uma situação comunicacional para se referir a: pessoas mais velhas e mais novas, pessoas do mesmo sexo e do sexo oposto, pessoas de status social elevado, entre outros. Ainda, Fergusson (1994, p.20) afirma que essas estruturas linguísticas com o tempo tornam-se marcas identificadoras, diferenciando-se do uso linguístico em outras situações comunicacionais.

Segundo Williams-van Klinken (2001), a formação do registro da igreja católica teve como principal responsável o ensino formal da renomada instituição católica: o colégio de Soibada, fundado em 1898. Esta instituição permaneceu durante muito tempo

demais variedades possuem um número maior de cognatos austronésios e um número menor de empréstimos lusófonos.

Ainda, uma comparação com o Proto-Malaio-Polinésio (PMP) apresenta mais evidências da filiação genética da língua Tetun a esta família e ao PAN, porém não faz parte do escopo da presente dissertação aprofundar esses temas relativos à linguística histórica.

como a única que fornecia educação a nível secundário e que facilitou a expansão dos seus ensinamentos. Esta variedade tem como traço principal o uso de elementos linguísticos do Tetun Terik, como:

- a manutenção dos fonema /w/ e /ʔ/, que foram substituídos pelo fonema /b/ e pelo fonema zero em Tetun Díli, respectivamente;
- evitar o uso de empréstimos lusófonos, recorrendo ao léxico nativo, como ao invés de ser usado o vocábulo lusófono *ajuda* utiliza-se o vocábulo nativo *tulun*;
- uso de algumas estruturas morfossintáticas nativas, principalmente o uso dos classificadores nominais, e o uso facultativo da flexão de 3ª pessoas do singular e do plural do verbo, os prefixo n- e r- respectivamente.

Outra variedade identificada pela autora para o Tetun Prasa é a do registro midiático. Nesta, há um número maior de empréstimos lusófonos em nível lexical e também em nível morfossintático, principalmente a marcação e a realização da concordância de número e gênero nos empréstimos lusófonos: *embaixadores brasileiros* e *primeira vise-secrétaria*. Segundo a autora, isto ocorre devido aos fatores de muitas jornalistas terem realizado seus estudos em Portugal, assim como, o apelo aos empréstimos lusófonos é uma espécie de reação para retirar do Tetun Prasa os malaísmos.

Somada às duas variedades mencionadas anteriormente, acrescento uma terceira, que alguns linguistas ora chamam de *Tetun literário* e resolvi manter esta terminologia. Esta variedade é pouquíssima usada pelos timorenses e, na maioria das vezes, é ininteligível para o falante tetumófono nativo. Trata-se na realidade de uma norma-padrão forjada pelo INL (Instituto Nacional de Linguística), instituição responsável pela promoção e manutenção do Tetun Prasa, assim como pela pesquisa e estudo das demais línguas nativas leste-timorenses.

Assim, o registro literário, algumas vezes identificado como padrão, do Tetun Prasa é extremamente artificial e desconhecido por grande parte da população até mesmo da parcela escolarizada. Este *Tetun literário* é conhecido somente por linguistas e pelos acadêmicos timorenses, usado somente em publicações técnicas, ou por poetas tetumófonos, mas ambas as publicações são escassas, de difícil acesso e em número reduzido, o que faz com que esse registro continue desconhecido para a população.

O Tetun Padrão foi elaborado por Hull e Correia (2005) e entre as características linguísticas notáveis são: o uso de um conjunto extenso de prefixos, circunfixos e sufixos nominais, adjetivais e verbais responsáveis por derivação e mudança de valência; regras específicas de uso dos empréstimos lusófonos, e vasto uso de marcadores aspectuais no verbo. Abaixo apresento duas regras de Hull e Correia (2005) a respeito da flexão nominal e do uso específico do empréstimo lusófono mencionado:

- o circunfixo substantival (RE) + a- -k deve ser usado para se derivar substantivos de verbos, como em *fahe* ‘dividir’ > *fafahek* ‘divisão’, *la’o* ‘andar’ > *lala’ok* ‘andamento, progresso’
- o empréstimo lusófono ‘serviço’ possui duas formas em Tetun Prasa *servisu* e *serbisu*, a primeira forma, *servisu*, deve ser usada com o sentido ‘serviço’, enquanto a segunda, *serbisu*, deve ser usada com o sentido de ‘trabalhar’ ou ‘trabalho’.

4.5 A variação da língua Tetun

De acordo com o mapa (3), a língua Tetun, além de ser a língua franca em Timor-Leste, é língua materna em diferentes regiões do território: chamadas de *zonas tetumófonas* destacadas em azul. Cada uma destas zonas fala variedades e subvariedades diferentes da língua Tetun. A seguir analisarei brevemente o que já foi dito sobre a variação da língua Tetun nas gramáticas existentes, que são: Hull e Eccles (2001), Hull (2002a), Klinken (1999), e Willians-van Klinken, Hajek e Nordlinger (2002).

A gramática de Hull e Eccles (2001) foi traduzida para o português pelo *Instituto Camões* e até os dias de hoje é a única gramática da língua Tetun existente em língua portuguesa. Este é um trabalho extenso, apresentando as diversas estruturas do Tétum-Praça, comparando-as com outras duas variedades da língua, chamadas pelos autores de *Tetun Terik* e *Tetun Belo*. Os autores identificam as variedades da língua Tétum como dialetos e apresentam as seguintes diferenciações: *Tetun Prasa* é usada como língua franca, possui maior influência da língua portuguesa, sofreu um grande processo de crioulização e possui falantes nativos somente em Díli; *Tetun Terik* é o dialeto que mantém as estruturas nativas tetumófonas e uma das únicas línguas de Timor-Leste que não sofreu processo de crioulização; *Tetun Belo* é o dialeto que sofreu crioulização

parcial, possui pouca influência da língua portuguesa, mas apresenta uma clara influência da língua Malaio, em sua variedade *bahasa indonesia*, por ser falada na região com a fronteira indonésia. Digno de nota é que essa divisão e denominação é a empregada pelos falantes tetumófonos de Dili, capital de Timor-Leste.

A gramática do *Tetun Dili* de Willians-van Klinken, Hajek e Nordlinger (2002) baseou-se em amplo *corpus* linguístico escrito coletado pelos autores. Uma das autoras elaborou também uma gramática de uma subvariedade do Tetun, intitulado *Tetun Fehan* (KLINKEN, 1999). A autora divide a língua Tetun em duas variedades: *Tetun Dili*, que corresponde ao Tetun Prasa, e *Tetun Terik*, variedade mais conservadora falada nas diferentes zonas rurais apontadas no mapa e até em território indonésio.

Outro fator importante para os estudos tetumófonos consiste na terminologia relacionada à variedade do Tetun utilizada como língua franca, ora esta variedade é referida como *Tetun Díli*, ora é referida como *Tetun Prasa*. Hull (2002a) diferencia as duas terminologias, considera *Tetun Prasa* como a variedade do *Tetun Língua-Franca* falada como segunda língua e, assim, sofrendo influências da língua materna dos falantes. Já o *Tetun Díli* corresponde a diferentes variantes do *Tetun Língua-Franca* faladas somente em Díli.

De maneira distinta, Willians-van Klinken (com. pes., Julho de 2009) considera *Tetun Praça* e *Tetun Díli* ambos sinônimos para se referir à língua franca de Timor-Leste. Ainda, a autora afirma haver uma diferença entre a língua franca falada na capital Díli, na maioria língua materna destes falantes, e a língua franca falada nos distritos mais rurais, apesar de a gramática ser basicamente a mesma, há uma grande diferença no léxico: nos distritos rurais há poucos empréstimos da língua portuguesa e completa ausência do léxico modernizado: terminologia técnico-científica, jurídico-administrativa etc.

Desta maneira, proponho com base no que Hull e Eccles (2001) e Willians-van Klinken, Hajek e Nordlinger (2002) afirmaram, assim como me baseando nos dados linguísticos coletados, a seguinte estrutura da língua Tetun e suas variedades:

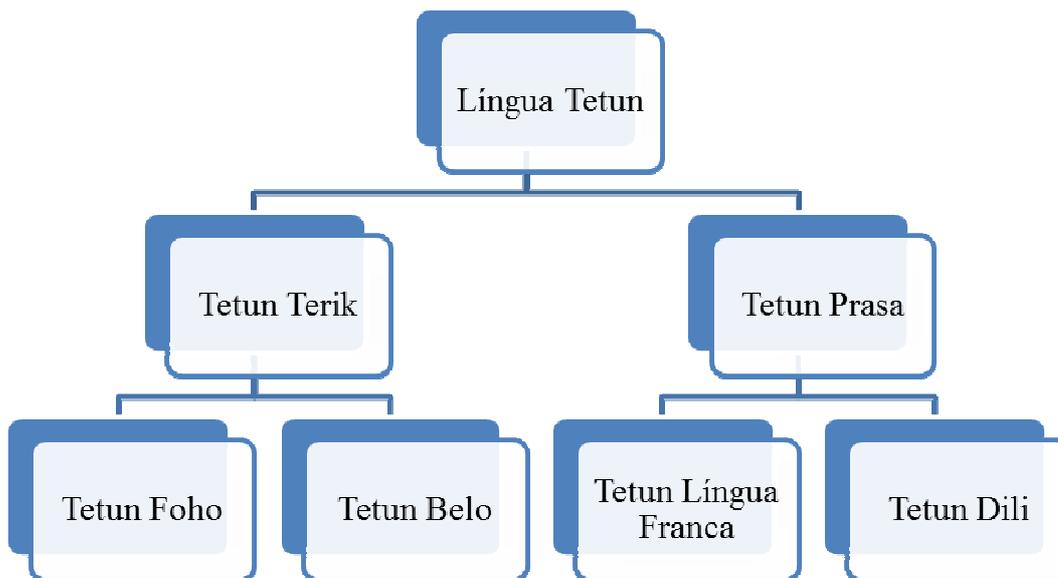


FIGURA 5. A língua Tetun, suas variedades e subvariedades

De acordo com a figura, procuro unificar as propostas de Hull e Eccles (2001) e Willians-van Klinken, Hajek e Nordlinger (2002)²³. Assim, a língua Tetun está dividida somente em duas variedades: o *Tetun Prasa* e o *Tetun Terik*. O Tetun Prasa está subdividido principalmente em duas subvariedades, a saber: *Tetun Língua-Franca* e *Tetun Dili*. O *Tetun Terik* também se encontra subdividido em duas subvariedades: *Tetun Foho* e *Tetun Belo*.

O Tetun Prasa, quando comparado ao *Tetun Fehan*, apresenta um grande número de exemplos que corroboram para a hipótese da reestruturação gramatical, devido a um conjunto de fatores linguísticos e sócio-históricos, entre eles 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 *Tetun Terik* (ALBUQUERQUE, 2010a).

²³ Conforme afirmei anteriormente, caso uma pesquisa futura revele que o Tetun Prasa, ou uma de suas subvariedades, seja uma língua mista ou entrelaçada, haverá necessidade de se rever tal subdivisão da língua Tetun.

Na fonologia, o Tetun Prasa apresenta um número maior de segmentos consonantais em relação ao *Tetun Terik*, que possui um número reduzido. Este número maior de fonemas do Tetun Prasa é fruto dos diversos empréstimos fonológicos, principalmente de origem lusófona:

TABELA 1. Os fonemas da variedade Tetun Terik²⁴

		Bilabial	Labio dental	Alveolar	Velar	Glotal
Oclusiva	surda			t	k	ʔ
	sonora	b		d		
Fricativa	surda		f	s		h
	sonora					
Lateral				l		
Vibrante				r		
Nasal		m		n		
Glide		w				

Em relação à morfologia e à sintaxe, o Tetun Prasa apresenta uma simplificação das estruturas morfológicas e morfossintáticas, que foram mantidas no *Tetun Terik*. Essas mudanças linguísticas, que simplificaram a morfossintaxe do Tetun Prasa, fizeram com que esse processo de contato linguístico fosse considerado como um processo de crioulização (HULL, 2001), e que esta variedade do Tetun fosse classificada como uma língua crioula pelo *Ethnologue* (LEWIS, 2009) e até como um pidgin (HAGÈGE, 2002 *apud* HAJEK, 2007). Em Albuquerque (2010a), apresentei argumentos contra essas classificações, verificando que os processos de mudanças sofridos pelas línguas de Timor-Leste, mesmo que simplificando muitas estruturas gramaticais, podem ser classificados como simples mudanças linguísticas.

As diferenças morfossintáticas entre as variedades *Tetun Terik* e Tetun Prasa que podem ser mencionadas são as seguintes: *Tetun Terik* apresenta uma série de prefixos marcadores de sujeito no verbo que varia de acordo com a pessoa, que se perdeu totalmente em Tetun Prasa (13) e (14); Tetun Prasa não apresenta a cópula *nii* (15); *Tetun Terik* possui dois verbos existenciais *iha* e *nó*, enquanto o Tetun Prasa manteve somente o *iha* (16) e (17); o sistema de negação complexo do *Tetun Terik* foi reduzido para a negação

²⁴ Os dados presentes nos exemplos do Tetun Terik basearam-se na subvariedade *Tetun Fehan* e foram extraídos de Klinken (1999). Os dados do Tetun Prasa foram adquiridos na pesquisa de campo feita pelo presente autor.

verbal *la=*, a negação nominal *la'ós* e a negação absoluta *lae* (18) e (19); *Tetun Terik* possui um sistema de classificação nominal marcada por classificadores numerais, o Tetun Prasa mantém somente três dos classificadores, porém com uso facultativo e limitado a situações formais (20):

13. TF²⁵ 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?'

14. TF sia r-alai r-ola r-ikar loro-sa'e=n baa.
3pl 3-correr 3-ir 3-voltar sol-subir=GEN DIST

TP sira halai ba lorosae
3pl correr CNTF leste
'eles fugiram em direção a leste'.

15. TF lale, tais oo=k nii nia
NEG roupa 2sg=GEN COP 3sg

TP lae, o=nia hena mak ida nee.
NEG 2sg=POS roupa TOP um este
'Esta roupa não é tua.'

16. TF ai kanoik nee hori rai moris noo kedas.
história este desde terra nascer EXI imediatamente

TP istoria nee iha hori rai moris.
história este EXI desde terra nascer

²⁵ Nos exemplos usam-se as siglas TF para a subvariedade do Tetun Terik chamada 'Tetun Fehan' e TP para 'Tetun Prasa'.

‘Esta história existe desde o início do mundo.’

17. TF Ei! noo feto ida. noo feto ida noo oa!
 EXCL EXI 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)

18. TF Oo! nuat e’e bei beur ha’i!
 oh coisa PRX Sr. enganar não
- TP Oo! buat nee lae senoor bosok.
 EXCL coisa este NEG senhor mentira
- Oh! Esta coisa (na realidade uma pessoa) não é o Sr. Enganador!

19. TF ne’e lahoos ema lian. manu lian.
 PRX realmente.não pessoa voz pássaro voz
- TP nee nee lae ema lian manu lian.
 ANA isto NEG pessoa voz pássaro voz
- Isto não é a voz de uma pessoa. (É a) voz de um pássaro.

20. TF feto kawa’ik na’in neen ne’e
 mulher velho CLF.HUM seis PRX
- TP ferik nee (nain) neen.
 velha este seis
- ‘estas seis velhas (senhoras).’

A seguir enumero algumas características sociolinguísticas e tipológicas como argumentos a favor da classificação da língua Tetun proposta aqui. Primeiramente, apontarei as diferenças entre *Tetun Língua-Franca* e *Tetun Dili*, e depois as diferenças entre *Tetun Foho* e *Tetun Belo*.

Tetun Língua-Franca:

- não possui falantes nativos;
- falada como segunda língua e em situações de diglossia;
- usada como língua franca nos diferentes distritos do país;
- influência do *bahasa indonesia*;
- ausência de léxico modernizado.

Tetun Dili:

- falada como língua materna da população de Dili;
- maior influência da língua portuguesa;
- léxico modernizado de origem lusófona, ou anglófona.

Tetun Foho:

- língua materna no distrito de Viqueque, zona tetumófona a sul no mapa;
- falada exclusivamente em Timor-Leste nas zonas montanhosas, por isso o nome *foho* ‘montanha’, proposto por Klinken (1999);
- apresenta um número menor de estruturas nativas, quando comparada ao *Tetun Belo*, pelo fato de sofrer influência das línguas austronésias vizinhas que perderam grande parte de suas marcações morfológicas e sintáticas, tornando-se altamente isolantes.

Tetun Belo:

- língua materna de uma ampla região que se estende de Timor-Leste à Indonésia;
- localizada na zona tetumófona a oeste do mapa;
- grande influência lexical e sintática do *bahasa indonesia*;

- nomenclatura proposta por Hull e Eccles (2001) pelo fato do território falado atualmente corresponder em grande parte à província dos *Belos*.

Desta maneira, pode-se afirmar que há evidências linguísticas suficientes para a classificação das variedades proposta acima. Porém, há a necessidade de pesquisa futura para mapear de maneira exata a ocorrência das subvariedades da língua Tetun, assim como o levantamento de um amplo corpus linguístico de cada uma destas para a realização de um estudo exaustivo. Estas tarefas, todavia, além de serem atividades a ser realizadas, fogem do escopo desta dissertação.

5. O LÉXICO

O léxico da língua Tetun até a atualidade quase não foi estudado, em termos da lexicologia, enquanto a lexicografia apresenta avanços significativos com a publicação de diversos dicionários e vocabulários. Dois grandes dicionários do TP que se destacam são: Hull (2002b) e Costa (2000). Entendo aqui *léxico* como um dos níveis de análise linguística em oposição aos demais níveis: fonologia, morfologia e sintaxe. Desta forma, *lexicologia* compreende o estudo da estrutura dos vocábulos de uma língua, ou variedade, e os avanços dos estudos do léxico, a lexicologia, são organizados e codificados pela lexicografia. Como a lexicografia tetumófono antecipou-se aos estudos da lexicologia²⁶, nas obras citadas acima, o dicionário Tetun-Inglês (HULL, 2002b) e o dicionário Tetun-Português (COSTA, 2000), podem ser encontradas diversas falhas, como no dicionário de Hull (2002b), que possui uma estrutura do verbete simplificada, muitas vezes com apenas a tradução, o que caracteriza uma lista de palavra, ou vocabulário, e não um dicionário; na obra de Costa (2000), que possui uma estrutura lexicográfica notável, mas falha ao apresentar entradas das diferentes variedades da língua Tetun sem identificá-las.

Neste capítulo, para se descrever de maneira sucinta o léxico tetumófono, adotou-se o conceito de *camadas* (ing. *layers*) que consiste no procedimento de identificar o impacto de línguas vizinhas em *camadas* subsequentes e separá-las do *núcleo genético* (AIKHENVALD, 2007, p.6). Dessa maneira, identificamos o *núcleo genético* austronésio (5.1), a influência do Malaio (5.2) em momentos distintos e, provavelmente, de variedades distintas, e a camada lusófona (5.3) que também apresenta influência no Tetun Prasa em períodos históricos distintos via variedades distintas do português, como os CPMal e CPMac, como foi discutido no capítulo 2. Ainda, conforme se apresentou no capítulo anterior, a análise linguística em camadas também se aplica aos estudos de estratificação linguística, já que diferentes registros e variedades da língua Tetun possuem diferentes camadas, conforme foi analisado em 4.4: o registro católico possui maior número de itens da camada austronésia, o núcleo genético, enquanto o registro mediático possui mais itens da camada lusófona.

²⁶ Há somente em Hull (2005) um estudo sobre os empréstimos malaios em Tetun, além dos estudos de natureza histórico-comparativa do mesmo autor.

5.1 O núcleo Austronésio

No capítulo anterior, foram apontadas retenções do léxico tetumófono, conforme a lista do **Apêndice** – Cognatos PAN e Tetun Prasa, que apresenta 122 itens lexicais reconstruídos para o PAN comparados com a variedade Tetun Prasa. Nesta foi possível identificar 84 cognatos regulares do TP com o PAN. O campo semântico destes itens lexicais escolhidos foi diverso, mas todos eles procuraram seguir o vocabulário básico para línguas austronésias, conforme foi proposto por Greenhill, Blust e Gray (2008). Assim, há nomes relativos a partes do corpo, fauna, flora e qualidades; verbos processuais e estativos; preposições; interrogativos; numerais.

Os demais itens lexicais restantes, estes em número de 38, possuem origens distintas. Identificou-se que 8 cognatos são construções recentes de itens culturais que não faziam parte da cultura tetumófona. Cerca de 4 itens podem ser de origem papuásica, ou de outra língua pré-austronésia existente que foi assimilada por algum povo leste-timorense em tempos pré-históricos, como já foi apontada algumas influências lexicais de povos pré-austronésios nas línguas da região central de Timor-Leste, como o Mambae e o Tokodede, no capítulo 2. Os 26 itens restantes, não identificados inicialmente como cognatos do PAN, provavelmente são de origem austronésia, mas de difícil identificação pelo fato de ter ocorrido mudanças fonéticas irregulares, fato que é comum nas línguas Austronésias, como nos exemplos abaixo, em relação a mudanças fonéticas do Proto-Malaio-Polinésio (PMP) para o estágio atual da língua (HULL, 2000):

21. Mudança irregular: *j > r

Proto-Timórico		Tetun Prasa
*pija ‘quanto’	>	hira
*ŋajan ‘nome’	>	naran
*maja ‘seco’	>	maran

22. Mudança irregular: *ə > u, o

*ndəkaw ‘mover’	>	duka
-----------------	---	------

*kəhendaw ‘plantar’ > kuda

*mabəke ‘molhado’ > bokon

Do vocabulário básico extraímos alguns exemplos claros de cogação entre o PAN e o TP, e dois exemplos de mudanças fonéticas irregulares do PAN para o Tetun:

23. Cognatos PAN > TP:

*(qa)lima > lima ‘mão’; *qaqay > ain ‘perna, pé’; *Nakaw > lao ‘andar’; *zalan > dalan ‘caminho’; *qaNiC > kulit ‘pele’; *Sikan > ikan ‘peixe’; *quzaN > udan ‘chuva’; *bali > anin ‘vento’; *bulaN > fulan ‘lua’; *bituqen > fitun ‘estrela’; *i-aku > ha’u ‘1sg’; *i-kita > ita ‘1pl.inc’; *kami > ami ‘1pl.exc’;

24. Outras mudanças irregulares (PAN > Tetun):

PAN *likud > Tetun *kotuk* ‘costas’

PAN *ɲusu > Tetun *ibun* ‘boca’.

Os 8 itens restantes que foram identificados como construçōes recentes da língua podem ser facilmente argumentados pelo fato destes poderem ser analisados em unidades menores portadoras de significado. Dessa forma, trata-se de um processo de formaçāo de palavras como em:

25. Compostos tetumófonos:

PAN *i-jan ‘quando’	> TP <i>bainhira</i>	> <i>bain</i> ‘tempo’ + <i>hira</i> ‘contar’
PAN *danaw ‘lago’	> TP <i>bee-lihun</i>	> <i>bee</i> ‘água’ + <i>lihun</i> ‘parado’
PAN *biRaq ‘folha’	> TP <i>ai-tahan</i>	> <i>ai</i> ‘árvore’ + <i>tahan</i> ‘aguentar’
PAN *kuja ‘como’	> TP <i>oinsá</i>	> <i>oin</i> ‘rosto’ + <i>sá</i> ‘o que’
PAN *Cinaqi ‘intestino’	> TP <i>tee-oan</i>	> <i>tee</i> ‘fezes’ + <i>oan</i> ‘cria’

Assim, pode-se concluir que, mesmo sofrendo um intenso contato linguístico e reestruturação gramatical no decorrer de sua história, a língua Tetun, em sua variedade Tetun Prasa, com toda a sua variação diastrática e diatópica, apresenta seu núcleo básico de origem Austronésia quase intacto. Esse é outro argumento contra a classificação da variedade Tetun Prasa como uma língua crioula, ou um pidgin. As influências malaia, portuguesa e outras de origem desconhecidas são, no vocabulário básico, em números reduzidos: 5 itens do Malaio, 3 da língua portuguesa (apresentadas abaixo) e 4 de origem desconhecida (provavelmente papuásica).

26. Itens de origem lusófona no vocabulário básico TP:

PAN *ka > TP *no, ho, i* ‘e’

PAN *ka > TP *si, bainhira* ‘se’

PAN *qumah > TP *servisu, traballu* ‘trabalhar’

5.2 A influência do Malaio

Os primeiros registros do chamado *Pazar Melayu*, variedade pidginizada do Malaio, datam do século VII da era cristã e no século XV essas diferentes variedades eram usadas como língua de comércio por todo o sudeste asiático (VERSTEEGH, 2008, p. 175). As variedades do *Pazar Melayu* posteriormente tornaram-se L1 de vários povos do arquipélago indonésio devido a grande influência e circulação que essa variedade teve durante os séculos citados, entre as regiões de que se tornou língua materna podem ser citadas: Malaio de Kupang, Malaio de Makassar e Malaio de Ambon.

A língua Malaia é de filiação genética próxima à língua Tetun (tronco austronésio e Família Malaio-Polinésia), o que faz com que seja tarefa difícil a identificação dos elementos da língua Tetun de influência do Malaio introduzidos no período proto-histórico, quando a ilha de Timor fazia parte das rotas comerciais asiáticas, dos elementos austronésios comuns às duas línguas, ou seja, as retenções linguísticas do PAN. Em Hull (2005), o autor mapeou as diferentes influências do Malaio na língua Tetun. Nossa descrição difere-se parcialmente da efetuada pelo autor, assim como nossos dados também se diferem dos apresentados pelo linguista.

Na presente análise, optamos por classificar a camada de origem malaia em duas. A primeira baseia-se em empréstimos antigos via *Pazar Melayu*, já adaptados aos

padrões fonológicos do Tetun Prasa e de uso frequente em todos os grupos sociolinguísticos. A seguir encontram-se alguns exemplos do primeiro subgrupo da camada Malaia:

27. Camada Malaia:

surat ‘documento’

barlaque ‘dote’

toples ‘jarro tradicional para guardar comida’

katuas ‘velho, ancião, homem de idade avançada conhecedor das traduções e histórias antigas, tratamento afetuoso da esposa em relação ao marido’ no Malaio *katuas* ‘velho’

adat ‘conjunto de tradições de um povo’

jambua ‘toranja’

katupa ‘arroz condimentado agridoce cozido em folhas de palmeira’

sate ‘pedaços de carne apimentados servidos em palitos’

botir ‘garrafa’ do Inglês *bottle* via Malaio *botol*

nona ‘garota, amante de Europeu’

kantor ‘escritório’ do Holandês *kantoor* via Malaio, há variação com o composto tetumófono *servisu fatin*

O segundo subgrupo da camada Malaia consiste nos itens lexicais que podem ser datados de um período mais recente – o período de dominação indonésia – através da variedade *bahasa indonesia*. Esses itens lexicais cobrem campos semânticos específicos do período da invasão, como: militarismo, política e alimentação. Outro fator notável é que muitos desses itens estão em alternância de código, sendo realizados na fala de grupos sociolinguísticos específicos, como os falantes de pouca escolaridade e/ou da faixa etária de 20-40 anos de idade. A seguir encontra-se uma pequena lista de alguns itens e suas alternâncias, quando houver:

28. Empréstimos do *bahasa indonesia*:

pangkat ‘hierarquia militar’

tentara ‘soldado Indonésio’

rakitan ‘arma tradicional feita em casa’
bahasa ‘língua, língua Indonésia’ variando com o TP *lian*
setengah ‘metade’ variando com o TP *balu*
kurang ‘falta’ variando com o Pt. *falta* ‘falta, ausência’
guru ‘professor de escola’ variando com o Pt. *mestre*
penjaga ‘zelador’ variando com o Pt. *servente* e Ing. *cleaner*
kan ‘afinal’
rendang ‘prato indonésio feito com carne temperada com leite de côco e pimenta’
bakso ‘sopa indonésia feita com vegetais e bolas de carne’
padang ‘Indonesian style for storing food using traditional pre-cook technique’
nasi goreng ‘prato feito com arroz frito’
mie goreng ‘prato feito com macarrão frito’
warung ‘restaurante indonésio, qualquer restaurante de estilo oriental’ contrasta com a palavra portuguesa *restaurante* que é usada para se referir a restaurantes com comida ocidental.
pomal ‘objeto sagrado’ do Malaio *pemali* variando com o TP *plulik*
suangue ‘feiticeiro, pessoa que faz mal a outras’ do Malaio *suang* ‘mago mal’,
bapa ‘tratamento para homem indonésio’ para mulher usa-se *ibu*
baba ‘tratamento carinhoso em relação ao pai’
hekte ‘grampeador’ variando com o Pt. *agrafador*
arsip ‘arquivo’ variando com o Ing. *file*
isolasi ‘fita isolante’ variando com o Pt. *fita cola*
amplop ‘envelope’ variando com o Pt. e Ing. *envelope*
fotokopi ‘xerox’ variando com o Pt. *fotocopiar*, *fotocópia*, Ing. *photocopy* e TP *fotokopia*
merdeka ‘independência’ variando com o Pt. *independência*, Ing. *independence* e TP *independensia*
integrasi ‘integração’ variando com o Pt. *integração*, Ing. *integration* e TP *integrasaun*
pos ‘posto military ou policial’ variando com o Pt. *posto* e Ing. *post*
milisi ‘milícia’ variando com o Pt. *milícia* e Ing. *militia*

5.3 A camada lusófona

A influência mais marcante no léxico tetumófono é a de origem lusófona. Porém, esta não é tão fácil de ser mapeada, assim como de seu próprio estudo por vários fatores, entre eles: a falta de documentação, as diferentes variedades da língua portuguesa que conviviam na região e o número elevado de variação linguística e alternância de código existente.

Contudo, apresentamos uma proposta de categorização da influência da língua portuguesa e suas variedades no léxico tetumófona com base em evidências linguísticas e na documentação histórica existente. Identificamos cinco períodos distintos da história com a influência de variedades diferentes da língua portuguesa, de acordo com o esquema a seguir:

- 1º período (XVI-XVIII): primeira influência portuguesa que provavelmente não ocorreu diretamente do português europeu, mas do CPMal, que foi usado como língua franca. Não há fontes primárias desse período;
- 2º período (XVIII-XIX): influência marcada do CPMac e bem documentada através de fontes históricas e primeiras fontes linguísticas;
- 3º período (1898-1974): a fundação do colégio de Soibada marca a implantação de uma política linguística de ensino do PE no Timor Português, que é interrompida pela invasão indonésia;
- 4º período (1975-1999): o período da invasão indonésia reprimiu totalmente o uso da língua portuguesa, relegou o TP à posição secundária, e iniciou a política de ensino do *bahasa indonesia*. A língua portuguesa continuou a ser usada apenas nas regiões rurais e montanhosas pelos grupos resistentes à invasão;
- 5º período (2002-atual): após constituição de 2002 elegeu a língua portuguesa como língua oficial, juntamente com o TP, uma nova política linguística de ensino foi implantada em Timor-Leste. A influência maior do português é marcada na renovação e modernização do léxico tetumófono.

Inicialmente, é possível dividir períodos distintos de influências de variedades distintas do léxico lusófono sobre o léxico do TP. Provavelmente, a primeira variedade da língua portuguesa que influenciou mais o léxico tetumófono foi o CPMal, pois foi o mais utilizado no sudeste asiático, substituindo o *pazar melayu* e funcionando também como língua franca (BAXTER, 1988).

O período de influência do CPMal no TP foi cerca do século XVI a XVIII, quando em 1769 a capital mudou-se de Oecussi para Dili. Esta mudança da capital foi administrada pela coroa portuguesa através de Macau, o que causou um grande fluxo de funcionários macaenses, falantes de CPMac, para Dili. O segundo período de influência portuguesa sobre o léxico tetumófono ocorreu nos séculos XVIII e XIX através da variedade do CPMac. Diferente do período anterior, esse período de influência do CPMac possui uma documentação nos dicionários de Silva (1889) e Dorés (1907), que apresentam certas formas tetumófonas de clara origem do CPMac, umas que se mantêm até os dias atuais e outras que não são mais faladas.

O terceiro período é marcado pela fundação do Colégio de Soibada, em 1898. Esse período foi o início de uma política de ensino de língua portuguesa em Timor, que no início do século XX começou a apresentar resultados com um aumento da escolaridade, assim como no número de falantes de português (THOMAZ, 2002). Porém, isso foi interrompido com a invasão indonésia, que marca o início do quarto período apontado no esquema anterior. Durante esse período, a língua portuguesa foi proibida, o que acarretou um decréscimo no número de falantes, assim como marcou de maneira definitiva as gerações nascidas nesse período, pois elas foram educadas somente em *bahasa indonesia*.

O quinto período, que se estende até os dias de hoje, é marcado pela língua portuguesa como língua oficial e também pela política de reintrodução do português em Timor-Leste. Uma das influências linguísticas que podem ser citadas é a elaboração de gramáticas, dicionários e guias de vocabulários para o TP, tendo como base de repositório lexical a língua portuguesa.

A seguir são apresentados alguns itens lexicais, que, segundo nossa análise, foram inseridos no TP há um período mais antigo, pois sofreram mudanças fonéticas e/ou mudanças semânticas:

29. Camada lusófona:

amu ‘padre católico, ou outra autoridade católica’

moradór ‘membro da milícia nativa’

topazes ‘mestiço’

bazár ‘mercado’

estilu ‘cerimônia tradicional de sacrifício de animais’

kolega ‘palavra usada para se referir a amigos, ou pessoas que se tenha uma intimidade e que sejam da mesma idade’

argoliña ‘brinco’

maun ‘irmão mais velho’

mestri ‘professor de escola’

dosenti ‘professor universitário’

alunu /a ‘estudante do ensino fundamental ou médio’

estudanti ‘estudante universitário’

valór ‘resultados de exames escolares’.

karreta ‘carro’

Os itens lexicais da camada lusófona inseridos recentemente no TP cobrem campos semânticos específicos, principalmente terminologia tecno-científica e jurídico-administrativa, não apresentam nenhum tipo de mudança fonética ou semântica, assim como um grande número deles, cerca de 6000 palavras, foram listados na publicação elaborada pelo INL (2003). A seguir, são elencados alguns desses itens adaptados somente à ortografia padrão do TP, pois seus respectivos significados em português são evidentes:

30. Itens lexicais lusófonos inseridos recentes no TP:

ekolojía, biolojía, linguístika, istória, siénsia, sientífiku, administrasaun, eleisaun, votu, governu, prezidenti, deputadu, embaixadór²⁷.

²⁷ Os exemplos seguem a ortografia padronizada do Tetun Prasa (INL, 2002), conforme será discutido no capítulo seguinte.

6. FONOLOGIA

Neste capítulo apresentarei uma proposta de análise fonológica do Tetun Prasa. Começarei apresentando a fonologia segmental e os pares mínimos (6.1), depois farei a análise da estrutura silábica (6.2). Um estudo do acento será feito em (6.3). Com base na análise fonológica proposta neste capítulo, discutirei brevemente aspectos das ortografias padronizadas propostas atualmente (6.4).

6.1 Fonologia segmental

A língua Tetun originalmente possuía um inventário fonológico simplificado, quando comparado com línguas indo-europeias, línguas tupi, entre outras. Este é um traço tipológico das línguas austronésias e papuásicas da região onde se localiza a ilha de Timor e ilhas vizinhas. Após uma grande influência originária do contato linguístico, principalmente com as variedades do malaio e da língua portuguesa, o Tetun Prasa adquiriu vários segmentos através de empréstimos que foram incorporados, outros segmentos limitam-se a empréstimos eruditos, ou pouco usados, que não foram incorporados ao TP, e, ainda, alguns registros e variedades da língua apresentam também alguns fonemas nativos que se perderam.

A fonologia segmental do TP será analisada separadamente, em 6.1.1, serão apresentados os segmentos consonantais do TP e serão discutidos problemas relacionados a alofonia, variação livre, diferentes usos e variedades da língua Tetun; em 6.1.2 os segmentos vocálicos serão analisados, juntamente com os problemas da altura e duração dos mesmos. Em seguida, os segmentos presentes em empréstimos eruditos e/ou pouco usados, as consoantes palatais e as vogais nasais ambas de origem portuguesa, serão analisadas separadamente em 6.1.3 e 6.1.4 respectivamente.

6.1.1 Os fonemas consonantais

O TP mesmo acrescentando uma série de fonemas de palavras emprestadas possui um quadro fonológico relativamente simples com apenas 15 segmentos consonantais, sendo realizados em 5 pontos de articulações diferentes: bilabiais, labiodentais, alveolares, velares e glotal, e também possuem 5 modos de articulação distintos, a saber: oclusivos, fricativos, vibrante, lateral e nasais, conforme tabela abaixo:

TABELA 2. Segmentos consonantais do Tetun Prasa

	bilabial	labiodental	alveolar	velar	glotal
oclusiva	p		t	k	
	b		d	g	
fricativa		f	s		h
		v	z		
vibrante			r		
lateral			l		
nasal	m		n		

6.1.1.1 Oclusivas

/p/ é consoante oclusiva bilabial surda; em alguns casos, é realizada aspirada [p^h] por influência da língua materna dos falantes:

/ka.'pa.as/ ~ [ka.'pa:s] ‘bonito’

/pro.'ble.ma/ ~ [pro.'ble.ma] ‘problema’

/'gru.pu/ ‘grupo’

/b/ apresenta o alofone fricativo bilabial [β] quando em posição entre vogais e, alguns casos, o alofone fricativo labiodental [w]:

/'bo.ot/ ~ [bot] ~ [bo:t] ‘grande’

/'be.le/ ~ ['be.le] ‘poder’

/'a.ban/ ~ ['a.βan] ‘amanhã’

/'ba.ba/ ~ ['ba. βa] ~ ['ba.va] ‘tipo de tambor cerimonial’

/t/ é consoante oclusiva alveolar surda, e, como as demais oclusivas surdas, possui uma realização aspirada em variação livre [t^h] e em posição de coda é realizada [t̚]:

/ˈma.tan/ ~ [ˈma.tã̃n] ‘olho’

/ˈfe.to/ ~ [ˈfe.tʰo] ‘mulher’

/ˈku.lit/ ~ [ˈkʰu.litʰ] ‘pele’

/ˈha.at/ ~ [hatʰ] ~ [ha:tʰ] ‘quatro’

/d/ é um dos segmentos consonantais mais problemáticos para análise. Apresenta a realização pós-alveolar na maioria dos dados [d̠], porém é realizado retroflexo [d̠] e alveolar [d] em alguns casos. Apresenta como alofone a vibrante [r], quando se encontra entre vogais:

/i.da/ ~ [i.ra] ‘um’

/di.ak/ ~ [di.akʰ] ‘bom, bem’

/u.dan/ ~ [u.ran] ‘chuva’

/dau.ˈda.uk/ ~ [d̠au.ˈd̠a.ukʰ] ‘ainda’

/da.lan/ ~ [d̠a.lã̃n] ~ [d̠a.lan] ‘estrada, caminho’

/mo.dok/ ~ [mɔ.d̠ɔkʰ] ~ [mɔ.rɔkʰ] ‘legumes, hortaliças’

/k/ como as demais oclusivas surdas apresenta a variante aspirada [kʰ]. Quando em posição de ataque pode ser realizado como oclusão glotal [ʔ] antes de outra consoante e em posição de coda como [kʰ]:

/ka/ ~ [ka] ~ [kʰa] ‘ou’

/ko.ta/ ~ [kɔ.ta] ~ [kʰɔ.ta] ‘bairro’

/ko.len/ ~ [kɔ.len] ~ [kʰɔ.len] ‘cansado’

/ki.ik/ ~ [kikʰ] ~ [ki:kʰ] ~ [ki.ikʰ] ‘pequeno’

/ka.rik/ ‘talvez’

/ta.ka/ ‘fechar’

/ba.rak/ ~ [ba.rakʰ] ‘muito’

/kbi.it/ ~ [ʔbi:tʰ] ‘força’

/g/ é consoante oclusiva velar sonora, não apresentando problemas para a análise:

/ˈgos.ta/ ~ [ˈgɔs.ta] ‘gostar’
/ˌga.do. ˈga.do/ ‘prato indonésio’
/ˈzo.ga/ ~ [ˈzɔ.ga] ~ [ˈzʲɔ.ga] ‘jogar’
/ˈgu.dan/ ~ [ˈgu.dan] ‘armazém’
/ˈgru.pu/ ‘grupo’

6.1.1.2 Fricativas

/f/ é fricativa labiodental surda e por influência da língua materna de alguns falantes pode ser realizada como a oclusiva bilabial surda aspirada [p^h], geralmente em início de palavra:

/ˈfo.o/ ~ [fɔ] ‘dar’
/ˈfa.li/ ‘novamente’
/ˈfi.la/ ~ [ˈp^hi.la] ‘voltar’
/ˈfi.ni/ ~ [ˈp^hi.ni] ‘semente’

/v/ é fricativa labiodental sonora em variação livre com a oclusiva bilabial sonora [b]:

/vi.ˈzi.nu/ ~ [bi.ˈzi.n̥u] ‘vizinho’
/ser.ˈvi.su/ ~ [ser.ˈvi.ʃu] ~ [ser.ˈbi.su] ‘trabalho’
/vi.ˈzi.ta/ ~ [bi.ˈzi.ta] ~ [bi.ˈzʲi.ta] ‘visitar’

/s/ é a fricativa alveolar surda e apresenta os alofones fricativo palatal [ʃ] por influência do português e o fricativo surdo mais retraído pós-alveolar [s̠] em final de palavra ou em posição de coda, assim como a oclusiva alveolar sonora /d/, no léxico tetumófono. Ainda, há casos de palatalização do /s/ quando próximo da vogal /i/ ou em empréstimos lusófonos, o que faz com que alguns falantes realizem a fricativa palatal [ʃ] ou fricativa alveolar palatalizada [sʲ]:

/ˈa.su/ ‘cão’
/ser.ˈvi.su/ ~ [ser.ˈvi.ʃu] ~ [ser.ˈbi.su] ‘trabalho’

/¹se.i/ ‘ainda’

/¹si.ra/ ~ [¹si.ra] ‘3 pl.’

/ka.¹pa.as/ ‘bonito’

/¹lo.os/ ~ [los] ~ [¹lo:s] ‘muito (intensidade)’

/ha.¹ta.is/ ~ [ha.¹ta.iʃ] ‘vestir-se’

/sa/ ~ [s¹a] ~ [ʃa] ‘chá’

/z/, a fricativa alveolar sonora, é realizada mais retraída, pós-alveolar, [z] em posição de coda, em final de palavra, e antes de segmentos [+soantes]. Ainda, nessas mesmas posições mencionadas anteriormente, os falantes mais escolarizados realizam a fricativa palatal [ʒ] por influência da língua portuguesa, e em empréstimos pode ser realizada palatalizada [zʲ] ou como a africada [dʒ] por influência do *bahasa indonesia*:

/¹u.za/ ~ [¹u.zʲa] ‘usar’

/¹me.za/ ‘meza’

/¹mu.zi.ka/ ~ [mu.¹zi.ka] ‘música’

/vi.¹zi.nu/ ~ [bi.¹zʲi.nu] ‘vizinho’

/a.¹zu.da/ ~ [a.¹dʒu.da] ‘ajudar’

/₁li.is.¹a.su/ ~ [₁li.iʒ.¹a.su] ~ [₁li.iʒ.¹a.su] ‘alho’

/h/ a fricativa glotal realiza-se como tal em início de palavra e de sílaba, porém em posição intervocálica apresenta o alofone [h̠] ou sofre deleção:

/ha.¹ka.rak/ ‘querer’

/hak.¹fo.dak/ ‘assustar-se’

/¹ti.ha/ ~ [¹ti.h̠a] ~ [¹ti.a] ‘já’

/¹o.hin/ ~ [¹o. h̠in] ‘hoje’

6.1.1.3 Vibrante

/r/ é a vibrante simples, que pode ser realizada como a vibrante múltipla [r] por influência da língua materna do falante, ou como [x] ou [ɣ] por causa da hipercorreção baseada na língua portuguesa e na ortografia padronizada do Tetun, que optou por

representar o fonema /r/ como *rr* dos empréstimos lusófonos, abaixo em itálico estão os exemplos da ortografia padronizada:

/*ra.i*/ ‘terra’

/*ru.in*/ ~ [*ru.ĩn*] ‘osso’

/*la.ran*/ ~ [*la.rã̃n*] ‘dentro’

/ka.*re.ta*/ ~ [ka.*xɛ.ta*] ‘carro’ > *karreta*

/bo.*ra.sa*/ ~ [bo.*xa.sja*] ~ [bu.*xa.sa*] ~ [bo.*ra.ʃa*] ‘borracha’ > *borraxa*

/*far.da*/ ~ [*fay.da*] ~ [*fay.dɐ*] ‘uniforme’

/*i.kan.sar.di.na*/ ~ [i.kan.say.*di.nja*] ~ [i.kan.say.*di.n a*] ‘sardinha’

6.1.1.4 Lateral

/l/ a lateral alveolar pode ser realizada velar [ɫ] em final de sílaba ou final de palavra por influência da língua portuguesa. Ainda, em empréstimos lusófonos com a presença da consoante lateral palatal realiza-se como: a palatal [ʎ] somente pelos falantes altamente escolarizados e/ou idosos; a lateral alveolar [l] por falantes não-escolarizados das zonas rurais, ou como lateral alveolar palatalizada por falantes não-escolarizados das zonas urbanas [lʝ]:

/la.*ko.hi*/ ~ [la.*ko.fii*] ~ [la.*ko.i*] ‘não.querer’

/la.*le.han*/ ~ [la.*le.fian*] ~ [la.*le.ã̃n*] ‘céu’

/na.*na.al*/ ~ [na.*na.ɐɫ*] ~ [na.*na.aʎ*] ‘língua’

/kon.*se.lu*/ ~ [kon.*se.lju*] ~ [kõ.*se.ʎu*] ‘conselho’

/*pi.la*/ ~ [*pi.lja*] ~ [*pi.ʎɐ*] ‘pilha’

/a.pa.*re.lu*/ ~ [a.pa.*re.lju*] ‘aparelho’

6.1.1.5 Nasais

/m/ é a nasal bilabial. Não apresenta problemas para análise, somente os fenômenos de pré e pós-nasalização, típico de falantes mais escolarizados e de influência lusófona:

/*ma.tan*/ ~ [*ma.tã̃n*] ‘olho’

/*u.ma*/ ‘casa’

/mak/ ~ [mak] ~ [maʔ] ~ [makʔ] ‘TOP’

/^lma.i/ ‘vir’

/^le.ma/ ~ [^lɛ.ma] ‘pessoa’

/ha. ^lmu.tuk/ ‘juntos’

/^lma.ne/ ~ [^lmã.nɛ] ‘homem’

/^la.man/ ~ [^lã.mãn] ~ [^la.mãn] ‘pai’

/n/ apresenta variação em posição de coda ou de final de palavra, sendo realizada como nasal alveolar [n], ou nasal velar [ŋ] por influência do *bahasa indonesia* ou em empréstimos de origem malaia. Ainda, é realizada como a palatal [ɲ] por falantes altamente escolarizados nos empréstimos de origem lusófona. Os falantes que possuem maior conhecimento da língua portuguesa também fazem a assimilação do traço nasal para a vogal:

/^lni.a/ ~ [^lnⁱ.a] ‘3sg’

/^lo.na/ ~ [^lɔ.na] ‘já’

/^lma.ne/ ~ [^lmã.nɛ] ‘homem’

/^lgu.dan/ ~ [^lgu.daŋ] ~ [^lgu.dãn] ‘armazém’

/^la.man/ ~ [^lã.mãn] ~ [^la.mãn] ‘pai’

/^li.nan/ [^li.nãn] ~ [^li.nãn] ‘mãe’

/^lba.nu/ ~ [^lba.n^u] ~ [^lbã.n^u] ~ [^lba.ɲu] ‘banho’

/se.^lno.or/ ~ [sɛ.^lnɔ:r] ~ [sɛ.^lnɔr] ~ [sɛ.^ln^or] ~ [sɛ.^lɲɔr] ‘senhor’

6.1.2 Os segmentos vocálicos

O inventário dos segmentos vocálicos TP consiste nas cinco vogais cardinais, de acordo com a tabela abaixo, que, de acordo com a posição silábica podem apresentar variação, como será descrito posteriormente. Ainda, vale lembrar que o fenômeno de lenição vocálica em sílabas não acentuadas não é comum nas línguas de Timor-Leste, sendo encontrado somente na fala de informantes que apresentam grande influência lusófona.

TABELA 3. Segmentos vocálicos do Tetun Prasa

	anterior	central	posterior
alto	i		u
médio	e		o
baixo		a	

6.1.2.1 Vogais anteriores

/i/ é vogal anterior alta não arredondada e não apresenta problemas para a análise:

/^hlo.ri/ ‘trazer’

/^hti.ha/ ~ [^hti.fia] ~ [^hti.a] ‘já’

/kin.^hta.al/ ~ [kĩn.^htaʔ] ‘quintal, varanda’

/^hfi.ni/ ~ [^hp^hi.ni] ‘semear’

/^hi.da.^hi.dak/ ~ [^hi.ra.^hi.rak^h] ‘um por um’

/e/ é vogal anterior média e apresenta uma grande variação nos dados entre a média aberta [ɛ] e a média fechada [e] com a predominância desta, enquanto aquela está mais ligada com a tonicidade e a altura das vogais próximas. Ainda, em certos empréstimos malaios e lusófonos pode ser realizada como o *schwa* [ə]:

/se.e/ ‘quem’

/ne.e/ ‘este (a), isto’

/e.ma/ ~ [ɛ.ma] ‘pessoa’

/go.^hver.nu/ ~ [go.^hber.nu] ‘governo’

/per.^hgun.ta/ ~ [pəɾ.^hgun.tə] ~ [pəɾ.^hgũn.tə] ‘perguntar’

/ge.^hla.as/ ~ [gə.^hlas] ~ [glas] ‘vidro, copo’

6.1.2.2 Vogal central

/a/ é a vogal central baixa e apresenta somente o alofone /ɐ/ quando em sílaba não acentuada pronunciada por falantes altamente escolarizados:

/ˈsa.a/ ‘o que’

/ˈa.mi/ ‘1pl.excl’

/ˈu.ma/ ‘casa’

/ˈpi.la/ ~ [ˈpi.l̩a] ~ [ˈpi.ɫɐ] ‘pilha’

/ˈfar.da/ ~ [ˈfaɾ.da] ~ [ˈfaɾ.dɐ] ‘uniforme’

/na.ˈna.al/ ~ [na.ˈna.ɐɫ] ~ [na.ˈna.aɫ] ‘língua’

6.1.2.3 Vogais posteriores

/u/, a vogal posterior alta, apresenta uma grande variação, sendo realizada como [ʊ] em sílabas não acentuadas e próxima da vogal alta /i/ e da baixa /a/. Ainda, pode ser realizada também como [ɯ] quando não acentuadas e próxima das vogais médias /e/ e /o/:

/ˈha.u/ ~ [haʊ] ‘1.sg’

/ˈa.su/ ‘cão’

/ˈfa.ru/ ~ [ˈfa.rʊ] ‘camisa’

/ˈto.hu/ ~ [ˈto.fɯ] ‘cana-de-açúcar’

/ˈe.tu/ ~ [ˈe.tɯ] ‘arroz cozido’

/o/ a vogal posterior média arredondada também apresenta variação com a predominância do fone [ɔ], porém os falantes tendem a realizar [o] quando em sílabas não acentuadas e as vogais próximas serem as vogais altas /i/ e /u/²⁸. Ainda, em lexemas trissílabos ou polissílabos, na maioria dos casos empréstimos, /o/ é realizado como [u], [ʊ], ou [ə] nas sílabas iniciais:

/o/ ~ [ɔ] ~ [ɔ:] ‘2.sg’

/ˈfo.o/ ~ [fɔ] ‘dar’

/ˈfo.ho/ ‘montanha’

²⁸ No caso das vogais médias /e/ e /o/ há indícios de que suas realizações distintas, conforme foram descritas aqui, estejam ligadas com o processo de harmonia vocálica, que necessita ser estudado separadamente.

/ko.tuk/ ~ [k^ho.tuk^h] ‘costas’

/so.ko.^hla.ti/ ~ [su.ko.^hla.ti] ~ [s^ho.ko.^hla.ti] ~ [ʃo.kə.^hla.ti] ‘chocolate’

/a.e.ro.^hpor.tu/ ~ [a.e.ru.^hpor.tu] ~ [a.e.ru.^hpox.tu] ‘aeroporto’

/ko.^hne.se/ ~ [ko.^hn^hɛ.se] ~ [ku.^hnɛ.se] ‘conhecer’

6.1.3 Fonemas incorporados ao TP

Há indícios sincrônicos e diacrônicos de que vários fonemas e fones mencionados não faziam, e não fazem, parte da fonologia do Tetun Prasa. Os indícios sincrônicos consistem na comparação das variedades da língua Tetun e da análise fonológica da variedade Tetun Prasa. A comparação do Tetun Prasa com a variedade Tetun Terik fornece exemplos de que o Tetun Terik não possui diversos segmentos consonantais presentes no Tetun Prasa.

Os estudos diacrônicos fornecem evidências de que o inventário fonológico do Proto-Timórico (proto língua que deu origem às línguas de origem austronesicas faladas em Timor-Leste) é simplificado, constando de 10 fonemas consonantais e 5 vocálicos, conforme estudo feito por Hull (2002b), aplicando o método histórico-comparativo às línguas austronésias faladas em Timor-Leste.

Argumenta-se aqui que alguns dos sons de línguas como o Malaio e o português, devido ao contato, foram incorporados à fonologia do TP, sendo produtivos e aparecendo em alguns vocábulos nativos, como é o caso de /p, v, z, g/. Será argumentado posteriormente também que, enquanto os fonemas citados foram incorporados ao TP, outros fones, no entanto, limitam-se a empréstimos e não são realizados pelos falantes tetumófonos, como é o caso das consoantes palatais da língua portuguesa e a nasal velar do malaio.

31. /p/

/ka.^hpa.as/ ‘bonito’

/^hsi.pu/ ‘concha’

/da.^hpu.ur/ ‘cozinha’

/^hro.pa/ ‘roupa’

32. /v/

/ser.'vi.su/ 'trabalho'

/vi.'zi.ta/ 'visitar'

/vi.'zi.nu/ 'vizinho'

33. /z/

/a.'zu.da/ 'ajudar'

/'u.za/ 'usar'

/'me.za/ 'mesa'

/'zo.ga/ 'jogar'

34. /g/

/'gu.dan/ 'armazém'

/'gru.pu/ 'grupo'

/'gos.ta/ 'gostar'

/'ga.do.'ga.do/ 'prato indonésio'

De acordo com os exemplos acima, os fonemas /p/ e /g/ já foram incorporados de maneira efetiva pelos falantes de TP, apresentando pouca, ou nenhuma alofonia, tanto em empréstimos de origem malaia, quanto de origem lusófona. Diferentemente, os fonemas /v/ e /z/ apesar de fazerem parte do quadro de fonemas do TP encontram-se limitados a empréstimos lusófonos recentes. Ainda, o fonema /z/ apresenta um alofone palatalizado que será analisado adiante (palatalização), e o fonema /v/ é o caso mais problemático que se encontra em variação livre, principalmente pelas diferentes tentativas de realização por parte dos falantes.

Os fonemas consonantais nativos que se perderam no TP, mas que se mantiveram nas demais variedades são: /ʔ/ e /w/. Ainda, falantes nativos das variedades rurais do Tetun o pronunciam em palavras que são comuns às duas variedades da língua, assim como ora realizam o fonema /b/ como [w]. Encontrou-se também nos dados linguísticos que em registros especiais de uso, como na igreja, na fala de alguns locutores de rádio e em discursos de políticos os falantes tetumófonos tendem a realizar a oclusão glotal [ʔ], pois a este segmento atribuem um valor de alto prestígio.

35. Ø < [ʔ]

/ha.u/ < ['ha.ʔu] '1sg'

/ko.a/ < ['kɔ.ʔa] 'cortar'

/di.ak/ < ['di.ʔak] 'bom'

36. /b/ < [w]

/ka.ben/ < ['ka.wen] 'casar'

/la.'ba.rik/ < [la.'wa.rik] 'criança'

/ba.in.'hi.ra/ < [wai.'ni.ra] 'quando, se'

37. Leitura da bíblia em missa na Catedral de Dili (Marcos: 28, 11-16)

[wai.ni.ra.si.ra la.ʔɔ.tu.ir.da.lan] (...)

Quando eles andar seguir caminho ...

'Quando eles seguirem seus caminhos'

[iʃ.ko.lã.te.sa.nu.lu.res.ni.da.ne.ʔe.la.ʔɔ.ba.ga.li.le.ja]

Discípulo dez CL DEI andar para Galiléia

'Os onze discípulos foram à Galiléia'

6.1.4 A palatalização

As consoantes palatais [ʃ, ʒ, ɲ, ʎ] foram emprestadas da língua portuguesa e não foram incorporadas à língua Tetun, já que permanecem limitadas somente aos itens lexicais de origem lusófona. Outro fator digno de nota é que esses fonemas apresentam diversos alofones e são articulados de acordo com a norma europeia somente pela parcela da população altamente escolarizada.

38. [ʃ]

xá /sa/ ~ [sʲa] ~ [ʃa]

xapeu /sa.'pe.u/ ~ ['sa.peu] ~ ['sa.peu] ~ [sʲa.'peu] ~ [sʲa.'pe.u] ~ [ʃa.'peu]

xave ~ ['sʲa.vɛ] ~ ['sa.bi] ~ ['sa.βi] ~ [ʃa.ve]

39. [ʒ]

janela /za.'ne.la/ ~ [za.'nɛ.la] ~ [zʲa.'ne.la] ~ [dʒa.'nɛ.la]

justisa /zus.'ti.sa/ ~ [zʲus.'ti.sa] ~ [ʒus.'ti.sa]

jornal /zor.'na.al/ ~ [ʲzɔr.nal] ~ [zʲɔr.'na.aʎ] ~ [ʒɔr.'naʎ]

40. [ʎ]

evanjellu /e.van.'ze.lu/ ~ [e.ban.'zɛ.lu] ~ [e.van.'zʲɛ.lʲu] ~ [e.vãn.'ʒɛ.ʎu]

jullu /'zu.lu/ ~ [ʲzʲu.lʲu] ~ [ʲzʲu.lu] ~ [ʲʒu.ʎu]

relijiaun /re.li.zi.'a.un/ ~ [re.li.zʲi.'a.un] ~ [re.li.zʲi.'aʊn] ~ [re.li.ʒi.'ãʊn]

41. [ɲ]

bañu /'ba.nu/ ~ ['ba.nʲu] ~ ['bã.nʲu] ~ ['ba.ɲu]

señór /se.'no.ɔr/ ~ [sɛ.'nɔ:r] ~ [se.'nɔr] ~ [sɛ.'nʲɔr] ~ [sɛ.'ɲɔr]

viziñu /vi.'zi.nu/ ~ [bi.'zi.nʲu]

De acordo com os dados apresentados acima, optou-se por considerar essas consoantes palatais apenas como alofones de suas respectivas consoantes alveolares, conforme foi analisado anteriormente. Ainda, com base na coleta de dados verificou-se que a realização das consoantes palatais está ligada a fatores sociolinguísticos como a idade [+ idoso] e o grau de escolaridade [+ escolarizado].

Dessa maneira, o fenômeno da palatalização encontrado nos dados apresentados acima está também ligado a fatores sociolinguísticos, pois se trata de uma tentativa dos

falantes mais jovens e menos escolarizados de aproximar suas realizações linguísticas aos falantes mais escolarizados, ou seja, buscando maior prestígio social.

6.1.5 A nasalização

O fenômeno da nasalização encontrado nos dados possui a motivação sociolinguística igual ao fenômeno da palatalização, sendo realizado pelos falantes com maior influência lusófona, ou seja, os falantes mais idosos e mais escolarizados. Porém há a motivação estrutural, que consiste na assimilação do traço [+ nasal] das consoantes nasais /m/ e /n/ pela vogal próxima.

42. Exemplos de pré-nasalização:

/ru.in/ ~ [ru.ĩn] ‘osso’

/la.ran/ ~ [la.rã̃n] ‘dentro, parte interna’

/kin.'ta.al/ ~ [kin.'ta.al] ~ [kĩn.'taʎ] ‘quintal, varanda’

/ma.ne/ ~ [mã̃.nɛ] ‘homem’

/gu.dan/ ~ [gu.daŋ] ~ [gu.dã̃n] ‘armazém’

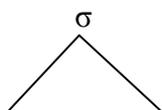
/a.man/ ~ [ã̃.mã̃n] ~ [a.mã̃n] ‘pai’

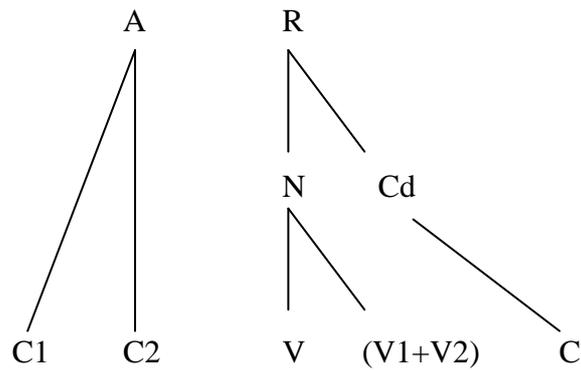
/i.nan/ [ĩ̃.nã̃n] ~ [i.nã̃n] ‘mãe’

6.2 A sílaba

A estrutura silábica do Tetun Prasa (σ) é baseada na sonoridade do segmento, ou seja, seu núcleo (N) consiste em uma vogal. Ainda, fazem parte da estrutura silábica segmentos de sonoridade ascendente, que fazem parte do Ataque (A), e segmentos de sonoridade descendente, que fazem parte da coda (Cd). A Coda (Cd) junta-se ao núcleo (N) na Rima (R) e esta, por sua vez, une-se ao Ataque (A) para constituir-se na sílaba tetumófona (σ).

O Ataque (A) pode ser composto somente por consoantes e no máximo duas, e a Coda (Cd) pode ser composta por consoante ou vogal. Dessa forma, a sílaba mais complexa do Tetun Prasa organiza-se da seguinte maneira:





O padrão silábico universal CV é o mais frequente, mas também há sílabas V, CVC, CCV, VC e VVC. Outro fato digno de nota na estrutura silábica do Tetun Prasa que a sílaba V é pouco frequente.

V

/o/ '2sg'

/'u.ma/ 'casa'

/'a.su/ 'cachorro'

CV

/ha.'te.ne/ 'saber'

/'si.ra/ '3pl'

/'ro.na/ 'ouvir'

CVC

/'mo.ras/ 'doença, doente'

/ham.'ro.ok/ 'sede'

/'ma.sin/ 'sal'

VC

/'o.in/ 'rosto'

/'o.an/ 'filho, cria'

/ha.'ta.is/ 'vestir-se'

CCV

/**kla**.ran/ ‘meio’

/**kma**.nek/ ‘maravilhoso’

/**kbi**.id/ ‘poder’

A fonotática tetumófona, conforme pode ser visto nos exemplos anteriores, apresenta várias restrições para as sílabas, como será apresentado a seguir. Na sílaba CCV, a C1 se restringe à série de oclusivas, como em /**kla**.ran/ ‘meio’, /**kma**.nek/ ‘maravilhoso’, /**pri**.mo/ ‘primo’, enquanto a C2 a possibilidade de ser preenchida pela maioria das consoantes, porém há uma tendência maior de aparecer as líquidas e nasais. Na sílaba VC, a vogal pode ser qualquer um dos cinco segmentos vocálicos, já a C se restringe somente aos segmentos /k, t, d, s, n, l, r/, seguem alguns exemplos: /**o**.in/ ‘rosto’, /**o**.an/ ‘filho, cria’ e /ha.'ta.**is**/ ‘vestir-se’. De maneira distinta, a sílaba CVC aceita uma série de segmentos consonantais tanto para C1 quanto para C2. /**mo**.ras/ ‘doença, doente’, /**ham**.ro.ok/ ‘sede’ e /**ma**.sin/ ‘sal’.

6.3 O acento

O acento em Tetun Prasa caracteriza-se como acento fixo, sendo previsível, de natureza não fonológica e caindo na penúltima sílaba. O acento do TP em palavras monossilábicas recai na única sílaba existente, porém existem alguns casos de monossílabos não acentuados. Seguem os exemplos:

43. Acento em monossílabos:

/o/ ‘2sg’

/ba/ ‘ir’

Seguindo a proposta de Hayes (2008) ao definir as regras acentuais do polonês, é possível aplicá-las ao TP, que possui um padrão acentual muito próximo ao dessa língua. Dessa maneira, a regra acentual completa do TP pode ser representada da seguinte maneira:

44. Regra do acento:

$\sigma \rightarrow$ [+acento] / ____ (σ)] palavra

O esquema acima, adaptado de Hayes (2008, p.274), pode ser lido como: atribua o acento à penúltima sílaba da palavra; quando esta tiver somente uma sílaba, atribua acento a esta última. No entanto, essa regra acentual do TP não é aplicada quando se trata de um grande número de empréstimos recentes de origem lusófona e de alguns compostos envolvendo elementos gramaticais.

O TP, como foi analisado no capítulo anterior, apresenta um grande número de vocabulário lusófono emprestado em relação a diversos itens culturais do mundo moderno que não faziam parte da cultura tetumófona. Os falantes tetumófonos escolarizados ou com algum conhecimento da língua portuguesa acabam por realizar esses vários empréstimos de acordo com seu grau de conhecimento da língua portuguesa:

45. Acento de empréstimos lusófonos (falantes escolarizados):

animál ‘animal’	[a.ni.'mal]
diretór ‘diretor’	[di.re.'tɔɾ]
fotokópia ‘xerox, cópia’	[fɔ.tɔ.'kɔ.pi.a]
kómiku ‘engraçado, cômico’	['kɔ.mi.ko]
ipóteze ‘hipótese’	[i.'pɔ.te.ze]

Contudo, a parcela da população não escolarizada, que sofreu pouca influência ou tem pouco conhecimento da língua portuguesa, adapta esses empréstimos, assim como os demais, à estrutura silábica e acentual do TP, conforme os exemplos abaixo:

46. Acento de empréstimos lusófonos (falantes escolarizados):

animál ‘animal’	[a.ni.'ma.al] ~ [a.'ni.mal]
diretór ‘diretor’	[di.re.'tɔ.ɔɾ]

fotokópia ‘xerox, cópia’ [fɔ.tɔ.'kɔ.pi]
 kómiku ‘engraçado, cômico’ [‘ko.mik]
 ipóteze ‘hipótese’ [i.po.'tɛ.ze]

O TP possui uma série de compostos gramaticais que se cristalizaram: há alguns que não podem ser usados separados e outros que perderam seu conteúdo semântico (tabela 4). Esses compostos gramaticais comportam-se de maneira distinta, aceitando duas interpretações fonológicas para o padrão acentual. Primeira, o primeiro elemento do composto pode ser considerado como um clítico que se apóia fonologicamente no elemento que o segue. Segunda, o elemento final, onde cai o acento, pode ser analisado como um hiato com duas vogais iguais, mas em sílabas separadas.

TABELA 4. Compostos gramaticais do TP

Composto ²⁹	Glossa	Significado
<i>ne'ebé</i>	<i>nee</i> ‘este(a), isto’ + <i>bee</i> ‘relativizador’	‘que (relativo)’
<i>ne'ebá</i>	<i>nee</i> ‘este(a), isto’ + <i>baa</i> ‘marcador de direção centrífuga, verbo <i>ir</i> ’	‘aquele(a), aquilo’,
<i>liubá</i>	<i>liu</i> ‘mas que (comparativo), verbo <i>exceder</i> ’ + <i>baa</i> ‘marcador de direção centrífuga, verbo <i>ir</i> ’	‘referência a tempo passado’ ex.: <i>fulan liubaa</i> ‘mês passado’
<i>kona-bá</i>	<i>kona</i> ‘verbo <i>encostar</i> ’ + <i>baa</i> ‘marcador de direção centrífuga, verbo <i>ir</i> ’	‘sobre, a respeito de’
<i>maibé</i>	<i>mai</i> ‘empréstimo lusófono <i>mas</i> (?)’ + <i>bee</i> ‘relativizador’	‘mas, porém’
<i>tanbasá</i>	<i>tan</i> ‘porque, mais’ + <i>ba</i> ‘para (preposição)’ + <i>saa</i> ‘qual, o que (interrogativo)’	‘por que (interrogativo)’

²⁹ Nesta tabela (Tab. 4), os compostos listados estão de acordo com a ortografia padronizada do Tetun Prasa. Somente as glosas estão em escrita fonologizada.

<i>oinsá</i>	<i>oin</i> ‘rosto’ + <i>saa</i> ‘qual, o que ‘como (interrogativo)’ (interrogativo)’
--------------	---

Optou-se pela segunda interpretação pelo fato de esta apresentar mais evidências linguísticas, já que o fenômeno de vogais iguais em sílabas manifesta-se por todo o léxico nativo tetumófono e na realidade consiste na queda das consoantes uvulares do proto-austronésio, e encontra-se nos exemplos abaixo (HULL, 2002a, p. 3):

47. Formação de vogais geminadas:

foos [ʼfɔ.ɔs] ‘arroz’ < * fɔas < *feas < *beRas

hamriik [ham.ʼri.ik] ‘ficar de pé’ < *hamrihi < *pamadiRi

kaan [ʼka.an] ‘cravar’ < *kahan < *kahen < *kaʼen

Ainda, a primeira interpretação não aparenta ser válida pelo fato de certos elementos não serem clíticos, como o *nee* /ʼne.e/ ‘isto, este, esta’ que é o pronome demonstrativo ‘este(a), isto’ e também um dêitico anafórico (DEI):

48. sa-ida mak nee?
 o.que-IND TOP isto
 ‘o que é isto?’

49. nee nee hau=nia faru
 isto ANA 1sg=POS roupa
 ‘Isto é minha roupa!’

Assim, os compostos acima podem ser interpretados fonologicamente da maneira que segue:

50. Intepretação fonológica de compostos:

/ne.ʼbe.e/ ‘que’

/ne.ʼba.a/ ‘aquele’

/liu.'ba.a/ 'antigamente'

/ko.na.'ba.a/ 'sobre, a respeito'

/mai.'be.e/ 'mas'

/tan.ba.'sa.a/ 'por que'

/o.in.'sa.a/ 'como'

Logo, a segunda interpretação, tratar os compostos gramaticais com o elemento final possuidor de vogais geminadas em sílabas diferentes, não fere o padrão acentual do TP, conforme foi apresentado. Assim, o acento do TP continua a ser fixo na penúltima sílaba.

6.3.1 O acento secundário

O TP apresenta também uma série de compostos lexicais e algumas reduplicações onde o acento secundário acaba por manifestar-se, diferentemente dos compostos gramaticais, que como foi analisado acima, possuem somente o acento principal. Esse acento secundário vale algumas palavras apenas, pois o padrão acentual fixo na penúltima sílaba é mantido, ou seja, o acento principal localiza-se sempre no último elemento do composto, e ainda o primeiro elemento do composto o acento secundário manifesta-se da mesma maneira do acento principal, na penúltima sílaba do elemento.

51. /₁tu.a.'si.in/ 'vinagre'

/tu.a/ 'vinho' + /'si.in/ 'azedo'

52. /₁bi.bi.ma.'la.e/ 'ovelha'

/bi.bi/ 'cabra' + /ma.'la.e/ 'estrangeiro'

53. /₁ba.in.'hi.ra/ 'quando'

/ba.in/ 'tempo' + /'hi.ra/ 'quanto, contar'

54. /₁lor.'lo.ron/ 'diariamente'

/'lo.ron/ 'dia'

55. /₁bai. 'ba.in/ 'sempre, todo o tempo'

/ba.in/ 'tempo'

6.4 A ortografia

A ortografia padronizada do Tetun Prasa foi oficializada no Decreto Lei No. 01 de 31 de Março de 2004 e baseou-se na publicação do Instituto Nacional de Linguística (INL) intitulada *Hakerek Tetun tuir Banati* (INL, 2002). Esta publicação, porém, foi liderada pelo linguista australiano Geoffrey Hull e teve como base seus estudos linguísticos muitos dos quais já foram citados na presente dissertação³⁰.

As críticas maiores que vêm sendo feitas contra essa proposta do INL são basicamente duas: o emprego de certos grafemas inadequados e a artificialidade da proposta diante dos falantes leste-timorenses.

Os grafemas presentes na ortografia padrão do Tetun Prasa que são contestados são vários, entre eles: *ll* para representar a palatal [ʎ], *ñ* para representar o a palatal [ɲ], *rr* para representar a vibrante /r/ dos empréstimos lusófonos, ' para representar a oclusão glotal [ʔ] do Tetun Terik, e o acento agudo, que ao mesmo tempo procura marcar o suposto acento fora da posição fixa (penúltima sílaba) e as vogais geminadas do Tetun Prasa.

56. Representações problemáticas da ortografia padrão:

evanjellu 'evangelho'

jullu 'julho'

bañu 'banho'

señór 'senhor'

karreta 'carro'

borraxa 'borracha'

ko'lia 'falar'

di'ak 'bom, bem'

animál 'animal'

³⁰ Para um histórico das diferentes propostas de ortografia para a língua Tetun realizadas desde o período colonial, ver INL (2002).

sá ‘o que’

bá ‘ir’

A proposta de ortografia padronizada apresenta-se extremamente artificial ao falante leste-timorense de Tetun Prasa, seja ele falante como L1 ou L2, devido aos problemas citados anteriormente. Ainda, mesmo os falantes tetumófonos escolarizados possuem dificuldades no emprego dos grafemas problemáticos. Um estudo de caso foi realizado por Williams-van Klinken (2007a) com cerca de 100 professores leste-timorense constatou que o uso dos grafemas *’* e *rr* apresenta grande dificuldade e o número de maior de acertos ocorreu somente com os lexemas usados com menor frequência, como *lian* ‘língua’, *liis* ‘cebola’, *to’o* ‘até’ e *sei* ‘ainda’, enquanto que lexemas usados com maior frequência, como: *di’ak* ‘bom, bem’ e *de’it* ‘apenas’, apresentaram a maior porcentagem de erros. Atualmente, os estudantes das séries iniciais e universitários, assim como os diversos funcionários públicos estão frequentando diferentes cursos oferecidos pelo governo de Timor-Leste que procuram ensinar somente a ortografia padronizada. Esses cursos, todavia, não surtem efeito, pois o grau de letramento dos estudantes e também a documentação redigida pelos funcionários públicos ainda apresentam uma grande alternância de escrita para várias palavras. Em documentos do governo analisados encontrei as diferentes grafias para os mesmos itens lexicais:

57. Variação na ortográfica

ne ~ *nee* ~ *ne’* ~ *ne’e* ‘isto’

lian ~ *li’an* ~ *lia* ~ *li’a* ‘língua’

fo ~ *fó* ~ *foo* ‘dar’

Há também mais duas propostas de ortografia padronizada para o Tetun Prasa, são elas: a da igreja católica e da linguista Catharina Williams-van Klinken. A ortografia da igreja católica além de seguir uma tradição que remonta a fundação do colégio secundário de Soibada, como já foi comentado no capítulo 2, tem certo apelo nas comunidades menos letradas e rurais, pelo fato da igreja católica ter sido um símbolo da resistência contra a invasão indonésia e pela maioria da população de Timor-Leste ser de religião católica. A *Diocese de Dili* publicou uma tradução do *Novo Testamento* que serve como base para muitos seguirem como modelo de ortografia,

mesmo não havendo uma ortografia padronizada explícita (COMISSÃO LITÚRGIA DA DIOCESE DE DILI, 1980). Já o caso da linguista australiana Catharina Williams-van Klinken é diferente, ela é diretora do *Dili Institute of Technology* (DIT), instituição responsável por elaborar materiais didáticos para a língua Tetun, assim como realizar pesquisas linguísticas e oferecer cursos de língua Tetun como Língua Estrangeira, e possui uma proposta ortográfica mais simplificada baseada em seus estudos linguísticos, também citados no decorrer deste trabalho, e na sua experiência e convivência com os alunos e professores leste-timorenses do DIT.

Finalmente, o que foi apresentado acima sobre a situação atual em torno da ortografia padronizada e suas propostas paralelas, procurou mostrar brevemente que esse problema está ligado com a política e o planejamento linguísticos de Timor-Leste, estando longe de ser resolvido a curto prazo e/ou por iniciativas individuais. O que foi possível verificar na pesquisa de campo junto com os falantes tetumófonos é que a ortografia oficial não é empregue pela maioria deles, independente do grau de escolaridade, e que ela não é aprovada pela maioria que necessita frequentar cursos de longa duração somente para aprendê-la, ou seja, desloca-se uma grande quantidade de recursos somente para realizar uma tentativa de ensiná-la. Por esses motivos é que não foi utilizada a ortografia oficial neste trabalho, salvo com as devidas menções a guisa de exemplo.

7. MORFOLOGIA

Neste capítulo serão descritas as categorias gramaticais do Tetun Prasa, de acordo com suas propriedades distribucionais, funcionais e semânticas. Entendendo-se aqui ‘propriedades distribucionais’ como a forma e ordem de distribuição dos morfemas dentro de um item lexical ou gramatical, ou seja, a morfologia; as ‘propriedades funcionais’ estão ligadas à função assumida dentro de um sintagma pela categoria gramatical específica, ou seja, a sintaxe; as ‘propriedades semânticas’ estão ligadas propriamente à semântica, ou seja, a(s) referência(s) que a categoria gramatical analisada faz ao mundo real e a que tipos de conceito.

Em 7.1 falarei brevemente da ausência de morfologia flexional e os processos de gramaticalização relacionados a esta ausência. A seguir, analisarei os processos da morfologia derivacional (7.2), e em (7.3) serão descritas as classes de palavras do Tetun Prasa.

7.1 A ausência de morfologia flexional e a gramaticalização

A morfologia flexional é o tipo de morfologia que marca diferentes propriedades no lexema, propriedades estas que possuem diferentes funções na morfologia e na sintaxe (BOOIJ, 2005, p.99). Como foi apontado anteriormente, no capítulo 4, a língua Tetun apresentava em um estágio prévio um amplo conjunto de afixos, formados por prefixos, sufixos e circunfixos, que possuem apenas algumas retenções na variedade Tetun Terik, enquanto que em Tetun Prasa perdeu-se totalmente.

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, argumentei em Albuquerque (2010a) que o Tetun Prasa 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; nos

últimos anos, o Tetun Prasa, em sua variedade Tetun Díli, como língua urbana vem recorrendo à língua portuguesa para realizar um grande número de empréstimos relativos ao mundo moderno.

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 esteja 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 CPMal e CPMac, 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).

7.1.1 As funções da gramaticalização

De acordo com o que foi exposto anteriormente, o TP desenvolveu diversas formas de transferência funcional, são elas: o gênero natural *mane* ‘masculino’ e *feto* ‘feminino’ e *aman* ‘macho’ e *inan* ‘fêmea’, a definição marcada por *ida*, o plural marcado por *sira*, o diminutivo marcado por *oan*, a marcação de TAM nos verbos *tiha*, *ona*, *sei*, *daudauk*, *atu*, *hela*, entre outros, e a marcação de direção com os verbos *baa* ‘ir’ e *mai* ‘vir’. Além de alguns resquícios de sua morfologia prévia, como é o caso do possessivo e a classe dos inalienáveis ambos marcados pelo sufixo *-n*, ou de certas estruturas que apresentam diferentes estágios de gramaticalização, como a negação verbal *la=*, a marcação de negação é feita por *la=*, *laos* e *lae*. A negação absoluta é feita com *lae*. Já o clítico *la=* é usado somente para negar a classe dos verbos e seus derivados, enquanto que *laos* é uma forma livre, menos gramaticalizada que *la=*, e nega as demais classes. A seguir encontram-se exemplos do emprego das diferentes formas:

57. **la=** iha ema ida iha nee!

NEG EXI pessoa IND LOC este

‘Não há ninguém aqui!’

58. ami **la=** bele tama iha klase agora.

1pl.exc NEG poder entrar LOC classe agora

‘Nós não podemos entrar em sala agora.’

59. ita =nia inan moras ka **lae?**

1pl.inc =POS mãe doente ou NEG

‘Sua mãe está doente (ou não)?’

60. **lae,** hau =nia vizinu maka moras.

NEG 1 =POS vizinho TOP doente

‘Não, é o meu vizinho que está doente.’

61. hau hakarak ai-fuan nee, **laos** (ai-fuan) nebaa.

1sg querer fruta esta não aquelas

‘Eu quero estas frutas, não aquelas.’

62. buat-ida nee **laos** ai-nanas!

coisa-IND esta não abacaxi

‘Isto (esta coisa) não é abacaxi!’

Assim, pode-se afirmar que a negação na classe dos verbos e dos adjetivos é parte da morfologia flexional, já que é realizada pelo proclítico *la=*, enquanto que nas demais classes é realizada pelo lexema *laos*.

O chamado ‘gênero natural’ é um fenômeno típico de línguas austronésias, consistindo na marcação de gênero sem nenhuma propriedade e reflexo na morfossintaxe da língua, apresentando somente propriedades e oposições semânticas, como [+ humano] x [+animal] e [+ masculino] x [+feminino], que são os traços utilizados para o TP.

Em TP, há as seguintes marcações de gênero natural: *mane* ‘homem’ e *feto* ‘mulher’ fazem a marcação de ‘masculino’ e ‘feminino’ respectivamente; *anan* ‘pai’ e *inan* ‘mãe’ marcam ‘macho’ e ‘fêmea’ respectivamente, segue os exemplos:

63. labarik **mane** sira lalika estuda.
 criança MSC PL não.quer estudar
 ‘Os meninos não querem estudar’

64. hau iha oan **feto** nain tolu ho oan **mane** nain rua.
 1sg EXI cria FEM CL três e cria MSC CL dois
 ‘Eu tenho três filhas e dois filhos’

65. toos-nain iha deit fahi **aman** ho manu **aman.**
 roça-dono EXI somente porco macho e pássaro macho
 ‘O agricultor tem somente porcos e galos’

66. nia tenke sosa fahi **inan** ho
 3sg tem comprar porco FEM e
 anu **inan** ho animal seluk tan.
 pássaro FEM e animal outro mais
 ‘Ele tem que comprar porcas, galinhas e outros animais.’

Os substantivos em TP possuem um traço semântico de [+ definido] e singular, quando não acompanhado por nenhum elemento. Assim, o quantificador numeral *ida* ‘um’ é gramaticalizado e colocado após os substantivos para expressar principalmente a ‘indefinição’, porém há algumas situações orientadas pelo discurso que *ida* pode marcar ‘singular’:

67. hau=nia aman foo ruin **ida** ba asu.
 1sg=POS pai dar osso IND para cão
 ‘Meu pai deu um osso ao cachorro’

68. buat-**ida** nee laos ai-nanas!

coisa-IND esta não abacaxi

‘Isto (esta coisa) não é abacaxi!’

69. labarik **ida**-nebee=nia bisikleta hau ferak ba uma
criança SG-REL=POS bicicleta 1sg quebrar para casa
‘A criança cuja bicicleta eu quebrei foi pra casa.’

A marcação de número é um tanto instável em TP, sendo realizada através da reduplicação, da posposição do pronome de terceira pessoa plural =*sira* e do *-s* final da língua portuguesa. O emprego dessas diferentes formas, além de estarem relacionadas com o grau de escolaridade do falante e as situações de uso da língua, aparentemente são selecionadas no léxico tetumófono, já que o falante, independente das características sociolinguísticas, tende a usar: a reduplicação, estratégia morfológica emprestada do *Pazar Melayu*, em empréstimos de origem malaia; o *-s* final, estratégia morfológica de origem portuguesa, em empréstimos lusófonos; e o pronome =*sira* no léxico tetumófono.

70. estudante **sira** hola buku-**buku** atu estuda
estudante PL ganhar livro-RED IRR estudar
‘Os estudantes ganharam livros para estudar.’

71. funionarius nasoes unidas nian too iha timoor ohin
funcionários nações unidas POS chegar LOC Timor hoje
‘Os funcionários das Nações Unidas chegaram em Timor hoje.’

O lexema *oan* ‘cria, descendente’ vem sendo gramaticalizado em TP para expressar a noção de diminutivo. Porém, a propriedade semântica expressa pelo lexema *oan* nos compostos não é de um diminutivo genuíno, já que mantem-se de certa forma a relação de *ascendente* > *descendente*, ou *criador* > *cria*, conforme os exemplos abaixo:

72. lafaek **oan** ida sai hosi nia fatin atu buka hahaan
crocodilo DIM IND sair de 3sg lugar IRR procurar comida
‘Um crocodilozinho saiu de sua toca para procurar comida.’

73. hau iha oan feto nain tolu ho oan mane nain rua.
 1sg EXI cria FEM CL três e cria MSC CL dois
 ‘Eu tenho três filhas e dois filhos’

74. malae sira ulun-bulak atu bosok timoor oan
 estrangeiro PL cabeça-maluco para enganar Timor cria
 ‘Os estrangeiros (agem como) malucos para enganar os timorenses’

O verbo TP pode ser acompanhado de inúmeros lexemas para indicar as informações de TAM, cujos verbos dessa variedade da língua Tetun, de certa forma, carecem. O verbo TP em sua forma simples, não acompanhado por nenhum lexema TAM, pode ser classificado como portador da noção de não futuro e modo *realis*, enquanto que para marcar demais informações no verbo, a variedade TP utiliza-se dos seguintes lexemas³¹:

- *tiha* ‘já’ é marcador do aspecto perfectivo (PERF), possui o traço semântico de algo que foi terminado, sua posição é após o verbo, e apresenta uma influência dos CP asiáticos que também utilizam a palavra lusófona ‘já’ para marcar este aspecto verbal:

75. haree **tiha** ema lubun boot, zezus sae
 ver PERF pessoa quantidade grande Jesus subir
 ba foho ida
 para montanha IND
 ‘Quando viu a multidão, Jesus subiu à montanha.’

76. hau mai atu hasai **tiha** ukun-fuan
 1sg vir IRR remover PERF comando
 no profeta sira
 e profeta PL
 ‘Eu vim para destruir leis e profetas.’

³¹ Para a análise dos lexemas gramaticalizados para marcar TAM aqui elaborada fez-se uso das análises e terminologias presentes em: Comrie (1976, 1985), Dik (1997), Palmer (2001) e Payne (1997).

- *ona* ‘já’ de maneira diferente de *taha* ‘já’ possui o traço semântico de algo que está em vias de iniciar, ou que iniciou em um tempo passado recente e continua a ter repercussão no momento da fala. Dessa forma, a gramaticalização de *ona* foi analisada aqui como marcadora de aspecto ingressivo (ING). Sua posição é após o verbo também:

77. *nia halo ona sala-foer ho nia iha nia laran.*
 3sg fazer ING pecado DAT 3sg LOC POS coração
 ‘Ele cometeu pecado em relação a ela (sua esposa) em seu interior.’

78. *hau dehan lia-los ba imi, sira simu ona*
 1sg dizer língua-certo para 2pl 3pl receber ING
sira=nia kolen
 3pl=POS pagar
 ‘Eu vos digo sinceramente que eles já receberam suas recompensas.’

79. *ba laron idaidak too ona ho*
 para dia um.por.um chegar ING COM
nia=rasik=nia servisu
 3sg=ENF=POS trabalho

(não se preocupe com o mal de amanhã, deixe que)... ‘cada dia chegue à sua própria maneira (com seu próprio trabalho)’

- *taha ona* ‘PF’ é o marcador realizado com mais frequência pelo falante leste-timorense, trata-se da junção dos dois marcadores anteriores e faz referência a qualquer evento acontecido no passado. Por esse motivo é analisado como um marcador de tempo-aspecto perfeito (PF):

80. *mestri haruka tiha-ona labariksira tuur ba rai.*
 professor mandar PF criança PL sentar para terra
 ‘O professor mandou as crianças sentarem no chão.’

- *hotu* ‘tudo, acabar, também’ das várias formas desse lexema é gramaticalizado o verbo intransitivo *hotu* ‘acabar, terminar’ para marcar o término de um evento. Assim, funciona como marcador de aspecto egressivo (EGR) que foca no tempo interno final do evento e é usado em posição pós-verbal:

81. maski lalehan no rai lakon, ukun-fuan nia
 enquanto céu e terra desaparecer comando 3sg
 buat ida sei la=lakon, sei kumpri **hotu**
 coisa IND FUT NEG=desaparecer FUT cumprir EGR
 ‘Enquanto o céu e terra existir, não se esconderá coisa alguma, tudo será cumprido.’

82. o sei la=sai husi nebaa too o selu **hotu**.
 2sg FUT NEG=sair ABL aquele chegar 2sg pagar EGR
 ‘Tu sairás de lá somente quando terminares de pagar.’

- *foin* ‘recentemente, apenas’ de acordo com o conteúdo semântico do lexema, o marcador *foin* indica que uma ação terminou em um tempo recente. Dessa forma, *foin* marca uma aspectualidade específica, chamada de aspectualidade ‘perspectival’, que marca um tempo recente em referência a perspectiva temporal interna do evento. O subtipo da aspectualidade perspectival marcado por esse lexema gramaticalizado em posição pré-verbal é a prospectiva (PROS), que indica que o evento ocorreu recentemente:

83. zudeu sira=nia liurai nebee **foin** moris ihanebee
 judeu PL=POS rei REL PROS nascer onde
 ‘Onde está o que acabou de nascer rei dos judeus?’

84. hau=nia oan feto **foin** mate, maibee ita-boot hiit
 1sg=POS cria FEM PROS morrer mas 2sg-grande erguer
 an mai tau liman ba nia, nia sei moris
 REF CNTP colocar mão LOC 3sg 3sg FUT nascer

‘Minha filha acabou de morrer, mas vós ergueis e repousais nela vossa mão (que) ela reviverá (nascer).’

- *kedas* ‘imediatamente’ é marcador também de aspectualidade perspectival e em posição pós-verbal. A diferença é que *kedas* marca a aspectualidade de perfeito recente (PF.R), ou seja, um evento que ocorre logo após outro ter terminado:

85. sira husik **kedas** dai hodi lao tuir nia

3pl abandonar PF.R rede dirigir andar seguir 3sg

‘Em seguida, eles abandonaram suas redes e deixaram-se serem conduzidos e seguiram-no.’

86. sira nee moris **kedas** basa rai ladun klean

3pl ANA nascer PF.R porque terra não.ter fundo

‘Logo nasceram porque não tinha terra funda.’

- *lai* ‘primeiramente, inicialmente’ expressa o aspecto pontual (PONT), ou demarcativo (DEM), pois o evento marcado por ele deve acontecer primeiro, antes dos demais eventos expressos nas sentenças. É usado em posição pós-verbal:

87. imi hela **lai** ihanee. hau sei baa iha.nebaa

2pl ficar DEM aqui 1sg FUT ir lá

halo orasaun

fazer oração

‘Vós fiquéis aqui. Eu irei até lá para orar.’

- *hela* ‘morar, estar fixo’ é um marcador aspectual que foca o tempo interno do evento como ainda em andamento. Pode ser analisado como marcador de aspecto contínuo (CONT) ou semelfactivo (SEM)³², pois o

³² Marcador para eventos pontuais, indicando que evento ocorre somente uma única vez ou momentaneamente.

evento em andamento ocorre apenas uma vez, ou poucas vezes, diferenciando-se assim do marcador *daudauk* que marca progressividade em eventos habituais. Ainda, o marcador *hela* é usado em posição pós-verbal:

88. sira soe **hela** dai baa tasi.
 3sg lançar SEM rede para mar
 ‘Eles lançaram suas redes ao mar.’

89. sira hadia **hela** dai iha roo laran hamutuk
 3pl consertar CONT redes LOC barco dentro junto
 ho sira=nia aman Zebedeu
 COM 3pl=POS pai Zebedeu
 ‘Eles consertaram as redes no interior do barco junto com o pai deles, Zebedeu.’

- *daudauk* ‘ainda’ é o marcador de aspecto progressivo (PROG), conforme o conteúdo semântico do lexema *daudauk* ‘ainda’, no ato da fala o evento expresso pelo verbo está acontecendo ‘ainda’, e sua posição é pós-verbal:

90. nia hanorin **daudauk** buat nee, bainhira nain=nia
 3sg pensar PROG coisa este quando senhor=POS
 anzu mosu ba nia iha mehi
 anjo aparecer para 3sg LOC sonho

‘Ele estava pensando nisso, quando um anjo do senhor apareceu para ele em seu sonho.’

91. lao **daudauk** tuir tasi-ibun Galileia nian,
 andar PROG ao.longo litoral Galileia POS
 nia haree maun-alin rua
 3sg ver irmãos dois

‘Andando ao longo da costa do mar da Galileia, ele viu dois irmãos.’

- *nafatin* ‘continuar’ é marcador de aspecto iterativo (ITE), pois foca na quantidade de vezes que o evento ocorre, ou seja, sua frequência. O próprio lexema *nafatin* ‘continuar’ é gramaticalizado com o sentido de o evento ocorrer ‘continuadamente’ e é usado em posição pós-verbal:

92. sira=nia anzu sira fihir **nafatin** hau
 3pl=POS anjo PL observar ITE 1sg
 aman lalehan nian futar oin
 pai céu POS CL.DIV face

‘Seus anjos sempre veem a face de meu pai do céu (que está no céu).’

93. efeito sira husi krize sei sente **nafatin**
 efeito PL ABL crise FUT sentir ITE
 iha timoor rai
 LOC timor terra

‘Os efeitos da crise serão sentidos continuamente em terras leste-timorenses’

- *sei* ‘ainda’ é marcador de tempo futuro em posição pré-verbal. Em TP, é possível afirmar que há a distinção entre o tempo futuro marcado por *sei* e não-futuro, não marcado:

94. hau **sei** halo kafee mai ami
 1sg FUT fazer café para 1exc
 ‘Eu farei café para nós’

95. efeito sira husi krize **sei** sente nafatin
 efeito PL ABL crise FUT sentir ITE
 iha timoor rai
 LOC timor terra

‘Os efeitos da crise serão sentidos continuamente em terras leste-timorenses’

- *atu* ‘para’ da mesma forma que o ‘tempo’, o ‘modo’ em TP diferencia somente duas categorias, marcadas de maneira análoga à categoria

‘tempo’. O modo *irrealis* é marcado por *atu* em posição pré-verbal e o modo *realis* não é marcado:

96. relatoriu nee **atu** kombina esforsu husi governu no
relatório este IRR combinar esforço ABL governo e
komunidade internasionaal **atu** hatuur dezvoltimentu.
comunidade internacional IRR iniciar desenvolvimento

‘Este relatório espera combinar os esforços do governo e da comunidade internacional para iniciar o desenvolvimento.’

97. nee hotu mosu, **atu** halo tuir lia
isto tudo acontecer IRR fazer de.acordo língua
nebee nai hatoo husi profeta
REL senhor passar ABL profeta

‘Tudo isto aconteceu para que se cumprisse as palavras do senhor através do profeta.’

Os marcadores de TAM no verbo TP apresentam grandes problemas para a análise linguística por diversos motivos, entre eles: a maioria desses marcadores é empregue somente nos textos escritos ou em registros formais e literários; o conteúdo semântico expresso por eles é muito próximo o que faz com que sejam intercambiáveis; há uma grande variação no uso desses marcadores; a própria terminologia na bibliografia linguística atual sobre estudos das categorias TAM, especialmente o aspecto, é ambígua ou há autores com diferentes propostas de análise muito diversas uma das outras, entre outros.

Alguns dos verbos TP, que podem ser classificados como verbos de movimento, ou verbos direcionais, que são verbos portadores da noção semântica de ação com movimento são usados juntamente com um complemento direcional expressos pelo verbo *baa* ‘ir’ que marca a direção centrífuga e *mai* ‘vir’ que marca a direção centrípeta.

98. mana foti **mai** ida tan, halo favoor
irmã trazer CNTP um mais fazer favor

‘Irmã, traga mais um, faz favor!’

99. lori **ba** bikan nee!
levar CNTF prato este
'Leve (para lá) estes pratos!'

100. nia fila fali **ba** uma
 3sg voltar novamente CNTF casa
'ele voltou para casa'

101. labarik mane, fila fali **mai!**
 criança masculino voltar novamente CNTP
'menino, volte aqui!'

7.2 Morfologia derivacional

A função básica da morfologia derivacional é criar lexemas novos a partir do inventário lexical da língua através de processos morfológicos, como afixação, justaposição de lexemas (composição), reduplicação, entre outros. Tetun Prasa apresenta esses três processos morfológicos: afixação, composição e reduplicação.

A seguir serão analisados cada um deles separadamente, iniciando pela afixação, seguida pela reduplicação e pela composição.

7.2.1 Afixação

TP apresenta os prefixos *ha-* 'causativizador, verbalizador' e *nak-* 'intransitivizador', e os sufixos adjetivadores *-n* e *-k*, e agentivos e/ou instrumental *-door*, *-ten* e *-nain*.

O prefixo *ha-*, derivado do PAN **pa-*, é produtivo em TP, sendo empregado em substantivos, adjetivos e verbos, possuindo funções diferentes. Quando usado nos substantivos e adjetivos possui função verbalizadora, transformando o substantivo ou o adjetivo em verbo. Nos verbos, é usado com função causativizadora. Nos exemplos abaixo o prefixo *ha-* é empregado no adjetivo *moos* 'limpo', gerando o verbo *hamoos* 'limpar', e no verbo intransitivo *mate* 'morrer', que se transforma em um novo verbo *hamate* 'acabar, encerrar':

102. nia **ha**-moos haris-fatin ninia uma
 3.sg CAU-limpo banheiro 3sg.POS casa
 ‘Ela limpou o banheiro da casa dela.’

103. mane nebaa **ha**-mate reuniaun horisehik
 homem aquele CAU-morrer reunião ontem
 ‘Aquele homem encerrou a reunião ontem’

O prefixo *nak-*, variante *nam-*, é usado somente em verbos com função intransitivizadora, ou seja, transforma os verbos transitivos em verbos intransitivos. Nos exemplos abaixo há o verbo sem o prefixo (a) e com o prefixo *nak-* que altera a transitividade dessas mesmas formas (b):

104. la=bele doko lata hanesan nee
 NEG=poder balançar lata igual isto
 ‘Não balance a lata desse jeito.’
 kareta **nak**-doko tanba dalan aat
 carro INT-balançar porque estrada mal
 ‘O carro treme porque a estrada é ruim.’

105. kolega, loke odamatan, halo favor.
 colega abrir porta fazer favor
 ‘colega, por favor, abra a porta!’
 odamatan **nak**-loke baa ema hotu.
 porta INT-abrir para pessoa tudo
 ‘a porta está aberta para todas as pessoas’

106. farda estudante nian lees bainhira monu
 uniforme estudante POS rasgar quando cair
 ‘o uniforme do estudante rasgou quando (ele) caiu’
 farda nebaa **nak**-lees tanba tuan liu
 uniforme aquele INT-rasgar porque velho demais
 ‘aquele uniforme rasgou porque estava muito velho’

Os sufixos *-n* e *-k* possuem função nominalizadora, transformando substantivos em adjetivos e verbos em substantivos. O sufixo *-n* é usado para derivar adjetivos a partir de verbos e *-k* para derivar adjetivos de substantivos. Os exemplos (107) e (108) mostram que sufixo *-n* é usado para derivar substantivos a partir de verbos: *rei* ‘beijar’ > *rein* ‘beijo’; *sosa* ‘comprar’ > *sosan* ‘compras’, e o sufixo *-k* para derivar adjetivos de substantivos: *tasi* ‘mar’ > *tasik* ‘marinho’; *mota* ‘rio’ > *motak* ‘ribeirinho’; *modo* ‘hortaliça’ > *modok* ‘verde’:

107. nia haruka **rein** ida baa ita.
 3sg enviar beijo IND para 2sg
 ‘ela mandou um beijo para você’
108. bele hela iha-nee ita=nia **sosan**
 poder deixar LOC-este 2sg=POS compra
 ‘Pode deixar aqui suas compras’
109. iha timoor bele hakail ikan **tasik** ho ikan **motak**.
 LOC Timor poder pescar peixe marinho e peixe ribeirinho
 ‘Em Timor, pode-se pescar peixe marinho e peixe ribeirinho.’
110. ai-tahan baibain **modok**, maibee bailoro kinur.
 árvore-folha constantemente verde mas estação.seca amarelo
 ‘As folhas das árvores geralmente são verdes, mas (na) época seca (ficam) amarelas.’

Um conjunto de sufixos que se mostra produtivo em TP é o conjunto de sufixos agentivos *-door*, *-ten* e *-nain*. O sufixo *-door* de origem lusófona (do sufixo agentivo e instrumental *-dor*) aparece desde as primeiras documentações da língua Tetun, que datam do século XIX. Nestas documentações, esse sufixo já se mostra produtivo em várias entradas do dicionário de Silva (1889), assim como é analisado pelo mesmo autor como parte da gramática Tetun (SILVA, 1889, p.3).

Hajek e Williams-van Klinken (2003) fizeram uma análise das funções desse sufixo no TP. Ele é empregado somente nos verbos intransitivos (*moe* ‘timidez’ >*moedoor* ‘pessoa muita tímida’; *koalia* ‘falar’ > *koaliadoor* ‘falador’; *dukur* ‘dormir’ >*dukurdoor* ‘pessoa que dorme demais’), derivando um adjetivo a partir desses verbos. O adjetivo tem os traços semânticos habitual (a ação expressa pelo verbo é praticada constantemente pelo agente [+humano]) e pejorativo (o hábito do agente praticar a ação constantemente é mal vista, como um vício, uma mania ruim etc.).

111. *funsonariu nebaa ministeriu edukasaun nian haluha-door*
 funcionário aquele ministério educação POS esquecer-AGT
tanba la=gosta servisu.
 porque NEG=gostar trabalhar

‘Aquele funcionário do ministério da educação é esquecedor (esquece constantemente) porque não gosta de trabalhar’

112. *zoven sira Timoor oan ruma hemu-door*
 jovem PL Timor cria algum beber-AGT
 ‘Alguns jovens timorenses são beberrões.’

113. *estudante balu hamnasa-door bainhira la=hatene responde*
 estudante alguns rir-AGT quando NEG=saber responder
 ‘Alguns estudantes ficam risonhos quando não sabem responder’

Vale a pena mencionar que o sufixo *-door* do Tetun não possui a função instrumental do sufixo lusófono *-dor*, já que esta função tornou-se produtiva em língua portuguesa por volta do século XIX (MARINHO, 2009, p.116.). Empréstimos lusófonos de palavras como *aseleradoor* ‘acelerador’, *radiador* ‘radiador’, *komputadoor* ‘computador’, entre outras, não se caracteriza como morfologia derivacional produtiva, já que se limita somente a esses tipos de empréstimos, diferentemente dos exemplos acima.

De maneira diferente, o sufixo *-teen* deriva da palavra *teen* ‘fezes’ e é usado para derivar um adjetivo com sentido também pejorativo, porém esse sufixo pode ser usado nas categorias substantivo, verbo e até em outros adjetivos:

114. nia hakarak deit hatoba lorloron, nia baruk-**teen**.
 3sg querer apenas deitar dia-RED 3sg preguiça-AGT
 ‘ele só quer ficar deitado todo dia, ele é muito preguiçoso.’
115. ...ema toman lia-bosok beibeik,
 pessoa acostumar língua-mentir sempre
 ami bolu ema nee bosok-**teen**.
 1pl.exc chamar pessoa esta mentir-AGT
 ‘...pessoas que estão acostumadas a mentir sempre, nós chamamo-las de mentirosas.’

Finalmente, o sufixo *-nain*, derivado do substantivo *nain* ‘dono, mestre’, pode ser usado tanto como agentivo, quanto como instrumental. Pode ser empregue em substantivos e em verbos:

116. aman nafatin **uma-nain**
 pai sempre casa-dono
 ‘o pai sempre é o chefe da casa (pai de família)’
117. toos-**nain** iha rai timoor kuda hare barak
 roça-dono LOC terra Timor plantar arroz muito
 ‘Os agricultores em Timor plantam muito arroz.’
118. lia-**nain** deit hatene konta ai-kananoik
 língua-dono apenas saber contar história
 ‘somente os senhores da língua sabem contar histórias.’

7. 2.2 Reduplicação

A reduplicação em Tetun foi um tema abordado de maneira introdutória por Esperança (2001, p.69) e analisado pormenorizadamente por Avram (2008), que dedicou um artigo a esse fenômeno em TP. Esperança (2001) elenca apenas as reduplicações encontradas em documentações escritas da língua e realiza uma breve

análise seguindo as teorias de Marantz (1982) e Kenstowicz (1994). Dessa forma, o linguista analisa que a reduplicação tetumófona é parcial e consiste em uma forma de afixação, com o afixo sem conteúdo fonético copiando uma camada segmental da base adjacente (ESPERANÇA, 2001, p.85). Já Avram (2008) analisa a reduplicação em Tetun seguindo a teoria otimalista, elencando um conjunto de parâmetros e restrições para que esta ocorra na gramática TP.

119. Exemplos de reduplicação:

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’

Assim, pode-se afirmar que a reduplicação em Tetun é preferencialmente parcial, como foi apontado acima, e difere-se de outros fenômenos linguísticos, como a repetição, pelo fato de ser produtiva, sendo aplicada nas classes dos advérbios principalmente, e alguns itens das classes dos substantivos, adjetivos e quantificadores.

Os conteúdos semânticos expressos pela reduplicação são vários: na classe dos advérbios a reduplicação serve como intensificação, ou iteração (120); nos substantivos funciona como formação de uma ação com sentido pejorativo em relação ao conteúdo semântico expresso pelo substantivo (121); ainda em substantivos que possuem conteúdo semântico de tempo (como dia, noite, ano, dias da semana, meses do ano, entre outros) a reduplicação funciona como iteração (122); nos adjetivos a função é somente de intensificação (123); nos quantificadores numerais serve como serialização, ou coletivização (124), e nos quantificadores não numerais serve como intensificação (125).

120. Reduplicação de advérbios:

hotu ‘todos’ > *hotu-hotu* ‘completamente todos (sem exclusão)’

kleur ‘muito tempo’ > *kleukleur* ‘em breve’

loos ‘sim, certamente’ > *loloos* ‘exatamente’

nafatin ‘constantemente’ > *nafatin-nafatin* ‘sempre’

121. Reduplicação de substantivos concretos:
ibun ‘boca’ > *ibuibun* ‘resmungar’
122. Reduplicação de substantivos temporais:
loron ‘dia’ > *lorloron* ‘todo dia’
kalan ‘noite’ > *kalkalan* ‘toda noite’
tersa ‘terça-feira’ > *ter tersa* ‘toda terça-feira’
fulan ‘mês’ > *fufulan* ‘mensalmente’
123. Reduplicação de adjetivos:
dook ‘longe’ > *dodook* ‘muito longe’
barak ‘muito’ > *barbarak* ‘muitos’
boot ‘grande’ > *boboot* ‘muito grande, grandes’
124. Reduplicação de quantificadores numerais:
ida ‘um’ > *idaidak* ‘um por um’
rua ‘dois’ > *ru rua* ‘em dupla, ou em grupo de dois’
tolu ‘três’ > *toltolu* ‘em grupo de três’
125. Reduplicação de quantificadores não numerais:
seluk ‘outro’ > *selseluk* ‘outros, vários outros’
hotu ‘todos’ > *hotu-hotu* ‘completamente todos (sem exclusão)’

A operação fonológica que ocorre na reduplicação em Tetun é a cópia de até duas primeiras sílabas adjacentes dos itens lexicais dissilábicos. Quando a palavra possui mais de duas sílabas, o falante opta pela reduplicação total. Ainda, não foram encontrados itens monossilábicos reduplicados nos dados coletados. A proposta de Avram (2008, p.433) para a explicação da reduplicação parcial em TP consiste na hipótese de que a língua aceita somente duas combinações de sílaba na reduplicação para gerar duas moras ($\mu\mu$). A primeira consiste na sequência de duas sílabas leves (uma sílaba leve possui uma mora: $\sigma = \mu$). A segunda possui somente uma sílaba pesada, que também é constituída de 2 moras ($\mu\mu$).

126. Estrutura da reduplicação em TP:

σ leve (μ) + σ leve (μ) = $\mu\mu$

σ pesada ($\mu\mu$) = $\mu\mu$

Seguem os exemplos da reduplicação, de acordo com a estrutura prosódica apontada acima:

127. Sílabas pesadas ($\sigma = \mu\mu$)

lo. ron ‘dia’ > *lor. lo. ron* ‘todo dia’

CV CVC CVC CV CVC

μ $\mu\mu$ $\mu\mu$ μ $\mu\mu$

128. Sílabas leves ($\sigma = \mu$)

bo. ot ‘grande’ > *bo. bo. ot* ‘muito grande, grandes’

CV CV CV CV CV

μ μ μ μ μ

129. Reduplicação total (itens lexicais com mais de duas sílabas):

na. fa. tin ‘constantemente’ > *na. fa. tin. na. fa. tin* ‘sempre’

CV CV CVC CV CV CVC CV CV CVC

Dessa forma, a maioria dos itens reduplicados encontra-se em variação, pois as duas estruturas são aceitas pelo falante. Somente alguns itens reduplicados específicos, que se encontram lexicalizados, aceitam apenas uma forma.

130. Variação de itens reduplicados:

hotu ‘todos’ > *ho. tu. ho. tu* ‘completamente todos (sem exclusão)’

CV CV CV CV

μ μ μ μ

hot. ho. tu

CVC CV CV

$\mu\mu$ μ μ

<i>barak</i> ‘muito’ >	<i>ba.</i>	<i>ra.</i>	<i>ba.</i>	<i>rak</i> ‘muitos’
	CV	CV	CV	CVC
	μ	μ	μ	μμ
	<i>bar.</i>	<i>ba.</i>	<i>rak</i>	
	CVC	CV	CVC	
	μμ	μ	μμ	

131. Itens cristalizados:

ida ‘um’ > *idaidak* ‘um por um’

daudauk ‘ainda, PROG’ > **dau* (?) forma não conhecida pelos falantes

7.2.3 Composição

A composição, segundo Booij (2005, p.75), é a forma mais frequente para se criar novos lexemas na maioria das línguas do mundo. Isto, porém, não acontece em Tetun, já que a variedade Tetun Terik utiliza um grande número de afixos em sua morfologia flexional e derivacional (KLINKEN, 1999) e a variedade Tetun Prasa não apresenta morfologia flexional e as composições forjadas para expressar muitos conceitos novos não são reconhecidas pelos falantes, que preferem os empréstimos³³.

132. Calques e empréstimos:

/naran hadook belek/ ‘nome que pode ser possuído’ > /substantivu alienavel/

/naran hadook labelek/ ‘nome que não pode ser possuído’ > /substantivu inalienavel/

/katetek/ > /kompostu/

/matuduk/ ‘o que mostra’ > /demonstrativu/

/mahusuk/ ‘o que pergunta’ > /interogativu/

/malenok/ ‘o que reflete’ > /refleksu/

/hahalok/ ‘ação’ > /asaun/

³³Alguns exemplos da terminologia do léxico tetumófono basearam-se na proposta de Hull e Correia (2005) e todos apresentados seguem a ortografia padrão oficial.

/lalaok/ ‘andamento, função’ > /funsau/

/tasik/ ‘marítimo, marinho’ > /marinu/

A estrutura da composição é basicamente binária (XY) com um lexema modificador (Y) e outro lexema-núcleo (X), que é o modificado. Ainda, o mesmo autor propõe um critério de classificação das composições que será adotado aqui, pois tal critério provou-se válido para a classificação das diferentes composições em Tetun Prasa. Segundo Booij (2005, p.79), as composições podem ser classificadas em quatro tipos de acordo com seu comportamento morfossintático e seus traços semânticos, a saber: *endocêntricas*, *exocêntricas*, *bahavrihi* e *copulativas*, ou *dvanda*.

- *endocêntricas*, composto com núcleo que mantém as propriedades morfossintáticas de um dos termos, geralmente do núcleo:

133. Compostos endocêntricos³⁴:

ain-fatin ‘pegada’ (/‘a.in/ ‘pé’ + /‘fa.tin/ ‘lugar’)

ain-fuan ‘dedo do pé’ (/‘a.in/ ‘pé’ + /‘fu.an/ ‘fruta’)

ain-tuur ‘joelho’ (/‘a.in/ ‘pé’ + /‘tu.ur/ ‘sentar’)

ai-tahan ‘folha’ (/ai/ ‘árvore’ + /‘ta.han/ ‘folha’)

ai-funan ‘flor’ (/ai/ ‘árvore’ + /‘fu.nan/ ‘flor’)

bibi-malae ‘ovelha’ (/‘bi.bi/ ‘cabra’ + /ma.‘la.e/ ‘estrangeiro’)

ibun-kulit ‘lábios’ (/‘i.bun/ ‘boca’ + /‘ku.lit/ ‘pele’)

haris-fatin ‘banheiro’ (/‘ha.ris/ ‘banhar-se’ + /‘fa.tin/ ‘lugar’)

manu-aman ‘galo’ (/‘ma.nu/ ‘pássaro’ + /‘a.man/ ‘pai’)

manu-knuuk ‘ninho’ (/‘ma.nu/ ‘pássaro’ + /‘knu.uk/ ‘ninho’)

manu-rade ‘pato’ (/‘ma.nu/ ‘pássaro’ + /‘ra.de/ ‘pato’)

manu-metan ‘corvo’ (/‘ma.nu/ ‘pássaro’ + /‘me.tan/ ‘preto’)

masin-midar ‘açúcar’ (/‘ma.sin/ ‘sal’ + /‘mi.dar/ ‘doce’)

nehan-isin ‘gengiva’ (/‘ne.han/ ‘dente’ + /‘i.sin/ ‘corpo’)

tasi-ibun ‘litoral’ (/‘ta.si/ ‘praia’ + /‘i.bun/ ‘boca’)

tua-siin ‘vinagre’ (/‘tu.a/ ‘vinho’ + /‘si.in/ ‘azedo’)

³⁴ Os núcleos, destacados em negrito, são aqueles que determinam semanticamente o composto.

As composições *endoncêntricas* são as mais produtivas em TP, pois são amplamente usadas para criar palavras que se referem a conceitos inseridos na cultura tetumófona com características mais concretas, ou seja, itens da cultura material, como exemplo: partes do corpo, tipos de animais e plantas e suas respectivas partes, lugares, alimentos, entre outros.

- *exocêntricas*, composto sem núcleo que os termos assumem propriedades morfossintáticas diferentes das originais:

134. Compostos exocêntricos:

foho-rai ‘pitão, tipo de cobra do Gênero *Python*’ (/‘fo.ho/ ‘montanha’ + /rai/ ‘terra’)

foin-sae ‘adolescente’ (/‘fo.in/ ‘recente’ + /‘sa.e/ ‘subir’)

manu-talin ‘fofoqueiro’ (/‘ma.nu/ ‘pássaro’ + /‘ta.lin/ ‘corda’)

matan-dook ‘curandeiro’ (/‘ma.tan/ ‘olho’ + /‘do.ok/ ‘longe’)

- *bahuvrihi* (do Sânscrito³⁵ *bahú* ‘muito’ + *vr̥hi* ‘arroz’), tipo especial de composição em que o composto se refere a uma pessoa/objeto através de referência a uma entidade/característica específica:

135. Compostos bahuvrihi:

isin-manas ‘febre’ (/‘i.sin/ ‘corpo’ + /‘ma.nas/ ‘quente’)

isin-rua ‘grávida’ (/‘i.sin/ ‘corpo’ + /‘ru.a/ ‘dois’)

Os exemplos de composição *bahuvrihi* em TP são poucos, mas justifica-se essa classificação tanto pelo critério semântico, nos exemplos *isin* ‘corpo’ não é um tipo de ‘corpo’, como pelo critério morfossintático, *isin* ‘corpo’ que é substantivo e *rua* ‘dois’ que é numeral comportam-se como adjetivo no composto *isin-rua* ‘grávida’.

³⁵ Termo empregado primeiramente pelos gramáticos do sânscrito para se referir ao tipo de composição citada, onde a entidade referida pelo composto não é mencionada em nenhum de seus elementos. O exemplo de *bahúvr̥hi* (muito + arroz) significava ‘rico’.

- *copulativas*, ou *dvanda*, composto sem núcleo que é caracterizado pelos elementos formadores estarem em uma relação de coordenação:

136. Compostos copulativos:

inan-aman ‘pais’ (/‘i.nan/ ‘mãe’ + /‘a.man/ ‘pai’)

baa-mai ‘lá e cá’ (/‘ba.a/ ‘ir’ + /mai/ ‘vir’)

ida-rua ‘uns, alguns’ (/‘i.da/ ‘um’ + /‘ru.a/ ‘dois’)

loron-kalan ‘dia e noite’ (/‘lo.ron/ ‘dia’ + /‘ka.lan/ ‘noite’)

7.3 As classes de palavras

Nesta seção, serão analisadas separadamente as classes de palavras do Tetun Prasa, identificados os traços semânticos específicos de cada uma delas, suas formas e respectivas funções. São elas: substantivo, pronome, verbo, demonstrativos, interrogativos, quantificadores, adjetivos, preposições, advérbios e conjunções.

7.3.1 Substantivo

A categoria gramatical ‘substantivo’ é aquela que se refere a conceitos estáveis no tempo (Givón, 1984, p.51). Morfossintaticamente, os substantivos podem ser modificados pelos sufixos da morfologia derivacional *-teen* ‘agentivo’, *-nain* ‘agentivo, instrumental’ e *-k* ‘adjetivador’ e pelos vários itens gramaticalizados com diferentes funções, como: indefinição *ida*, gênero natural *mane x feto* e *anan x inan*, plural *sira*, e diminutivo *oan*; Ainda, o substantivo e os pronomes são as únicas classes que podem assumir a posição e função de núcleo do SN e de outros sintagmas que se formam a partir deste.

137.	okupasaun	indonesia	nian	autoriza	violensia	nuudar
	ocupação	Indonésia	POS	autorizar	violência	como
forma	lejitimu	ba	rezistensia			
forma	legítima	para	resistência			

‘A ocupação da Indonésia autorizou a violência como forma legítima de lidar com a resistência’.

138. fetu nebaa kapaas
mulher aquela bonita
‘Aqueela mulher é bonita.’

139. hau=nia aman foo ruin ida ba asu.
1sg=POS pai dar osso IND para cão
‘Meu pai deu um osso ao cachorro.’

140. uma=nia laran iha sala rua,iha kuartu tolu
casa=POS dentro EXI sala dois EXI quarto três
ho iha haris-fatin rua.
e EXI lugar-banhar.se dois

‘Em minha casa (já mencionada anteriormente) há duas salas, três quartos e dois banheiros.’

A justificativa para se diferenciar uma classe de ‘substantivos’ de outra de ‘adjetivos’ encontra-se pelas propriedades estruturais e distribucionais de ambas. A classe dos substantivos é única que aceita os sufixos *-teen*, *-nain* e *-k*, assim como as gramaticalizações mencionadas. Ainda, os adjetivos não assumem a posição e a função de núcleo do SN pelo fato de possuírem características funcionais próximas da classe dos verbos, posição esta que os substantivos assumem dentro do SN:

141. iha timoor bele hakail ikan tasik ho ikan motak.
LOC Timor poder pescar peixe marinho e peixe ribeirinho
‘Em Timor, pode-se pescar peixe marinho e peixe ribeirinho.’

142. ai-tahan baibain modok, maibee bailoro kinur.
árvore-folha constantemente verde mas estação.seca amarelo
‘As folhas das árvores geralmente são verdes, mas (na) época seca (ficam) amarelas.’

143. toos-nain iha rai Timoor kuda hare barak
roça-dono LOC terra Timor plantar arroz muito

‘Os agricultores em Timor plantam muito arroz.’

7.3.2 Pronomes

Os pronomes em Tetun Prasa ocorrem como formas livres, sem modificação ou modificados pelo clítico =*nia* para indicar posse – propriedade distribucional; eles podem fazer referência a diversos conceitos estáveis no tempo, como pessoas e objetos, e, assim, comportam-se da mesma maneira que os nomes – propriedade funcional. Ainda, podem substituir os nomes, e expressar os diversos argumentos dos papéis sintático-semânticos – propriedade semântica.

O paradigma dos pronomes aparece na tabela abaixo:

TABELA 5. O paradigma pronominal do TP

	SINGULAR	PLURAL	OUTRAS INFORMAÇÕES
1ª INCL	<i>hau</i>	<i>ita</i>	O pronome <i>ita</i> é utilizado para se referir a 2ª pessoa formal.
1ª EXCL		<i>ami</i>	
2ª	<i>o</i>	<i>imi</i>	O pronome <i>o</i> é usado somente de maneira informal, ou para se referir a crianças.
3ª	<i>nia</i>	<i>sira</i>	O pronome <i>sira</i> é colocado após os substantivos com a função de plural.

Os traços notáveis do paradigma pronominal do Tetun Prasa são os seguintes: não há diferença de gênero; há uma distinção de 1ª pessoa do plural inclusiva *ita* e exclusiva *ami*; o pronome *ita* é usado para se referir a 2ª pessoa do singular de maneira um pouco mais formal do que a 2ª pessoa do singular *o*, que é usada somente para se referir a crianças; a palavra *boot* ‘grande’ é indexada após qualquer pronome pessoal como forma de tratamento a pessoas que fazem parte dos status sociais mais altos da sociedade leste-timorense, como: chefes de aldeias, anciãos, pessoas de origem nobre, assim como profissões: juiz, advogado, político, professor.

144. **hau**=*nia* suporter *sira* sei ataka
 1sg=POS patrocinador PL FUT atakar

ita=nia supporter sira
2sg=POS patrocinador PL

‘Meu patrocinador atacará o seu.’

145. **sira** organizadu diak loos
3pl organizados bom certo

‘Eles são muito organizados’

146. kolisensa, **hau** bele husu ba imi-boot
com licença 1sg poder perguntar para 2pl-grande

‘Com licença, posso perguntar algo a vós (ou vossa senhoria)?’

147. **ami** la=bele tama iha klase agora.
1pl.exc NEG=poder entrar LOC classe agora

‘Nós não podemos entrar em sala agora.’

148. **ita** tenke sai agora, tanba la=iha tempu.
1pl.inc tem sair agora porque NEG=EXI tempo

‘Nós temos que sair agora, porque não há muito tempo.’

7.3.3 Verbo

A categoria ‘verbo’ é aquela que faz referência a conceitos mutáveis no tempo, como eventos, ações, estados (GIVÓN, 1984, p.55), e que assume a posição e função de núcleo do SV e de outros sintagmas que se formam a partir deste. O verbo diferencia-se de outras classes por aceitar os afixos: *ha-* ‘causativizador, verbalizador’ e *nak-* ‘intransitivizador’ e os agentivos e/ou instrumental *-door* e *-nain*. Ainda, é acompanhado dos diversos lexemas gramaticalizados marcadores de TAM e dos marcadores diretivos *ba* ‘centrífugo’ e *mai* ‘centrípeto’.

149. sira halai **ba** foho
3pl correr para montanha

‘Eles fugiram para a montanha.’

150. lori **ba** bikan nee!

levar CNTF prato este
'Leve (para lá) estes pratos!'

151. estudante balu hamnasa-**door** bainhira la=hatene responde
estudante alguns rir-AGT quando NEG=saber responder
'Alguns estudantes ficam risonhos quando não sabem responder'

152. hau **sei** halo kafee mai ami
1sg FUT fazer café para 1exc
'Eu farei café para nós'

7.3.3.1 Os verbos seriais

O Tetun Prasa apresenta algumas construções de verbos seriais, conforme foi analisado em Hajek (2006), Klinken (2000) e Williams-van Klinken, Hajek e Nordlinger (2001). Essas construções verbais serializadas, porém vêm sofrendo ampla redução tanto em seu inventário, quanto em seu uso, devido ao contato e influência da língua portuguesa.

Entre os fatores que justificam a presença de construções com verbos seriais em Tetun Prasa são: os verbos seriais encontram-se na mesma oração, compartilham o mesmo contorno entonacional, representam semanticamente um evento apenas, os verbos possuem as mesmas propriedades gramaticais, como marcação de aspecto, negação, entre outras características (HAJEK, 2006, p. 240).

A presente análise dos verbos seriais para o Tetun Prasa trabalha com a hipótese do contato dos falantes tetumófonos com as variedades crioulas da língua portuguesa (CPMal e CPMac) em um período anterior, que fez tornar esse tipo de construção verbal mais produtiva. Algumas influências dos CPs asiáticos nesse tipo de construção podem ser vistas em:

153. nia **baa** **dansa** iha kolega uma nian .
3sg ir dançar LOC colega casa POS
'ele foi e dançou na casa do colega' ou 'ele(a) vai dançar na casa do colega'

154. ita **baa** **selu** nia ka

2sg ir pagar 3sg INTG
'Você foi e pagou-o?' ou 'você vai pagá-lo?'

155. hau tiha **foo han** labarik
1sg já dar comer criança
'Eu já alimentei a criança' ou 'eu já dei de comer a criança'

156. dotoor nee **foo hemu** baa nia
médico este dar beber para 3sg
'Este (o) médico deu de beber a ele (ao paciente).'

Conforme, pode ser visto nos exemplos acima, há uma clara origem lusófona das construções de verbos seriais em TP, como *baa dansa* 'ir dançar', *baa selu* 'ir pagar', *foo han* 'dar de comer' e *foo hemu* 'dar de beber'. Argumento que, além da origem lusófona, esse tipo de construção com verbos seriais foram produtivas em um período anterior em TP, pois deram origens a construções com a mesma estrutura, mas que não são usadas em língua portuguesa, assim como possuem conteúdos semânticos diferentes:

157. malae nebaa **foo aluga** hau=nia uma
estrangeiro aquele dar alugar 1sg=POS casa
'Aquele estrangeiro alugou minha casa'

158. amu **foo sarani** labarik sira
padre dar cristão criança PL
'O padre batizou as crianças.'

159. kolega nee **foo empresta** ita=nia livru mai hau
colega este dar empresta 2sg=POS livro CNTP 1sg
'O colega me emprestou o livro dele.'

A análise adotada no presente trabalho baseia-se em parte na proposta de Hajek (2006). Porém, considero que o autor ampliou demais o inventário de verbos seriais do

Tetun, incluindo os complementos direcionais *baa* ‘CNTF’ e *mai* ‘CNTP’ dos verbos de movimento e locomoção, assim como certas construções que fazem uso do verbo *baa* ‘ir’, *foo* ‘dar’ e *halo* ‘fazer’ apenas como complementos para expressar noções semânticas inexistentes em Tetun. Ainda, apesar de apresentar também a influência da língua portuguesa nos verbos seriais do Tetun Prasa (HAJEK, 2006, p. 252), o autor deixa escapar um grande número de construções lusófonas que podem ter influenciado a língua Tetun em tempos passados.

O trabalho de Hajek (2006), apesar de apresentar uma diferença da análise aqui apresentada, é o mais completo sobre as construções com verbos seriais em Tetun Prasa, revisando a bibliografia existente e analisando as diferentes estruturas desse tipo de construção. Ainda, outro fator de grande importância elencado pelo autor é que a frequência de uso dos verbos seriais está diminuindo devido ao contato atual com a língua portuguesa, principalmente em sua variedade padrão do Português Europeu.

7.3.3.2 As classes de verbos

Nesta subseção procurar-se-á agrupar os verbos TP em classes abertas e classes fechadas, e em subclasses menores de acordo com critérios semânticos para verificar como as diferentes classes de verbo estão distribuídas. As classes abertas podem ser divididas semanticamente em verbos de movimento, locomoção, impacto, postura, comunicação, cognição e afins, atividades, estados. Já as classes fechadas são menores e divididas principalmente em: modal, dêitico, posicional, existencial e cópula.

A justificativa para esse agrupamento consiste no fato de que alguns verbos de diferentes classes apresentam comportamentos morfossintáticos diferentes. Entre as diferentes classes, há: as classes dos verbos de movimento e locomoção, que exigem os complementos direcionais; os verbos de comunicação e cognição, que apresentam formação mais recente na língua, sendo derivados com o prefixo causativo cristalizado ou formados a partir de construções com verbos seriais, conforme foi apresentado anteriormente; os verbos modais, que a maioria surgiu com o contato com o português.

A classe dos verbos de movimento e locomoção, como *lori* ‘levar’, *foti* ‘trazer, levantar’, *tuda* ‘arremessar’ entre outros, exigem um complemento direcional derivado dos verbos *ba* ‘ir’ e *mai* ‘vir’ para marcar a direção centrífuga (CNTF) e centrípeta (CNTP) respectivamente, como nos exemplos abaixo:

160. mana foti **mai** ida tan, halo favor!
 irmã trazer CNTP um mais fazer favor
 ‘Irmã, traga mais um, faz favor!’

161. lori **ba** bikan nee!
 levar CNTF prato este
 ‘Leve (para lá) estes pratos!’

Os verbos de comunicação e cognição, como *hatene* ‘saber’, *hanoin* ‘pensar’, *hanorin* ‘ensinar’, *koalia* ‘falar’, em sua maioria apresentam em sua estrutura elementos cristalizados e/ou que são passíveis de decompor em unidades menores. Nos exemplos citados anteriormente é possível identificar o prefixo causativo *ha-*, mas as raízes como *-tene*, *-noin* e demais, não existem mais na língua, assim como o verbo *koalia* é, formado a partir do verbo *koa* ‘cortar’ mais o substantivo *lian* ‘língua’.

162. lia-nain deit hatene konta ai-kananoik
 língua-dono apenas saber contar história
 ‘Somente os senhores da língua sabem contar histórias.’

163. kolega bele hanorin lia-tetun mai hau ka
 colega poder ensinar língua-tetun para 1sg INTG
 ‘O colega pode ensinar a língua tetun para mim?’

Em relação às classes fechadas o TP não apresenta cópula, verbos posicionais ou dêiticos. O verbo existencial é a única classe existente tipicamente tetumófono com o verbo *iha*. Já os verbos modais apresentam-se como construções recentes e empréstimos, com exceção do verbo *bele* ‘poder’:

164. Formação de alguns verbos modais:

tenke ‘obrigação’ > do português *tem que*;
 bele ‘poder’ > lexema tetumófono;
 hakarak ‘querer’ > ha ‘CAU’ + *karak ‘querer’

lakohi ‘não.querer’ > la ‘NEG’ + kohi (?)

lalika ‘não.precisar’ > la ‘NEG’ + lika (?)

165. hau hakarak ai-fuan nee, laos (ai-fuan) nebaa.
1sg querer fruta esta não aquelas
‘Eu quero estas frutas, não aquelas.’

166. labarik mane sira lalika estuda.
criança MSC PL não.quer estudar
‘Os meninos não querem estudar’

7.3.4 Demonstrativos

Os demonstrativos em TP são uma classe que localizam um referente no espaço ou no discurso de acordo com a posição deste em relação ao falante. Os demonstrativos que localizam um referente no espaço são os demonstrativos dêiticos, o demonstrativo que localiza o referente no discurso é o demonstrativo anafórico, e há a subclasse dos demonstrativos que fazem uma localização pouco especificada, ou indireta, de seu referente, que são chamados de demonstrativos não dêiticos. Os demonstrativos são empregados sempre pospostos ao SN; desta maneira, podem ser empregados pospostos à classe dos substantivos, dos adjetivos e das preposições.

- Demonstrativos dêiticos: em TP são dois que marcam o referente próximo ao falante *nee* ‘isto, este(a)’ e distante do falante *nebaa* ‘aquilo, aquele(a)’. Eles podem ser acompanhados de *ida* para marcar indefinição, *sira* para marcar plural, e também por *ema* ‘pessoa’ e *buat* ‘coisa’ para marcar que o referente é [+ humano] ou [- humano] respectivamente.

167. hau hakarak ai-fuan **nee**, laos (ai-fuan) **nebaa**.
1sg querer fruta esta não aquelas
‘Eu quero estas frutas, não aquelas.’

168. buat ida **nee** laos ai-nanas!
coisa IND esta não abacaxi

Isto (esta coisa) não é abacaxi!

- Demonstrativo anafórico: *nee* é empregado para funcionar como anáfora direta, e também como anáfora distante, ou algum referente com que o ouvinte esteja familiarizado; anáfora referente ao foco pragmático; marcador de um elemento inicial topicalizado:

169. ema nebaa mak laos azuda,
 pessoa aquele TOP NEG ajudar
nee nee mak azuda daudauk hau.
ANA este TOP ajudar PROG 1sg

‘Aqueles são as pessoas que não estão me ajudando, mas estas estão.’

170. ...ema toman baibain lia-bosok beibeik,
 pessoa acostumar constantemente língua-mentir sempre
ami bolu ema **nee** bosok-teen.
1sg.exc chamar pessoa esta mentir-AGT

‘Pessoas que estão acostumadas a mentir sempre, nós chamamo-las de mentirosas.’

- Demonstrativos não dêiticos: comportam-se como adjetivos ou pronomes de acordo com sua forma e/ou função no SN. Localizam-se também posposto ao SN e na maior parte do corpus consistia nos lexemas *ruma* ‘alguns’ e *seluk* ‘outro’.

171. nia tenke sosa fahi, manu inan ho
 3sg tem comprar porco pássaro mãe e
animaal **seluk**.
animal outro

‘Ele tem que comprar porcos, galinhas e outros animais.’

7.3.5 Interrogativos

A classe dos interrogativos é formada por lexemas usados para questionar algo em relação ao seu referente. Sintaticamente encontram-se no final da sentença, podendo

ser deslocados para a posição inicial, ou medial, com a função de foco pragmático e assim marcado pelo lexema *mak* ‘TOP’. Alguns lexemas da classe dos interrogativos podem ser marcados para a definição com *-ida* e para o número plural *-sira*.

172. *ola irmaun agora halo saa-ida*
olá irmão agora fazer o.que-IND
‘Olá irmão, o que você está fazendo agora?’

173. *kolega baa nebee*
colega ir (a)onde
‘Colega, aonde vai?’

174. *mana, hira mak nee*
irmã quantoTOP isto
‘Moça, isto custa quanto?’

175. *ita hakarak sosa sira-nebee*
2sg querer comprar PL-qual
‘Quais (destes) você quer comprar?’

Semanticamente, além de funcionar como itens que possuem traços semânticos relativos à diversas questões, a classe dos interrogativos ainda traz consigo traços semânticos referentes ao grau de animacidade e referência temporal: *see* ‘quem’ [+humano] e *saa* ‘o que’ [-humano], *bainhira* ‘quando’ [tempo futuro] e *horibainhira* ‘quando’ [tempo passado]:

176. *ema see*
pessoa quem
‘Quem é (este)?’

177. *ita koalia ho see*
2sg falar com quem
‘Com quem você falou?’

178. ita=nia naran **saa**
 2sg=POS chamar o.que
 ‘Como você se chama?’
179. nee **saa**-ida
 isto o.que-IND
 ‘O que é isto?’
180. apa, senoor mai **horibainhira**
 pai senhor chegar quando
 ‘Pai, quando o senhor chegou?’
181. **bainhira** mak nia fila-fali mai hosi Baukau
 quando TOP 3sg voltar CNTP de Baucau
 ‘Ele voltará de Baucau quando?’
182. tuku **hira** mak ami aranka
 hora quanto TOP 1pl.exc partir
 ‘Nós partimos a que horas?’
183. o hela iha **nebee**
 2sg morar LOC onde
 ‘Tu moras onde?’

7.3.6 Quantificadores

Os quantificadores são os lexemas do TP que expressam a informação da quantidade em relação ao núcleo a que se referem. Os quantificadores podem ser determinantes tanto do núcleo do SN, marcando a quantidade propriamente dita do substantivo núcleo do sintagma, como do SV, marcando a frequência ou localização temporal do evento expresso pelo verbo núcleo do sintagma.

184. uma=nia laran iha sala **rua,iha** kuartu **tolu**

casa=POS dentro EXI sala dois EXI quarto três
 ho iha haris-fatin **rua.**
 e EXI lugar-banhar.se dois

‘Em minha casa (já mencionada anteriormente) há duas salas, três quartos e dois banheiros.’

185. ema **balu** desempregadu tanba la=hatene
 pessoa alguns desempregado porque NEG=saber
 koalia portuges
 falar português

‘Algumas pessoas estão desempregadas porque não sabem falar português’

Em TP, os quantificadores podem ser divididos em numerais e não numerais. Ambos podem sofrer reduplicação para expressar outros traços semânticos, ou um traço semântico quantificador mais específico, como será visto abaixo:

- Numerais: os numerais tetumófonos seguem um sistema decimal baseado na contagem dos dedos da mão, fator comum nas línguas do mundo e das línguas de origem austronésia. Morfologicamente, eles podem ser antecidos dos pelo classificador humano *nain*, que é mais usado, e também podem ser antecidos pelos classificadores *lolon* ‘objetos cilíndricos’, *tahan* ‘objetos chatos’, *fuan* ‘objetos redondos’ e *musan* ‘objetos redondos pequenos’; pelo prefixo *da-* para indicar ordem, nesse caso recebe também o sufixo *-k* por se tratar de uma forma adjetiva³⁶; sofrer reduplicação para expressar iteração, hábito, ou distribuição. Sintaticamente, os numerais funcionam como determinantes e podem modificar os substantivos e os verbos.

186. ferik nee **nain** neen.
 velha este CL.HUM seis
 ‘Estas seis senhoras.’

³⁶Essa forma é pouco usada coloquialmente e provavelmente trata-se da redução da expressão *baa + dala + NUM* usada para expressar iteração, como: *baa dala rua* ‘segunda vez’ > *dalarua* > *darua* > *daruak* ‘segundo’ e *baa dala tolu* ‘terceira vez’ > *dalatolu* > *datolu* > *datoluk* ‘terceiro’.

187. masu L.A. iha tabaku **lolon** sanulu-rua
 maço L.A. EXI cigarro CL.CIL doze
 ‘O maço de LA (marca indonésia de cigarro) tem doze cigarros.’

TABELA 6. Classificadores numerais

Classificador	Significado	Classificador para
nain	nobre, dono	pessoas
lolon	galho	objetos cilíndricos grandes, ex. vela, peixes.
tahan	folha	objetos chatos e finos, ex. papel, roupa.
fuan	fruta, coração	objetos arredondados grandes, ex. côco, ovo.
musan	semente	objetos arredondados pequenos, ex. remédios.

188. (...) **datoluk**, sei hatudu konabaa rekonstrusaun
 Terceiro FUT mostrar a.respeito reconstrução
 husi pasadu Timoor Lorosae nian.
 de passado Timor leste POS
 ‘(...) em terceiro, será mostrada uma reconstrução do passado de Timor-Leste.’

189. **daruak** introdusaun hau sei hakerek
 segundo introdução 1sg FUT escrever
 uitoan konabaa linguistika teoretika.
 pouco a.respeito linguística teórica
 ‘Segundo (após a introdução), eu escreverei um pouco sobre linguística teórica.’

190. katuas nebaa hakarak koalial **dahituk** depois hotu-hotu.
 velho aquele querer falar sétimo depois todos-RED
 ‘Aquele senhor quer ser o sétimo a falar, depois de todos’

191. estudante sira tenke baa profesoos **ru-rua**
 estudante PL tem.que ir professor dois-RED
 atu hatudu tarefa.
 IRR mostrar tarefa
 ‘Os estudantes devem ir ao professor de dois em dois para mostrar a tarefa.’

192. depois eskola estudante sira tenke lao baa
 depois escola estudante PL tem.que andar para
 uma **toltolu**
 casa três-RED

‘Depois da escola (das aulas), os estudantes devem andar (ir) para casa de três em três (em grupos de três)’.

Os numerais complexos são formados a partir dos numerais simples, indicados na tabela abaixo:

TABELA 7. Os quantificadores numerais

ida	1
rua	2
tolu	3
haat	4
lima	5 (> <i>liman</i> ‘mão’)
neen	6
hitu	7
ualu	8
sia	9
sanulu	10 (<i>nulu</i> ‘dezena’)
atus	100, ou ‘centena’
rihun	1000, ‘milhar’

A ordem dos numerais complexos é distinta, as dezenas são formadas pela flexão de *nulu* ‘dezena’ (20 *rua-nulu* ‘2-dezena’, 30 *tolu-nulu* ‘três-dezena’) somado ao lexema *resin* ‘mais’ e o numeral simples; as centenas são formadas por *atus* ‘centena’ seguido do numeral simples para marcar a quantidade das centenas (100 *atus-ida* ‘cem-1’, 200 *atus-rua* ‘cem-dois’); os milhares são formados com o lexema *rihun* ‘milhar’ seguidos pela mesma ordem dos numerais das dezenas para identificar a quantidade de

milhares (2000 *rihun-rua* ‘mil-dois’, 15000 *rihun-sanulu-resin-lima* ‘1000-10-mais-5’).
 Abaixo seguem exemplos de cada um dos numerais³⁷:

13 >	sanulu resin	tolu		
	10	mais	3	
25 >	rua-nulu	resin	lima	
	2-dezena	mais	5	
32 >	tolu-nulu	resin	rua	
	3-dezena	mais	2	
146 >	atus-ida	haat-nulu	resin	neen
	centena-1	4-dezena	mais	6
255 >	atus-rua	lima-nulu	resin	lima
	centena-2	5-dezena	mais	5
12000 >	rihun	sanulu resin	rua	
	1000	10	mais	2

- Não numerais: expressam uma quantidade nãoespecífica em relação ao seu referente e dependendo de sua forma e/ou função assume características de outras classes, principalmente de adjetivos e de pronomes. Entre eles estão os seguintes itens: *uitoan* ‘pouco’ *balu* ‘alguns (pouca quantidade)’ *sira-balu* ‘alguns deles’ *ruma* ‘alguns (muita quantidade)’ *oioin* ‘vários’ *barak* ‘muito’ *barbarak* ‘grande quantidade’ *natoon* ‘suficiente’ *hotu* ‘tudo’ *ida* ‘um’ *idaidak* ‘cada um’ *tomak* ‘o todo (grupo)’ *seluk* ‘outro’ *ida-seluk* ‘o outro’:

193. depoisde fulan **balu**, ami=nia oan filafali.
 depois mês alguns 1pl.exc=POS filho voltar
 ‘Depois de alguns meses, nosso filho voltou.’

194. hau=nia ferik nee halo kafee **natoon** ba ita.

³⁷Vale lembrar, conforme foi dito no capítulo 2, os numerais complexos tetumófonos raramente são usados, a população escolarizada utiliza os numerais lusófonos, enquanto que a população mais jovem e a população menos escolarizada utilizam os numerais do *bahasa indonesia*.

1sg=POS velha este fazer café suficiente para 1pl.inc
'Minha esposa aqui fez café suficiente para nós.'

7.3.7 Adjetivos

A classe dos adjetivos em muitas línguas austronésias é difícil de ser identificada. Segundo Himmelmann (2005, p.128), nas línguas da região da Nusantara Oriental a classe dos adjetivos é ainda mais problemática para ser identificada e, em muitos casos, várias línguas não possuem esta classe: as palavras com significados adjetivais apresentam um comportamento morfossintático análogo ao da classe dos verbos.

Em Tetun Prasa, há evidências suficientes que justifiquem a existência desta classe, todavia a origem verbal dessa classe permanece presente no comportamento semântico e morfossintático dos adjetivos. Semanticamente, muitos adjetivos derivam dos verbos intransitivos, mantendo o significado da raiz verbal: *araska* 'difícil, ser difícil de obter', *hakneak* 'ajoelhado, ajoelhar', *hamulak* 'tipo de canto animista, rezar (durante ritual animista)', *hamlaha* 'fome, esfomeado, ter fome'. Morfossintaticamente, os adjetivos podem receber os afixos verbais, como o prefixo causativo *ha-* transformando-se em verbos: *manas* 'quente' > *hamanas* 'esquentar', *mate* 'morto, morrer' > *hamate* 'encerrar', e ainda assumir a posição de modificador do núcleo do SN (típica do adjetivo), ou pode assumir a posição de núcleo do SV, que é exclusiva do verbo:

195. agora susubeen **manas** demais
agora leite quente demais
'Agora o leite está quente demais!'

196. hamanas susubeen ba bebee.
esquentar leite para bebê
'Esquenta o leite para o bebê.'

Separo aqui a classe dos adjetivos em duas subclasses distintas, as duas possuem o traço sintático-semântico básico de funcionar como modificador do substantivo. A primeira subclasse trata-se dos 'adjetivos genuínos' que consistem em um pequeno

grupo de itens lexicais que se comportam como verdadeiros adjetivos, formados a partir da sufixação, possuem características funcionais diferentes dos verbos e são usados em funções atributivas e predicativas.

197. ami iha uma **boot** iha foho.
 1pl.exc EXI casa grande LOC montanha
 ‘Nós temos uma casa grande nas montanhas.’

198. malae australia **aas** demais.
 estrangeiro Austrália alto demais
 ‘Os australianos são muito altos.’

Já a segunda subclasse, denomino de ‘adjetivos verbais’, que é uma classe maior e mais fluída de adjetivos que se comportam sintaticamente igual aos verbos e/ou são derivados desta mesma classe.

199. kolega nee **sala**
 colega este errar
 ‘O colega aqui está errado’ ou ‘O colega errou.’

200. moras balu bele daet bainhira ema,
 doença alguns poder espalhar quando pessoa
 moras mear besik ema isin **diak** seluk
 doença tossir perto pessoa corpo bom outro

‘Algumas doenças podem ser transmitidas (espalhadas), quando o doente tossir próximo a pessoas saudáveis.’

As propriedades semânticas principais expressas pelos adjetivos tetumófonos são cor, tamanho e qualidade. Digno de nota, é que os adjetivos eram em número reduzido em um estágio anterior da língua, sendo os adjetivos relativos à cor, tamanho e qualidade categorizando apenas as noções básicas dessas propriedades. A expansão lexical da classe dos adjetivos ocorreu devido à necessidades culturais e os adjetivos inseridos posteriormente no léxico são facilmente identificados, pois trata-se de

empréstimos do malaio e do português, ou de uso de verbos e substantivos em posição e função adjetiva sem nenhuma marcação morfológica.

Ainda, há compostos que atribuem certas propriedades semânticas prototípicas ao substantivo e que possuem características funcionais de adjetivos. Apesar de muitos desses compostos sofrerem grande variação pelo fato de terem se formado recentemente na variedade TP e por mesclar traços linguísticos e culturais de origem austronésia e lusófona, pode se identificar sua estrutura [parte do corpo/local + substantivo/adjetivo/verbo]. Estes compostos servem para indicar conceitos de emoção ou caráter (WILLIAMS-VAN KLINKEN, 2007b):

201. distritu Ainaru **rai-malirin** liu.
 distrito Ainaro terra-gelada passar
 ‘O distrito de Ainaro é muito frio.’
202. uluk bainhira tempu funu, iha mane barak **aten-boot**³⁸
 antes quando época guerra EXI homem muito fígado-grande
 ‘Antigamente, na época da guerra, existiam muitos homens corajosos.’
203. iha **foin-sae** mane barbarak
 EXI novo-subir MSC muito-RED
 iha universidade ho fuan-moras. (Laran moras)
 LOC universidade com coração-doente
 ‘Há muitos (mesmo) jovens (homens) na universidade apaixonados’
204. hau la=baa servisu ohin, tanba
 1sg NEG=ir trabalhar hoje porque
 hau **ulun-moras** (ulun fatuk moras)
 1sg cabeça-doente
 ‘Eu não fui trabalhar, porque eu estava com enxaqueca’³⁹

³⁸Aqui se trata de um exemplo de variação, pois também pode ser usado *aten-barani* (fígado + bravo) ‘corajoso’.

³⁹Comparar com: hau=nia ulun moras
 1sg=POS cabeça doente
 ‘minha cabeça está doendo (está com dor)’

205. malae sira **ulun-bulak** atu bosok timooroan
 estrangeiro PL cabeça-maluco para enganar Timor cria
 ‘Os estrangeiros (agem como) malucos para enganar os timorenses’

206. iha Bekora iha tais **folin-kmaan**
 LOC Becora EXI tais preço-baixo
 Em Becora, há tais⁴⁰ barato.

Os adjetivos podem ser modificados em relação ao grau através do uso de certos lexemas que assumem funções específicas. O verbo *liu* ‘passar, superar’ quando usado posposto ao adjetivo marca o grau comparativo ou superlativo, *hanesan* ‘igual, como’ marca o grau de igualdade. Além de *liu*, várias outras formas podem ser usadas para intensificar e marcar o grau superlativo do adjetivo tetumófono, entre elas: *loos* ‘certo’, *lahalimar* ‘sem brincadeira’, *atumate* ‘de morrer’.

207. feto nebaa kapaas **loos**
 mulher aquela bonita muito
 ‘Aquele mulher é bonita demais!’

208. hahaan manas **la=halimar**
 comida quente NEG=brincar
 ‘A comida está quentíssima!’ ou ‘sem brincadeira, a comida está muito quente!’

209. nia beik **atumate**
 3sg burro para.morrer
 ‘Ele é burríssimo!’ ou ‘ele é burro de morrer!’

7.3.8 Preposições

O uso das preposições em TP varia consideravelmente, assim como o das conjunções, por se tratar de uma classe que sofre forte influência lusófona tanto na forma, quanto na função. Sintaticamente, como o TP possui uma tendência a sintaxe

⁴⁰Tipo de faixa usada como vestimenta pelos nativos timorenses, que posteriormente tornou-se um símbolo cultural do país. Na atualidade, consiste em uma faixa cerimonial entregue como um presente, ou homenagem.

paratática, as preposições muitas vezes são omitidas pelos falantes, e quando usadas, possuem a função, principalmente, de modificadores espaciais e temporais.

As preposições podem ser divididas em preposições simples, que em sua maioria são derivadas de verbos, e locuções prepositivas, formadas principalmente a partir de uma preposição simples somada a uma parte do corpo.

210. hau=nia aman foo ruin ida **ba** asu.
 1sg=POS pai dar osso IND para cão
 ‘Meu pai deu um osso ao cachorro.’

211. nia tanis daudauk **iha** kuartu nian.
 3sg chorar PROG LOC quarto POS
 ‘Ela está chorando em seu quarto.’

212. ema nebaa dosente **husi** Aileu.
 homem aquele docente de Aileu.
 ‘Aquele é o professor de Aileu.’

213. sira la=bele baku malu **ho** katana.
 3pl NEG=poder lutar REC com espada
 ‘Eles não podem brigar (entre si) com espada’

214. azente-polisia sira halai **tuir** tasi-ibun.
 agente-polícia PL correr por praia-boca
 ‘Os policiais correm ao longo do litoral’

215. turma nebaa hotu-hotu iha servisu
 turma aquela RED-todos EXI trabalhar
nuudar profesoos biolozia
 como professor biologia

‘Naquela turma todos trabalham como professores de biologia.’

As preposições simples, como foi dito, derivam de verbos. Os argumentos principais para tal afirmação são dois principais: o primeiro é que algumas preposições ainda guardam sua clara origem verbal, às vezes até estão em relação de sinonímia, conforme os exemplos abaixo:

216. Origem de algumas preposições:

baa ‘ir’	>	baa ‘para’
iha ‘EXI’	>	iha ‘em’
la= iha	>	laiha ‘sem’
NEG= EXI		
‘não tem, não existe.’		

O segundo argumento baseia-se no fato de haver formas cristalizadas de preposições que apresentam evidências de formas verbais com a presença do prefixo *ha-*, que faz parte da morfologia derivacional verbal, em formas como *hasoru* ‘contra’, *hanesan* ‘igual’ e *haleu* ‘em volta, ao redor de’:

217. iha taxi barak **haleu** Katedral Vila-Verde nian
 EXI táxi muito ao.redor catedral Vila-Verde POS
 ‘Há muitos táxis ao redor da catedral de Vila-Verde.’

218. sira halo funu **hasoru** inimigu baibain.
 3pl fazer guerra contra inimigo dia-RED
 ‘Eles guerrearam contra o inimigo comum’

Outras preposições são empréstimos lusófonos, como *kontra* ‘contra’⁴¹, *depoizde* ‘após’, *antezde* ‘antes’, *embora* ‘embora’, *komesa* ‘até, desde’, *duranti* ‘durante, por’.

219. hau hela iha rai liur **duranti** indonesia tempu nian
 1sg morar LOC terra fora durante Indonésia tempo POS

⁴¹Há uma grande variação no emprego de empréstimos lusófonos, e no caso de *kontra* ‘contra’ isso se mantém. No corpus analisado, apesar de serem encontrados exemplos de *kontra* ‘contra’ como preposição, na maioria *kontra* ‘contra’ é usado como um verbo estativo ‘estar.contra, estar.em.oposição’.

‘Eu morei fora durante o tempo indonésio (período da dominação indonésia).’

220. **antezde** ita too, ami koalia konaba
antes 2sg chegar 1pl.excl falar a.respeito
ita=nia peskiza lia-Tetun
2sg=POS pesquisa língua-Tetun

‘Antes de você chegar, falávamos de sua pesquisa sobre a língua Tetun.’

As locuções prepositivas são construções formadas através de uma preposição simples, que em quase todo o corpus analisado tratava-se da preposição *iha* ‘LOC’⁴², seguida por um SN mais um substantivo com traços semânticos espaciais, principalmente partes do corpo, como *laran* ‘coração, dentro’, *leten* ‘topo, cima’, *okos* ‘abaixo’, *oin* ‘rosto’, *kotuk* ‘costas’, *soirin* ‘lado’, *klaran* ‘meio’, *leet* ‘entre’. A estrutura fica da seguinte maneira representada abaixo:

221. Estrutura das locuções prepositivas:
iha + (SN) + partes do corpo/palavras espaciais

Vale lembrar que o SN encontra-se marcado entre parênteses pelo fato de ser opcional. Dessa forma, as preposições podem ser empregues também sem a presença do SN, sendo formada apenas por *iha* + substantivo:

222. bele tama **iha laran,** halo favoor.
poder entrar LOC dentro fazer favor
‘Pode entrar (dentro de casa), por favor’ ou ‘entre, por favor’.

223. inan timoor nian tau labarik **iha kotuk**
mãe Timor POS carregar criança LOC costas
‘As mães timorenses carregam as crianças nas costas’

7.3.9 Advérbios

⁴²Há somente alguns exemplos no corpus coletado de que no lugar da preposição *iha* ‘LOC’ verificou-se o uso de *husi* ‘de’.

A classe dos advérbios tradicionalmente consiste em uma categoria muito ampla, onde se classificam itens com valor semântico que não são bem definidos como pertencentes a classes como do nome, verbo, ou adjetivo (PAYNE, 1997). Porém a justificativa maior para a existência de uma classe de advérbios consiste no critério sintático e semântico. No nível da sintaxe, a classe dos advérbios é a única que geralmente possui uma distribuição livre pela sentença (GIVÓN, 1984). No nível da semântica existem subclasses prototípicas já consagradas nos estudos linguísticos descritivos e na gramática tradicional, que são facilmente identificadas. São elas: modal, temporal, local e evidencial.

Em TP, os advérbios, além de possuírem as características sintático-semânticas mencionadas (excetuando a presença de uma subclasse modal), sofrem processo de reduplicação e seu conteúdo semântico pode ser aplicado ora sobre uma categoria gramatical apenas, ora sobre um sintagma, e em alguns casos sobre uma sentença inteira. Como será mostrado nos exemplos abaixo, muitos advérbios TP na realidade consistem em preposições, demonstrativos, interrogativos, ou adjetivos, sendo usados com função adverbial:

224. hau la=hatene. koalia **neineik**, halo favoor
 1sg NEG=entender falar devagar fazer favor
 ‘Eu não estou entendendo. Fale mais devagar, por favor!’

225. **dala ruma** timoor rai hela manas demais ba malae
 às.vezes Timor terra ficar quente demais para estrangeiro
 ‘Às vezes Timor fica quente demais para os estrangeiros’

226. labarik sira baku malu **beibeik**, depois eskola
 criança PL bater REC constantemente depois escola
 ‘As crianças brigam entre elas constantemente, depois da escola.’

As subclasses dos advérbios TP são formadas por lexemas tetumófonos, compostos e empréstimos lusófonos. Seguem alguns exemplos de cada subclasse: modal *hamutuk* ‘junto’, *leet* ‘em vão’, *nunee* ‘por esse motivo’, *ate* ‘até’; temporal *ohin* ‘hoje’, *dalaruma* ‘às vezes’, *nafatin* ‘constantemente’, *aban* ‘amanhã’, *horiseik* ‘ontem’,

ulukliu ‘inicialmente’, *ikusliu* ‘finalmente’, *oras nee* ‘agora’, *nunka* ‘nunca’, *sedu* ‘cedo’, *tardi* ‘tarde’; local *baa* ‘além’, *mai* ‘aquém’, *baa oin* ‘a frente’, *iha oin* ‘em frente’, *iha leten* ‘no topo’; quantidade *uitoan* ‘pouco’, *tan* ‘mais’, *kala* ‘aproximadamente’, *demais* ‘demais’, *menus* ‘menos’.

227. nia isin manas **too** horisehik
 3sg pele quente desde ontem
 ‘Ele está com febre (pele quente) desde ontem.’

228. funsonarius nasoos unidas nian too **iha** timoor ohin
 funcionários nações unidas POS chegar LOC Timor hoje
 ‘Os funcionários das Nações Unidas chegaram em Timor hoje.’

229. hau halo viazen ba estranjeiru. **nunee,** hau=nia
 1sg fazer viagem para estrangeiro por.isso 1sg=POS
 família tenki tau matan ami=nia oan
 família tem colocar olho 1pl.exc=POS filho
 ‘Eu viajarei para o estrangeiro. Por isso, minha família deverá cuidar dos nossos filhos.’

230. **dala ruma** malae sira halo viazen ba Bali
 às.vezes estrangeiro PL fazer viagem para Bali
 ‘Às vezes, os estrangeiros viajam para Bali nas férias.’

7.3.10 Conjunções

O Tetun Prasa, juntamente com as demais línguas de origem austronésia faladas em Timor-Leste, possui uma sintaxe preferencialmente paratática (como será visto no capítulo 9). Assim, grande parte das conjunções são empréstimos, ou adaptações de modelos lusófonos de locuções conjuntivas. Isso ocorre pelo fato de o TP sofrer o processo de específico de mudança linguística, chamada de ‘metatipização’.

O termo ‘metatipo’ foi criado por Ross (1996) e consiste em um tipo especial de mudança linguística devido ao contato, onde a ‘língua modificada’, de menor prestígio, copia as estruturas sintáticas da ‘língua modelo’, de maior prestígio. A variedade Tetun

Prasa aqui é classificada como a ‘língua modificada’ e a língua portuguesa como ‘língua modelo’. Abaixo, seguem exemplos de estruturas morfossintáticas lusófonas copiadas pelo TP:

231. ele **tem que** fazer o trabalho dele.
 nia **tenke** halo ninia servisu
 3sg tem fazer 3POS trabalho

232. ela **tem que** sair agora.
 nia **tenke** sai agora
 3sg tem sair agora

233. este médico **deu de beber** para ele
 dotoor nee **foo hemu** ba nia.
 médico este dar beber para 3sg

Conforme os exemplos anteriores, em TP ocorreram não apenas empréstimos de conjunções lusófonas, mas também de conceitos relacionais expressos pelas conjunções da língua portuguesa e de estruturas sintáticas desta mesma língua. Podem-se dividir as conjunções TP em três subgrupos distintos, a saber: conjunções tetumófonas, calque de conceitos lusófonos e empréstimos linguísticos:

- Conjunções tetumófonas: *ho* ‘e, com’, *no* ‘e’, *hodi* ‘de’, *ka* ‘ou’, *katak* ‘que’ e *atu* ‘para’:

234. nia dehan **katak** la=gosta estuda
 3sg dizer que NEG=gostar estudar
 ‘Ele disse que não gosta de estudar’

235. ami presiza atu estuda pasta **ho** livru
 1pl.exc precisar IRR estuda mochila e livro
 ‘Nós precisamos de mochila e livros para estudar.’

- Calques: *ho mós* ‘e também’, *selae* ‘senão’, *tanba* ‘porque’, *bainhira* ‘se, quando’, *depoizde* ‘depois de’:

236. ita la=bele hemu bee foer, **selae** ita sei hela moras
 2sg NEG=poder bebu água sujo senão 2sg FUT ficar doente
 ‘Você não pode beber água suja, senão ficará doente.’

237. kareta nakdoko **tanba** dalan la= diak.
 carro balançar porque estrada NEG=bom
 ‘O carro treme porque a estrada não é boa.’

- Empréstimos: *i* ‘e’, *maibee* ‘mais que’, *maski* ‘mas que’, *mas* ‘mas’, *mais* ‘mas’, *se* ‘se’.

238. hau la=gosta estuda **maibee** presiza ba
 1sg NEG=gostar estudar mas precisar para
 servisu diak liu
 trabalho bom passar
 ‘Eu não gosto de estudar, mas preciso (se quiser) um trabalho melhor.’

239. **se** nia sei sai, hau mos
 se 3sg FUT sair 1sg também
 ‘Se ele sair, eu sairei também.’

8. SINTAXE

Neste capítulo descreverei a sintaxe do TP enfatizando o sintagma, a estrutura dos principais sintagmas e outras construções elaboradas pelos falantes. Dessa forma, analisarei questões sobre a ordem dos constituintes e as construções paratáticas em (8.1) para depois descrever a estrutura dos SNs (8.2), dos SVs (8.3) e dos demais sintagmas (8.4).

O TP possui a ordem canônica SVO com os constituintes sendo marcados principalmente pela sua posição na sentença. Porém, há casos em que outras ordens de constituintes também sejam consideradas gramaticais pelo falante por motivos de natureza semântico-pragmático, são elas: VO, OSV e SOV, conforme será analisado em (8.1).

A estrutura do SN é complexa com uma série de elementos que podem assumir a posição posterior ao núcleo, mas com poucos elementos que assumem a posição anterior (8.2). Já o SV aceita poucos elementos tanto localizados em posição anterior ao núcleo, como em posição posterior a ele (8.3).

8.1 Parataxe

A sintaxe do TP é tipicamente de ordem SVO e paratática, ou seja, a relação entre os constituintes é marcada somente pela justaposição deles. Há grande diferença entre a modalidade escrita e a modalidade oral quando se analisa a sintaxe, pois os principais fenômenos de mudança da ordem canônica de SVO para VO, OSV e SOV são típicos da fala, assim como a sintaxe paratática está ligada também a essa modalidade do TP. Isso não exclui, porém, que esses fenômenos ocorram em registros escritos.

O TP, conforme foi descrito em (7.3.10), apresenta um pequeno conjunto de conjunções nativas e uma tipologia de língua analítica. Dessa forma, isso reflete na sintaxe com os constituintes e as relações gramaticais geralmente sendo marcadas somente pela posição que assumem na sentença, assim como na justaposição entre os constituintes, como já foi dito. Segue o exemplo:

	S	V	O	Conj.	Comp.
240.	[sira]	[halai]	[(ba)	Suai]	(ho) [(depois?) indonesia].

família tenki tau matan ami=nia oan
 família tem colocar olho 1pl.exc=POS filho

‘Eu viajarei para o estrangeiro. Assim/por isso/por esse motivo, minha família deverá cuidar dos nossos filhos.’

245. ita la=bele hemu bee foer, ita sei hela moras
 2sg NEG=poder beber água sujo 2sg FUT ficar doente
 ‘Você não pode beber água suja, senão ficará doente.’

246. hau husu tudik ida tan (mai) ami.
 1sg pedir faca uma mais CNTP 1exc
 ‘Eu pedi mais uma faca para nós.’

247. (dala ruma) malae sira halo viazen Bali
 às.vezes estrangeiro PL fazer viagem Bali
 ‘Às vezes, os estrangeiros viajam para Bali nas férias.’

A ordem SOV em TP é típica da fala e ocorre quando o falante realiza uma exclamação ou uma resposta surpresa contrária ao que se falava. Dessa forma, o que ocorre é uma incorporação do objeto, pois o objeto (O) é realizado fonologicamente apoiado no sujeito (S), assim como não há a inserção do lexema *mak* ‘TOP’ marcador de topicalização, que é inserido em qualquer operação sintática de mudança na ordem dos constituintes.

	S		O		V
247.	[so	ida	nee]	[feizuada]	[han]
	só	um	este	feijoada	comer

‘Eu só comi feijoada dessa vez!’

	S	O	V
248.	[hau]	[portuges]	[koalia]
	1sg	português	falar

‘Eu falo português (sim)!’

A ordem VO (ou simplesmente V) é comum em TP, pois ela é realizada com os verbos modais, e outros tipos de verbos, que não exigem um sujeito do evento/ação. Entre os verbos modais, há *iha* ‘EXI’, *bele* ‘poder’, *tenke* ‘dever, obrigação’:

	V	O		Comp.
249.	[iha]	[treze distritus]	[iha	timoor]
	EXI	13 distritos	LOC	Timor
	‘Há 13 distritos em Timor’			
	V			
250.	[agora]	[bele han]		
	agora	poder comer		
	‘...agora pode comer!’			

Já a ordem OSV é usada para realizar a ênfase no objeto (O), a topicalização de um elemento (neste caso há a necessidade do emprego do lexema topicalizador *mak* ‘TOP’), e quando se refere a um complemento que já foi mencionado anteriormente e é de conhecimento comum do(s) falante(s) e ouvinte(s):

	O		S	V	
251.	[buat]	nebee mak	[nia]	[hakarak]	
	coisa	REL TOP	3sg	querer	
	‘Isso é o que ele quer!’				
	O		S	V	Comp.
252.	[odamatan	zanela](ami)	[taka]	[metin]	[uluk nebaa]
	porta	janela	1pl.exc	fechar apertar	antes aquele
	‘Naquele tempo, nós fechamos bem as portas e janelas’.				

8.2 O sintagma nominal

O sintagma nominal em TP pode ser formado pelos seguintes elementos: o núcleo (N), o marcador de posse (POS), a negação (NEG), advérbios (Adv.), determinantes (Det.), modificadores (Mod.) e complementos (Comp.). O único elemento obrigatório é o núcleo (N), que pode ser assumido pela classe dos substantivos, dos pronomes pessoais e demonstrativos.

A posição dos elementos no sintagma nominal (SN) do TP pode ser anterior (Ant.) ao N, havendo somente uma posição, e três posições para os elementos posteriores (Pos.) ao N, de acordo com o diagrama (1) abaixo. A seguir será analisado cada um dos elementos do SN:

DIAGRAMA 1. Estrutura do SN

[(Ant.)	Núc.	(Pos.1)	(Pos.2)	(Pos.3)]
POS	Sub	POS	Mod.	Comp.
Adv.	Pron.	Det.	Adv.	
NEG				

8.2.1 A posse (POS) e Sintagma Possessivo (SPos)

A posse (POS) em TP pode ser marcada em posição anterior ou posterior ao N. A POS marcada por *nia* que em posição anterior é um enclítico ligado ao possuidor =*nia*, e em posição posterior é uma forma livre flexionada como *nian*:

	SN	SV	SN	
253.	[ema]	[sunu]	[[amu= nia uma]	
	pessoa	queimar	padre= POS	casa
	[ema]	[sunu]	[amu [uma nian]]	
	pessoa	queimar	padre casa	POS
	‘Alguém queimou a casa do padre’.			

Quando em posição anterior, o sintagma que expressa a posse é outro sintagma nominal, sendo este formado pela estrutura [N (possuidor) + =*nia*]. No exemplo acima, trata-se do Sintagma Possessivo (SPos) [*amu=nia*] ‘do padre’. A outra estrutura do SPos em posição posterior é [N (possuído) + *nian*], ou seja, o núcleo é o nome do objeto possuído mais a forma livre *nian*, no exemplo acima [uma *nian*] ‘casa POS’.

Ainda, o enclítico =*nia* quando usado anexado aos pronomes pessoais possui a função de pronome possessivo. Seguem outros exemplos:

254.	hau=nia	vizinu	maka	moras.
	1=POS	vizinho	TOP	doente

‘Meu vizinho é que está doente.’

255. ita=**nia** naran saa?
2sg=POS chamar o.que
‘Como você se chama?’

256. hau dehan lia-los ba imi, sira simu ona
1sg dizer língua-certo para 2pl 3pl receber ING
sira=**nia** kolen
3pl=POS pagar
‘Eu vos digo sinceramente que eles já receberam suas recompensas.’

257. ba loron idaidak too ona ho
para dia um.por.um chegar ING COM
nia=rasik=**nia** servisu
3sg=ENF=POS trabalho

(não se preocupe com o mal de amanhã, deixe que) ... ‘cada dia chegue a sua própria maneira (com seu próprio trabalho)’

8.2.2 A negação (NEG)

A negação (NEG) de qualquer elemento em TP é sempre em posição anterior ao elemento negado. A negação de elementos nominais em TP é feita por *laos*, que como foi dito, é colocado em posição anterior, no caso do SN, ao N. Nos exemplos abaixo há a NEG do N *nebaa* ‘aquele’ referindo-se a ‘aquelas frutas’ e *ai-nanas* ‘abacaxi’:

SN SV SN SN
258. [hau] [hakarak] [ai-fuan nee], [**laos** nebaa]
1sg querer fruta este NEG aquele
‘Eu quero estas frutas, não aquelas.’

SN SN
259. [buat ida nee] [**laos** ai-nanas]

coisaIND esta NEG abacaxi
 Isto (esta coisa) não é abacaxi!

8.2.3 Os determinantes (Det.)

Os determinantes assumem a posição logo após o núcleo N do SN. As classes que podem assumir essa posição são várias, porém com maior frequência se encontram os demonstrativos e os quantificadores. Ainda, a função dos determinantes basicamente é marcar o nome como um argumento, portanto as muitas estratégias de gramaticalização, descritas em (7.1.1), como indefinição, plural, diminutivo, gênero, entre outras, são também classificadas como Det.

Nos exemplos abaixo pode ser vista a posição de Det. no interior do SN sendo assumida pelas classes de demonstrativos *nee* ‘este, esta, isto’ e *nebaa* ‘aquele, aquela, aquilo’ e quantificadores:

	SN		SV		SN	
260.	[nia]	[hanoin	daudauk]	[buat	nee],	bainhira...
	3sg	pensar	PROG	coisa	este	quando
	‘Ele estava pensando nisso, quando...’					

	SV		SN	
261.	[lori	ba]	[bikan	nee]!
	levar	CNTF	prato	este
	‘Leve (para lá) estes pratos!’			

	SN		SV		SN	
262.	[malae	nebaa]	[foo	aluga]	[hau=nia	uma]
	estrangeiro	aquele	dar	alugar	1sg=POS	casa
	‘Aquele estrangeiro alugou minha casa.’					

		SN		SN		SV
263.	[depoizde	fulan	balu],	[ami=nia	oan]	[filafali]
	depois	mês	alguns	1pl.exc=POS	filho	voltar
	‘Depois de alguns meses, nosso filho voltou.’					

- | | | | | | | | |
|------|----------|-------|------------|--------|--------|---------------|--------------|
| | SN | | SV | | SN | | SN |
| 264. | [hau=nia | ferik | nee | [halo] | [kafee | natoon | [ba ita] |
| | 1sg=POS | velha | este | fazer | café | suficiente | para 1pl.inc |
- ‘Minha esposa aqui fez café suficiente para nós.’

Ainda, seguem exemplos dos elementos gramaticalizados (*mane* ‘MSC’, *feto* ‘FEM’, *oan* ‘DIM’, *ida* ‘IND’, *sira* ‘PL’) assumindo a posição de Det.:

- | | | | |
|------|-----------------------------|--|--------------------|
| | SN | | SV |
| 265. | [labarik mane sira] | | [lalika estuda]. |
| | criança MSC PL | | não.querer estudar |
- ‘Os meninos não querem estudar.’

- | | | | | | | | |
|------|--------------------------|--|--------|--|-------------|---------|---------|
| | SN | | SV | | SN | | SV |
| 266. | [estudante sira] | | [hola] | | [livru][atu | estuda] | |
| | estudante PL | | ganhar | | livro | IRR | estudar |
- ‘Os estudantes ganharam livros para estudar.’

- | | | | | | | |
|------|-------|-------|-----------|--------------|------|-------|
| | SN | SV | | SN | | Conj. |
| 267. | [hau] | [iha] | [oan feto | [nain tolu]] | ho | |
| | 1sg | EXI | cria FEM | CL | três | e |
- SN
[oan **mane** [nain rua]].
cria MSC CL dois

‘Eu tenho três filhas e dois filhos’

- | | | | | | | | |
|------|-----------|----------------|-------|-------|------|---------|--|
| | SN | | SV | | SN | | |
| 268. | [lafaek | oan ida | [sai] | [hosi | [nia | fatin]] | |
| | crocodilo | DIM IND | sair | de | 3sg | lugar | |
- [atu buka] [hahaan]
IRR procurar comida
- ‘Um crocodilozinho saiu de sua toca para procurar comida.’

8.2.4 Os modificadores (Mod.)

Optei por diferenciar modificadores (Mod.) de determinantes (Det.) por diversos fatores que se apresentaram no decorrer da análise dos dados. O primeiro deles é a clara diferença em relação à posição que eles assumem dentro do SN, sendo o(s) Det. colocado logo após o N, enquanto o(s) Mod. assumem a posição após o(s) Det. e somente são realizados após o N na ausência de Det. O segundo fator consiste na estrutura dos Mod. e Det.: os Mod. são limitados somente a classe dos adjetivos e quantificadores numerais, assim como seus respectivos sintagmas, o Sintagma Adjetival (SAdj) e o Sintagma Numeral (SNum); os Det. podem ser das mais variadas classes em TP e assumem a estrutura de sintagma somente em um único caso específico o do Sintagma Relativo (SRel), que será analisado mais abaixo.

No exemplo abaixo, o SN [*malae sira*] ‘os estrangeiros’ apresenta o Det. *sira* ‘PL’ e os SNs [*kabun boot*] ‘barriga grande’ e [*ema timoor*] ‘timorense’ apresentam o adjetivo *boot* e o substantivo *timoor* (como adjetivo) ambos com função de Mod.:

	SN		SV	SN		SV	SN
269.	[malae	sira	[iha]	[kabun boot]	[liu]	[ema	timoor]
	estrangeiro	PL	EXI	barriga grande	passar	pessoa	Timor
	‘Os estrangeiros têm barriga maior do que a dos Timorenses’.						

Outro exemplo que serve para corroborar o que foi dito a respeito das diferenças entre Det. e Mod. é o que segue, onde fica evidente a posição, a forma e a função diferentes que eles assumem dentro do SN:

	N	Det.	Mod.	
270.	[feto	[nebaa]	[kapaas	loos]]
	mulher	aquela	bonita	muito
	‘aquela mulher é bonita demais’			

Ainda, os Mod. podem ser representados por sintagmas menores dentro do SN, sendo os SAdj e SNum, conforme os exemplos abaixo:

	SAdj		SAdj	
271.	[[malae	australia]	[aas demais]]	
	estrangeiro	Austrália	alto	demais

‘Os australianos são muito altos.’

	SN	SV	SNum	Conj.	SNum
272.	[[uma=nia] laran]	[iha]	[sala rua]	ho	[kuartu tolu]
	casa=POS	dentro	EXI sala dois	e	quarto três
	‘Dentro da minha casa há duas salas e três quartos.’				

	SN	SAdj		
273.	[ferik nee [nain neen]].			
	velha este	CL.HUM	seis	
	‘Estas seis senhoras.’			

8.2.5 Os complementos (Comp.)

Os complementos (Comp.) são SNs menores inseridos dentro de um SN maior, fazendo uma referência ao núcleo N do SN maior. Os Comp., assim como os demais elementos do SN, não são obrigatórios. Em muitos casos, consistem em outros tipos de sintagmas que são analisados como elemento interno do SN pelo fato de terem sido realizados sem os elementos de ligação como preposição, conjunção, ou até um adverbio.

	SN	SV	SN	Comp.
274.	[sira]	[halai]	[ba Suai]	[(ho) Indonesia]
	3pl	correr	para Suai	e Indonésia
	‘Eles correram para Suai e para Indonésia’.			

	SN	SV		SN	Comp.
275.	[nia]	[tenke sosa]		[fahi manu [animaal seluk tan]]	
	3sg	tem	comprar	porco pássaro animal	outro mais
	‘Ele tem que comprar porcos, pássaros e outros animais.’				

8.3 O sintagma verbal

O sintagma verbal em TP pode ser formado pelos seguintes elementos: o núcleo (V), os marcadores de tempo, aspecto e modo (TAM), a negação (NEG), verbos

auxiliares das construções de verbos seriais, os marcadores de reflexividade (Refl.) e reciprocidade (Rec.), marcadores direcionais (Dir.) e advérbios (Adv.). O único elemento obrigatório é o núcleo (V), que pode ser assumido somente pela classe dos verbos.

A posição dos elementos no sintagma verbal (SV) do TP pode ser anterior (Ant.) ao V, havendo somente uma posição, e duas posições para os elementos posteriores (Pos.) ao V, de acordo com o diagrama (2) abaixo. A seguir será analisado cada um dos elementos do SV:

DIAGRAMA 2. Estrutura do SV

[(Ant.)	Núcleo	(Pos.1)	(Pos.2)]
TAM	V	TAM	Adv.
NEG		VSer.	Dir.
VSer.		Reflex.	
		Recip.	

8.3.1 Marcadores pré-verbais de TAM

Conforme foi analisado em (7.1.1), os elementos que marcam TAM em TP são lexemas que estão em vias de gramaticalização. Esses marcadores são usados tanto antes do verbo, como após verbo. Os marcadores pré-verbais de TAM são: o marcador de tempo futuro *sei* ‘FUT’; o marcador de aspecto prospectivo *foin* ‘PROS’; o marcador de modo irrealis *atu* ‘IRR’:

	SN		SV
276.	[efeitu sira	[husi krize]]	[sei sente nafatin]
	efeito PL	ABL crise	FUT sentir ITE

SN

[iha timoor rai]

LOC timor terra

‘Os efeitos da crise serão sentidos continuamente em terras leste-timorenses.’

	SN	SV	SN
277.	[hau]	[sei halo]	[kafee [mai ami]]

1sg FUT fazer café para 1pl.excl
 ‘Eu farei café para nós.’

	SN		SV	Conj.	SN
278.	[hau=nia	oan feto]	[foin	mate],	maibee [ita-boot]
	1sg=POS	cria FEM	PROS	morrer	mas 2sg-grande
	SV		SN	SPrep	SN SV
	[hiit=an	mai tau]	[liman]	[ba nia],	[nia] [sei moris]
	erguer=REF	CNTP	colocar mão	LOC 3sg	3sg FUT nascer

‘Minha filha acabou de morrer, mas vós ergueis e repousais nela vossa mão (que) ela reviverá (nascer).’

279.	[[estudante sira]	[tenke	baa]	[profesoor]	ru-rua]
	estudante PL	tem	ir	professor	dois-RED
	[atu hatudu]	[tarefa]			
	IRR	mostrar	tarefa		

‘Os estudantes devem ir ao professor de dois em dois para mostrar a tarefa.’

280.	[ami]	[presiza	atu	estuda][pasta	ho	livru]
	1pl.exc	precisar	IRR	estuda mochila	e	livro

Nós precisamos de mochila e livros para estudar.’

8.3.2 A negação verbal

A negação (NEG) de qualquer elemento em TP, como foi analisado anteriormente, é sempre em posição anterior ao elemento negado. A negação verbal é marcada pelo proclítico =*la*. Nos exemplos abaixo há a NEG de somente um verbo (281) e de verbos seriais e modais (282 e 283):

	SN	SV		SV		SV	SN
281.	[hau]	[la=hatene]		[koalia neineik],		[halo	[favor]]
	1sg	NEG=entender		falar devagar		fazer	favor

‘Eu não estou entendendo. Fale mais devagar, por favor!’

	SV		SN		
282.	[la=bele	doko]	[lata	hanesan	nee]
	NEG=poder	balançar	lata	igual	isto
	'Não balance a lata desse jeito.'				
	SN	SV	Conj.	SV	SN
283.	[ita]	[tenke sai	agora], tanba	[la=iha]	[tempu]
	1pl.inc	tem	sair	agora	porque
				NEG=EXI	tempo
	'Nós temos que sair agora, porque não há muito tempo.'				

8.3.3 A posição dos verbos seriais e marcação de direção

Conforme analisei em (7.3.3.1) os verbos seriais, há uma série de construções em TP que se utilizam da serialização verbal para expressar diversos tipos de relações. Williams-van Klinken, Hajek e Nordlinger (2001, 2002) classificam essas relações em movimento-ação, movimento-direção, instrumental, e outras duas que os autores citam, mas não classificam. Essas duas podem ser classificadas como copulativa, com dois verbos compartilhando o mesmo sujeito, e temporal, com os verbos expressando eventos simultâneos ou que ocorreram um após o outro.

O sintagma verbal (SV) das construções com verbos seriais do TP pode possuir, desta forma, dois verbos dentro de sua estrutura. Desses dois verbos, somente um ou os dois podem assumir a função de núcleo V do SV, de acordo com a relação semântica entre os verbos seriais.

	SN	SV	SV	Int.	
284.	[ita]	[baa	selu]	[nia] ka	
	2sg	ir	pagar	3sg INTG	
	'Você foi e pagou-o?' ou 'você vai pagá-lo?'				
	SN	SV	SN		
285.	[hau]	[tiha	foo	han]	[labarik]
	1sg	já	dar	comer	criança
	'Eu já alimentei a criança' ou 'eu já dei de comer a criança'				

- | | | | | | | |
|------|-----------|----------------------|------|----------|-----------|--------|
| | SN | SV | | SPrep. | | SN |
| 286. | [nia] | [halai sae] | [too | [foho | laelaku]] | [besik |
| | 3sg | correr subir | até | montanha | Laelaco | perto |
| | [frontera | indonesia]] | | | | |
| | fronteira | Indonésia | | | | |
- ‘Ele subiu (correndo) até a montanha Laelaco, (que fica) perto da fronteira indonésia.’

- | | | | | | | | |
|------|--------|---------------------|------|-------|--------|---------|----|
| | SN | SV | | SN | | SV | SN |
| 287. | [mana] | [foti mai] | [ida | tan], | [halo] | [favor] | |
| | irmã | trazer CNTP | um | mais | fazer | favor | |
- ‘Irmã, traga mais um, faz favor!’

- | | | | |
|------|--------------------|--------|------------|
| | SV | | SN |
| 288. | [lori ba] | [bikan | nee] |
| | levar | CNTF | prato este |
- ‘Leve (para lá) estes pratos!’

8.3.4 Marcadores pós-verbais de TAM

Conforme foi analisado em (7.1.1) e comentado acima, os elementos que marcam TAM em TP são lexemas que estão gramaticalizados ou em vias de gramaticalização. Esses marcadores são usados tanto antes do verbo, como após verbo. Após a análise feita anteriormente dos marcadores pré-verbais de TAM, serão apontados os marcadores pós-verbais de TAM, que se limitam na realidade a uma série de marcadores aspectuais e o tempo ‘perfeito’: perfectivo *tiha* ‘PERF’, ingressivo *ona* ‘ING’, há a junção dos dois anteriores *tiha ona* que pode ser analisado como perfeito ‘PF’, egressivo *hotu* ‘EGR’, perfeito recente *kedas* ‘PF.R’, demarcativo *lai* ‘DEM’, semelfactivo *hela* ‘SEM’, progressivo *daudauk* ‘PROG’, iterativo *nafatin* ‘ITE’:

- | | | | | | | |
|------|--------|---------------|--------|------------|--------------|-------------|
| | SV | | SN | | SN | SV |
| 289. | [haree | tiha] | [ema | lubun | boot], | [zezus][sae |
| | ver | PERF | pessoa | quantidade | grande Jesus | subir |
- SV

- ba] [foho ida]
 CNTF montanha IND
 ‘Quando viu a multidão, Jesus subiu à montanha.’
290. SN SV SN SN SPrep
 [nia] [halo **ona**] [sala-foer] [ho nia] [iha [laran nian]]
 3sg fazer ING pecado DAT 3sg LOC coração POS
 ‘Ele cometeu pecado em relação a ela (sua esposa) em seu interior.’
291. SN SV SN SV SN
 [mestri] [haruka **tiha-ona**] [labarik sira][tuur ba] [rai]
 professor mandar PF criança PL sentar para terra
 ‘O professor mandou as crianças sentarem no chão.’
292. [nia] [tanis **daudauk**] [iha [kuartu nian]]
 3sg chorar PROG LOC quarto POS
 ‘Ela está chorando em seu quarto.’
293. SN SV SPrep SV SN SV
 [o] [sei la=sai] [husi nebaa][too] [o] [selu **hotu**].
 2sg FUT NEG=sair ABL aquele chegar 2sg pagar EGR
 ‘Tu sairás de lá somente quando terminares de pagar.’
294. SN SV SN SV SPrep
 [sira] [husik **kedas**][dai] [hodi lao] [tuir nia]
 3pl abandonar PF.R rede dirigir andar seguir 3sg
 ‘Em seguida, eles abandonaram suas redes e deixaram-se serem conduzidos e seguiram-no.’
295. SN SV SPrep SN SV SPrep
 [imi] [hela **lai**] [ihanee]. [hau] [sei baa] [iha nebaa]
 2pl ficar DEM aqui 1sg FUT ir lá
 SPrep

[halo orasaun]

fazer oração

‘Vós fiqueis aqui. Eu irei até lá para orar.’

296. SN SV SN SPrep
[sira] [soe **hela**] [dai] [ba tasi]
3sg lançar SEM rede para mar
‘Eles lançaram suas redes ao mar.’

297. SN SPrep SV
[efeitu sira] [husi krize] [sei sente **nafatin**]
efeito PL ABL crise FUT sentir ITE

SPrep

[iha timoor rai]

LOC Timor terra

‘Os efeitos da crise serão sentidos continuamente em terras leste-timorenses.’

8.3.5 Reflexivos e Recíprocos

A marcação de reflexividade e reciprocidade em TP são usadas com pouca frequência, pois na maioria das sentenças, de acordo com o contexto, é possível para o falante inferir semanticamente se o verbo está sendo usado com sentidos reflexivos, recíprocos ou não.

A marcação de reflexividade é feita pelo enclítico =*an*, que pode ser usado em duas posições: após o verbo marcado como reflexivo (298), ou no final da oração (299). Já o marcador de reciprocidade *malu* somente é usado após o verbo (300-301):

298. SN SV Conj. SN
[hau=**nia** oan feto] [foin mate], maibee [ita-boot]
1sg=POS cria FEM PROS morrer mas 2sg-grande
SV SN SPrep SN SV
[hiit=**an** mai tau] [liman] [ba nia], [nia] [sei moris]

erguer=REF CNTP colocar mão para 3sg 3sg FUT nascer
 ‘Minha filha acabou de morrer, mas vós ergueis e repousais nela vossa
 mão (que) ela reviverá (nascer).’

	SN		SV		SV	
299.	[mane	nebaa]	[tenta	tiha ona]	[koa	rasik=nia= an]
	homen	aquele	tentar	PF	cortar	ENF=POS=REF

‘Aquele homem tentou se cortar (a si mesmo).’

	SN		SV		Conj.	SN
300.	[sira]	[la=bele	baku	malu]	ho	[katana]
	3pl	NEG=poder	lutar	REC	com	espada

‘Eles não podem brigar (entre si) com espada.’

	SN		SV		SN	
301.	[dosenti sira]	[diskuti	malu]	[depois	klase]	
	professor	PL	discutir	REC	depois	aula

‘Os professores discutiram (entre eles) depois da aula.’

8.4 Demais sintagmas

Há diversas outras construções em TP que são marcadas para expressar diversas estratégias discursivo-pragmáticas. Algumas foram analisadas anteriormente, como a mudança na ordem dos constituintes e a omissão de um deles, outras, porém, não serão contempladas no presente trabalho. Nesta seção, analisarei certas construções que são, de certa forma, orientadas pelo discurso. São elas: as estratégias de relativização, enfatizando as orações relativas (8.4.1); as orações interrogativas (8.4.2); e a modificação adverbial (8.4.3).

8.4.1 As orações relativas

O único trabalho que já contemplou de alguma maneira a relativização em língua Tetun foi o artigo de Eccles (2001) que faz uma análise do relativizador *nebee*, elaborando uma série de testes linguísticos para tal classificação devido à dificuldade de se classificar certos elementos da gramática do TP. Porém, pouco foi dito sobre os tipos

de orações relativas, ou estratégias de relativização, e possíveis testes para se identificá-las.

Para a análise das estratégias de relativização em TP partiu-se do princípio básico de que a função primeira da oração relativa, então, é a restrição (modificador) de um núcleo N específico. As formas de codificar essa restrição, porém, nas línguas do mundo são as mais variadas. As línguas podem codificar a relativização através de estratégias perifrásticas, morfológicas, sintáticas e/ou morfossintáticas, sendo que ainda as línguas podem apresentar estruturas específicas para a relativização, ou utilizar outras estruturas já presentes na língua, o que torna a relativização ambígua.

A bibliografia utilizada foi Comrie (1989), Comrie e Horie (1995), Givón (2001), Keenan e Comrie (1977) e Keenan (1985). Segundo Givón (2001), as orações relativas são aquelas que possuem a função de modificador do nome (o N de um SN) e, ainda, essa mesma oração codifica um estado ou evento com um participante correferencial com o N por ela modificado. Dessa maneira, a oração relativa possui uma função também de indicar para o ouvinte um referencial e, assim, facilitar a identificação cognitiva do referente. Comrie (1989) propõe que a oração relativa é uma estrutura que possui uma função de restringir as possibilidades de referência de um núcleo nominal específico, e em Comrie e Horie (1995) os autores, analisando o Japonês e o Khmer⁴³, afirmam, baseados nos dados dessas duas línguas, que nem todas as línguas do mundo fazem uma distinção entre Sintagma Relativo (SRel) e orações complementos. Ainda, Keenan e Comrie (1977) e Keenan (1985) definem a oração relativa como uma oração que funciona como modificador atributivo de algum nome.

Segundo os dados coletados, foi possível identificar em TP duas estratégias de relativização: o emprego do *nebee* 'REL' como relativizador no início do Sintagma Relativo (SRel) para marcar as orações relativas simples⁴⁴ (302-303), o REL *nebee* pode ainda ser acompanhado de uma série de elementos gramaticalizados a ele para marcar outros tipos de relativas, como a relativa genitiva e relativa sem sujeito; e simplesmente orações justapostas com traços semânticos de relativas também expressando o conteúdo semântico das relativas simples (304-305):

302. [[asu] [nebee tata] [mane nee]] [toba ona]

⁴³ Língua nacional do Camboja e falada por mais de 13 milhões de pessoas, pertencente à família linguística Mon-Khmer (Goddard, 2005, p.32).

⁴⁴Há variação no uso do *nebee* 'REL', sendo realizado como *be* 'REL' em registros formais, e como *nebee-ke* 'REL' ou simplesmente *ke* 'REL' em registros informais, devido à influência lusófona.

cão REL morder homem este dormir
 ‘O cão que mordeu o homem começou a dormir’.

303. [hau] [haree][[mane][nebee sunu]] [hau=nia uma foun]
 1sg ver homem REL queimar 1sg=POS casa nova
 ‘Eu vi o homem que queimou minha casa nova’.

304. [asu] [tata] [mane nee] [toba ona]
 cão morder homem este dormir EGR
 ‘O cão que mordeu o homem começou a dormir’.

305. [hau] [haree] [mane] [sunu] [hau=nia uma foun]
 1sg ver homem queimar 1sg=POS casa nova
 ‘Eu vi o homem que queimou minha casa nova’.

No caso das orações justapostas, não há orações relativas. Apenas duas, ou mais, orações dependentes sendo interpretadas pelo falante como havendo uma relação de relativização entre elas. Desta maneira, as orações justapostas caracterizam-se como uma estratégia de relativização. Somente os sintagmas introduzidos pelo relativizador *nebee* é que consistem em orações relativas, e conseqüentemente o sintagma introduzido pelo REL como SRel.

Ainda, outros tipos de orações relativas, como relativas sem sujeito e relativas genitivas, são marcadas pela inserção do *nebee* ‘REL’ juntamente com elementos gramaticalizados a ele, como *buat* ‘coisa’ (308) ou *ema* ‘pessoa’ (309) para expressar o sujeito correferencial da relativa sem sujeito, e o enclítico =*nia* que marca posse nas relativas genitivas (310):

306. hau hemu daudauk bee [nebee manas].
 1sg beber PROG água REL quente.
 ‘Eu estou bebendo a água que está quente’.

307. mane [nebee asu tata] nee toba tihaona.
 homem REL cão morder este dormir PF

‘O homem que o cachorro mordeu estava dormindo’.

308. hau han daudaun buat-nebee hau=nia inan tunu
1sg comer PROG coisa-REL 1sg=POS mãe assar
‘Eu estou comendo o que a minha mãe assou’.

309. hau haree ona ema-neebee nia oho
1sg ver INC pessoa-REL 3sg matar
‘Eu vi quem ele matou’.

310. labarik ida-nebee=nia bisikleta hau ferak ba uma
criança SG-REL=POS bicicleta 1sg quebrar para casa
‘A criança cuja bicicleta eu quebrei foi pra casa’.

No exemplo (308) não está expresso o que foi comido, apenas sabe-se que foi algo que *hau=nia inan tunu* ‘minha mãe assou’. Assim como no exemplo (309), não está expresso também quem foi que sofreu a ação, já que *nia oho* ‘ele matou’, sabe-se somente que *nia* ‘ele’ matou alguém. É possível argumentar-se que a posição de sujeito é preenchida pelos nomes *buat* ‘coisa’ e *ema* ‘pessoa’, nos exemplos (308) e (309) respectivamente, mas tal argumento não é válido, pois as relativas em Tetun Prasa são pós-nominais, a posição de sujeito realmente encontra-se vazia e, finalmente, a análise da função desses dois nomes nos revela que se tratam na realidade de nomes que assumem uma função classificadora. Seguem outros exemplos:

311. hau hemu buat-nebee ita haruka.
1sg beber coisa-REL 2sg mandar
‘Eu bebi o que você mandou’.

312. nia hakoi ema-neebee mate tihaona.
3sg enterrar pessoa-REL morrer PF
‘Ele enterrou aquele que morreu’.

As relativas genitivas possuem o mesmo comportamento sintático que as relativas sem sujeito, ou seja, elas são pós-nominais, introduzidas pelo relativizador *nebee* e este recebe, além da marcação =*nia* ‘POS’, como foi dito acima, os lexemas *ida* ‘um’ e *sira* ‘3pl’ gramaticalizados como marcadores de ‘singular’ e ‘plural’ respectivamente, conforme o exemplo abaixo:

313. mane ida-nebee=*nia* uma ami sunu
 homem SG-REL=POS casa 1pl.exc queimar
 mai tiha-ona
 chegar PF

‘O homem cuja casa nós queimamos chegou’.

As relativas genitivas têm a função de expressar que o NRel da oração principal é possuidor do N a que se refere na construção relativa. No exemplo (313), o NRel é *mane* ‘homem’ e este NRel é possuidor do N referente uma ‘casa’. As relativas nos exemplos acima, sempre pós-nominais e introduzidas pelo relativizador, apresentam outros processos morfológicos que se aplicam ao *nebee*, esses processos, porém, são similares ao das relativas sem sujeito.

Sobre as funções das orações relativas em Tetun é possível afirmar que não há distinção entre o N relativizado com função de sujeito e o N relativizado com função de argumento de um verbo. Elas são pós-nominais e possuem também função atributiva, além da função restritiva.

As relativas justapostas são explicadas quando analisadas no nível discursivo-pragmático, assim como muitos outros fenômenos da sintaxe do TP. A estratégia de parataxe para expressar a relativização é orientada pelo discurso e pela pragmática, mas ela é usada de maneiras distintas no nível pragmático e no nível discursivo.

314. mane nebee asu tata nee toba tihaona
 homem REL cão morder ANA dormir PF
 ‘O homem que o cachorro mordeu dormiu.’

315. mane asu tata toba tihaona.
 homem cão morder dormir PF

‘O homem que o cachorro mordeu dormiu’.

Em (314) há a necessidade de se usar o REL pelo falante para especificar ao ouvinte quem é *mane* ‘homem’, ou seja, sobre quem é que se está falando. Assim como, faz-se necessário também falar de maneira específica sobre a relativa inteira *asu tata* ‘o cachorro mordeu’, já que o ouvinte não sabe a que o falante está se referindo, por isso minha análise do *nee* em (314) ser de dêitico anafórico ‘ANA’, pois é necessário fazer a referência discursiva *nee* ‘ANA’ como algo que já foi mencionado, em outras palavras, aponta-se ‘este, ou aquele’ a que se refere’. Dessa maneira, há uma diferença quando *nee* é usado com a função de pronome demonstrativo equivalente a português ‘este, esta, isto’, e de quando possui a função de um dêitico. Já em (315) o falante e o ouvinte sabem quem, ou o que, está sendo referido no diálogo, desta maneira não há necessidade de se fazer nenhuma referência, logo não se usa o *nebee* ‘REL’, nem *onee* ‘ANA’.

Pode-se concluir, então, que o uso das estratégias de relativização em TP está orientado pelo discurso, já que o relativizador é usado somente quando se insere uma informação nova e quando o ouvinte não sabe a quem, ou ao que, esta informação nova se refere. Caso seja uma informação antiga e/ou algo que já foi mencionado no discurso não se usa o REL para introduzir a relativa.

8.4.2 As orações interrogativas

As orações interrogativas em TP são formadas pela inserção dos lexemas *ka lae* ‘ou não’, que sofre variações como *ka ~ ga* ‘ou’, no final do sintagma. Ainda, as orações interrogativas podem ser formadas pela com a presença de um elemento da classe dos interrogativos, que possuem uma sintaxe de ordem diferente da língua.

316. ita=nia inan moras **ka** **lae**
1pl.inc =POS mãe doente ou NEG
‘Sua mãe está doente (ou não)?’
317. ita baa selu nia **ka**
2sg ir pagar 3sg ou
‘Você foi e pagou-o?’ ou ‘você vai pagá-lo?’
318. nee **saa**-ida

isto o.que-IND

‘O que é isto?’

319. apa, senoor mai **horibainhira**
pai senhor vir quando (passado)
‘Pai, quando o senhor chegou?’

A variação nas orações interrogativas básicas consiste na redução do composto *ka lae* ‘ou não’ para simplesmente *ka* ‘ou’ ou sua forma mais vernácula realizada sonorizada e de forma enclítica ao elemento final =*ga* ‘ou’⁴⁵.

320. kolisensa, hau bele husu buat-ida ba ita **ka**
com licença 1sg poder perguntar coisa-IND para 2sg ou
‘Com licença, posso perguntar algo a você?’

321. kolega bele hanorin lia-tetun mai hau =**ga**
colega poder ensinar língua-tetun para 1sg=INTG
‘O colega pode ensinar a língua tetun para mim?’

As orações interrogativas formadas com elementos da classe dos interrogativos (as perguntas QU), como *saa* ‘o que’, *see* ‘quem’, *hira* ‘quanto’, *bainhira* ‘quando’, possuem a ordem invertida com o lexema pertencente a esta classe sendo colocada no final da oração:

322. ita koalia ho **see**
2sg falar com quem
‘Com quem você falou?’

323. ita=nia naran **saa**
2sg=POS chamar o.que
‘Como você se chama?’

⁴⁵ A frequência e distribuição exata das três variantes (*ka lae* ~ *ka* ~ *ga*), assim como os fatores linguísticos e não linguísticos que as condicionam, necessitam um estudo separado que não foi realizado até a atualidade.

324. apa, senoor mai **horibainhira**
 pai senhor chegar quando
 ‘Pai, quando o senhor chegou?’
325. o hela iha **nebee**
 2sg morar LOC onde
 ‘Tu moras onde?’
326. ola irmaun agora halo **saa-ida**
 olá irmão agora fazer o.que-IND
 ‘Olá irmão, o que você está fazendo agora?’
327. kolega baa **nebee**
 colega ir (a)onde
 ‘Colega, aonde vai?’

Há a possibilidade de movimento do lexema interrogativo para o meio ou início da oração como estratégia pragmático-discursiva para enfatizar, ou topicalizar alguma das categorias sintáticas, ou ainda para marcar a pergunta como uma surpresa ou espanto do falante em relação à informação nova recebida durante o discurso. Esse movimento do interrogativo, porém, é marcado com o topicalizador *mak* ‘TOP’ colocado logo após a ele:

328. bainhira **mak** nia fila-fali mai hosi Baukau
 quando TOP 3sg voltar CNTP de Baucau
 ‘Ele voltará de Baucau quando?’
329. tuku hira **mak** ami aranka
 hora quanto TOP 1pl.exc partir
 ‘Nós partimos a que horas?’
330. mana, hira **mak** nee
 irmã quanto TOP isto
 ‘Moça, isto custa quanto?’

8.4.3 A modificação adverbial

A análise do advérbio, e como é feita a modificação adverbial, foi deixada por último pelo fator funcional de que a modificação adverbial é relativamente complexa, pois o conteúdo semântico do advérbio pode modificar os mais diversos elementos, como: o núcleo dos sintagmas, um sintagma inteiro, mais de um sintagma e, ainda, marcar relações entre sintagmas e/ou orações.

A análise de Williams-van Klinken, Hajek e Nordlinger (2002, p.110) é que o advérbio apresenta-se como uma ‘oração adverbial’ (ing. *adverbial clause*), simples ou composta, que modifica somente outras orações de acordo com conteúdos semânticos prototípicos, como tempo, finalidade, concessão e condição. Dessa forma, minha análise difere da anterior, pois analiso o advérbio como modificador de uma série de elementos sintáticos, assim como parte dos sintagmas, sejam eles SN ou SV. Por isso, inseri o advérbio também como parte do interior do SN e do SV, nos diagramas apresentados acima, a seguir apresento argumentos para justificar minha análise.

No exemplo (331), o advérbio *nafatin* ‘sempre’ faz parte do SN e modifica o núcleo deste *uma-nain* ‘pai de família’, sendo posicionado antes do N; em (332) *beibeik* ‘sempre’ está posicionado depois do N *lia-bosok* ‘mentira’; já no exemplo (333) abaixo há o sintagma adverbial (SAdv) modificando o SN anterior *isin manas* ‘febre’⁴⁶:

	SN		SN	
331.	[aman]	[nafatin	uma-nain]	
	pai	sempre	casa-dono	
	‘O pai sempre é o chefe da casa (pai de família)’			
	SN	SV		SN
332.	...[ema]	[toman]	[lia-bosok	beibeik],
	pessoa	acostumar	língua-mentir	sempre
	SN	SV	SN	SN
	[ami]	[bolu]	[ema	nee] [bosok-teen]
	1pl.exc	chamar	pessoa	esta mentir-AGT

⁴⁶ Apesar de apresentar alguns argumentos e os dados linguísticos que os corroboram, há certos traços da classe dos advérbios que permanecem um tanto problemáticos e necessitam de um estudo mais aprofundado separadamente, como o caso da função do advérbio no interior dos SNs.

‘...pessoas que estão acostumadas a mentir, nós chamamo-las de mentirosas.’

- | | | | | |
|------|---|---------------|--------------------|------|
| | SN | SN | | SAdv |
| 333. | [nia] | [[isin manas] | [too horisehik]] | |
| | 3sg | pele | quente desde ontem | |
| | ‘Ele está com febre (pele quente) desde ontem.’ | | | |

A modificação do SV feita pelo advérbio pode ser vista nos exemplos (334) e (335). Em (334), o advérbio *neineik* ‘devagar’ está modificando o verbo *koalia* ‘falar’, e em (335) o advérbio *beibeik* ‘constantemente’ modifica o verbo *baku* ‘bater’.

- | | | | | |
|------|---|--------------|---------|-----------------|
| | SN | SV | | SV |
| 334. | [hau] | [la=hatene]. | [koalia | neineik] |
| | 1sg | NEG=entender | falar | devagar |
| | ‘Eu não estou entendendo. Fale mais devagar.’ | | | |

- | | | | | | |
|------|---|------------|------------------|---------|---------|
| | SN | | SV | | SAdv |
| 335. | [labarik sira] | [baku malu | beibeik], | [depois | eskola] |
| | criança PL | bater REC | constantemente | depois | escola |
| | ‘As crianças brigam entre elas constantemente, depois da escola.’ | | | | |

O advérbio possui a função também de modificar a oração inteira, ou seja, um conjunto de sintagmas ou um sintagma como um todo. Em (336), o advérbio *uluk* ‘antes, antigamente’ pode ser interpretado como modificando o SN posterior a ele *tempu funu* ‘tempo de guerra’ ou modificando a oração inteira; em (337) o advérbio *dala ruma* ‘às vezes’ está modificando a oração inteira; e em (338) o advérbio *ohin* ‘hoje’ pode ser interpretado como modificador do SV *too* ‘chegar’ ou como modificador da oração inteira, porém nesse exemplo a interpretação tende a ser de modificador da oração devido a sua posição no final da sentença, caso a interpretação seja de modificador do SV o advérbio estaria logo após ao núcleo V. Ainda, quando o advérbio modifica o sintagma inteiro ou orações distintas muitas vezes ele é independente, sendo

analisado sintaticamente somente o advérbio como compondo o sintagma adverbial (SAdv):

	SAdv	SN	SV	SN	
336.	[uluk]	[tempu	funu], [iha]	[mane	sarani]
	antes	época	guerra EXI	homem	coragem
	‘Antigamente, no tempo das guerras, havia homens corajosos.’				

	SAdv	SN	SV	SN	SPrep
337.	[dala ruma]	[timoorrai]	[hela]	[manasdemais]	[ba malae]
	às.vezes	Timor terra	ficar	quente demais	para estrangeiro
	‘Às vezes Timor fica quente demais para os estrangeiros.’				

	SN	SPos	SV	SPrep	SAdv
338.	[funsionarius [nasoos unidas nian]]	[too]	[iha timoor]	[ohin]	
	funcionários	nações unidas POS	chegar LOC	Timor	hoje
	‘Os funcionários das Nações Unidas chegaram em Timor hoje.’				

Finalmente, o último tipo de modificação que o advérbio pode realizar em TP é a modificação entre orações. Esta função do advérbio pode ser um tanto ambígua para a classificação, pois há alguns casos, como em (341), que conjunções prototípicas podem funcionar como advérbios, assim como os demonstrativos e os quantificadores que, dependendo de sua função e posição na sentença, também podem ser analisados como advérbios. No exemplo (339), *nunee* ‘por isso, por este motivo’ é um advérbio prototípico que marca a relação entre orações, devido ao seu significado específico. Já nos exemplos (340) e (341) o SAdv é formado somente por um elemento, o próprio advérbio, porém por lexemas que fazem parte de outras classes, como *bainhira* ‘quando’ e *tanba* ‘porque’ e marca a relação entre as orações:

339.	hau	halo	viazen	ba	estranjeiru.	nunee, hau=nia
	1sg	fazer	viagem	para	estrangeiro	por.isso1sg=POS
familia	tenki	tau	matan	ami=nia	oan	
família	tem	colocar	olho	1pl.exc=POS	filho	

‘Eu viajarei para o estrangeiro. Por isso, minha família deverá cuidar dos nossos filhos.’

	SN	SV	SAdv	SPos
340.	[hau]	[bele toba deit],	[bainhira]	[hau=nia
	1sg	poder dormir somente	quando	1sg=POS
	SV	SN		
	oan]	[mai] [uma]		
	filho	vir casa		

‘Eu consigo dormir somente quando meu filho chega em casa.’

	SN	SV	SAdv	SV	SN
341.	[ita]	[tenke sai agora],	[tanba]	[la=iha]	[tempu]
	1pl.inc	tem sair agora	porque	NEG=EXI	tempo

‘Nós temos que sair agora, porque não há muito tempo.’

9. CONCLUSÕES

No presente trabalho, procurei elaborar um esboço gramatical da variedade Tetun Prasa, língua oficial de Timor-Leste, considerando os dados coletados por mim em campo, juntamente com as análises anteriores feitas por Hull (2002), Hull e Eccles (2001) e Williams-van Klinken, Hajek e Nordlinger (2002). Porém, a presente descrição gramatical foi além das contribuições anteriores, identificando problemas específicos da fonologia, analisando de maneira mais esquemática a morfologia, e apontando a estrutura dos sintagmas e a ocorrência destes sendo orientada discursiva e pragmaticamente.

Antes de entrar nos níveis de análise linguística do Tetun Prasa, elaborei um breve histórico do contato entre os diversos povos na ilha de Timor. Esses vários contatos são de fundamental importância para compreensão do estado atual das línguas faladas em Timor-Leste, pois devido a eles verificaram-se fenômenos de reestruturação gramatical e da ascensão da língua Tetun, em sua variedade Tetun Prasa, à língua franca de grande parte do território da ilha de Timor. Após este histórico, elenquei questões sociolinguísticas básicas sobre Timor-Leste, como o status das línguas nativas faladas atualmente, a educação formal, as situações de multilinguismo e a ecolinguística. Para finalizar, analisei as variedades da língua Tetun faladas em Timor, assim como os diferentes registros e usos da língua nos mais diversos tipos de interações sociais.

O léxico do Tetun Prasa apresenta três camadas distintas. A primeira delas, a camada austronésia, trata-se do núcleo genético da língua que, apesar de intenso contato no decorrer de sua história, manteve-se relativamente intacto com grande parte do vocabulário consistindo de retenções da proto-língua. A camada malaia consiste de empréstimos antigos, via *pazar melayu*, e recentes, via *bahasa indonesia*, de alguma das variedades citadas da língua malaia. Já a camada lusófona revelou-se extensa e também com relativa profundidade temporal, revelando influências tanto dos crioulos de base lexical portuguesa, como o CPMal e CPMac, quanto do Português Europeu padrão.

Na fonologia, o Tetun Prasa apresenta uma grande variação na realização dos segmentos consonantais e vocálicos. Ainda, na análise efetuada aqui, não se verificou o status fonológico dos segmentos palatais lusófonos, sendo estes descritos apenas como fones, o que difere da análise proposta por Hull e Eccles (2001) e Williams-van Klinken, Hajek e Nordlinger (2002). O que se verificou no corpus foi a realização da

palatalização estando limitada a empréstimos linguísticos do português e a grupos sociolinguísticos altamente escolarizados. Ainda, argumentei o mesmo para a nasalização de segmentos vocálicos, como influência da língua portuguesa sobre a língua Tetun, estando limitada a grupos sociolinguísticos mais escolarizados e não possuindo valor distintivo. O acento em Tetun Prasa é fixo recaindo sempre na penúltima sílaba, incluindo os empréstimos lusófonos que são adaptados a esse padrão acentual.

Na morfologia, o Tetun Prasa destaca-se pela completa ausência de morfologia flexional, dando espaço a essa variedade da língua para gramaticalizar uma série de lexemas com funções que foram perdidas na reestruturação gramatical que gerou a perda da morfologia flexional. Juntamente com a gramaticalização, a morfologia derivacional é produtiva, com os falantes fazendo uso de diversas estratégias de afixação, composição e reduplicação. Analisei também as classes gramaticais existentes em Tetun Prasa, sendo elas: substantivo, adjetivo, verbo, advérbio, demonstrativo, quantificador, classificador, preposição e conjunção.

A sintaxe do Tetun Prasa revelou-se com o alinhamento neutro, nenhum dos argumentos (S, A, O) sendo marcados. A ordem canônica das categorias são SVO, porém há uma série de estratégias discursivo-pragmáticas que possibilitam a mudança para outras ordens, como: V, VO, SOV e OSV. Ainda, as categorias sintáticas são reconhecidas pelo falante através de sua posição na sentença, e por estratégias discursivas, como anáfora, reciprocidade e reflexividade. O sintagma nominal (SN) apresenta a possibilidade de uma série de elementos serem anexados ao núcleo N do SN, enquanto o núcleo V do sintagma verbal (SV) pode ser acompanhado apenas por um número menor e limitado de elementos.

Finalmente, esta dissertação é a primeira gramática do Tetun Prasa em língua portuguesa, sendo, desta maneira, uma contribuição de valor tanto ao público acadêmico interessado, com uma análise das estruturas de uma língua natural, como também ao público lusófono em geral que tenha interesse nas questões de Timor-Leste. Ainda, este material único disponibilizado em língua portuguesa aos cidadãos leste-timorenses pode ajuda-los como um instrumento de ensino, aprendizagem e reflexão sobre a língua oficial do país.

10. REFERÊNCIAS BIBLIOGRÁFICAS

- ADELAAR, ALEXANDER; HIMMELMANN, NIKOLAUS P. (eds.). *The Austronesian Languages of Asia and Madagascar*. Londres: Routledge, 2005.
- AIKHENVALD, A. Grammars in Contact: a Cross-Linguistic Perspective. In AIKHENVALD, A.; DIXON, R. (eds.). *Grammars in Contact: A Cross-Linguistic Typology*. Nova York: Oxford University Press, 2007. p.1-66.
- ALBUQUERQUE, D. B. Pré-história, história e contato lingüístico em Timor Leste. *Domínios de Lingu@agem*, v.6, n.2, p.75-93, 2009.
- _____. Contatos lingüísticos em Timor Leste: mudanças e reestruturação gramatical. Comunicação apresentada ao VI Encontro da Associação Brasileira de Estudos Crioulos e Similares, Salvador, 2010a.
- _____. Peculiaridades prosódicas do português falado em Timor Leste. *ReVEL*, v.8, n.15, p.270-285, 2010.b
- _____. Elementos para o estudo da ecolinguística de Timor Leste. *Domínios de Lingu@agem*, v.7, p.21-36, 2010c.
- _____. As línguas de Timor Leste: perspectivas e prospectivas. *Língua e Literatura (USP)*, v. 27, p. 313-335, 2010d.
- _____. O Português de Timor Leste: contribuição para o estudo de uma variedade emergente. *Papia*, v. 21, n.1, p.65-82, 2011.
- ALVES, S. B. *O tétum-praça e a construção da identidade de Timór Lorosa'e*. Dissertação de Mestrado. Brasília: UnB, 2005.
- 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.
- BATALHA, G. N. Estado actual do dialecto macaense. *Revista Portuguesa de Filologia*, v. 9, p.177-213, 1958.
- BAXTER, A. *A grammar of Kristang (Malacca Creole Portuguese)*. Camberra: Pacific Linguistics, 1988.
- Notes on the Creole Portuguese of Bidau, Timor. *Journal of Pidgin and Creole Languages*, v.5, n.1, p.1-38, 1990.

- BLUST, R. Central and Central-Eastern Malayo-Polynesian. *Oceanic Linguistics*, v. 32, p.241-293, 1993.
- _____. Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics. In ZEITOUN, E.; LI, P. J-K. *Selected Papers From the 8th International Conference on Austronesian Linguistics*. Taiwan: Academica Sinica, 1999.
- BOOIJ, G. *The Grammar of Words*. An introduction to linguistic morphology. Oxford: Oxford University Press, 2005.
- CALVET, L-J. *As políticas lingüísticas*. São Paulo: Parábola Editorial/IPOL, 2007.
- CAPELL, A. Peoples and Languages of Timor (I). *Oceania*, v.14, n.3, p.191-219, 1943 a.
- _____. Peoples and Languages of Timor (II). *Oceania*, v.14, n.44, p.311-337, 1943b.
- _____. Peoples and Languages of Timor (III). *Oceania*, v.15, n.1, p.19-48, 1944.
- CLAMAGIRAND-RENARD, B. The Social Organization of the Ema of Timor. In: FOX, J. J. (ed.). *The Flow of Life: Essays on Eastern Indonesia*, p. 231-247. Cambridge: Harvard University Press, 1980.
- COMRIE, B. *Aspect: An Introduction to the Study of verbal aspect and related problems*. Cambridge: Cambridge University Press, 1976.
- _____. *Tense*. Cambridge: Cambridge University Press, 1985.
- _____. *Language Universals and Linguistic Typology: Syntax and Morphology*. Chicago: The University of Chicago Press, 1989.
- COMRIE, B.; HORIE, K. Complement Clauses versus Relative Clauses: some Khmer evidence. In: ABRAHAM, W; GIVÓN, T; THOMPSON, S. (Ed.). *Discourse Grammar and Typology*. Amsterdam: John Benjamins, 1995.
- COSTA, L. *Dicionário de Tétum-Português*. Lisboa: Colibri, 2000.
- COUTO, H. H. *Ecolingüística: estudo das relações entre língua e meio ambiente*. Brasília: Thesaurus, 2007.
- COMISSÃO LITÚRGICA DA DIOCÊSE DE DILI. *Ordinário da Missa: Texto Oficial Tétum*. Dili: Diocese de Dili, 1980.
- DIK, S. C. *The theory of functional grammar*. Part 1: The Structure of the clause. Berlin/New York: Mouton de Gruyter, 1997.
- DORES, R. *Diccionario teto-português*. Lisboa: Imprensa Nacional, 1907.

- ECCLES, L. East Timorese Language Policy and the Policies of Other Small Pacific Nations. *Studies in Languages and Cultures of East Timor*, v. 3, p.1-30, 2000.
- _____. A Tetum Reference Grammar: Some Problems on Grammatical Classification. *Studies in Languages and Cultures of East Timor*, vol. 4, p 37-47, 2001.
- _____. Early Chinese accounts of Timor. *Studies in Languages and Cultures of East Timor*, v.6, 2004. p. 178-187.
- EDWARDS, J. *Multilingualism*. Londres: Routledge, 1994
- ENGELENHOVEN, 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*. Utrecht: LOT, 2009. p. 333-362.
- ESPERANÇA, J. P. T. *Estudos de Lingüística Timorese*. Aveiro: SUL, 2001.
- FERGUSON, C. Dialects, Register and Genre: Working Assumptions About Conventionalization. In BIBER, D.; FINEGAN, E. (Eds.). *Sociolinguistics Perspectives on Register*. p. 1-31. Nova York: Oxford University Press, 1994.
- FERNANDES, A. *Método prático para aprender o tétum*. Macau/Dili: Comando Territorial Independente de Timor, 1964.
- FOLEY, W. *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press, 1986.
- FOX, J. J. Tracing the path, recounting the path: historical perspectives on Timor. In: FOX, J. J.; SOARES, D. B. (eds.) *Out of the ashes: destruction and reconstruction of East Timor*, p. 1-27. Hindmarsh: Crawford House Publishing, 2000.
- _____. Drawing from the past to prepare for the future: responding to the challenges of food security in East Timor. In: COSTA, H. et al. (Eds.). *Agriculture: New Directions for a New Nation-East Timor (Timor-Leste)*, p.105-114. Canberra: The Australian National University, 2003.
- _____. Ritual languages, special registers and speech decorum in Austronesian languages. In: ADELAAR, ALEXANDER; HIMMELMANN, NIKOLAUS P. (eds.). *The Austronesian Languages of Asia and Madagascar*. Londres: Routledge, 2005. Adelaar and Himmelmann (eds), p. 87-109.
- GIVÓN, T. *Syntax: a Typological Functional Introduction*. Vol. I. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1984.

- _____. *Syntax: a Functional-Typological Introduction*. Vol. II. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2001.
- GODDARD, C. *The Languages of East and Southeast Asia. An introduction*. Oxford: Oxford University Press, 2005.
- GOMES, F. A. *Os Fataluku*. Lisboa: Universidade Técnica de Lisboa, 1972.
- GREENHILL, Simon J.; BLUST, Robert; GRAY, Russell D. The Austronesian Basic Vocabulary Database: From Bioinformatics to Lexomics. *Evolutionary Bioinformatics*, v. 4, p.271-283, 2008.
- HAGÈGE, C. *Morte e Rinascita delle Lingue*. Milão: Feltrinelli, 2002.
- HAJEK, J. Language planning and the sociolinguistic environment of East Timor: colonial practices and changing language ecologies. *Current Issues in Language Planning*, v. 1, p. 400-413, 2000.
- _____. Serial Verbs in Tetun Dili. In AIKHENVALD, A.; DIXON, R. (eds.). *Serial Verbs Constructions. A Cross-Linguistic Typology*. p.239-253. Nova York: Oxford University Press, 2006.
- _____. 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*. p. 163-178. Nova York: Oxford University Press, 2007.
- HAJEK, J; WILLIAMS VAN-KLINKEN, C. Um sufixo românico numa língua austronésia: -dór em Tetum. *Revue de linguistique romane*, n. 67, p. 55-65, 2003.
- HAUGEN, E. *The Norwegian Language in America*. Philadelphia: University of Pennsylvania Press, 1953.
- _____. *The Ecology of Language*. California: Stanford University Press, 1972.
- HAYES, B. *Introductory Phonology*. Oxford: Blackwell, 2008.
- HIMMELMANN, N. The Austronesian languages of Asia and Madagascar: typological characteristics. In *The Austronesian languages of Asia and Madagascar* ADELAAR, K. A.; HIMMELMANN, N. (eds.). p. 110-181. Londres: Routledge, 2005.
- HOLM, J. *Languages in Contact. The Partial Restructuring of Vernaculars*. Nova York: Cambridge University Press, 2004.
- HOPPER, P. J.; THOMPSON, S. A. Transitivity in Grammar and Discourse. *Language*, v.56, p.251-99, 1980.
- _____. The discourse basis for lexical categories in Universal Grammar. *Language*, v.60, p.703-52, 1984.

HULL, G. Historical phonology of Tetum. *Studies in Languages and Cultures of East Timor*, v.4, p.158-212, 2000.

_____. O Mapa Linguístico de Timor Leste: Uma Orientação Dialectológica. *Studies in Languages and Cultures of East Timor*, v.4, p.1-11, 2001 a.

_____. A Morphological Overview of the Timoric Sprachbund. *Studies in Languages and Cultures of East Timor*, v.4, p.98-205, 2001b.

_____. *Dili Tetum*. Sydney/Dili: Sebastião Aparício da Silva Project/Instituto Nacional de Linguística/Universidade Nacional de Timor Lorosa'e, 2002a.

_____. *Standard Tetum-English Dictionary*. 3ª ed. Sydney: Sebastião Aparício da Silva Project, Dili: Instituto Nacional de Linguística/Universidade Nacional de Timor Lorosa'e, 2002b.

_____. The Papuan Languages of Timor. *Studies in Languages and Cultures of East Timor*, v.6, p.23-99, 2004.

_____. The Malay Lexical Element in Tetum. *Studies in Languages and Cultures of East Timor*, v.7, p.84-124, 2005.

HULL, G.; CORREIA, A. J. *Kursu Gramátika Tetun*. Dili: Instituto Nacional de Linguística, 2005.

HULL, G.; ECCLES, L. *Tetum Reference Grammar*. Sydney/Dili: Sebastião Aparício da Silva Project/Instituto Nacional de Linguística/Universidade Nacional de Timor Lorosa'e, 2001.

INSTITUTO NACIONAL DE LINGUÍSTICA (INL). *Hakerek Tetun tuir Banati, Kursu Ortografia Patronizada nian*. Dili: Instituto Nacional de Linguística, Universidade Nacional Timor Lorosa'e, 2002.

_____. *Matadalan Ortográfiku ba Tetun-Prasa*. Dili: Instituto Nacional de Linguística, Universidade Nacional Timor Lorosa'e, 2003.

_____. *The Standard Orthography of the Tetum Language, 115 Years in the Making*. Dili: Instituto Nacional de Linguística, Universidade Nacional Timor Lorosa'e, 2004.

KEENAN, E. L. Relative clauses. In: SHOPEN, T. (Ed.). *Language typology and syntactic description: complex constructions*, Vol. 2. Cambridge: CUP, 1985. P. 140-170.

KEENAN, E. L & COMRIE, B. Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry*, v. 8, n.1, p. 63-99, 1977.

KENSTOWICZ, M. *Phonology in Generative grammar*. Cambridge: Blackwell, 1994.

- KLINKEN, C. *A Grammar of the Fehan Dialect of Tetun, an Austronesian Language of West Timor*. Canberra: Pacific Linguistics, 1999.
- _____. From Verb to Coordinator in Tetun. *Oceanic Linguistics*, v. 39, p. 350-262, 2000.
- LANZA, E. Multilingualism and the family. In: Wei, L.; Auer, P. (eds.). *Handbook of Multilingualism and Multilingual communication*. Berlin: Mouton de Gruyter, 2007.
- LARANJEIRA, M. M. *Cartilha Tétum*. Díli: Imprensa Nacional, 1916.
- LENCASTRE, J. G. Algumas regras gramaticais da língua tétum e vocabulário. In: *Boletim da Agência Geral das Colônias*, v. 5, n.54, p.82-92, 1929.
- LEWIS, P. (ed.). *Ethnologue: Languages of the World*. 16a edição. Dallas: SIL International, 2009.
- LOBATO, M. Malaca. In: *História dos Portugueses no Extremo Oriente*. Vol.I. Tomo II. p. 13-74. Lisboa: Fundação Oriente, 2000a.
- _____. Timor. In: *História dos Portugueses no Extremo Oriente*. Vol.I. Tomo II. . p. 349-374. Lisboa: Fundação Oriente, 2000b
- LOUREIRO, R. M. Os Portugueses em Timor – relance histórico. In: *Onde Nasce o Sândalo: Os Portugueses em Timor nos Séculos XVI e XVII*. p. 29-44. Lisboa: Grupo de Trabalho do Ministério da Educação para as Comemorações dos Descobrimentos Portugueses, 1995.
- 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.
- MARANTZ, A. Re-reduplication. *Linguistic inquiry*, v.13, n.3, p.435-482, 1982.
- MARINHO, M. A. *Do Latim ao português: percurso histórico dos sufixos –dor e –nte*. Tese de Doutorado. Rio de Janeiro: UFRJ, 2009.
- MCWHORTER, J. 2007. *Language Interrupted: Signs of Non-Native Acquisition in Standard Language Grammars*. Nova York: Oxford University Press.
- _____. 2008. Why does a language undress? Strange cases in Indonesia. In Miestamo, M., K. Sinnemäki & F. Karlsson (eds.). *Language Complexity: typology, contact, change*. p. 167-190. Amsterdam: John Benjamins.
- NATIONAL BOARD OF STATISTICS. *Timor-Leste Census of Population and Housing 2004*. Priority Tables Editions: National Board of Statistics and the United Nation Fund for Population, 2006.

- O' CONNOR, S. "Unpacking the Island Southeast Asian Neolithic Cultural Package, and Finding Local Complexity". I. C. GLOVER; E. A. BACUS & V. C. PIGOTT (eds.). *Uncovering Southeast Asia's Past*. Selected Papers from the 10th International Conference of the European Association of Southeast Asian Archaeologists. Cingapura: National University of Singapore, 2006. p. 74-87.
- O'CONNOR, S.; VETH, P. Early Holocene shell fish hooks from Lene Hara Cave, East Timor establish complex fishing technology was in use in Island Southeast Asia five thousand years before Austronesian settlement. *Antiquity*, v.79, p. 1-8, 2005.
- O'CONNOR, S; SPRIGGS, M.; VETH, P. Excavation at Lene Hara Cave establishes occupation in East Timor at least 30,000–35,000 years ago. *Antiquity*, v.76, p. 45-50, 2002.
- OLIVEIRA, L. *Timor na História de Portugal*. Vol. I. Lisboa: Fundação Oriente, 2004.
- PALMER, F. R. *Mood and Modality*. Cambridge: Cambridge University Press, 2001.
- PAYNE, T. E. *Describing Morphosyntax. A guide for field linguistics*. Cambridge: Cambridge University Press, 1997.
- PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO. *Relatório do Desenvolvimento Humano de Timor Leste*. Dili: UN Agency House, 2002.
- ROMAINE, S. *Bilingualism*. Oxford: Blackwell, 1995.
- ROSS, M. Contact-induced change and the comparative method: cases from Papua New Guinea. In: DURIE, M.; ROSS, M. (eds.). *The comparative method reviewed: regularity and irregularity in language change*. p.180-217. New York: Oxford University Press, 1996.
- SÁ, A. B. Notas sobre linguística timorense: Sistema de representação fonética. *Estudos Coloniais*, v. 3, n.1-2, p. 39-60, 1952.
- _____. *Textos em teto da literatura oral timorense*. Lisboa: Junta de Investigações do Ultramar, 1961.
- SALIWANGI, et al. *Sistem morfologi kata kerja (verba) bahasa Tetun*. Jakarta: Departemen Pendidikan dan Kebudayaan, 1991.
- SIEGEL, J. *The Emergence of Pidgin and Creoles Languages*. Nova York: Oxford University Press, 2008.
- SILVA, S. M. A. *Diccionario Português-Tétum*. Macau: Typographia do Seminario, 1889.

TAYLOR-LEECH, K. The ecology of language planning in Timor-Leste. *Development Bulletin*, n 63, p. 116-120, 2005.

THOMAZ, L. F. F. R. *De Ceuta a Timor*. Carnaxide: Difel, 1994.

_____. Elementos para um glossário luso-timorense. In: *Onde Nasce o Sândalo: Os Portugueses em Timor nos Séculos XVI e XVII*. Lisboa: Grupo de Trabalho do Ministério da Educação para as Comemorações dos Descobrimentos Portugueses, 1995. p. 157-179.

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

TRAUBE, E. *Cosmology and Social Life: Ritual Exchange among the Mambai of Timor*. Chicago: The University of Chicago Press, 1986.

TROEBOES et al. *Struktur bahasa Tetum*. Jakarta: Departemen Pendidikan dan Kebudayaan, 1987.

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.

WENDEL, J. N. Notes on the Ecology of Language. *Bunkyo Gakuin University Academic Journal*, n. 5, p. 51-76, 2005.

WILLIAMS-VAN KLINKEN, C. High Registers of Tetun Dili: Portuguese Press and Purist Priests. Trabalho apresentado a *The 2001 Conference of the Australian Linguistics Society*, 2001.

_____. *Tetun Language Course*. Dili: Peace Corps East Timor, 2003.

_____. *Spelling options in Tetun Dili: Awareness of 'rr' and glottal stop amongst tertiary students*. MS. 2007a. Disponível em:

<http://www.tetundit.tl/Publications/Awareness%20of%20glottals%20and%20rr.pdf>.

(Acesso: 10/06/2010).

_____. Is he hot-blooded or hot inside? Expression of emotion and character in Tetun Dili. Trabalho apresentado à *The 5th ENUS Conference on Language and Culture*.

Kupang: The University of Nusa Cendana, 2007b. Disponível em: <http://www.tetundit.tl/Publications/Emotion%20in%20Tetun%20Dili.pdf>. (Acesso:

15/08/2010).

WILLIAMS-VAN KLINKEN, C., HAJEK, J. e NORDLINGER, R. Serial Verbs in Tetun-Dili: A Preliminary Account. *Studies in Languages and Cultures of East Timor*, v.4, p.48-60, 2001.

_____. *Tetun Dili: A grammar of an East Timorese language*. Canberra: Pacific Linguistics, 2002.

APÊNDICE

Cognatos PAN e Tetun Prasa⁴⁷

GLOSSA	PROTO- AUSTRONÉSIO	TETUN PRASA
1. mão	*(qa)lima	lima
2. perna/pé	*qaqay	ain
3. andar	*Nakaw	lao
4. caminho	*zalan	dalan
5. pele	*qaNiC	kulit
6. costas	*likud	kotuk
7. barriga/ventre	*tiaN	kabun
8. osso	*CuqelaN	ruin
9. intestino	*Cinaqi	tee-oan
10. fígado	*qaCay	tein
11. ombro	*qabara	kabaas
12. temer/medo	*ma-takut	tauk
13. sangue	*daraq	raan
14. cabeça	*qulu	ulun
15. cabelo	*bukeS	fuuk

⁴⁷ As reconstruções do proto-austronésio foram retiradas de Blust (1999) e encontram-se também disponíveis no sítio: <http://language.psy.auckland.ac.nz/austronesian/language.php?id=280>.

16. nariz	*mujiŋ	inus
17. boca	*ŋusu	ibun
18. dente	*nipen	nehan
19. língua	*Sema	nanal
20. rir	*Cawa	hamnasa
21. chorar	*Caŋis	tanis
22. vomitar	*utaq	muta
23. comer	*kaen	han
24. cozinhar	*tanek	tein
25. beber	*mimah	hemu
26. morder	*karat	tata
27. sugar	*sepsep	susu
28. orelha	*Caliŋa	tilun
29. olho	*maCa	matan
30. dormir	*tudur	toba
31. pessoa	*Cau	ema
32. homem	*ma-ruqaNay	mane
33. mulher	*bahi	feto
34. criança	*aNak	labarik
35. mãe	*t-ina	inan
36. pai	*t-ama	aman

37. casa	*Rumaq	uma
38. nome	*ŋajan	naran
39. corda	*CaliS	tali
40. agulha	*zarum	daun
41. caçar	*qaNup	duni
42. bater	*palu	baku
43. roubar	*Cakaw	naok
44. matar	*p-aCay	oho
45. morrer	*m-aCay	mate
46. arranhar	*karaw	kamat
47. trabalhar	*qumah	servisu, tralalu
48. plantar	*mula	kuda
49. espremer	*pereq	kumu
50. segurar	*gemgem	kaer
51. cavar	*kalih	kee
52. comprar	*beli	sosa
53. cachorro	*asu	asu
54. pássaro	*qayam	manu
55. ovo	*qiCelur	tolun
56. rato	*labaw	laho
57. carne	*isi	naan

58. cobra	*Sular	samea
59. piolho	*kuCu	utu
60. mosquito	*likeS	susuk
61. aranha	*kakaCu	labadain
62. peixe	*Sikan	ikan
63. folha	*biraq	ai-tahan
64. raiz	*RameC	abut
65. terra	*dareq	rai
66. pedra	*batu	fatuk
67. areia	*qenay	rai-henek
68. água	*daNum	bee
69. mar	*tenem	tasi
70. sal	*timus	masin
71. lago	*danaw	bee-lihun
72. árvore	*kaSiw-kaSiw-an	ai
73. céu	*laŋiC	lalehan
74. lua	*bulaN	fulan
75. estrela	*bituqen	fitun
76. nuvem	*lemlem	kalohan
77. chuva	*quzaN	udan
78. vento	*bali	anin

79. fogo	*Sapuy	ahi
80. queimar	*cuNuh	sunu
81. fumaça	*CebuN	suar
82. preto	*CeņeN	metan
83. branco	*ma-puNi	mutin
84. vermelho	*ma-taNah	mean
85. verde	*mataq	modok
86. pequeno	*kedi	kiik
87. grande	*ma- <i>raya</i>	boot
88. velho	*ma-tuqaS	katuas
89. novo	*ma-baqeru	foun
90. mal	*kuya	aat
91. noite	*berņi	kalan
92. dia	*qalejaw	loron
93. ano	*kawaS	tinan
94. esconder	*buNi	subar
95. escalar	*dakiS	kadopa
96. em	*di	iha
97. dentro	*i-dalem	laran
98. acima	*i taqas	(iha) leten
99. este	*i-ni	nee

100.	perto	*ma-azaNih	besik
101.	longe	*ma-dawiN	dook
102.	1sg	*i-aku	hau
103.	2sg	*i-kaSu	oo
104.	3sg	*si-ia	nia
105.	1pl.inc	*i-kita	ita
106.	1pl.exc	*kami	ami
107.	2pl	*i-kamu	imi
108.	o que?	*n-anu	saa
109.	quem?	*si-ima	see
110.	quando?	*ija-n	bainhira
111.	onde?	*i-nu	(iha) neebec
112.	outro	*duma	seluk
113.	tudo	*amin	hotu
114.	e	*ka	no, ho, i
115.	se	*ka	si, bainhira
116.	como?	*kuja	oinsaa
117.	não	*ini	la
118.	um	*isa	ida
119.	dois	*duSa	rua
120.	três	*telu	tolu

121.	quatro	*Sepat	haat
122.	cinco	*lima	lima
