

*Natalia Czopek*

Universidade Jaguelónica

TENTATIVAS  
DE ESTANDARDIZAÇÃO  
ORTOGRÁFICA DOS  
VOCÁBULOS DE ORIGEM  
PORTUGUESA  
EM TÉTUM-PRAÇA

**Attempts of standardization of the spelling of the words of Portuguese origin in Tetum-Prasa**

ABSTRACT

In this paper we aim to put a few comments on the first attempts of standardization of the spelling of numerous words of Portuguese origin in Tetum-Prasa, followed by the literacy campaign organized by Fretilin after which, due to its growing importance as a language of resistance and writing, was elaborated the current phonemic model. The corpus of examples that we analyse is based on texts of scientific nature written in Tetum-Prasa.

KEYWORDS: Tetum-Prasa, orthography, standardization, East Timor, phonemic.

Um termo peregrino  
não pode incorporar-se numa língua  
sem se despir do seu carácter heterogéneo,  
do mesmo modo que um estrangeiro  
não pode naturalizar-se sem se sujeitar  
às leis do país (Dalgado 1919)

No presente estudo, pretendemos colocar algumas observações sobre os ajustamentos ortográficos ao sistema do tétum que sofreram os vocábulos de origem portuguesa ao serem emprestados, contribuindo para a formação de uma língua de índole crioulanizada, o chamado tétum-praça ou tétum-Díli<sup>1</sup>. Os exemplos que servirão de base na nossa análise serão extraídos principalmente da palestra dirigida ao Congresso Nacional pelo Prof. Dr. Geoffrey Hull *Identidade, Lian no Política Edukasionál* (2001a) e da obra de Frédéric Durand *Istória Timor-Leste Nian. Husi Pre-istória to 'o Atualidade* (2009), sendo o *corpus* enriquecido por exemplos do *Dicionário de Tétum-Português* de Luís

---

<sup>1</sup> Cf. Thomaz (2002: 103): a cidade de Díli costuma ser designada por Praça ('fortaleza, praça de armas') pelos naturais de Timor.

Costa (2000)<sup>2</sup>. Na parte introdutória, a título de ilustração, adicionamos alguns comentários sobre a ortografia empregada no registo escrito de lendas timorenses de tradição oral recolhidas por Artur Basílio de Sá na coletânea *Textos em Teto da Literatura Oral Timorense* (1961) e por Davi Borges de Albuquerque no seu trabalho *Língua e meio ambiente na literatura oral em língua Tétun, Timor Leste* (2012). Umhas considerações sobre os textos religiosos e poéticos descritos por Luís Thomaz em *Babel Loro Sa'e. O problema linguístico de Timor-Leste* (2002), enriquecerão esta parte. Embora o trabalho não pretenda de forma nenhuma ser exaustivo, esperamos que as nossas observações ajudem os leitores a ficar com uma imagem bem definida da evolução da ortografia timorense.

## ENQUADRAMENTO HISTÓRICO E TEÓRICO

Descrever e analisar a história timorense pode resultar uma tarefa bastante complicada, já que as primeiras fontes escritas acessíveis provêm do século XVI, ou seja, foram criadas após a chegada dos portugueses. Até ao fim do século XVII, a documentação foi escassa e descontínua. Mesmo na escrita de Sumatra e Java, anterior e muito melhor desenvolvida, as referências à ilha eram quase nulas (Thomaz 2002: 22). Timor apareceu pela primeira vez na cartografia europeia num mapa desenhado por um cartógrafo jovem Francisco Rodrigues, por volta do ano 1512. No entanto, como afirma Horta Schmidt (1983: 358), o estudo dos empréstimos, que trazem em si a marca da sua origem, pode ajudar muito a reconstruir a história de um povo sem escrita. As primeiras migrações incluíam povos de origem melanésia e, mais tarde, austronésia (Durand 2009: 23). A base linguística criada por estes povos ficou enriquecida, nos séculos seguintes, graças ao comércio e intensos contactos com os outros povos da zona, com influências malaias<sup>3</sup>, javanesas, goesas<sup>4</sup>, africanas, portuguesas, indonésias, macaenses e árabes. A presença portuguesa naquela parte do mundo começou com a tomada de Malaca, em 1511, após a qual a procura da preciosa madeira de sândalo, mel, cera e escravos continuou rumo a Timor, ilha “onde nasce o sândalo”. A primeira viagem portuguesa à parte ocidental da ilha, chamada de Sevião, teve lugar provavelmente em 1514 ou 1515 e durante a primeira metade do século XVI, mantiveram-se os contactos puramente comerciais, baseados, como era de costume nas situações desse tipo, no comércio silencioso com auxílio da linguagem gestual (Loureiro 2001: 93–104). Com o passar do tempo e com o crescente perigo da invasão dos povos islamizados das Celebes, de outros povos asiáticos e dos

<sup>2</sup> No nosso trabalho, exceto algumas referências soltas, não vamos fazer distinção entre as pronúncias das pessoas de diferentes graus de escolarização. Precisamos de sublinhar que o padrão para a ortografia dos empréstimos lexicais é a pronúncia acroletal, ou seja, a das elites urbanas, normalmente fluentes em português. Assinalemos aqui, apenas como curiosidade, que Luís Costa (2000: 279) inclui no seu dicionário a pronúncia dos falantes não-escolarizados, não fluentes em português, por exemplo *palasi* em vez de *palásiu*.

<sup>3</sup> A descrição das influências malaias e árabes em Timor pode ser consultada em Thomaz (2002: 81–102, 108–109, 145). O próprio nome do país, Timor ou Timur, deriva do malaio com o significado de ‘Oriente’.

<sup>4</sup> Mais informações sobre a influência indo-javanesa em Timor podem encontrar-se em Thomaz (2002: 77–79). O autor sublinha que os poucos vocábulos indianos em tétum não devem ser empréstimos diretos e provavelmente entraram por intermédio do malaio.

holandeses, alguns dos sessenta pequenos reinos timorenses (pertencentes ora à confederação dos Belos, a leste, ora à dos Baiquenos, a oeste) admitiram a soberania do rei português, aceitando, ao mesmo tempo, a fé cristã (Loureiro *et al.* 2002: 36). Os portugueses, encorajados para casarem com as *liurai*, mulheres da realeza nativa, começaram a estabelecer-se na ilha de forma significativa desde 1640, mas o primeiro governador nomeado por Lisboa chegou apenas em 1703 e a colonização portuguesa nessa zona nunca se pôde chamar de maciça. Os chamados assimilados, “portugueses pretos” ou topazes, filhos dos portugueses casados com as mulheres nativas, de sangue misto, pertencentes às classes sociais que receberam a sua educação em português, ou seja, à elite colonial, desempenhavam o papel primordial na administração. Apesar das dificuldades no uso dos manuais, que poucas vezes correspondiam à realidade na qual viviam, na maioria dos casos eram pessoas bilíngues cujo léxico era suscetível a interferências entre as línguas nativas e o português. Como os portugueses habitavam maioritariamente a zona de Díli<sup>5</sup>, a língua nativa sujeita a mais influências era o tétum.

Após a segunda Guerra Mundial, o país ficou bastante destruído e de uma certa forma “abandonado” por Portugal ocupado com a sua política colonial africana. O caminho à descolonização e à introdução de reformas políticas e sociais ficou aberto com o movimento de 25 de abril de 1974 (Loureiro *et al.* 2002: 37–40). O novo governo da Fretilin reconheceu a importância da escolha de uma língua veicular que servisse como ferramenta de comunicação interétnica. Escolheu-se o português como língua oficial<sup>6</sup> e decidiu-se elevar gradualmente a posição do tétum-praça<sup>7</sup>, uma língua crioualizada de tradição oral<sup>8</sup> que continha o elemento português muito forte mas precisava de standardização ortográfica para assumir o estatuto da língua

---

<sup>5</sup> Como resultado da presença holandesa cada vez mais forte, em 1769 os portugueses tiveram de mudar a capital de Lifau para Díli. A fronteira oficial entre os territórios sob o domínio português e holandês fixou-se só em 1915. Os holandeses anexaram os enclaves portugueses em Timor-Oeste, menos o de Oecússi. Thomaz (2002: 68–69) realça a hipótese de o tétum não se ter tornado língua franca por se falar em Díli, mas de se ter começado a falar em Díli por ser, precisamente, língua franca da província (o que podem confirmar p. ex. os topónimos fora das regiões da predominância do tétum) e compara a sua situação à da língua geral do Brasil, usada lá até ao século XVIII.

<sup>6</sup> Durante a ocupação indonésia, que começou dez dias depois da declaração da independência, o uso da língua portuguesa foi visto como ato de espionagem e colaboração com o movimento de resistência. Como consequência, o tétum evoluiu para o símbolo da identidade nacional.

<sup>7</sup> Timor é um país de poliglossia, com cerca de 20 grupos linguísticos principais e um número mais reduzido de dialetos reconhecíveis, a maioria proveniente da família austronésia ou malaio-polinésia (Loureiro *et al.* 2002: 42). O tétum, que se formou provavelmente na Formosa e, talvez, no sul da China continental, tem duas variedades principais: o tétum-térique, tradicional, “mais alto”, designado também como o tétum clássico, verdadeiro ou *los* (‘correto’), que tinha sido usado como língua veicular antes da chegada dos portugueses; e o tétum-praça, ou tétum-Díli, uma língua simplificada que tem como base o léxico do tétum-térique com várias palavras portuguesas e malaias incorporadas. A maior influência portuguesa efetuou-se após a mudança da capital para Díli em 1769. Como de costume, os primeiros empréstimos designavam realidades novas ou termos abstratos que, às vezes, tinham as suas correspondentes nativas. A facilidade em aceitar tais empréstimos foi criticada por exemplo por Raphael de Dorez, no seu dicionário tétum-português de 1891. Hull (2001b: 87) afirma que, atualmente, nessa língua híbrida, os elementos neolatinos predominam sobre os indígenas. É muito importante salientar que Timor-Leste é a única das antigas possessões portuguesas onde uma língua nativa foi elevada ao estatuto de língua cooficial, sendo também língua nacional.

<sup>8</sup> Podemos mencionar aqui os poucos materiais escritos daquela altura, como alguns textos religiosos ou jornalísticos.

cooficial. O tétum-praça tinha facilitado a comunicação entre os membros do movimento de resistência que eram, na maioria dos casos, camponeses não educados em português, que falavam as suas línguas nativas. O jornal da Fretilin era impresso em português, língua usada pela classe dirigente, e, pela primeira vez num contexto que não fosse religioso, em tétum romanizado, posteriormente indicado como idioma de alfabetização do povo timorense independente. No entanto, em 1975, este plano foi adiado pelo motivo da invasão do ocupante indonésio que, no final das contas, fortaleceu a posição do tétum como língua de resistência e, mais tarde, como a língua mais falada do país.

Durante muito tempo a sua forma escrita permaneceu imperfeita (Osório de Castro 1996: 91–94)<sup>9</sup>. As primeiras tentativas de registar os sons portugueses devem-se ao “fundador dos estudos sobre o tétum oriental, Padre Sebastião Maria Aparício da Silva, que em 1887, publicou o seu dicionário português-tétum, revisto mais tarde pelo Padre Manuel Mendes<sup>10</sup>. Nesse primeiro dicionário, reconheceu-se a existência das vogais nasais, marcadas com o til. No entanto, muitos sons portugueses ainda não tinham sido incluídos no sistema do tétum como fonemas separados, sendo a sua pronúncia assimilada aos fonemas existentes. No século XX, a maior irregularidade foi a prática de usar a ortografia portuguesa, com ou sem diacríticos, ao registar empréstimos portugueses, contribuindo à formação de um sistema macarrónico e colonial, no qual a ortografia não correspondia à fonética nativa. Também não era raro encontrar ortografia mista com elementos portugueses e nativos, ex. *reconciliaun* (Van Klinken 1999: 19). Como já se mencionou, o efémero governo da Fretilin promoveu, durante a sua campanha de alfabetização de 1975, uma ortografia integrada e estandardizada. Nos anos seguintes, porém, muitos autores resistiram a reconhecer a natureza fonémica dos sons portugueses em tétum e a unificar ortograficamente as variantes acro, meso e basileta (Hull, Eccles 2001: 221–222). Assim, o problema principal do tétum moderno era a estandardização da sua ortografia que acabou por ser fonémica, isto é, correspondente em grande parte à pronúncia real das palavras. Sendo uma língua com grande variedade de falantes, dos quais muitos não a usam em casa, a sua situação era especial<sup>11</sup>. Os falantes familiarizados com o português tinham esta língua como a principal fonte de empréstimos, entre os jovens de formação indonésia notava-se uma tendência a usar palavras indonésias, podiam-se ouvir também interferências com outras línguas nativas usadas na ilha. Desde 1999, após a libertação do país, o Conselho Nacional da Resistência Timorense começou a im-

<sup>9</sup> Golden (2004) defende a opinião de o tétum ainda não estar em condições para poder funcionar como língua de todos os domínios políticos e sociais por não dispor de estruturas linguísticas adequadas. Por outro lado, Esperança (2001: 166) insiste que cada língua está perfeitamente equipada para as necessidades comunicativas da comunidade que a usa. Observem-se mais considerações sobre a situação linguística atual de Timor-Leste em Hull (1999: 1–7).

<sup>10</sup> Os seguintes dicionários com muito impacto na ortografia e pronúncia, com inclusão da diversificação dialetal, eram o dicionário inglês-tétum de Raphael das Dores (1906) e tétum-português do Padre Manuel Mendes (1934). O enriquecimento dialetal foi, porém, muitas vezes visto, como retrocesso no processo da estandardização ortográfica da língua franca (Hull, Eccles 2001: 221–222).

<sup>11</sup> Van Klinken (1999: 16–17) enumera as dificuldades que os falantes não familiarizados com o português e falantes do tétum rural podem ter usando o tétum-praça, ex. dificuldade no emprego dos vocábulos emprestados, sobretudo abstratos, ou erros de colocação.

plementar a política linguística anteriormente planeada (Hull 2000a:1)<sup>12</sup>. De acordo com o Decreto N.º 1 de 2004, tétum-praça é a base da língua literária, os empréstimos indonésios que se podem ainda encontrar no tétum coloquial das pessoas de formação indonésia devem ser excluídos da língua padrão a favor dos termos nativos ou portugueses e todos os empréstimos têm de ser adotados de acordo com as regras da ortografia tétum. No entanto, nos registos escritos é possível encontrar, às vezes, hesitações quanto à ortografia. Neste estudo, vamos propor um quadro dos ajustamentos mais importantes que sofreram os vocábulos de origem portuguesa sendo adaptados ao sistema do tétum.

## ANÁLISE DOS EXEMPLOS ENCONTRADOS NAS FONTES ESCOLHIDAS

Como já se afirmou, a fonte mais comum de empréstimos para designar novas realidades foi o português (Hull, Eccles 2001: XIV). Parece-nos interessante mencionar neste ponto as opiniões dos falantes nativos de tétum que sublinham a importância dos empréstimos portugueses como elemento enriquecedor e modernizador que completa o sistema numa língua vista ainda por muitos como deficiente (Taylor-Leech 2008: 166–173)<sup>13</sup>.

A título de comparação, coloquemos algumas considerações breves sobre as influências portuguesas encontradas na coletânea de Artur Basílio de Sá, no trabalho de Davi Borges de Albuquerque e na monografia de Luís Thomaz. Como já foi dito, os textos, sendo registos da tradição oral timorense em tétum-térique, dificilmente admitem empréstimos portugueses.

Os textos que se podem encontrar no trabalho de Borges de Albuquerque representam os três géneros textuais mais importantes da tradição oral tetumófona: o *hamulak* (oração narrativa em versos usada em cerimónias ritualísticas pelos mestres da palavra, os *lian-na'in*), *ai knananuk* (poesia oral usada em festas tradicionais) e *ai knanoik* (contos populares, mitos de origem ou fábulas com fins educativos). Como os primeiros dois géneros abrangem orações ritualísticas ou poesia tradicional, a sua linguagem é um tétum-térique muito arcaico, conservador, atualmente menos usado no dia-a-dia e, por conseguinte, privado de quaisquer empréstimos portugueses<sup>14</sup>. Nos fragmentos dos cinco contos que

---

<sup>12</sup> Ainda em 1997, J.P. Esperança afirmou que “o tétum continua a ser escrito de forma variável, e as tentativas de estabelecimento de regras que se têm feito carecem normalmente de base científica (Esperança 1997: s.p.). Em 1999, C. Van Klinken disse que:

Tetum-Prasa's history as a written language of mass communication is rather short. Most written materials at this stage are religious, including Scripture portions, liturgies, and some articles published in the Catholic diocesan bulletin *Seara*. In addition, other East Timorese newsletters and the newspaper include articles in Tetum, and some organizations are using it in publications such as health booklets (Van Klinken 1999: 14).

<sup>13</sup> Hull e Eccles (2001: xiii) põem em realce a forte europeização do tétum-praça, chamando-o até de *Europeanoid language*, o que reflete a natureza híbrida da civilização timorense.

<sup>14</sup> No entanto, Hull (1999: 283) afirma que não se deve exagerar chamando o tétum-térique de puro, já que inclui empréstimos malaios e, atualmente, algumas palavras portuguesas assimiladas, como *oras*.

representam o terceiro género<sup>15</sup> encontram-se umas poucas palavras de origem portuguesa cuja presença se explica pelo seu carácter menos formal e pelas possíveis influências dos idioletos das pessoas que os transmitiam. Vejamos a sua ortografia:

- agora/agora tempu (duas variantes do mesmo advérbio, com o -o final com valor fonémico); animaal/animal (hesitação ortográfica resultante, provavelmente, da vontade de sublinhar na fala a vogal acentuada pronunciada como mais longa); orariu (o *h* mudo nunca fica grafado e o *o* final tem valor fonémico); komesa; kon-sege; kontente (como se comentará mais pormenorizadamente a seguir, o alfabeto timorense carece da letra *c* que fica representada pelo *k*; o fonema [s] corresponde sempre à única representação gráfica *s*; o *g* tem sempre o valor de [g], o que torna desnecessário o uso do fonema mudo *u* no grupo *gue*; na pronúncia popular o -e final é pronunciado como [i], como no Brasil, ex. *xave* [fave] pronunciado como [sabi]); entaun (o ditongo nasal -ão grafado como -aun).

O carácter accidental destas interferências pode ser confirmado pelo facto de não designarem nenhuma realidade nova nem noções completamente abstratas e difíceis de serem descritas em tétum, formando até duplas com as palavras nativas correspondentes (ex. *daudaun/agora*), o que as torna empréstimos não imprescindíveis.

Também nos *ai knanoik* registados por Artur Basílio de Sá, os estrangeirismos são raros e aparecem grafados em itálico. O autor explica que optou pelo alfabeto de representação fonética, com cada letra a corresponder a um só fonema. Os grafemas *c*, *ç*, *g*, *j*, *p*, *v*, *x*, *z*, *y*, *nh*, *lh* encontram-se apenas nos estrangeirismos grafados em itálico. Vejamos os exemplos do primeiro volume da coletânea: *nia força* ('a sua força'; o termo português usa-se apesar da existente palavra tétum *kbiit*; mantém-se a grafia original se ca-lhar por se tratar de uma palavra introduzida acidentalmente pelo autor da transmissão oral<sup>16</sup>); *nusu licença* ('pedir licença'; expressão híbrida com grafia original); *marinheiro/conselho* (grafia original); *selu favor* ('paga o favor', palavra portuguesa com grafia original, empréstimo desnecessário); *guarda liurai* ('casa da guarda', sem alteração fonética); *reino-reino* (um empréstimo português reduplicado para indicar multiplicidade, grafia original); *amo sira* ('senhores'; o vocábulo português com grafia original); *sentido, morador, costuma, ordem, sinal* (palavras importadas sem alteração fonética); *sorte/sorti/sortin* (hesitação ortográfica no registo da palavra portuguesa 'sorte'); *bensa* ('bênção', grafia simplificada que não reflete o ditongo nasal final, de acordo com a grafia estandardizada seria *bensaun*); *histori* ('historia', palavra parcialmente adaptada

<sup>15</sup> Trata-se dos *ai knanoik* titulados *Manumatadador*, *Buibabukulasak*, *O macaco e o rato*, *O macaco e o crocodilo* e *O cachorro e o macaco*, sendo os primeiros dois antropónimos.

<sup>16</sup> Em 1999, Hull sublinhava que, apesar do carácter fonémico e inclusivo (admite variantes sociais e fonéticas) da ortografia padronizada, ainda se podiam encontrar vários exemplos de ortografia não estandardizada, macarrónica, não assimilada aos padrões do tétum, o que é mais uma prova do seu hibridismo, de uma certa arbitrariedade e da índole criulizada. Comparem-se os dois trechos de um Evangelho citados pelo autor:

<sup>(A)</sup> Iha loron hirak nee duni mak Jesus mai housi Nazare Galileia nian, hodi simu baptismo hosi Joao iha Jordao. (...) Nia hela ho balada fuik sira, no anjo sira serbi Nia. (...) "Tempo too ona," Nia dehan, "Maromak Nia reino besik ona". (...) Wainhira Jesus sai housi sinagoga, Nia, hamutuk ho Tiago no Joao, ba los Simao no Andre nia uma.

<sup>(B)</sup> Iha loron hirak-ne'e duni maka Jezus mai housi Nazare Galileia nian hodi simu batizmu hosi Juaun iha Jordaun. (...) Nia hela ho balada fuik sira, no anju sira serbi Nia. (...) "Tempu to'o ona," Nia dehan, "Maromak Nia reinu besik ona". (...) Bainhira Jezus sai hosi sinagoga, Nia, hamutuk ho Tiagu no Juaun, bá loos Simaun no Andre nia uma (Hull 1999: xxx).

à fonética do tétum); *oras wai loro/iha oras ne'e be/oras mei dia/oras ne'e* (verão/quando/meio-dia/agora mesmo; o empréstimo português *oras* faz parte de diferentes compostos híbridos, parece bastante enraizado no tétum, sendo a sua grafia assimilada às normas deste). Notam-se, então, ao mesmo tempo, palavras portuguesas sem nenhuma alteração ortográfica e casos de assimilação ortográfica parcial ou total ao tétum, sendo a grafia em muitos casos macarrónica.

Os textos citados por Luís Thomaz podem-se agrupar em três categorias: os que representam os primeiros dois géneros da tradição oral, a saber, oração em verso e poesia tradicional ritualística; traduções de textos religiosos católicos para tétum, dos princípios do século XX; e textos políticos e revolucionários da segunda metade do mesmo século. No primeiro grupo, como anteriormente se disse, é bastante difícil encontrar empréstimos portugueses. O único exemplo que se pode citar aqui é a palavra portuguesa *mor* usada, como afirma o próprio autor, talvez por necessidade de paralelismo e métrica. Nos textos do segundo grupo, como era de prever, encontramos uma vasta gama de palavras portuguesas e latinas, registadas sempre de acordo com a ortografia original ou latinizante: *Adão, Jesus, Simão Pedro, João, Maria Virgem, Santa, Christo, escolante, cruz, salve*.<sup>17</sup> A justificação deste facto é, obviamente, a inexistência de palavras nativas para designarem a nova realidade do catolicismo. Nos textos do último grupo, já que provêm do período em que a ortografia ainda não tinha sido standardizada, há exemplos de grafia original e também assimilada aos padrões do tétum (apenas um caso): *contrato, escola, avião, empregado, doutor, liberdade, independente, Demokrasia*.

O ponto central do nosso *corpus* é constituído, no entanto, por dois textos de índole científica anteriormente mencionados. Encontramos neles um leque de exemplos que representam diferentes processos do ajustamento ortográfico ao sistema tétum. Como já se observou, a ortografia moderna do tétum-praça é fonémica e a representação dos sons corresponde à sua pronúncia. Vejamos os fenómenos que achamos os mais interessantes e representativos:

1. Na ortografia do tétum-praça usa-se o alfabeto latino privado das letras *c, q e y*<sup>18</sup>: *kuartu* (quarto), *moskiteiro* (mosquiteiro), *konkorda* (concorda), *kartaun-krédito* (cartão de crédito), *porke* (porque), *kualkér* (qualquer), *teknolojia* (tecnologia), *ekolójiku* (ecológico).
2. As letras *g, j, h e x* nem sempre têm os mesmos valores do português, por exemplo, o *g* é sempre pronunciado como velar [g], por conseguinte não é preciso introduzir o *u* mudo nos grupos *gue* e *gui*: *portugés* (português). O *h* é sempre aspirado [h], então o *h* mudo das palavras portuguesas não fica grafado: *otél* (hotel), *ospitál* (hospital), *ijiene* (higiene), *iena* (hiena), *elikóptero* (helicóptero), *olandeza* (holandesa). O *x* tem duas pronúncias, uma culta [ʃ] e uma popular [s], ex. *kaixa* [kaiʃa] vs. [kaisa]<sup>19</sup>; nas palavras portuguesas onde o *x* tem outro

<sup>17</sup> É interessante comparar duas versões da Salve Rainha citadas pelo autor. A versão mais moderna, de 1972, inclui muitos vocábulos portugueses de grafia original, como *esperança, piedosa, desterro, Virgem, doce*, etc.

<sup>18</sup> Nos anos 80-90 do século XX, os membros da resistência timorense sublinharam a necessidade do emprego de um alfabeto unificado para todas as línguas faladas na ilha. O grafema *c* existe, por exemplo em fataluku, e o *y* no dialeto kawaimina. O *q*, por sua vez, não se usa no alfabeto de nenhuma das línguas timorenses. No caso do tétum-praça, os três grafemas podem-se encontrar em estrangeirismos não assimilados.

<sup>19</sup> O sistema das consoantes pode ser dividido em duas categorias: as consoantes de um valor fonológico e as de dois valores, o que tem a ver com as diferenças na pronúncia culta (europeizada) e popular. As consoantes de dois valores são: *v* [v] vs. [b], *j* [ʃ] vs. [z], *x* [ʃ] vs. [s], *r* [rr] vs. [r], *ll* [ʎ] vs. [il] e *ñ* [ɲ] vs. [in]. Existem também divergências entre as pronúncias urbana e rural *p* [p] vs. [b], *g* [g] vs. [k], *z* [z]

valor fonético, fica substituído pelos grafemas correspondentes, por exemplo, *ezatu* (exato), *ezije* (exige), *ezemplo* (exemplo); de acordo com a regra de um fonema corresponder a uma representação gráfica, as palavras portuguesas que contêm o *ch* ficam grafadas com o *x*: *xuveiro* (chuveiro), *xefe* (chefe), *xave* (chave), *xá* (chá), *xumbu* (chumbo), *xeke* (cheque), *dixiplina* (disciplina), *faxista* (fascista). O *j* substitui o *g* na representação de [ʒ]: *urjente* (urgente), *jeleira* (geleira), *jerasaun* (geração), *rejime* (regime), *liturjia* (liturgia).

3. O grafema *ɥ* nos empréstimos portugueses pode ter uma leitura mesoletal como [b], o que às vezes parece ter consequências ortográficas: *serbisu* (serviço), *abó* (escrito também como avó).

4. Os timorenses estão acostumados à grafia na qual a cada fonema corresponde um grafema, por isso a praticada uniformização da pronúncia do *ʂ* como [s] em diferentes contextos linguísticos (ex. *prosesu* como [prosesu] e não [prozezu]; uma exceção resultante da influência portuguesa no tétum-Dili é a pronúncia acroletal do *s* como [ʃ] na posição final e antes de uma consoante surda). De acordo com a mesma regra o *ce*, *ci* e *ç* são grafados como [s]: *pedasu* (pedaço), *diferensa* (diferença), *ameasa* (ameaça), *eskosés* (escocês), *koinside* (coincide), *pasaporte* (passaporte), *asasinu* (assassino), *susesu* (sucesso). O *s* português com o valor de [z] é grafado como *z*: *Brazil* (Brasil), *propóziito* (propósito), *gazolina* (gasolina), *fraze* (frase), *dioseze* (diocese), *uza* (usa).

5. Como as sequências com os sons aspirados /lh/ e /nh/ já existiam em tétum com a sua pronúncia específica diferente da portuguesa (ex. *bainhira* – ‘quando’, com o *h* sonoro heterossilábico), para evitar confusões adotaram-se os dígrafos do antigo português e do galego *ll* e *ñ*, que representam duas variantes – a pronúncia culta como dos sons portugueses *lh* e *nh* e a popular [il] e [in]: *folletu* (folheto), *jullo* (julho), *barullu* (barulho), *señora* (senhora), *bañu* (banho), *koñese* (conhece), *dezeñu* (desenho), *koñesimentu* (conhecimento).

6. Nenhuma das consoantes portuguesas que não existiam no sistema nativo, como *p*, *g*, *v*, *j*, etc.<sup>20</sup>, (exceto o [ʃ] da pronúncia acroletal) aparece na posição final nas palavras adaptadas ao sistema do tétum, ex. o *-m* final português é representado pelo *-n* que nasaliza a vogal anterior: *jejún* (jejum), *sondajen* (sondagem), *sin* (sim), *bon* (bom), *tenke* (tem que), *konlisensa* (com licença), *bondia* (bom dia), *benvindo* (bem-vindo), *bagajen* (bagagem), *viajen* (viagem). Encontramos um caso onde o *-n* marcava uma vogal fechada precedente: *bisavón* (bisavô). Achamos necessário mencionar, a este ponto, que, de acordo com esta regra, a forma vernácula do nome da língua analisada é *tétun*<sup>21</sup>.

7. O *o* não acentuado pronuncia-se como [o] ou [u], o que, no segundo caso, se reflete na ortografia: *ventuñna* (ventoinha), *feijuada* (feijoadá), *konsuada* (consoada), *autucarro* (autocarro).

8. O sistema das vogais segue as regras da ortografia fonémica. As vogais de pronúncia prolongada (normalmente acentuadas) são marcadas, na maioria dos casos, com uma vogal com acento agudo ou, mais raramente, com uma vogal duplicada. O prolongamento tem função distintiva nas palavras homónimas: *se* [se] (do português *se*) e *se(e)* [se:] (‘quem, direto’). As vogais acentuadas na última sílaba são sempre longas, ex. *perú* (peru), *atór* (ator), *fransés* (francês), *perfil* (perfil).

9. As palavras nativas do tétum-praça carecem de ditongos, as vogais normalmente são justapostas e heterossilábicas, com a primeira delas acentuada; o mesmo acontece nas palavras

vs. [s], *j* [ʒ/z] vs. [d], *w* [b] vs. [w] (Hull 1999: xxv–xviii). Pode-se acrescentar aqui a pronúncia de *iu* final não acentuado como [ju] e [i] e de *ia* final não acentuado como [ja] e [i] (Hull, Eccles 2001: 228). Além disso, existem diferenças entre os falantes expostos mais ao contacto com o português que, por exemplo, fecham mais as vogais, perdem o *h* aspirado, pronunciam o *-s* final como [ʃ].

<sup>20</sup> No dicionário de 1907, Raphael de Dóres, empregando a sua *ortographia sónica*, prescinde dos grafemas *c*, *g*, *j*, *p*, *q*, *v*, *w*, *x*, *y*, *z* por não lhe parecerem necessários para escrever as palavras nativas ou por nunca os ter ouvido na conversação com os nativos, que têm problemas da pronúncia de alguns deles (*g*, *j*) (Esperança 1997: s.p.).

<sup>21</sup> O nome deriva possivelmente do verbo *tetu* (‘pesar’), o que pode indicar que era a língua do comércio e, assim, língua veicular de longa data (Thomaz 2002: 72).



oxítonas de origem portuguesa: *eroi* (herói). Os ditongos portugueses registam-se de acordo com a transcrição fonémica: *jeitu* (jeito), *noivu* (noivo), *baunilla* (baunilha), *lousa* (louça). A nasalidade do ditongo *-ão* é sempre representada pelo *-n* final: *televizaun* (televisão), *alemaun* (alemão), *konfirmasaun* (confirmação), *kondisaun* (condição), *aviaun* (avião), *paixaun* (paixão), *relijiun* (religião). Na pronúncia popular, os ditongos costumam monotongar-se, ex. *doutór* [dooto:r] vs. [doto:r].

10. A pausa glotal, típica do tétum-térique, fica marcada ou não (*di'ak/diak*), mas normalmente, por não existir em português, também não se marca no tétum-Díli. No entanto, os falantes de outras línguas timorenses com a pausa glotal presente tendem a introduzi-la no tétum-praça onde seja historicamente justificada (Hull, Eccles 2001: 229).

11. Apesar de a ortografia ter sido normalizada, é frequente escrever os nomes próprios, apelidos e topónimos de acordo com as regras da ortografia portuguesa: Pedro (e não *Pedru*), Martinho (e não *Martiñu*), Adão (e não *Adaun*), Carvalho (e não *Karvallu*), Liquiçá (e não *Likisá*). Os católicos timorenses adotaram apelidos portugueses e muitas pessoas têm dois nomes: o nativo e o português, recebido no batismo (Hull 1999: 10–11).

12. Há alterações prosódicas: a maioria dos vocábulos são graves, o acento gráfico agudo usa-se apenas nas palavras agudas e esdrúxulas<sup>22</sup>, mesmo que não sejam acentuadas graficamente nas palavras portuguesas de origem. O acento circunflexo não se emprega. Os monossílabos portugueses com vogal nasal não são acentuados, pois a vogal é pronunciada como curta: *sin* (sim). Estas regras criam muitas diferenças entre as duas ortografias: *profesór* (professor), *invazór* (invasor), *fasil* (fácil), *konsul* (cônsul), *perú* (peru), *juri* (júri), *taksi* (táxi), *alfândega* (alfândega), *admiravel* (admirável), *natural* (natural), *abdomen* (abdómen), *amañá* (amanhã), *azúl* (azul), *kór* (cor), *kris* (cruz), *ofisiál eleitorál* (oficial eleitoral), *util* (útil), *rezistênsia* (resistência).

13. Vocábulos compostos escrevem-se como uma palavra única (*kazadebañu* – casa de banho, *Kabuverde* – Cabo Verde, *meiudia* – meio-dia, *korderoza* – cor-de-rosa, *tankedegera* – tanque de guerra, *derrepente* – de repente, *enzezde* – em vez de), com hífen ou como duas palavras separadas, embora a forma tradicional seja a segunda: *karta-kondusaun* (carta de condução), *xeke-viajen* (cheque de viagem), *previzaun-tempu* (previsão meteorológica). As palavras compostas dos elementos portugueses e nativos são uma verdadeira prova do hibridismo do tétum-praça: *Portugál-lian* (*lian* – ‘voz’, ‘língua portuguesa’), *amu-lulik/amlulik* (‘padre, sacerdote’, *amo* – ‘senhor’ + *lulik*-sagrado), *estrutura-laran* (*laran* – ‘interior’), *oras ne'e* (horas – esta(s) – ‘agora’), *materiál-hanorin* (*hanorin* – ‘ensinar’, ‘material pedagógico’), *sarani* – *katóliku* (*sarani* – ‘cristão, batismo’, uma palavra de origem árabe), *hahán-meiudia* (*hahán* – ‘comida’, ‘almoço’), *fó valór* (‘dar valor’), *loron-terça* (terça-feira).

14. Os gentílicos indicando nacionalidade ou grupo étnico em termos gerais podem ser grafados com maiúsculas: *Timoroan sira* (os timorenses); escrevem-se da mesma forma as palavras *señór* e *dona* antes do nome: *Señór Afonsu*, *Dona Adelina*.

Hoje em dia, após muitos anos de tentativas de standardização ortográfica, o tétum-praça parece poder ser designado como língua de escrita, com uma tradição bastante recente, mas com um número de textos crescente a cada ano. Nos textos em tétum-praça que analisámos, não encontramos nenhum caso de palavras com grafia portuguesa nem exemplos de hesitação ortográfica ou grafia macarrónica. A língua veicular de Timor

<sup>22</sup> As palavras esdrúxulas são todas emprestadas do português, do mesmo modo que a maior parte das graves (Hull, Eccles 2001: 230). Uma exceção a esta regra de acentuação constituem os vocábulos paroxítonos terminados em *-ia* e *-iu*, nos quais se mantém o acento português: *istória* (história); *kópia* (cópia); *polisia* (polícia); *disionáriu* (dicionário); *armáriu* (armário); *negósiu* (negócio); *relójiu* (relógio). A mesma situação observa-se no caso de hiatos: *saúde* (saúde); *jesuíta* (jesuíta); *proteína* (proteína), excetuando-se *saida* pronunciada como [sa'ida] ou [ˈsajda].

permanece cheia de empréstimos portugueses para o registo dos quais se seguem as normas ortográficas bem estabelecidas. Fazendo referência às opiniões de acordo com as quais o tétum-praça é uma língua deficiente, privada de meios para funcionar a cada nível da vida, achamos oportuno resumir a nossa pesquisa constatando o contrário.

#### REFERÊNCIAS BIBLIOGRÁFICAS:

- ALBARRAN DE CARVALHO Maria José, 2001, Panorama Linguístico de Timor, *Camões. Revista de Letras e Culturas Lusófonas* 14: 65–79.
- BORGES DE ALBUQUERQUE Davi, 2012, Língua e meio ambiente na literatura oral em língua Tetun, Timor Leste, *Language and Ecology* 3(3), [http://www.academia.edu/2121129/Lingua\\_e\\_meio\\_ambiente\\_na\\_literatura\\_oral\\_em\\_lingua\\_Tetun\\_Timor\\_Leste](http://www.academia.edu/2121129/Lingua_e_meio_ambiente_na_literatura_oral_em_lingua_Tetun_Timor_Leste) (consultado em 12.11.2013).
- COSTA Luís, 2000, *Dicionário de Tétum-Português*, Lisboa: Edições Colibri.
- DALGADO Sebastião Rodolfo, 1919, *Glossário luso-asiático*, vol. 1, New Delhi: Asian Educational Services.
- DURAND Frédéric, 2009, *História de Timor-Leste. Da pré-história à atualidade / Istória Timor-Lestean. Husi pre-istória to 'o atualidade*, Lisboa: Lidel.
- ESPERANÇA TAVARES João Paulo, 1997, *Política linguística em Timor Leste. Alfabetos e tentativas de normalização ortográfica do tétum*, Lisboa: Associação Luso-Timorense de Informação e Cultura.
- ESPERANÇA TAVARES João Paulo, 2001, *Estudos de linguística timorense*, Aveiro: SUL.
- GOLDEN Jill, 2004, *When the Diaspora Returns: Language Choices in Post-Independence Timor Lorosa'e*, (in:) *The Regenerative Spirit: (Un)settling, (Dis)location, (Post-)colonial, (Re)Presentations Australian Post-Colonial Reflections*, Sue Williams et al. (eds.), Adelaide: Lythrum Press, 115–126.
- HORTA SCHMIDT Korinna Mathilde, 1983, *O léxico português numa ilha remota da Oceânia*, (in:) *Atas do Congresso sobre a situação atual da língua portuguesa no mundo*, Lisboa: Instituto de Cultura e Língua Portuguesa, 350–359.
- HULL Geoffrey, 1999, *Towards a Language Policy for an independent East Timor*, (in:) *Studies in Languages and Cultures of East Timor*, Geoffrey Hull, Lance Eccles (eds.), Sydney: Academy of East Timor Studies, 1–7.
- HULL Geoffrey, 2000, *Short English-Tetum Dictionary*, Sydney: Sebastião Aparício da Silva Project.
- HULL Geoffrey, 2001a, *Timor-Leste. Identidade, língua e política educacional / Identidade, lian no política edukasionál*, Lisboa: Ministério dos Negócios Estrangeiros/Instituto Camões.
- HULL Geoffrey, 2001b, Língua, Identidade e Resistência, *Camões. Revista de Letras e Culturas Lusófonas* 14: 80–92.
- HULL Geoffrey, ECCLES Lance, 2001, *Tetum Reference Grammar*, Sydney: Sebastião Aparício da Silva Project.
- LOUREIRO José Luís et al. (coords), 2002, *Atlas de Timor Leste*, Lisboa: Lidel.
- LOUREIRO Rui Manuel, 2001, Onde nasce o sândalo: os portugueses em Timor nos séculos XVI e XVII, *Camões. Revista de Letras e Culturas Lusófonas* 14: 93–104.
- OSÓRIO DE CASTRO Alberto, 1996, *A Ilha Verde e Vermelha de Timor*, Lisboa: Cotovia.
- SÁ Artur Basílio de, 1961, *Textos em Teto da Literatura Oral Timorense*, vol. 1, Lisboa: Junta de Investigações do Ultramar.
- TAYLOR-LEECH Kerry, 2008, Language and identity in East Timor, *Language Problems and Language Planning* 32(2): 153–180.
- THOMAZ Luís, 2002, *Babel Loro Sa'e. O problema linguístico de Timor-Leste*, Lisboa: Instituto Camões.
- VAN KLINKEN Catharina, 1999, *Writing Tetum-Prasa: A Communication Challenge*, (in:) *Studies in Languages and Cultures of East Timor*, Geoffrey Hull, Lance Eccles (eds.), Sydney: Academy of East Timor Studies, 14–21.