Escreve um dicionário

Catharina Williams-van Klinken

Dili Institute of Technology

6/7/2016

## Porquê escrever um dicionário? [[1]](#footnote-1)

1. Pode ajudar as pessoas a aprender a(s) lingua(s).
2. Dá mais prestígio à língua.
3. Demonstra à geração futura bem como ao mundo exterior as riquezas da língua.
4. Contribui para a base de dados do mundo linguístico.

## Tipos de dicionários

* monolingue / bilingue / multilingue
* geral / específico (ex.: turismo, engenharia civil)
* referência / para crianças (ex.: dicionário escolar, dicionário ilustrado)
* livro / electrónico

## Tamanhos

* listas de palavras: 200-1000 palvras, inclui apenas palavras com apenas uma simples nota breve de explicação, ex. *lao – andar*.
* vocabulário / glossário: 500-2000 palavras, pode incluir mais de uma nota breve de explicação, ex. *lao – andar, função*. Alguns incluem partes do discurso.
* dicionário: 2000 ou mais palavras, incluí partes do discurso, definição completa, todos os significados de uma palavra; incluí compostos; alguns incluem exemplos.

## Informação a incluir

Uma definição completa pode incluir muita informação.

**hatene**

 *vt vcomp.*

 • know (something), understand, aware of; know how to, be good at

 *Category:* think

 *Similar:*

 kompriende *understand*

hatene tuir *understand well*

 Isaias: Nona hosi nebee? Marcos: Hau la hatene.

 *“Isaias: Where is Nona from? Marcos: I don't know.”*

 Nia hatene hanorin.

 *“She teaches well.*

|  |  |
| --- | --- |
|  |  |

**reimbolsa**

*vt.*

reimburse

 *Category:* economics

 *Usage:* educated

 *From:* Port *reimbolsar*

Informações podem incluir o seguinte dos quais os primeiros dois são obrigatórios:

1. a palavra na língua em que a descrição é feita
2. o seu significado
3. partes do discurso: nome, verbo, ...
4. informação gramatical: ex.: plural, género
5. pronunciação
6. ortografia alternativa: e.g. *kedas / kedan*
7. a origem da palavra: ex.: do português, do japonês, do latim
8. informação cultural: Normalmente esta informação não é incluída no dicionário. Contudo, se quiser documentar a língua que ainda não foi antes descrita com pormenor, pode incluir essa informação, em especial em artigos e práticas culturais, plantas e animais (ex.: os seus usos e as crenças culturais associadas)
9. examplo de frases ( com tradução, se não se tratar de um dicionário monolingue).
10. imagens. Se não for o autor das imagens, vai precisar de permissão para as utilizar. “A arte de ler” [The art of reading] (<http://bloomlibrary.org/artofreading>) tem muitas imagens que podem ser utilizadas na produção de imagens.
11. nomes científicos (para plantas e animais)
12. palavras que têm uma relação semântica com essa palavra. Por exemplo:
	* sinónimos: ex.: *hananu - canta*
	* quase sinónimos (ou palavras que são sinónimos em apenas alguns contextos): ex.: *rekursus – rikusoin*
	* antónimos: ex.: *diak – aat*
	* palavras com significados mais gerais: ex.: *gazoel - mina*
	* palavras com significados mais específicos: ex.: *Eskritura – Biblia*
13. uso: ex.: obsoleto, juventude, técnico, formal, educado
14. gravação da palavra, gravação de uma frase-exemplo.

## Onde encontrar as palavras

1. Traduzir uma lista de plavras. Ex.: SIL-CAWL no programa WeSay tem 1700 palavras para traduzir para inglês, português e indonésio.
	* Vantagem: poder escrever rapidamente o seu dicionário
	* Desvantagem: a sua língua pode conter algumas palavras que outras línguas não têm. Por exemplo, em inglês e português não existem as palavras *mak, fetosaa-umane*, ou *tisi*.
2. Procura por palavras em textos.
	* Onde pode encontrar os textos? Talvez haja textos escritos como no Facebook?
	* Se houver textos na Internet, pode fazer o seu download usando o programa gratuito Teleport Pro (<http://teleport-pro.en.softonic.com/>). Coloca a morada incial, e o programa procura por todos os textos que aparecem sob aquela morada. Por exemplo, pode baixar todos os artigos do website do STL ou Timor Post. Para prevenir que o programa baixe artigos em português, indonésio ou inglês que possam estar no website, pode especificar que palavras devem aparecer no texto (ex.: *iha, hanesan*).
	* Como encontra novas palavras nos textos?
		+ Pode ler os textos e escolher as palavras que quer acrescentar ao seu dicionário.
		+ Pode usar um programa do computador para fazer a lista de todas as palavras que aparecem nos seus textos.

Toolbox: Adicione textos no Toolbox (isto não é assim tão fácil). Despois escolha Tools – Word list para obter a lista de todas as palavras e o número de vezes em que cada uma aparece. Comece por adicionar as palavras que aparecem com mais frequência. Se elas aparecerem apenas uma ou duas vezes, é porque talvez tenha havido um erro na sua escrita ou variação na ortografia (ex.: *demokrsia*, *dadersan*) e não por a palavra ser nova.

dtSearch: Cria automaticamente um indíce de todos os ficheiros na sua língua. Depois escolhe *Search* e irá obter uma lista de todas as palavras dos seus textos e o número de vezes em que cada uma aparece

* + Não se esqueça de procurar pelas palavras compostas como por exemplo *matan dukur* e *ain kudeik*. Se estas forem escritas como palavras separadas, então o Toolbox e o dtSearch não as listará.
1. Procure por imagens em livros ou da internet, por ex.: livros sobre animais ou plantas.

Tenha cuidado para não forçar as pessoas a nomear cada animal ou planta; alguns podem ter os mesmos nomes ou não ter um nome e podem não existir em Timor.

1. Procure por palavras por domínio semântico, por ex.:palavras para cozinhar, partes do corpo ou do carro. Pode usar a opção do “Domínio semântico” no programa WeSay. É mais efectivo se fizer isso num grupo.

Um exemplo do WeSay:

**3.5.1.1: Voz: Que palavras se referem à voz da pessoa? (voz)**

* 1. Que palavras se referem a falar em voz alta? (gritar, falar alto, berrar, bradar, falar em voz alta, projectar a voz, exclamar, vociferar)
	2. Que palavras se referem a falar baixo/quietly? (falar baixo, falar com delicadeza, sussurar, murmurar)
	3. Que palavras se referm a falar muito? (falar muito, falador, loquaz, palavroso)
	4. Que palavras se referem a falar pouco? (falar pouco, taciturno, tímido, homem de poucas palavras, pouco comunicativo, reticente, segurar a língua)

# Programas de lexicografia

É mais fácil adicionar palavras se utilizar Word ou Excel para preparar o dicionário. Porém, quando quiser imprimir o dicionário, irá ter muitos problemas, como por exemplo na formatação de cada palavra, de cada página e fazer uma versão inversa do dicionário (ex.: converter o dicionário de Tokodede-Tétum num dicionário de Tétum-Tokodede)

 É portando muito melhor usar o programa desenhado para fazer dicionários. (Os programas seguintes são todos gratuitos)

* WeSay (<http://www.sil.org/resources/software_fonts/wesay>). É fácil para adicionar palavras, pode escolher a parte do discurso da lista. Existe uma lista de 1700 palavras em inglês, português e bahasa-indonésio que pode traduzir para a sua língua para inciar o dicionário de modo rápido. Existe também uma lista de questões bastante detalhada sobre domínios semânticos para ajudar a preparar um dicionário mais completo.
* LexiquePro ([www.lexiquepro.com/](http://www.lexiquepro.com/)). Pode exportar dados do programa WeSay para LexiquePro e usar o LexiquePro para produzir um dicionário interactivo e/ou um dicionário impresso.
* FLEx: FieldWorks Language Explorer (<http://fieldworks.sil.org/flex/>). Este prgrama foi escrito para linguistas e pode ser usado para produzir dicionários, analisar a morfologia das palavras, inter-linearizar textos e estudar discursos.

# References

* van den Berg, R. (2012). *Making dictionaries: A course for national translators*. Ukarumpa: SIL Papua New Guinea.

# Dictionaries in local languages

### Atauro languages

Duarte, J. B. (1990). *Vocabulário Ataúro-Português Português-Ataúro*. Lisbon, Instituto Português do Oriente. (Listed in Hull 1998)

Boarccaech, Alessandro (2013). *Dicionário Hresuk-Português*. In: A Diferença entre os Iguais. São Paulo: Editora Porto de Ideias. http://issuu.com/kirstygusmao/docs/dicionario\_hresuk\_-\_portugues/1

### Baikenu

Hull, Geoffrey (2006). *Disionáriu Badak / Disionari Pala' Baikenu-Tetun Tetun-Baikenu*. Dili: Insituto Nacional de Linguística.

### Bunak

Bele, A. (2009). *Kamus bahasa Buna'-Indonesia*. Kupang: Gita Kasih.

### Fataluku

Hull, G. (2006). *Disionáriu Badak / Disionari Kava Fataluku-Tetun Tetun-Fataluku*. Dili, Insituto Nacional de Linguística.

Nacher, A. (2003). "Léxico Fataluco-Português; Primeira Parte." *Studies in Languages and Cultures of East Timor* **5**: 138-196.

Nacher, A. (2004). "Léxico Fataluco-Português; Segunda Parte." *Studies in Languages and Cultures of East Timor* **6**: 119-177.

Valentim, J. (2002). *Fata-lukunu i disionariu / Dicionário Fataluku / Dictionary*, Timor Lorosa'e Nippon Culture Centre. (monolingual Fataluku dictionary)

Valentim, J. (2006). Disionariu Tetun Fataluku Ingles / Dictionary: Tetun - Fataluku - English, Timor Lorosa'e Nippon Culture Centre.

### Galolen

da Silva, M. M. A. (1905). *Diccionario Portuguez-galoli*. Macao, Typographia Mercantil.

Viegas, V. S. (unpublished). *Dicionario Classico Elementar Tetum/Galolem/Portugûes*.

### Idate

Ferreira, Luisa (2005). *Ma ta su’ar idate [Idaté-English dictionary]*, Timor Lorosa’e Nippon Culture Center.

### Kemak

Gomes, M. (2004). *Kemak*. Hawaii, University of Hawaii Documentation Project. (Phonemes, vocabulary for 300 word Swadesh list.)

### Makalero

Pinto, T. M. (2004). *Makalero Disionariu*. Dili, Timor Loro Sa’e-Nippon Culture Center.

### Makasae

da Costa Correia, J. and D. Morais da Silva (2006). *Vocabulário Makasae - Tetum - Português – Inglê Simposio Makasae 13th Agostu 2005*. Timor Lorosa'e Nippon Culture Center.

Hull, G. (2006). *Disionáriu Badak / Disionariu gi Digara Makasae-Tetun Tetun-Makasae*. Dili, Insituto Nacional de Linguística.

Nacher, A. (unpublished). *(Fataluku-Makasai-Tetum dictionary)*.

### Makuva

Fataluku Language Project. <http://www.fataluku.com/makuva/> [includes a text, and some scanned pages of a 1952 wordlist by by António de Almeida]

### Naueti

Menezes, António, and António Adelino Pires (2006). *Nauweti Tetun English: Mai hira eyi/nenehe te'e le'a wa'a nauweti ta*. Timor Loro Sa'e-Nippon Culture Center.

### Tetun Terik

Monteiro, F. (1985). *Kamus Tetun-Indonesia*. Jakarta, Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan. (Examples are Tetun Terik. Unsure which dialect.)

Morris, C. (1984). *Tetun - English dictionary*. Canberra, Pacific Linguistics, C-83. (Words from several dialects. Grammatical notes on Tetun-Los (south coast). Hull claims this is a translation and amplification of Mendes' dictionary, which was prepared by Paulo Quintao da Costa).

### Tokodede

Lekede'e Study Group (2006). *Disionáriu Lia Tokodede - Tetun - Ingles*. Timor Loro Sa'e-Nippon Culture Center.

Lekede'e Study Group (2006). *Lia fuan ou termu adat ema Tokodede Likisá*. Timor Loro Sa'e-Nippon Culture Center.

# Disionariu tekniku: Tetun

Iha organizasaun barak nebee prepara ona lista lia fuan, maibee maioria nunka foo sai ba publiku.

(2002). Glossário de termos seleccionados da Formação Proffissional em Português, Tetum e Inglês / Vocational training: a glossary of selected terms in Portuguese, Tetum and English. Dili: Secretaria de Estado do Trabalho e Solidariedade.

(2002). Glosáriu Edukasaun Sívika ho lian Tetun, Portugés, Inglés, Indonézia. Dili: UNDP.

ARD, MOI, & UNTL. (2004). Glossário de Terras. Retrieved from http://pdf.usaid.gov/pdf\_docs/Pnadb833.pdf

Independent Electoral Commission and UNDP (2002). *Glosáriu eleisaun nian: lia fuan kona ba eleisaun*. Dili, East Timor. [Developed by Catharina Williams-van Klinken, in conjunction with Nuno Gomes, Fransisco Luis Fernandes and Bernardino Menezes. www.tetundit.tl

Qintão, G. da Silva (2007). *Disionáriu jerál agrikultura*. Dili, Timor Aid.

—— (2008). *Disionáriu tékniku pekuária no veterinária*. Dili, Timor Aid.

de Souza, I., & Morgan de Queiroz, J. (2012). *Glossário multilíngue de ferramentas e instrumentos: Português – Tétum – Indonésio - Inglês*. Hera: Faculdade de Engenharia, Ciências e Tecnologia, Universidade Nacional Timor Lorosa’e.

UNTAET Human Rights Unit (1999). Human rights glossary. Dili.

USAID Land law Program II (no date). *Glosáriu agrikultura nian*. Dili.

USAID Land law Program II (no date). *Glosáriu lei rai nian*. Dili.

USAID Land law Program II (no date). *Glosáriu ekonomia nian*. Dili.

Williams-van Klinken, C., Cham, M. F. X., & Alkatiri, S. A. (2013). *Glosariu informatika ho lian Tetun, Portuges, Ingles ho Indonezia*. Dili: Dili Institute of Technology.

Actividades: Escrita do dicionário

Catharina Williams-van Klinken

Dili Institute of Technology

6/7/2016

1. Porque quer escrever um dicionário?
2. Que tipo de dicionário quer produzir?
	1. Língua(s): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
	2. Geral / área específica. Se específica, em que área? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
	3. Referência / crianças
	4. Livro / electrónico / ambos
	5. Tamanho: \_\_\_\_\_\_\_\_\_ palavras
3. Que informação gostaria de incluir?
	1. a palavra na língua em que está a descrever
	2. o seu significiado
	3. as partes do discurso
	4. a informação gramatical: ex. plural, género
	5. a pronunciação
	6. a ortografia alternativa: ex.: *kedas / kedan*
	7. a origem da palavra: ex.: do português, japonês, latim
	8. a informação cultural
	9. o exemplo de frases (com tradução, se não se tratar de um dicionário monolingue)
	10. imagens. De onde irá buscar as imagens?
	11. o nome científico (para plantas e animais)
	12. plavras semanticamente relacionadas
	13. o uso
	14. a gravação da palavra, gravação de uma frase de exemplo
4. Onde irá procurar pelas palavras?
	1. Traduza de uma lista de palavras.
	2. Procure por palavras em textos. Será que os textos existem? Onde? (Facebook? Jornais? Gravações? Livros?)
	3. Use o “dicionário em branco”.
	4. Imagens. Onde irá encontrá-las?
	5. Procure as palavras pelo domínio semântico. Use o WeSay?
5. Trabalho em pares ou em grupos pequenos e escreva o maior número de palavras e expressões que saiba na sua língua para cada um dos domínios semânticos seguintes. Depois de fazer a lista longa quanto possível tenta descobrir quais as diferenças entre algumas das palavras e expressões.
	1. ‘cortar’: (ex.: cortar, trinchar, aparar)

‘praise’:

* + 1. ‘elogiar’: Que palavras se referem para elogiar alguém? (elogiar, aclamar, enaltecer, cumprimentar, exaltar, glorificar, honrar, dar crédito, reconhecer, fazer elogios, falar bem, expressar aprovação)
		2. ‘agradecer’: Que palavras se referem para agradecer alguém? (agradecer, dizer obrigado, expressar agradecimento, apreciar, mostrar apreciação, expressar gratitude, demonstrar gratitude, reconhecer, em reconhecimento de)
		3. Que palavras se referem para lisojear alguém? (lisonjear, cair nas boas graças de alguém, adular, elogiar)
		4. Que palavras se referem para alguém gabar-se de si próprio? (gabar-se, orgulhar-se, blasonar, bravatear, vangloriar-se, tocar o seu próprio chifre)
1. Se tiver um portátil consigo, instale o programa WeSay e introduza alguns dados usando o “a lista de palavras da CAWL” e depois use os “domínios semânticos”.
1. Muita da informação apresentada baseia-se no trabalho de van den Berg (2012). [↑](#footnote-ref-1)