

Adequação do vocabulário DeCS à nova ortografia da língua portuguesa

Arthur Alberto Corrêa Treuherz;
Ana Cristina Espindola Campos;
Heitor Barbieri;
Olga Pedroza Ribeiro.

BIREME - OPAS - OMS

Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde

O Acordo Ortográfico da Língua Portuguesa de 1990 é um tratado internacional firmado em 1990 com o objetivo de unificar a ortografia portuguesa, a ser usada por todos os países de língua oficial portuguesa. Foi assinado por representantes oficiais de Angola, Brasil, Cabo Verde, Guiné-Bissau, Moçambique, Portugal e São Tomé e Príncipe em Lisboa, em 16 de dezembro de 1990. Timor-Leste aderiu ao Acordo em 2004, após a independência.

O DeCS (Descritores em Ciências da Saúde) é um vocabulário trilingue usado para a indexação de documentos técnico-científicos das bases LILACS e MEDLINE, dentre outras.

Objetivo

O presente trabalho visa atender a determinação do Decreto 6583/08 do governo brasileiro de atualização da ortografia da língua portuguesa até 31 de dezembro de 2012.

Assim, as versões em português do DeCS a partir de 2013 compreenderão a nova ortografia. Em virtude do grande número de vocábulos em português no DeCS, o objetivo do trabalho foi permitir o máximo de automatização das alterações necessárias, com o mínimo de erro.

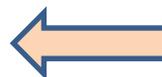
1. Transformação dos conteúdos dos campos de termos, sinônimos e notas do DeCS em português em palavras



2. Eliminação de repetições



3. Processamento das palavras no corretor ortográfico do LibreOffice



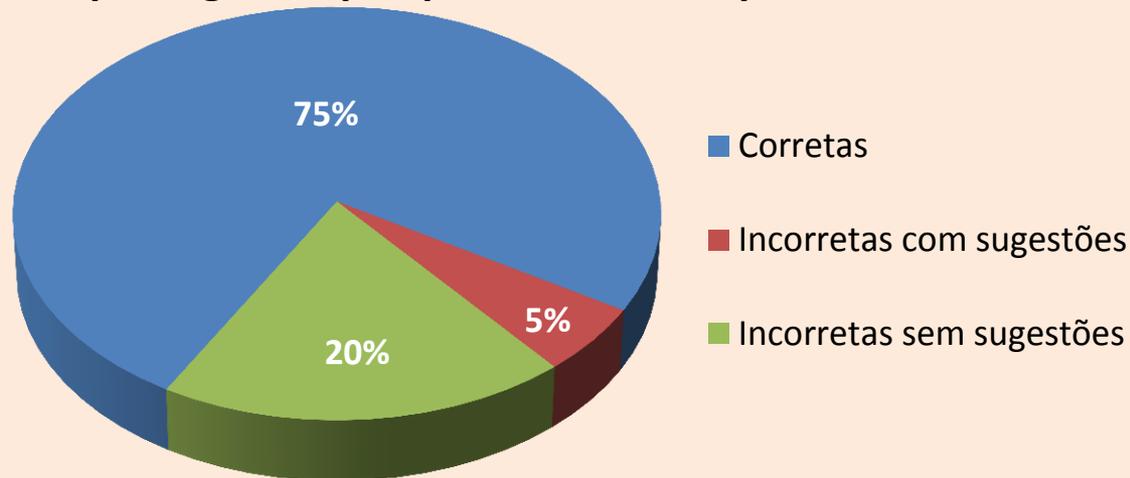
4. Separação do resultado da análise do corretor em três distintos arquivos

a) Palavras corretas

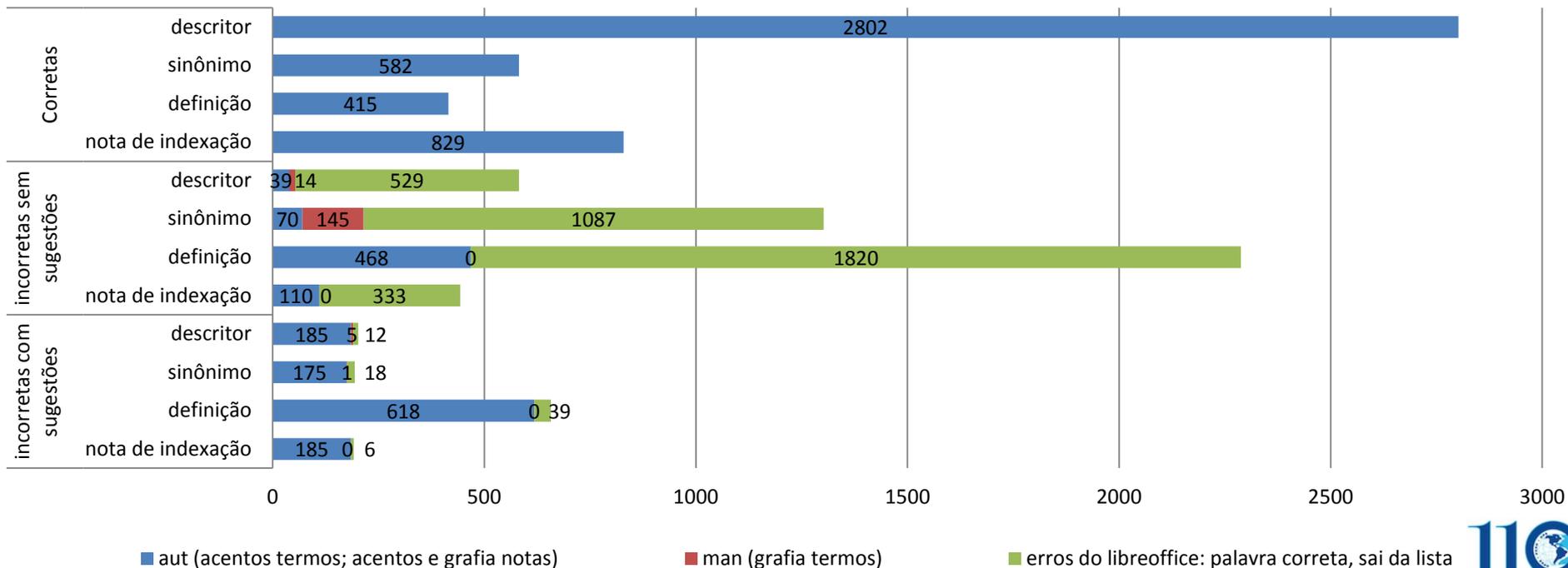
b) Palavras incorretas com sugestões de correção pelo LibreOffice

c) Palavras incorretas sem sugestões de correção pelo LibreOffice

Percentual de palavras “Corretas”, “Incorretas com sugestões” e “Incorretas sem sugestões” do universo contendo acento e/ou hífen, nos campos de descritores, sinônimos, definições e notas de indexação em português, após processamento pelo LibreOffice



Palavras identificadas pelo LibreOffice e alteradas de forma automática ou manual após análise



Exemplos de casos

A. Casos tratados pelo procedimento

1. Acentos e grafia - sugestões **acertadas** do LibreOffice:

Intra-Ocular → Intraocular

Intra-Radicular → Intrarradicular

Amilóide → Amiloide

2. Acentos – **sugestões errôneas** do LibreOffice (linhas excluídas após análise):

Braquiúros → Braquiuros (erro)

Intermarés → Intermares (outro conceito)

Moiré (palavra francesa) → Moire (outro conceito)

Subaracnóideo → Subaracnoideo (erro)

Áxis → Axis (erro)

Ésteres → Esteres (outro conceito)

Exemplos de casos (análise)

3. Grafia - sugestões do LibreOffice e **ajuste** manual para padrão MeSH / DeCS:
Aminooxiacético → Amino-oxiacético (LibreOffice) → Amino-Oxiacético

4. Incorretas sem sugestão (palavras **técnicas**)

Acentos, hífen – regras: éia(s), éica(s), éico(s), óica(s) e óico(s), -hexa, etc. :

Docosahexaenóicos → Docosa-Hexaenoicos

Feniltiazoliltiouréia → Feniltiazoliltioureia

B. Casos não tratados pelo procedimento, resolvidos manualmente

1. Hífen inserido:

Antiidiotipos → Anti-Idiotipos

Diazo~~o~~onorleucina → Diazo-Oxo-Norleucina

Eletro~~o~~smose → Eletro-Osmose

Formiltetra~~id~~rofolatos → Formiltetra-Hidrofolatos

Exceções (grafia mantida):

Coniinum; Biiodatus; Aliivibrio

Discussão

- O porquê da escolha do LibreOffice

Ao começar a revisão, os editores LibreOffice juntamente com o OpenOffice foram os primeiros a oferecer corretores ortográficos adaptados às novas regras. Além disso, o Libre/Open Office era o único que possuía uma aplicação com a linguagem de programação Java (utilizada na Bireme) disponível para a manipulação do dicionário interno.

Observação: O LibreOffice surgiu como um projeto que se originou do OpenOffice, por isso o que vale para um vale para outro, mas como o projeto do LibreOffice ficou mais ativo, disseminado e “moderno” que o OpenOffice então este foi citado como o editor utilizado, embora o dicionário e a biblioteca de manipulação sejam os mesmos nos dois editores.

- Trema foi eliminado automaticamente após análise visual das palavras (em nomes próprios e seus derivados, o trema deve ser mantido)

Fontes Consultadas

[Academia Brasileira de Letras](#)

[VOLP \(Vocabulário Ortográfico da Língua Portuguesa, 2009\)](#)

[Publicação da Folha](#)

[Artigo da Folha](#)

[Artigo do Jornal Folha de São Paulo](#)

[Wikipedia](#)

[Fapesp](#)

[Guia Melhoramentos](#)

[BROffice](#)

[LibreOffice](#)

[Artigo da Teresa Avalos , bibliotecária e profa de inglês da BC da UNIFESP](#)

[Arquivo do curso de 29/4/2009, da Teresa Avalos](#)

[Arquivo com exercícios de ortografia, da Teresa Avalos](#)

[Referência sobre hífen](#)