

Léxico em foco: dicionários com que sonhamos

**Cristina Martins Fargetti
Clotilde de Almeida Azevedo Murakawa
Odair Luiz Nadin
(Org.)**

**CULTURA
ACADÊMICA**
Editora

**LÉXICO EM FOCO:
DICIONÁRIOS COM QUE
SONHAMOS**

SÉRIE
TRILHAS LINGUÍSTICAS
n° 32 – 2019

**Faculdade de Ciências e Letras, UNESP – Univ Estadual Paulista,
Campus Araraquara**

Reitor: Sandro Roberto Valentini

Vice-Reitor: Sergio Roberto Nobre

Diretor: Cláudio Cesar de Paiva

Vice-Diretora: Rosa Fátima de Souza Chaloba

Programa de Pós-Graduação em Linguística e Língua Portuguesa

Coordenador: Prof. Dr. Jean Cristtus Portela

Vice-Coordenadora: Profa. Dra. Alessandra Del Ré

SÉRIE TRILHAS LINGUÍSTICAS Nº 32

Comissão Editorial da Pós-graduação em Linguística e Língua Portuguesa

Alessandra Del Ré

Angélica Terezinha Carmo Rodrigues

Cristina Martins Fargetti

Jean Cristtus Portela

Marina Célia Mendonça

Nildicéia Aparecida Rocha

Odair Luiz Nadin da Silva

Rosane de Andrade Berlinck

Diagramação: Eron Pedroso Januskevictz

Normalização: Biblioteca da Faculdade de Ciências e Letras

**LÉXICO EM FOCO:
DICIONÁRIOS COM QUE
SONHAMOS**

Cristina Martins Fargetti
Clotilde de Almeida Azevedo Murakawa
Odair Luiz Nadin
(Org.)

CULTURA
ACADÊMICA 

Editora

Copyright © 2019 by FCL-UNESP Laboratório Editorial
Direitos de publicação reservados a:
Laboratório Editorial da FCL

Rod. Araraquara-Jaú, km 1
14800-901 – Araraquara – SP
Tel.: (16) 3334-6275

E-mail: laboratorioeditorial.fclar@unesp.br
Site: <http://www.fclar.unesp.br/laboratorioeditorial>

Obra disponível em formato impresso e eletrônico
(consultar endereço acima).

L591 Léxico em foco: dicionários com que sonhamos / Organizado por:
Cristina Martins Fargetti, Clotilde de Almeida Azevedo
Murakawa e Odair Luiz Nadin. –
São Paulo, SP : Cultura Acadêmica, 2019.
272 p. ; 14x21 cm. – (Série Trilhas Linguísticas; 32)

ISBN 978-85-7249-018-4

1. Lexicografia. 2. Enciclopédias e Dicionários. 3. Lexicologia.
I. Fargetti, Cristina Martins. II. Murakawa, Clotilde de Almeida Azevedo.
III. Nadin, Odair Luiz. VI. Série.

CDD 410

SUMÁRIO

Apresentação	
<i>Cristina Martins Fargetti</i>	7
A pesquisa reversa: dicionários codificadores e dicionários onomasiológicos	
<i>Álvaro Iriarte Sanromán</i>	15
Verbetes que sonhamos... Verbetes que fazemos: a equivalência em recentes trabalhos lexicográficos	
<i>Claudia Zavaglia</i>	31
Dicionários que fizemos: algumas reflexões sobre o Dicionário Histórico do Português do Brasil – séculos XVI, XVII e XVIII (DHPB)	
<i>Clotilde de Almeida Azevedo Murakawa</i>	49
Léxico histórico do português brasileiro: caminhos percorridos, metas a atingir	
<i>Vanderci de Andrade Aguilera e Fabiane Cristina Altino</i>	67
O desafio da retrodatação para os estudos etimológicos de língua portuguesa	
<i>Mário Eduardo Viaro, Aldo Luiz Bizzocchi, Mariana Giacomini Botta, Marco Dimas Gubitoso e Gustavo Luiz Vieira</i>	83
Dicionários para línguas indígenas do Brasil	
<i>Jorge Domingues Lopes e Ana Suelly Arruda Câmara Cabral</i>	101

Dicionário do Nordeste <i>Nelly Carvalho</i>	115
Quando a elaboração de dicionários pode auxiliar na reafirmação de uma identidade étnica: o caso dos professores Mura de Autazes <i>Cristina Borella e Eneida Alice Gonzaga dos Santos</i>	127
A elaboração do dicionário crítico da obra de Domingos Vieira Filho <i>Conceição de Maria de Araujo Ramos, José de Ribamar Mendes Bezerra, Luís Henrique Serra e Maria de Fátima Sopas Rocha</i>	143
Brincadeiras infantis e o léxico regional: um estudo de dados geolinguísticos <i>Láisa Bauermeister Stelo e Aparecida Negri Isquerdo</i>	155
O vocabulário rural da Serra da Canastra/MG: um estudo linguístico na nascente do rio São Francisco <i>Gisele Aparecida Ribeiro</i>	167
O regional e o global nos nomes próprios de pessoas no Brasil <i>Jéssica Paula Vescovi e Márcia Sipavicius Seide</i>	183
A regionalidade na seleção lexical de informantes do Atlas Linguístico do Estado de Alagoas (ALEAL) <i>Maranúbia Pereira Barbosa Doiron</i>	199
Criações lexicais literárias: a irreverência expressa pelas criações lexicais na poética de João Cabral de Melo Neto <i>Rosana Maria Sant'Ana Cotrim</i>	215
Informatização e trabalho humano em Terminologia: extrair terminologias em <i>corpus</i> é o mesmo que identificá-las? <i>Maria José Bocorny Finatto</i>	231
Anotação de <i>corpus</i> do grego antigo: projeto de tradução alinhada, <i>treebank</i> e léxico grego-português em serviços <i>web</i> <i>Anise D'Orange Ferreira</i>	247

A PESQUISA REVERSA: DICIONÁRIOS CODIFICADORES E DICIONÁRIOS ONOMASIOLÓGICOS

Álvaro IRIARTE SANROMÁN

Dicionários codificadores e dicionários onomasiológicos

Apesar do princípio de arbitrariedade do signo linguístico, estabelecido por Saussure (1999), há um século, no seu *Curso de Linguística Geral*, muito frequentemente esquecemos que o significado de uma palavra (isto é, o seu sentido) é, de facto, o significado dessa palavra quando combinada com outras palavras. Com efeito, Saussure entende este carácter arbitrário do signo não só como oposto à motivação onomatopeica, no sentido óbvio de que não existe uma relação direta entre o significante e o significado mas apresenta também uma muito mais importante concepção de arbitrariedade que o leva a conceber a língua como uma estrutura em que se estabelecem relações. Saussure fala em arbitrariedade porque, em português, “**passeio**” combina com “**dar**” (dar um passeio), em francês, “*promenade*” combina com “*faire*” (*faire une promenade*), em inglês, “*walk*” combina com “*go for*” ou com “*take*” (*go for a walk / take a walk*).

É essa concepção discreta, ou atomista, da linguagem (uma palavra ligada a um significado; um conjunto de significantes ligados, um a um, a outros tantos significados) que leva, muito frequentemente, os dicionários a apresentarem como aceção de uma

palavra o que, em rigor, é o significado dessa palavra combinada com outras. Veja-se, por exemplo, na entrada **ouvido**¹ do dicionário Aurélio, como a aceção **3** tem o mesmo valor que a combinação “**ter bom ouvido**”, também registrada na parte sintagmática do dicionário:

ouvido. [Part. de *ouvir*] S. m. **1.** Faculdade de ouvir, [...] **2.** Anat. Cada um dos dois [...]. **3.** Aptidão para captar com relativa precisão sons musicais ou não, e de reproduzir aqueles sem o auxílio de partitura: *ter bom ouvido* [...]

• [...] **Ter bom ouvido.** Ter fácil percepção de sons, especialmente musicais. (Aurélio. FERREIRA, 1999).

Neste caso (que não é único nos dicionários portugueses) regista-se uma contradição, pois o mesmo significado (‘aptidão para captar com relativa precisão sons musicais’) é atribuído tanto à palavra **ouvido** como à expressão **ter bom ouvido**, que aparece como locução no fim do artigo.

Tal valor corresponde, de facto, à combinação lexical **ter bom ouvido**, devendo ser registado, portanto, na parte sintagmática ou combinatória do dicionário. É o que acontece, por exemplo, no *Caldas Aulete*, um grande dicionário quanto ao tratamento da combinatória lexical, como já indicava Rodrigues Lapa (1984)².

Se o adjetivo **cego** só adquire determinado valor quando combinado com o substantivo **amor** e o adjetivo **mortal** só adquire esse mesmo valor quando combinado com o substantivo **ódio**, não deveríamos explicitar isso nas entradas correspondentes? Contudo, nos dicionários, continuamos a deparar-nos com entradas que recolhem um suposto sentido literal e outro sentido figurado das palavras, consideradas isoladamente:

¹ O exemplo é adaptado de Calderón Campos (1994, p. 58).

² “É precisamente neste capítulo da fraseologia, muito importante, que os dicionários correntes deixam mais a desejar. O mais celebrado de entre eles e o mais moderno dos grandes dicionários, o de Cândido de Figueiredo, é muito pobre em grupos fraseológicos, o que constitui um grave defeito, porque é nessas locuções que se imprime o chamado génio da língua. Como repositório de fraseologia, nada há que possa substituir entre nós o *Dicionário Contemporâneo* de Caldas Aulete.” (LAPA, 1984, p. 83).

cego [...] **1.** Que não pode ver: ... **2.** FIG. em alto grau, intenso, descontrolado, ... [...];

mortal [...] **1.** letal: ... **2.** FIG. em alto grau, intenso, descontrolado, ... [...].

Repare-se que o adjetivo **cego** só tem o significado recolhido na aceção **2** quando combinado com a palavra **amor** (e, ocasionalmente, duas ou três palavras mais). Trata-se de um fenómeno lexical, não semântico ou sintáctico, que devemos assumir como claramente irregular. Por isso, não é possível recorrer aqui ao uso de um “etc”: “**cego** [...] **2.** FIG. em alto grau, intenso, descontrolado [amor, **etc.**]” ou ao uso de um eventual hiperónimo ou nome genérico: “**cego** [...] **2.** FIG. em alto grau, intenso, descontrolado [aplicado a **sentimentos**]”.

Não podemos generalizar dizendo que o adjetivo *cego* tem este significado quando aplicado a “sentimentos”. Não é possível usar estas estratégias simplesmente porque outros substantivos, que também expressam sentimentos, seleccionam adjetivos diferentes para significar “em alto grau, intenso...”, como é o caso de:

admiração ⇒ extrema;

afeto ⇒ enorme;

agonia ⇒ profunda;

alegria ⇒ imensa;

amargura ⇒ profunda;

apreço ⇒ especial;

beleza ⇒ infinita;

consolo ⇒ enorme;

culpa ⇒ grande;

deceção ⇒ forte;

desejo ⇒ ardente;

desgosto ⇒ grande;

desinteresse ⇒ total;

desprezo ⇒ absoluto;

dor ⇒ viva;

dor ⇒ forte;

dor ⇒ intensa;

entusiasmo ⇒ caloroso;

felicidade ⇒ indescritível;

frustração ⇒ enorme;

infelicidade ⇒ grande;

interesse ⇒ vivo;

ódio ⇒ figadal;

ódio ⇒ mortal;

pânico ⇒ absoluto;

pesar ⇒ grande;

prazer ⇒ sumo;

preocupação ⇒ enorme;

saudade ⇒ imensa;

saudade ⇒ infinita;

tristeza ⇒ profunda;

vontade ⇒ louca;

etc.

Da mesma maneira, não é possível encontrar um hiperónimo que abranja todos os nomes que selecionam, por exemplo, a adjetivo **forte** para significar “em alto grau, intenso ...”:

alto +

preço

qualidade

conceito

voz

traição

matemática

inteligência

estilo

façanha

hiperónimo ?

Falámos acima na necessidade de explicitar, na entrada correspondente, esta informação (que um adjetivo como **cego** só adquire determinado valor quando combinado com o substantivo **amor**). Mas em que entrada é que deveríamos fornecer esta informação? Sob o lema **cego** ou sob o lema **amor**? Que palavra é que devo procurar (ou deveria poder procurar) quando o que desconheço é que, em português, **passeio** combina com **dar** ou que é **cego** o adjetivo que devo usar para exprimir a ideia de “amor muito intenso, ou não controlado pela razão”? O utilizador não poderá procurar informação no artigo **dar** ou no artigo **cego** porque desconhece que são estas as palavras que em português se combinam com **passeio** e com **amor**, respetivamente, para exprimir os sentidos desejados. Mas é isso o que acontece na maior parte dos dicionários.

O problema coloca-se quando nos situamos na direção de síntese, ou codificação ou produção textual, devido a que as possibilidades colocacionais (combinatórias) variam de uma língua para outra (e, em geral, os usos léxico-sintático e pragmáticos).

Um falante espanhol não terá problemas para entender os significados das combinações “**tirar uma fotografia**” e “**tirar conclusões**”, mas poderá ter dificuldades para produzir tais combinações em português (“tirar una fotografia”, em espanhol, significa “deitar

fora uma fotografia” e as “*conclusiones*”, em espanhol, “*se sacan*”). Neste caso, este utilizador não poderá procurar a informação que necessita sob o lema **tirar** porque desconhece que é esta a palavra que em português se combina com **fotografia** e com **conclusão**. Deveria poder consultar as entradas **fotografia** e **conclusão** para saber qual o verbo que combina com estes substantivos para exprimir o sentido pretendido. Mas, tal como estão elaborados a maior parte dos nossos dicionários, o utilizador só poderá consultar essa informação na entrada **tirar** (justamente a palavra que ele desconhece!):

tirar [...] v. [...] **II**. Como verbo suporte de predicção, combina-se com nomes ... **1**. Actos físicos. ≈ FAZER. ... *Tirar uma fotografia* ... **2**. Actos morais ≈ OBTER. *Tirar conclusões* (Academia. CASTELEIRO, 2001).

tirar [...] v. **31**. captar (imagem), ... fotografando, ...: *t. uma foto de uma pessoa*. (HOUAISS, 2001).

tirar V. t. d. [...] **15**. Fazer (uma fotografia [2]); *Fique aí quieto, vou tirar a fotografia*. **16**. Fazer tirar, parar para tirar (uma fotografia [2]); *Aprontou-se toda para tirar o retrato*. [...] (Aurélio. FERREIRA, 1999).

TIRAR V. tr. [...] *Tirar (alguém) o retrato, fazer-se retratar: Fui tirar o retrato para a carteira de identidade.* } || *Tirar o retrato a alguém, fazer-lhe o retrato: [...]* (Caldas Aulete. AULETE, 1987).

tirar V. t. [...] Derivar: *tirar conclusões*. [...] (Cândido de Figueiredo. FIGUEIREDO, 1982).

Só no dicionário *Houaiss*³ e no *iDicionário Aulete* (AULETE DIGITAL, 2008) é que temos também a informação na entra-

³ Aliás, no dicionário Houaiss (2001), sob a informação gramatical da entrada **dar** podemos ler: “**a. 3**) por sua importância, diversas acepções de *dar*, usado como verbo-suporte, estão registradas no corpo deste verbete; diversas outras devem ser procuradas pelo substantivo que faz parte do objeto direto, como de hábito no restante dicionário”.

da fotografia (mas nada na entrada “conclusão”): “**fotografia** [...] **tirar f.**: acionar uma câmara para obter uma imagem [...] (HOUAISS, 2001).”, “**fotografia** [...] **Tirar fotografia**: 1. Fotografar [...] (AULETE DIGITAL, 2008).”

Outra exceção é o pequeno *Dicionário Básico da Língua Portuguesa* (VILELA, 1991), com um magnífico tratamento da questão:

tirar [...]

S. 8 *tirar + nome, sentido VIII*, equivale a um verbo simples: CONCLUIR (FRASE 1), [...] (VILELA, 1991).

conclusão [...]

[...] // (pessoa) **tirar conclusões**: (5) – Que conclusões podemos tirar da sua atitude? • (6) - Não quero tirar conclusões erradas do caso. [...]

S. • *Tirar conclusões* (frases 5, 6) tem como sins.: CONCLUIR, TIRAR ILAÇÕES OU DEDUÇÕES ... (VILELA, 1991).

Afirmar que o significado de uma palavra é, de facto, o significado dessa palavra quando combinada com outras não implica que estes conjuntos de palavras devam ser considerados como lemas nos nossos dicionários. Há várias razões para que as entradas dos dicionários (lemas) devam continuar a ser palavras individuais (no sentido em que as pessoas instintivamente entendem por palavra: um conjunto de letras delimitadas por dois espaços em branco, um espaço em branco e um signo de pontuação, etc.), nomeadamente, as expectativas conservadoras dos utilizadores de dicionários (COWIE, 1983) e o bem-fazer da tradição lexicográfica, com excelentes amostras de dicionários “*user-friendly*” (BERGENHOLTZ; TARP, 1995).

Num dicionário tradicional (em papel ou digitalizado), podemos procurar informação sobre uma palavra percorrendo os lemas incluídos na nomenclatura, normalmente ordenados por ordem alfabética. Mas o que acontece quando não sei que palavra é que devo procurar? Por exemplo, como se chama a caixa em que se

transportam e comercializam as garrafas de cerveja? Que eufemismo ou termo (politicamente) correto posso usar em vez de um eufemismo? Neste caso usamos os chamados dicionários onomasiológicos, que são dicionários orientados ao conceito, não à palavra. Quando usamos um dicionário onomasiológico, o que procuramos é uma ideia, um conceito e, normalmente, os resultados são as maneiras de exprimir essa ideia em palavras. Para isso, navegamos por lista estruturada de conceitos, ordenada por assuntos, com a ajuda de listas de hiperónimos ou termos mais genéricos, que nos conduzem à palavra que procuramos.

Implementação de ferramentas de pesquisa reversa nos dicionários eletrônicos

A implementação de ferramentas de pesquisa reversa nos dicionários eletrônicos transforma estas obras em dicionários codificadores e dicionários onomasiológicos, permitindo-nos encontrar não apenas as palavras que correspondem a uma ideia (como nos dicionários onomasiológicos) como também que palavra pode ser combinada com outra para exprimir uma ideia (como nos dicionários codificadores).

É comum encontrarmos ferramentas de pesquisa avançada nos dicionários eletrônicos (embora, lamentavelmente, não tanto nas versões *on-line* dos dicionários). Essas ferramentas permitem, por exemplo, pesquisar por formas de palavra, e não apenas por lema; pesquisar palavras ortograficamente semelhantes; pesquisar na microestrutura, e não apenas navegar pela macroestrutura, utilizando mais do que um termo de pesquisa⁴; pesquisar partes do lema, permitindo estudar, por exemplo, a produtividade de determinados afixos; etc.

⁴ Fazendo, deste modo, um uso onomasiológico e codificador do dicionário:

– “Como se diz endurecer o aço ?” ⇒ (metal + endurecer ⇒ *temperar*)

– “O que acontece à água com o frio?” ⇒ (água + frio ⇒ *desnevado, fresco, gelo, neve*)

– “Quem é o médico dos olhos?” ⇒ (médico + olhos ⇒ *oculista, oftalmiatra, oftalmologista*).

Como qualquer dicionário eletrônico, o *Dicionário Aberto* (DA)⁵ possui ferramentas de pesquisa como as que acabamos de apresentar, com a particularidade de que é possível descarregar os resultados das pesquisas.

Figura 1 – Adjetivos em **-vel** e de advérbios em **-velmente** descarregados do *Dicionário Aberto*

...		
Agitável	Abominavelmente	
Aglutinável	Admiravelmente	
Agradável	Adoravelmente	
Agradecível	Afavelmente	
Agricultável	Affavelmente	
Ajuntável	Agradavelmente	
Alcançável	Amavelmente	
Alcoolizável	Amigavelmente	
Alheável	Amoravelmente	
Aliável	Aprazivelmente	
Alienável	Civilmente	
Alliável	Comendavelmente	
Alterável	Commendavelmente	
Amável	Compativelmente	
Amigável	Compreensivelmente	
Amisível	Compreensivelmente	
Amoedável	Consideravelmente	
Amoldável	Defensavelmente	
Amolgável	Deleitavelmente	
Amorável	Deploravelmente	
Amortizável	Desagradavelmente	
Amotinável	Desamoravelmente	
Amovível	Desconversavelmente	
Amparável	Desculpavelmente	
...	...	

Fonte: Elaboração própria.

Esta característica faz do DA um recurso importante para a investigação linguística, bem como um excelente auxiliar para a elaboração de gramáticas e de outros dicionários. Utilizamos esta capacidade, por exemplo, num exercício em que perguntamos aos

⁵ O Dicionário Aberto está disponível na rede, para consulta e para extração automática de informação, em <<http://www.dicionario-aberto.net>>, mas também para uso local, de modo aberto e gratuito. O projeto iniciou-se em junho de 2005, com a transcrição da edição de 1913 dos dois volumes do *Novo Dicionário da Língua Portuguesa*, de Cândido de Figueiredo. Sobre o Dicionário Aberto ver Simões e Farinha (2011), Simões, Iriarte e Almeida (2012, 2014).

alunos se todos os adjetivos formados com o sufixo **-vel** podem, em português, formar advérbios em **-mente**. Após a extração de todos os adjetivos acabados em **-vel** e de todos os advérbios acabados em **-velmente** (vd. Figura 1) e após descarregar todos resultados, estes podem ser facilmente alinhados (vd. Figura 2), respondendo à pergunta formulada de se todos os adjetivos em **-vel** podem formar advérbios em **-mente**, e ajudando a verificar hipóteses que expliquem quais podem e quais não.

Figura 2 – Adjetivos em **-vel** e advérbios em **-velmente** do *Dicionário Aberto*, alinhados

...	...	
Agitável	-	
Aglutinável	-	
Agradável	Agradavelmente	
Agradecível	-	
Agricultável	-	
Ajuntável	-	
Alcançável	-	
Alcoolizável	-	
Alheável	-	
Aliável	-	
Alienável	-	
Alliável	-	
Alterável	-	
Amável	Amavelmente	
Amigável	Amigavelmente	
Amissível	-	
Amoedável	-	
Amoldável	-	
Amolgável	-	
Amorável	Amoravelmente	
Amortizável	-	
Amotinável	-	
Amovível	-	
Amparável	-	
...	...	

Fonte: Elaboração própria.

Mas muito mais interessantes são as funcionalidades avançadas de pesquisa, que poderão transformar o DA num sofisticado dicionário onomasiológico e codificador ao mesmo tempo. Utilizando estas funcionalidades o utilizador poderá procurar, por exemplo, unidades lexicais relacionadas (sinónimos, quase-sinónimos, hiperónimos, hipónimos, merónimos, holónimos, co-ocorrentes, etc.) a partir de uma ou de várias palavras. Estas funcionalidades avançadas, disponíveis para utilizadores registados, podem ser, como dissemos, ferramentas muito úteis para linguistas e investigadores em Processamento da Linguagem Natural.

O caso da “pesquisa ontológica” utiliza a experiência na extração de relações léxico-semânticas mediante técnicas de Processamento de Linguagem Natural (PLN) a partir de padrões léxico-semânticos e léxico-sintáticos, ou padrões de conhecimento (*knowledge patterns*) (HEARST, 1992), com aplicações na análise automática de dados textuais em áreas como a medicina, etc. No nosso caso, e graças à regularidade da estrutura das definições lexicográficas, e com o recurso aos dicionários em formato eletrónico (*Machine-Readable Dictionaries*), podem ser estabelecidos um conjunto de regras ou padrões (HEARST, 1992) utilizando as sequências de palavras que se pretendem encontrar nas definições e que indicam a grande probabilidade de a palavra que segue o padrão estar relacionada com o lema da respetiva definição.

Para o reconhecimento e a extração de relações léxico-conceptuais foram usadas estruturas como (SIMÕES; IRIARTE; ALMEIDA, 2012, p. 297, grifo do autor):

- *o mesmo (ou melhor) que ...* ⇒ Sinonímia (SYN);
- *que não é ...* ⇒ Antonímia (ANT);
- *espécie de ...* ⇒ Hiponímia (HIPO);
- *que tem por tipo...* ⇒ Hiperonímia (HIPER);
- *cada uma das partes que formam ...* ⇒ Meronímia (MERO);
- *composto por...* ⇒ Holonímia (HOLO).

Também se usaram relações calculadas (por exemplo, utilizando a transitividade da relação de hiperonímia), inferindo novos relacionamentos a partir das relações iniciais:

– A propriedade de simetria entre sinónimos: se **a** é sinónimo de **b**, então **b** também é sinónimo de **a** (a SYN b \Rightarrow b SYN a). Esta relação é bastante produtiva, permitindo recuperar sinónimos que apenas foram recolhidos pelo lexicógrafo numa das entradas das palavras envolvidas.

– A relação de co-hiponímia pode ser calculada a partir das relações de hiponímia: se duas palavras **a** e **c** são hipónimos da mesma palavra **b**, então **a** e **c** são co-hipónimos (a HIPO b \wedge c HIPO b \Rightarrow a COHIPO c).

– A transitividade das relações hierárquicas, como é o caso da hiperonímia/hiponímia, permite que se possa procurar termos genéricos com base num termo mais específico ou termos específicos usando termos mais genéricos (a HIPO b \wedge b HIPO c \Rightarrow a HIPO c). Assim, podemos encontrar entradas referentes a “animais” a partir da pesquisa de termos mais específicos como “mamífero” ou “peixe” (hipónimos de “animal”).

Evidentemente, o reconhecimento e a extração automáticos de relações léxico-conceptuais não é fácil. Somos conscientes de que algumas regras podem ser problemáticas (por exemplo, no caso dos sinónimos e quase-sinónimos). Em todo o caso, parece-nos preferível apresentar um conjunto de possíveis falsos resultados do que garantirmos a correção e diminuir drasticamente o número de relações da ontologia resultante.

Na Figura 3 temos uma amostra de pesquisa reversa normal, em que todos os resultados contêm o termo introduzido na janela de pesquisa (“**mamífero**”).

Figura 3 – Pesquisa reversa a partir do termo “mamífero”

Dicionário Aberto
de a a zythógala (128 521 entradas)

mamífero

Prefixo Infixo Sufixo **Pesquisa Reversa** Pesquisa Ontológica

Procurar

Descarregar
TEI TXT

aguti¹ – *m.*; Mamífero americano, da ordem dos roedores.
almiqui¹ – *m.*; Mamífero roedor da ilha de Cuba.
anomalu¹ – *m.*; Mamífero roedor, de cauda anômala, encontrado em Fernando-Pé. (Do gr. *anoma...*
antamba¹ – *f.*; Espécie de leopardo, referida por Bluteau. Mamífero de Madagáscar.
aperca¹ – *f.* Bras; Pequeno mamífero roedor.
ariranha¹ – *f.* Bras; Mamífero, semelhante à lontra.
arminho¹ – *m.*; Mamífero, da fam. dos mustelídeos, cuja pele macia é, no inverno, alvíssim...
arvícola¹ – *m.*; Habitante do campo. Lavrador. *Zool.*; Mamífero roedor, da fam. dos murídeo...
baleia¹ – *f.*; Corpolento mamífero, da ordem dos cetáceos. Constelação austral. (Lat. *bae...*
cabra¹ – *f.*; Animal mamífero, da ordem dos ruminantes, fêmea do bode. Guindaste. Espécie ...
caititu¹ – *m.* Bras; Mamífero paquíderme, indígena da América. Engenho para fazer farinha, ...
callocéfalo¹ – *adj.*; Que tem cabeça formosa. *M. Zool.*; Mamífero, da ordem das focas.
calocéfalo¹ – *adj.*; Que tem cabeça formosa. *M. Zool.*; Mamífero, da ordem das focas.
canguru¹ – *m.* *Zool.*; O maior mamífero da ordem dos marsupiais. Designação genérica dos mar...
capivara¹ – *f.* Bras; Mamífero roedor. (T. tupi)
carneiro-almiscarado¹ – *m.*; Mamífero fóssil, (*ocibes moschatus*, Blainv.), procedente das regiões árti...
castor¹ – *m.*; Animal mamífero roedor. Pêlo deste animal. Estrela dupla da constelação dos ...
caxinguelé¹ – *m.* Bras; Mamífero roedor. (Talvez t. afr.)
cazu¹ – *m.*; Árvore africana, de frutos comestíveis. Mamífero da Guiné.

Fonte: Dicionário Aberto.

Ao contrário do que acontece com a pesquisa reversa normal, graças à extração automática de relações lexicais, na pesquisa ontológica (Figura 4), alguns dos resultados são entradas que não contêm nenhum dos termos introduzidos pelo utilizador na janela de pesquisa, neste caso, “**mamífero**”:

Figura 4 – Pesquisa ontológica a partir do termo “mamífero”



The screenshot shows the 'Dicionário Aberto' interface. At the top, there is a search bar containing the word 'mamífero'. Below the search bar, there are radio buttons for different search types: 'Prefixo', 'Infixo', 'Sufixo', 'Pesquisa Reversa', and 'Pesquisa Ontológica' (which is selected). To the right of the search bar is a 'Procurar' button. Below the search bar, there is a list of search results for 'mamífero'. The results include terms like 'marsupiais', 'perissodactylos', 'felídeos', 'otária', 'lontra', 'primates', 'golfinho', 'macaco', 'capiguará', 'zibellina', 'armadillo', 'gutuoso', 'nictalo', 'gutuoso', 'focáceos', 'leporídeos', 'marsupiais', 'oryxter', and 'zolanço', each followed by a brief definition. On the right side of the results, there is a 'Descarregar' button with 'TXT' and 'PDF' options.

Fonte: Dicionário Aberto.

Conclusões

Estamos convictos de que o DA é uma excelente ferramenta que pode ser usada como um dicionário tradicional, mas também como um recurso para tarefas de processamento de linguagem natural, como um auxílio para a elaboração de gramáticas e de outros dicionários bem como uma ferramenta para ajudar na verificação de hipóteses colocadas na investigação linguística.

O facto de ser um recurso aberto poderá também significar que as suas funcionalidades poderão vir a aumentar em quantidade e em qualidade.

Por outro lado, e embora a evolução do conteúdo linguístico do dicionário seja lenta, a implementação de algoritmos e a execução de experiências sobre o DA têm sido bastante proveitosa, demonstrando que é possível criar funcionalidades úteis a partir de dicionários convencionais.

REFERÊNCIAS

AULETE, F. J. C. **Dicionário Contemporâneo da Língua Portuguesa Caldas Aulete**. 5. ed. Rio de Janeiro: Delta, 1987.

AULETE DIGITAL. 2008. Disponível em: <<http://www.aulete.com.br/>>. Acesso em: 25 ago. 2014.

BERGENHOLTZ, H.; TARP, S. (Ed.). **Manual of Specialised Lexicography: The Preparation of Specialised Dictionaries**. Amsterdam; Philadelphia: John Benjamins Publishing Company, 1995.

CALDERÓN CAMPOS, M. **Sobre la elaboración de diccionarios monolingües de producción: Las definiciones, los ejemplos y las colocaciones léxicas**. Granada: Ed. de la Universidad de Granada, 1994.

CASTELEIRO, J. M. (Coord.). **Dicionário da Língua Portuguesa Contemporânea da Academia das Ciências de Lisboa**. Lisboa: Academia das Ciências de Lisboa: Editorial Verbo, 2001.

COWIE, A. P. On Specifying Grammar: On Specifying Grammatical Form and Function. In: HARTMANN, R. R. K. (Ed.). **Lexicography: Principles and Practice**. London: Academic Press, 1983. p. 99-107.

FERREIRA, A. B. de H. **Novo Aurélio Século XXI: O Dicionário da Língua Portuguesa**. Rio de Janeiro: Nova Fronteira, 1999.

FIGUEIREDO, C. de. **Dicionário da Língua Portuguesa**. 16. ed. Lisboa: Livraria Bertrand, 1982.

HEARST, M. Automatic acquisition of hyponyms from large text corpora. **Proceedings of the Fourteenth International Conference on Computational Linguistics**, Nantes, France, v.2, p.539-545, 1992.

HOUAISS, A. (Coord.). **Dicionário Houaiss da Língua Portuguesa**. Rio de Janeiro: Objectiva, 2001.

LAPA, M. R. **Estilística da Língua Portuguesa**. Coimbra: Coimbra Editora, 1984.

SAUSSURE, F. **Curso de linguística geral**. Lisboa: Publicações Dom Quixote, 1999.

SIMÕES, A.; FARINHA, R. Dicionário Aberto: Um novo recurso para PLN. **Vice-versa**, Vigo, Espanha, v. 16, p.159-171, 2011.

SIMÕES, A.; IRIARTE, Á.; ALMEIDA, J. J. Dicionário-Aberto: Construção semiautomática de uma funcionalidade codificadora. In: CONGRES INTERNATIONAL DE LINGUISTIQUE ET DE PHILOLOGIE ROMANES, 27., Nancy. **Actes...** Nancy : ATILF, 2014. Section 16: Projets en cours; ressources et outils nouveaux. Disponible en: <<http://www.atilf.fr/cilpr2013/actes.php>>. Acesso em: 25 ago. 2014.

SIMÕES, A.; IRIARTE, Á.; ALMEIDA, J. J. Dicionário aberto: a source of resources for the portuguese language processing. In: CASELI, H. et al. (Ed.). **Computational Processing of the Portuguese Language, Lecture Notes for Artificial Intelligence**. Berlim: Springer, 2012. p.121-127.

VILELA, M. **Dicionário do Português Básico**. Porto: Edições Asa, 1991.