

Estudo Semântico e Aplicação Computacional de Adjetivos do Português do Brasil

¹Cláudia Dias de Barros

¹Programa de Pós-Graduação em Linguística – Universidade Federal de São Carlos
(UFSCar)

Caixa Postal 676 – 13.565-905 – São Carlos – SP – Brasil

claudias84@gmail.com

Abstract. *This paper describes the semantic study and a computational resource of Brazilian Portuguese adjectives. We describe some features of adjectives, such as their subclasses: descriptive, relational and determinative. We present the semantic relation called antonymy and its two types: direct and indirect. The research data were inserted into TeP, providing the representation of antonymy improvement in this computational resource.*

Resumo. *Este artigo descreve o estudo semântico e a aplicação computacional de adjetivos do Português do Brasil. São apresentadas algumas características dos adjetivos, como sua divisão nas subclasses dos descritivos, relacionais e determinativos. Também é descrita a relação semântica chamada antonímia, com seus dois tipos: direta e indireta. Os dados da pesquisa foram inseridos no TeP, proporcionando, assim, um aprimoramento da representação da antonímia nesse recurso computacional.*

1. Adjetivos

Neste artigo é descrita a pesquisa realizada sobre adjetivos do Português do Brasil (PB), na qual é abordado principalmente o aspecto semântico relacionado à relação mais importante para os adjetivos: a antonímia (oposição de sentidos).

Os adjetivos podem ser definidos como itens lexicais que acompanham um substantivo, indicando-lhe uma qualidade (“Ela é uma moça **gentil**”), o modo de ser (“Ele é uma pessoa **hábil**”), o aspecto ou aparência (“Este jardim está **florido**”), o estado (“A criança está **enferma**”) [Cunha 1981]. Eles apresentam duas funções sintáticas principais: a função atributiva, em que o adjetivo se relaciona diretamente ao substantivo (“Que muro **alto!**”) e a função predicativa, em que se relaciona com o substantivo por meio de um verbo copulativo (“O muro **é alto**”).

A função semântica básica dos adjetivos é atribuir uma propriedade a uma entidade, que pode ser qualificadora ou subcategorizadora dos substantivos [Borba 1996].

Os adjetivos na função qualificadora (qualificativos ou descritivos) são os que atribuem ao substantivo que acompanham uma característica que não é necessariamente inerente a este. Eles também apresentam antônimos (Rapaz **bonito**/Rapaz **feio**) e sinônimos (Rapaz **bonito**/Rapaz **belo**) e podem ocorrer tanto em função atributiva

(“Aqui há paisagens **calmas**”), quanto predicativa (“Esta paisagem é **calma**”). Esse tipo de adjetivos foi o alvo da pesquisa, por apresentarem antônimos.

A atribuição subcategorizadora dos substantivos é realizada pelos adjetivos classificadores ou relacionais, aqueles que colocam os substantivos em uma subclasse (“Há indústrias **alimentícias** nesta região”) [Neves 2000]. Eles são relacionados semântica e morfológicamente aos substantivos (Instrumento para fazer música→Instrumento musical) [Fellbaum 1998]. Esse tipo de adjetivo ocorre principalmente em função atributiva, posposto ao substantivo (“Tenho um problema **dental**”). Eles não possuem antônimos (Cidadão **brasileiro**/Cidadão ?) e podem ser parafraseados em ‘preposição+nome’ (“Este é o sistema **digestivo**”=“Este é o sistema **de digestão**”).

Há uma terceira subclasse de adjetivos, chamada de determinativos, que ocorrem apenas em função atributiva, basicamente em posição pré-nominal e não possuem antônimos (“Disseram **várias** palavras”).

Existem outros trabalhos que abordaram o estudo sintático-semântico da classe dos adjetivos, como Dalla Pria (2005), que realizou um estudo comparativo da distribuição dos adjetivos no sintagma nominal do inglês e do português, visando aplicações para o Processamento de Língua Natural (PLN) e a pesquisa de Di Felippo e Dias-da-Silva (2005) sobre a estrutura argumental dos adjetivos valenciais do PB e sua representação linguístico-computacional.

Ambos os trabalhos foram muito importantes no desenvolvimento da pesquisa abordada neste artigo.

2. Antonímia

A antonímia pode ser definida como a oposição entre os sentidos e é considerada a relação semântica básica entre os adjetivos descritivos [Fellbaum 1998]. Isso pode ser comprovado por meio dos testes associativos (*word association tests*), que são muito utilizados pelos psicolinguistas, para observar como as palavras estão organizadas na mente de um falante nativo. Nesses testes, quando é pedido um adjetivo familiar a outro, a resposta mais comum é seu antônimo, por exemplo, ‘**bom/mau**’.

A antonímia pode ser formada por pares não relacionados morfológicamente, como ‘Prédio **alto**/Prédio **baixo**’, ‘Menina **alegre**/Menina **triste**’ que, segundo Lyons (1977), é o caso mais comum em muitas línguas. Outra forma de apresentação da antonímia é por pares relacionados morfológicamente, em que um dos membros do par deriva do outro, por exemplo, pela adição de um prefixo de negação, como ‘in-’ (‘Filho **legítimo**/Filho **ilegítimo**’).

Segundo Murphy (2003), a antonímia possui algumas propriedades como simetria (se A é antônimo de B, então B é antônimo de A), marca (um dos termos é marcado) e binaridade (ocorre preferencialmente entre pares de palavras).

Para Miller *et al.* (1993), os adjetivos antônimos expressam valores opostos (polos) de um atributo. Por exemplo, para o atributo TAMANHO, os antônimos são ‘**grande**’ e ‘**pequeno**’.

A antonímia pode ser classificada em dois tipos: direta ou lexical – oposição que ocorre entre unidades lexicais (relação léxico-semântica), como ‘Carro **grande**/Carro

pequeno’; indireta ou conceitual – ocorre entre conceitos opostos, mesmo que as unidades lexicais não sejam um par de antônimos. Esse tipo de antonímia se dá por meio da sinonímia, uma vez que o adjetivo que não apresenta um antônimo direto herda o antônimo de um adjetivo sinônimo, como no caso de ‘Carro **enorme**’, que é similar a ‘Carro **grande**’ e tem como antônimo indireto ‘Carro **pequeno**’.

A similaridade entre os adjetivos indica um tipo de especialização, ou seja, os substantivos modificados por ‘**enorme**’ estão incluídos naqueles que podem ser modificados por ‘**grande**’.

A figura 1 ilustra a representação da antonímia direta entre o par ‘**grande/pequeno**’ e os adjetivos que formam com eles, através da sinonímia, a antonímia indireta (‘**enorme**=**grande/pequeno**’, **infinitésimo**=**pequeno/grande**), por exemplo.

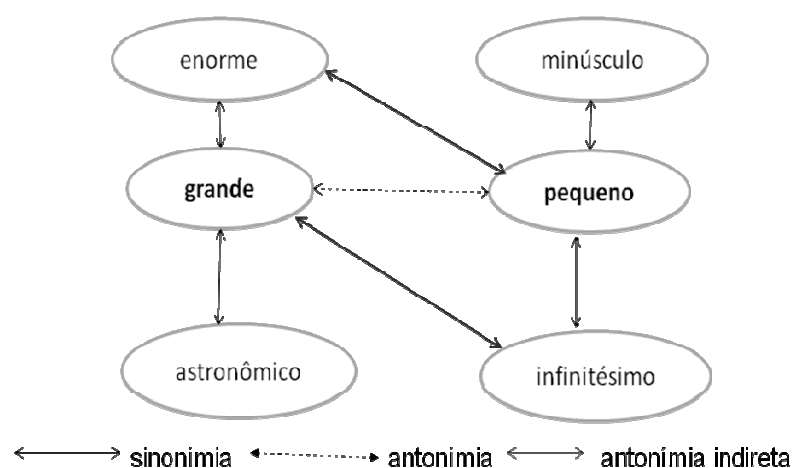


Figura 1. Representação de antônimos diretos e indiretos

Mediante o conceito de antonímia indireta, pode-se afirmar que é possível atribuir um antônimo a todos os adjetivos que não têm antônimos diretos, como ‘obeso’, ‘enorme’, entre outros.

3. Aplicação Computacional: TeP

Após a exposição das principais características pesquisadas dos adjetivos e da antonímia, passa-se a relatar a aplicação computacional feita com os dados da pesquisa.

O trabalho teve como ponto de partida os 100 adjetivos mais frequentes extraídos do corpus Mac-Morpho, do Projeto LacioWeb [Aluísio *et al.* 2003], que contém 1.167.183 ocorrências e é composto de artigos da Folha de São Paulo, do ano de 1994. Foi utilizada a ferramenta Unitex [Paumier 2002] para a extração dos adjetivos. Dentre os 100 mais frequentes havia 76 descritivos, 23 relacionais e 1 determinativo.

O objetivo da pesquisa foi tornar a representação da antonímia em alguns adjetivos presentes no Thesaurus Eletrônico do Português do Brasil (TeP) [Dias-da-Silva *et al.* 2000; Dias-da-Silva e Moraes 2003; Maziero *et al.* 2008] (um dicionário eletrônico que apresenta sinônimos e antônimos de substantivos, adjetivos, verbos e

advérbios), parecida com a representação utilizada pela WordNet de Princeton [Miller *et al.* 1993] (uma base de dados para o inglês norte-americano que apresenta as mesmas classes de palavras do TeP agrupadas em synsets (conjuntos de sinônimos que definem um conceito), relacionados por relações semânticas (p.ex. antonímia, hiponímia e meronímia), definidos por glosas (definições informais do conceito) e exemplificados por frases-exemplo (contextos mínimos de uso)), por meio da inserção da antonímia indireta.

A partir dos 76 adjetivos descritivos extraídos do corpus, foram formados os pares de antônimos, totalizando 108 adjetivos. Alguns pares de antônimos já estavam presentes no corpus (17 deles) e o restante foi formado com o auxílio de dicionários de sinônimos e antônimos [Barbosa 1999]. Esse dicionário apresenta os sinônimos de palavras, acompanhados de frases exemplificativas, seguidos dos respectivos antônimos, como: **JEITOSO** – 1. Destro, hábil, apto: *Operário jeitoso. Ant.* Desajeitado.

Por meio da análise da antonímia nos adjetivos, pôde-se perceber que os pares de antônimos mais frequentes foram os pares lexicalizados (47 deles - ‘verdadeiro/falso’) e não os pares com prefixos de negação (13 pares - ‘disponível/indisponível’). O tipo de antonímia mais frequente encontrada no corpus foi a direta (61 adjetivos - ‘alto/baixo’). Percebeu-se, também, que dos adjetivos trabalhados, apenas 16 não apresentavam antônimos no TeP.

4. Contribuições e Conclusão

Uma das contribuições do trabalho foi a inserção de 322 frases-exemplo no TeP, sendo que 226 foram extraídas do corpus e 96 de dicionários de sinônimos e antônimos.

No TeP também foram relacionados 47 pares de antônimos diretos (direto/indireto); 172 synsets foram relacionados por antonímia indireta ({grande, grandioso, magnífico}/{inferior, insignificante, secundário}) e foi criado o rótulo ‘Antônimo indireto’.

Conclui-se, assim, que esta pesquisa foi importante devido ao estudo aprofundado dos adjetivos, observando, principalmente, a divisão em três subclasses e ao estudo da antonímia direta e indireta. Houve também o aprimoramento da base de dados do TeP, por meio da inserção dos resultados (antônimos de adjetivos e frases-exemplo).

Referências

- Aluísio, S. M. et al. (2003) “An account of the challenge of tagging a reference corpus of Brazilian Portuguese”. São Carlos: NILC, 20 p. Relatório técnico NILC-TR-03-04.<<http://www.nilc.icmc.usp.br/lacioweb/downloads/NILC-TR-03-04.zip>>.
- Barbosa, O. (1999) “Grande dicionário de sinônimos e antônimos”. Rio de Janeiro: Ediouro, 568 p.
- Borba, F. S. (1996) “Uma gramática de valências para o português”. São Paulo: Ática. 199p.
- Cunha, C. (1981) “Gramática de base”. 2. ed. Rio de Janeiro: FENAME – Fundação Nacional de Material Escolar. 371 p.

- Dalla Pria, A. (2005) “Estudo da distribuição dos adjetivos no sintagma nominal do inglês e do português com vistas ao processamento automático de línguas naturais”. 2005. 110 p. Dissertação (Mestrado em Letras) - Faculdade de Ciências e Letras, Universidade Estadual Paulista, Araraquara.
- Dias-da-Silva, B.C. et al (2000) “Construção de um thesaurus eletrônico para o português do Brasil”. In: Processamento Computacional do Português Escrito e Falado (PROPOR), 4., 2000, Atibaia. Anais..., São Carlos: USP. p. 1-10.
- Dias-da-Silva, B.C.; Moraes, H.R. (2003) “A construção de um thesaurus eletrônico para o português do Brasil”. ALFA, v. 47, n. 2, p. 101-115.
- Di Felippo, A.; Dias-da-Silva, B. C. (2005) “Modelo linguístico-computacional da estrutura argumental de adjetivos valenciais do português do Brasil”. Estudos Linguísticos, v. 34. São Paulo, São Paulo, Brasil. 1 CD ROM.
- Fellbaum, C. (Ed.) (1998). “WordNet: an electronic lexical database”. Cambridge, MA: MIT Press, 423 p.
- Lyons, J. (1977) “Semantics”. Cambridge: Cambridge University Press. 2 v.
- Maziero, E. G. et al. (2008) “A base de dados lexical e a interface web do TeP 2.0 – Thesaurus eletrônico para o português do Brasil”. In: Workshop em Tecnologia da Informação e da Linguagem Humana (TIL), 6., 2008, Vila Velha: Universidade Federal do Espírito Santo, p. 390-392.
- Miller, G. A. et al. (1993) “Five papers on WordNet”. <<http://www.cogsci.princeton.edu/~wn>>.
- Murphy, M. L. (2003) “Semantic relations and the lexicon: antonymy, synonymy, and other paradigms”. Cambridge: Cambridge University Press. 292 p.
- Neves, M. H.M. (2000) “Gramática de usos do português”. São Paulo: Ed. UNESP, 1037p.
- Paumier, S. (2002) “Unitex: manuel d'utilisation, research report”. França: University of Marne-la-Vallée, 200 p.