



HAL
open science

Tipologia das expressões adjetivais do português do Brasil

Amanda dos Santos Carneiro, Oto Araújo Vale

► **To cite this version:**

Amanda dos Santos Carneiro, Oto Araújo Vale. Tipologia das expressões adjetivais do português do Brasil. Workshop de Iniciação Científica em Tecnologia da Informação e da Linguagem Humana, Oct 2013, Fortaleza, Brazil. pp.28 - 30. hal-01410226

HAL Id: hal-01410226

<https://hal.science/hal-01410226>

Submitted on 6 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Tipologia das Expressões Adjetivais do Português do Brasil

Amanda dos Santos Carneiro¹, Oto Araújo Vale²

¹Universidade Federal de São Carlos (UFSCar)

²UFSCar/ NILC

amanda.scarneiro@gmail.com
otovale@ufscar.br

1. Introdução

Dentro dos estudos do léxico, podemos destacar a importância das expressões multipalavras, que englobam tanto as expressões idiomáticas quanto as chamadas colocações. Isso porque essas são amplamente utilizadas na língua cotidiana e em diferentes gêneros do discurso e devem ser tomadas como uma unidade lexical fechada, uma vez que as palavras que as compõem não podem ter seu significado interpretado separadamente.

Assim, esse trabalho tem como objetivo estudar as regularidades e propriedades das expressões cristalizadas (doravante EC) do tipo adjetivais, adotando o conceito de EC como sendo “uma expressão formada por mais de um segmento cujo significado total não pode ser deduzido pelo significado das partes que a compõem.” (VALE, 2001). Mais especificamente as que exercem a função de adjetivo na oração, por isso, adjetivais. Por exemplo, na ocorrência (extraída do corpus):

A verdade é que o Novo “Carrie” é um **abacaxi dos diabos**

o significado total da expressão **abacaxi dos diabos** nada tem a ver com a soma dos significados das unidades lexicais **abacaxi** e **diabo**. Por isso, tais construções precisam ser entendidas como uma unidade lexical composta que possui significado próprio.

Nessa perspectiva, podem ser acrescentadas em léxicos computacionais, evitando que sejam interpretadas como fragmentos sintáticos. Um exemplo dessa aplicação é o caso dos tradutores automáticos de texto que, ao traduzirem palavra por palavra separadamente, descaracterizam a expressão, gerando orações com significado diferente da original. Além disso, o não reconhecimento da unidade lexical faz com que analisadores sintáticos façam uma análise equivocada, uma vez que uma EC pode conter palavras de diferentes classes gramaticais, mas assumir outra função na oração.

Além da aplicação computacional, a problemática também pode ser aplicada ao âmbito do estudo de linguagens de especialidade e o ensino de línguas estrangeiras. Entretanto, apesar dessa importância, essas expressões, sabidamente numerosas no português do Brasil, não tiveram até o momento uma descrição específica e exaustiva, sendo essa a razão da escolha do tema.

Dentre os trabalhos que abordaram as expressões adjetivais, poucos são aqueles que têm buscado fazer uma tipologia sistemática no português do Brasil. Existem trabalhos de cunho lexicográfico, como Riva (2009), que fez um dicionário onomasiológico de expressões idiomáticas usuais no português do Brasil, em que se arrolam diversas expressões adjetivais. Ainda assim de grande contribuição, estas estão catalogadas juntamente com outros tipos de expressões cristalizadas como as expressões

verbais, nominais e adverbiais, não havendo um foco específico em um tipo de expressão.

Propomos, então, uma tipologia para esse tipo de EC, adotando a perspectiva teórico-metodológica do Léxico-Gramática (GROSS, 1975, CASTELEIRO, 1981) associada à linguística de corpus. O léxico-gramática inspira-se nas ciências experimentais e dá enfoque à coleta de fatos e, conseqüentemente, à conferência com a realidade dos usos da língua, do ponto de vista quantitativo e qualitativo. Também prevê uma formalização específica em que os resultados devem ser formais o suficiente para que possam ser utilizados no processamento automático das línguas. Adota-se assim, o modelo de tabelas de dupla entrada, chamadas de tabelas ou matrizes, que cruzam os itens lexicais nas linhas com as propriedades sintático-semânticas a serem verificadas nas colunas.

Quanto à questão dos adjetivos, é importante melhor delimitá-los, já que esses possuem um comportamento sintático-semântico muito peculiar. Adotaremos aqui a perspectiva de Borba (1996) na qual a conceituação do adjetivo pressupõe-se que: (i) a adjetivação implica um processo mental de diferenciação, discriminação e seleção, (ii) todo adjetivo é palavra de natureza abstrata e (iii) é próprio do adjetivo não incidir sobre si mesmo.

Mais especificamente, Casteleiro (1981) os classifica como predicativos ou não predicativos. Os adjetivos predicativos são aqueles que “são introduzidos na seqüência de base em posição pós-cópula, por meio da frase relativa”, como em *A menina é bela*, enquanto que os não predicativos “são introduzidos nas estruturas sintáticas de base em posição pós-nominal”, como em *A menina bela*. No presente projeto, buscou-se expressões adjetivas predicativas, ou seja, aquelas introduzidas na seqüência do verbo *ser*.

2. Metodologia

Para este trabalho, partimos do uso dessas expressões em corpus, pois, somente a procura em dicionários não seria suficiente, uma vez que não há acesso ao uso real. O corpus utilizado foi o do projeto PLN-BR, constituído por textos selecionados em dez anos do jornal Folha de S. Paulo de diferentes seções. Já para a busca, foi utilizado o software Unitex, que permite a procura por categorias gramaticais e a indexação por frequência das seqüências encontradas. Portanto, buscamos primeiramente as expressões que entrariam nas seguintes estruturas:

<ser> (um+uma+E) <N> <prep> <N>

<ser> (um+uma+E) <N> (do+da) <N>

<ser> (um+uma+E) <N> <PREP> <N> <N>

<ser> (um+uma+E) <N> <PREP> <N> <A>

em que <ser> é qualquer forma flexionada do verbo *ser*, seguido por artigo definido *um*, *uma* ou sem artigo (simbolizado por E), em que o sinal + indica *ou*; <N> sendo qualquer substantivo flexionado, <PREP> para qualquer preposição e <A> para as flexões de adjetivo.

Estas seqüências foram escolhidas inicialmente para delimitar o escopo da pesquisa e também por se revelarem num primeiro momento da observação como seqüências produtivas para os adjetivos compostos.

A partir dessas buscas foram construídas listas das ocorrências candidatas a EC em planilhas de Excel, ordenadas em ordem alfabética. Dessas listas foram selecionadas as ocorrências das construções adjetivais que serão devidamente separadas em classes

de acordo com sua constituição sintática interna e outras propriedades. Ao todo, *a priori*, foram selecionadas 260 ECs.

3. Resultados

Após a busca através do Unitex, todas as ECs passaram a compor uma única planilha que foi ordenada em ordem alfabética. Entretanto, a cada uma delas foi associada uma propriedade nomeada *Estr. Interna*, na qual era informada a sequência que foi usada para busca, a saber: NprepN, NprepNN e NprepNA.

Em seguida, já na construção da matriz, foi feita a separação das unidades que compunham as expressões e a inserção de uma frase exemplo retirada do corpus ou da internet. Com esses procedimentos, pudemos aprimorar a classificação “Estr. Interna”; verificar que algumas na verdade englobam mais unidades do que as selecionadas e eliminar ocorrências que, ao serem verificadas no contexto e buscadas em grande escala não poderiam ser classificadas como expressões cristalizadas adjetivais, ou por não terem recorrência ou por se inserirem em uma estrutura maior que não se encaixa na definição de EC.

Após essas constatações, permanecendo ao final 228 ECs adjetivais, partiu-se para a determinação das propriedades a serem verificadas, pois, para estabelecer uma tipologia, é necessário verificar características em comuns a tais estruturas.

Nesse momento, o trabalho encontra-se na parte de construção da matriz binária, composta pelas expressões e pelas propriedades pertinentes para o estabelecimento das classes a que o trabalho se propõe, sendo que as propriedades são marcadas com um sinal positivo (“+”) quando se verificarem e negativo (“-“) quando não se verificam.

4. Conclusões

A partir dessa caracterização já será possível uma melhor visão dessas expressões. Sabemos, entretanto, que o número de unidades levantadas é insuficiente para uma cobertura abrangente do fenômeno, mas esperamos que a partir desse trabalho seja possível reconhecer novas expressões adjetivais em grande corpora. Portanto, consideramos que esse já é um primeiro passo para uma classificação mais geral dessas expressões.

Referências Bibliográficas

Borba, F.S. Uma gramática de valências para o português. Editora ática, 1996.

Casteleiro, J. M. Sintaxe transformacional do adjectivo: regência das construções completivas. Lisboa: Instituto Nacional de Investigação Científica, 1981

Gross, M. Méthodes empiriques em syntaxe. Paris : Hermann, 1975.

Gross, G. Les Expressions figées en français. Noms composés et autres locutions, Gap: Ophrys, 1996.

Riva, H. C. Dicionário onomasiológico de expressões idiomáticas da língua portuguesa usuais na língua portuguesa do Brasil. (Tese de Doutorado) São José do Rio Preto: UNESP, 2009.

Vale, O.A. Expressões cristalizadas do português do Brasil: uma proposta de tipologia. (Tese de Doutorado) Araraquara: UNESP, 2001.