

DESCRIÇÃO DOS PREDICADOS NOMINAIS RELACIONADOS AO VERBO “DAR” NO CORPUS PLN.BR FULL

Doutoranda: Amanda Pontes Rassi

Orientador: Prof. Dr. Oto Araujo Vale (UFSCar)

LP: Linguagem Humana e Tecnologia

Projeto de Doutorado mar/2012 – mar/2016

Este trabalho investiga as ocorrências do verbo “dar” como verbo-suporte, bem como os nomes predicativos a ele associados. Esse verbo, assim como “ter”, “fazer”, “levar”, “pôr” e outros, possui uma característica que o distingue sintático-semânticamente de todos os outros verbos plenos: é o fato de funcionar prototipicamente como verbo-suporte. Os verbos-suporte ou leves são “verbos semanticamente vazios que permitem construir um *SN* com *V-n*” (GROSS, 1981), servindo apenas para suportar as categorias de tempo, modo, número e pessoa, as quais não podem ser marcadas em nomes. Essa relação pressupõe um *V-n*, que é sempre um nome predicativo que seleciona os argumentos exigidos pela nominalização. Assim, a pesquisa parte das seguintes questões: como funcionam os verbos-suporte e como eles se comportam com cada substantivo predicativo?, quais deles possuem equivalência com um verbo pleno?, quais nominalizações admitem construção conversa com “dar/levar” ou “dar/receber”?, dentre outras que possam garantir a fidelidade à descrição. Os fundamentos teórico-metodológicos se pautam na noção de predicados e argumentos (HARRIS, 1965) e na Teoria do Léxico-Gramática (GROSS, 1975, 1981). As ocorrências dos nomes predicativos foram retiradas do corpus PLN.BR FULL (MUNIZ et al., 2007), com o auxílio do software Unitex 3.0 (PAUMIER, 2002), e representadas em tábuas do Léxico-Gramática (GROSS, 1975), ou seja, tabelas binárias cujas linhas correspondem às entradas lexicais dos nomes predicativos, e em cujas colunas são colocadas as propriedades sintático-semânticas das nominalizações. Dessa forma, pode-se representar os dados formalmente, marcando com sinais de “+” e “-” a aceitabilidade ou inaceitabilidade de determinadas propriedades sintáticas e semânticas previamente identificadas nas Tábuas. Esse tipo de formalização permite uma posterior implementação dos resultados em sistemas, programas, ferramentas e recursos computacionais. Trabalhos de descrição linguística de nomes predicativos e de verbos-suporte já foram realizados para o francês, o italiano, o português de Portugal e várias outras línguas. Em português do Brasil, algumas descrições também vêm sendo feitas desde a década de 90, porém o objetivo desta pesquisa não é apenas descrever os nomes predicativos, mas também representá-los formalmente para poder subsidiar o treinamento de sistemas de PLN que utilizem aprendizado de máquina para identificação dos verbos-suporte. Espera-se, com isso, subsidiar a criação do recurso léxico NomBank.Br, uma base de dados de nomes predicativos, já que são eles que selecionam os argumentos da sentença, e não o verbo.

GROSS, Maurice. *Méthodes en Syntaxe*. Paris: Hermann, 1975.

_____. *Les bases empiriques de la notion de prédicat sémantique*. *Langages*, n.63, p.7-52, 1981.

HARRIS, Zellig Sabbetai. Transformational Theory. *Language* 41, p. 363-401, 1965.

MUNIZ, Marcelo; et al. Taming the tiger topic: an XCES compliant corpus Portal to generate subcorpus based on automatic text topic identification. In: *CORPUS LINGUISTICS CONFERENCE*, 2007, Birmingham. Proceedings of the Corpus Linguistics 2007 Conference. Birmingham: University of Birmingham, 2007. Disponível em: <<http://ucrel.lancs.ac.uk/publications/CL2007/>>. Acesso em: fev. 2012

PAUMIER, S. *Unitex: manuel d'utilisation, research report*. França: University of Marne-la-Vallée, 2002. 200 p.