



Claudia Maria Garcia Medeiros de Oliveira

O Substantivo-suporte

Critérios Operacionais de Caracterização

Tese de Doutorado

Tese apresentada ao Programa de Pós-graduação em Letras do Departamento de Letras da PUC-Rio como requisito parcial para obtenção Do título de Doutor em Letras

Orientador: Prof. Margarida Basilio

Rio de Janeiro
Setembro de 2006



Claudia Maria Garcia Medeiros de Oliveira

O Substantivo-suporte

Critérios Operacionais de Caracterização

Tese apresentada ao Programa de Pós-graduação em Letras do Departamento de Letras do Centro de Teologia e Ciências Humanas da PUC-Rio como requisito parcial para obtenção Do título de Doutor em Letras. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Margarida Basilio

Orientador
Departamento de Letras - PUC-Rio

Prof. Violeta de San Tiago Dantas Barbosa Quental

Departamento de Letras - PUC-Rio

Prof. Helena Franco Martins

Departamento de Letras - PUC-Rio

Prof. Heronides Maurílio de Melo Moura

Departamento de Língua e Literatura Vernáculas - UFSC

Prof. Maria Carlota Amaral Paixão Rosa

Departamento de Letras Vernáculas - UFRJ

Prof. Paulo Fernando Carneiro de Andrade

Coordenador Setorial do Centro de Teologia e Ciências
Humanas - PUC-Rio

Rio de Janeiro, 1 de Setembro de 2006

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Claudia Maria Garcia Medeiros de Oliveira

Graduou-se em Computação no ITA (São José dos Campos, SP) em 1986. Obteve o título de Mestre em Sistemas e Computação em 1990, pelo Instituto Militar de Engenharia (Rio de Janeiro); e de PhD em Computação em 1995, pelo Imperial College (Londres, Reino Unido). Áreas de interesse: Estudos do Léxico, Lingüística Computacional e Inteligência Artificial.

Ficha Catalográfica

Oliveira, Claudia

O Substantivo-suporte: Critérios Operacionais de Caracterização / Claudia Maria Garcia Medeiros de Oliveira; orientador: Margarida Basilio. — 2006.

118 f: il. ; 30 cm

1. Tese (Doutorado em Letras) - Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2006.

Inclui bibliografia

1. Letras – Teses. 2. Substantivo-suporte. 3. Substantivo. 4. Adjetivo denominal. 5. Classe de palavras. 6. Lexicografia de corpus. 7. Lingüística. I. Basilio, Margarida. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Letras. III. Título.

CDD: 400

Agradecimentos

À professora Margarida Basilio, pela confiança e pelo respeito que sempre teve por minhas idéias e minha pesquisa. Considero um privilégio ter trabalhado com essa lingüista influente, extremamente produtiva e inovadora, que fez com que eu me apaixonasse pelos estudos do léxico.

À PUC-Rio pela bolsa de estudos concedida e pela infra-estrutura do campus que me ofereceu.

Ao Departamento de Letras, onde descobri a Lingüística como disciplina científica, tendo sido guiada por professores e pesquisadores admiráveis.

Aos amigos do Clic, Claudia Freitas, Milena Garrão, Luiz Alexandre Amaral, Cícero Santos, Violeta Quental e Maria Carmelita Dias, pelas discussões e idéias compartilhadas, e pelas parcerias na pesquisa em Lingüística Computacional, que contribuíram imensamente para esta tese.

Aos meus alunos, todos os que já tive, por perguntarem. Em especial, a Gilberto Silva por ter-me proposto um tema de pesquisa que fez com que eu buscasse o conhecimento lingüístico e que assim iniciasse esse novo caminho.

Mais uma vez, ao Departamento de Letras e sua infra-estrutura administrativa, em particular à Chiquinha, por suavizar as dificuldades e tornar quase imperceptível a burocracia inerente à gerência de uma unidade de ensino e pesquisa.

A Roger Troth, meu marido, pelo apoio, incentivo e por todas as pequenas coisas que são tudo. Não devo agradecer à Clara e à Raquel, minhas filhas, pois chegaram no meio do trabalho e atrapalharam muito, por terem se tornado o novo centro do meu mundo.

Resumo

Oliveira, Claudia; Basilio, Margarida. **O Substantivo-suporte: Critérios Operacionais de Caracterização**. Rio de Janeiro, 2006. 118p. Tese de Doutorado — Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro.

Este trabalho tem por objetivo prover um critério operacional para caracterizar substantivos em combinações de substantivo seguido de adjetivo, em que o substantivo se apresenta em situação análoga à dos chamados verbos leves ou verbos-suporte, largamente estudados em Lingüística e Processamento de Linguagem Natural nos últimos anos. O trabalho se situa na confluência entre estudos lingüísticos, lexicográficos e computacionais e pretende explorar a potencialidade da análise automática de corpora e instrumentos quantitativos em busca de uma maior objetividade na fundamentação de conceitos que norteiam a atividade de análise lingüística. O desenvolvimento da pesquisa alia a pesquisa em corpus ao dicionário tradicional para realizar o levantamento das principais propriedades das combinações *S-Adj* particularizado para o caso de ocorrência de adjetivos denominais. A partir das informações lexicográficas e contextuais demonstra-se a existência de um conjunto de substantivos que participam das construções estudadas de maneira semelhante aos verbos-suporte em combinação *V-SN*. Um método automático de reconhecimento dos substantivos-suporte em textos é elaborado com o objetivo de fornecer aos estudiosos um instrumento capaz de produzir evidências convincentes dada a insuficiência de julgamentos intuitivos para justificar a delimitação de expressões de aparente irregularidade.

Palavras-chave

Substantivo-suporte; Substantivo; Adjetivo denominal; Classe de palavras; Lexicografia de corpus; Lingüística.

Abstract

Oliveira, Claudia; Basilio, Margarida. **Support Nouns: Operational Criteria for Characterization**. Rio de Janeiro, 2006. 118p. PhD Thesis — Department of Letras, Pontifícia Universidade Católica do Rio de Janeiro.

The main goal of this work is to provide operational criteria for characterizing nouns in Noun - Adjective combinations, in which the noun occurs in an analogous way to so called light verbs or support verbs, widely studied in recent years in both Linguistics and Natural Language Processing. In the work, linguistic, lexicographic and computational studies converge in order to explore the potential for automatic analysis of corpora, whose aim is to provide quantitative tools and methods which would lead to a more objective way of establishing concepts which underlie linguistic analysis. The work unites corpus-based research with traditional lexicography in order to elicit the main properties of the N-Adj combinations occurring with denominal adjectives. The lexicographic and contextual data reveal the existence of a set of nouns that occur in the studied constructions in a way similar to light verbs in V-Noun phrasal combinations. An automatic method for recognizing support nouns in texts is developed, which will provide language specialists with an instrument capable of bringing solid evidence to add to intuitive judgments in the task of justifying the delimitation of expressions that are apparently irregular.

Keywords

Support noun; Noun; Denominal adjective; Part of speech ; Corpus lexicography; Linguistics.

Sumário

1	Introdução	12
1.1	Caracterização do problema	13
1.2	Objetivos	15
1.3	Posicionamento Metodológico	16
1.4	Organização do texto	17
2	Quadro Teórico	19
2.1	Introdução	19
2.2	Modelos lexicais	19
2.3	Polissemia	30
3	Classes de Palavras	42
3.1	Introdução	42
3.2	Propostas de Classificação	43
3.3	Considerações sobre a classe dos substantivos	46
3.4	Considerações sobre a classe dos adjetivos	52
3.5	Fronteiras entre classes de palavras	62
4	A Conceituação do Substantivo-suporte	67
4.1	Introdução	67
4.2	Lexicografia do Substantivo-suporte	67
4.3	A Função de Suporte: o Verbo e o Substantivo	83
4.4	Exemplário	93
5	O substantivo-suporte no corpus	96
5.1	Introdução	96
5.2	Composicionalidade semântica	96
5.3	O Método	98
5.4	Experimentos	104
6	Conclusões	108
6.1	Contribuições	108
6.2	Desdobramentos	109
6.3	Considerações Finais	110
	Referências Bibliográficas	111

Lista de figuras

2.1	Rede de conexões para <i>cat, cats, rats, mats</i> e <i>caps</i>	23
2.2	Frame para QUESTION no FrameNet	28
2.3	Linhas contínuas unindo entre homonímia/monossemia e polilexia/vagueza	34
4.1	Entrada do dicionário Houaiss para COISA	73
4.2	Diagrama das referências cruzadas entre definições de substantivos-suporte	76
4.3	Entrada do dicionário Houaiss para FORMA	77
4.3	Entrada do dicionário Houaiss para FORMA (cont.)	78
4.4	Diagrama das referências cruzadas entre definições de substantivos-suporte (reiteração da figura 4.2)	90
5.1	Resumo do método de identificação de substantivos-suporte em contexto	104

Lista de tabelas

1.1	Linguística baseada em corpus vs. linguística dirigida por corpus	17
3.1	Classes de palavras na tradição grega (esq.) e latina (dir.)	43
3.2	Motivações para as classes de palavras	44
3.3	Lista de substantivos genéricos do inglês	49
4.1	Lista de substantivos-suporte do Português	69
4.2	Número de adjetivos distintos em sintagmas <i>N Adj</i> com substantivos-suporte	72
4.3	Número de sentidos dicionarizados	73
4.4	Número de expressões multi-vocabulares dicionarizadas	74
4.5	Referências cruzadas em entradas do dicionário para substantivos-suporte	75
4.6	Acepções de <i>FORMA</i> , com exemplos do corpus	78
4.6	Acepções de <i>FORMA</i> , com exemplos do corpus (cont.)	79
4.7	Acepções de <i>FORMA</i> , com categorias de distinção	80
4.8	Resultados de concordância	82
4.9	Paráfrases para <i>S-Adj</i> com substantivos-suporte	92
4.9	Paráfrases para <i>S-Adj</i> com substantivos-suporte (cont.)	93
5.1	Expressões <i>S Adj</i> testadas	105
5.2	Resultados com as expressões <i>S Adj</i>	105
5.3	Resultados com <i>biológico</i>	106
5.4	Resultados com <i>cultural</i>	106

Convenções Tipográficas

MAIÚSCULAS PEQUENAS

Para exemplos de palavras e expressões.

Itálico

Para exemplos de contextos.

Negrito

Para terminologia, quando for introduzida.

“Aspas duplas”

Para citações de outros autores.

‘Aspas simples’

Para exemplos de sentidos.

[Colchetes]

Para exemplos não retirados do corpus.

? Interrogação

Para estranheza semântica ou sintática.

* Asterisco

Para impossibilidade semântica ou sintática.

O físico Leo Szilard certa vez anunciou ao amigo Hans Bethe que estava pensando em escrever um diário: “Não pretendo publicar. Só vou registrar os fatos para a informação de Deus”. “Você não acha que Deus sabe dos fatos?”, Bethe perguntou. “Sim”, respondeu Szilard. “Ele sabe dos fatos mas não desta versão dos fatos”

Hans Christian von Baeyer, *Taming the Atom* (1992).