

# Substantivos-suporte: uma visão lexicográfica

Claudia Oliveira\*  
Maria Claudia de Freitas\*\*

## RESUMO

Este trabalho apresenta um estudo lexicográfico dos substantivos-suporte. Esses nominais possuem características análogas aos chamados verbos leves ou verbos-suporte, apresentando propriedades lexicográficas distintivas, como o elevado número de acepções em suas entradas, o elevado número de expressões multi-vocabulares em que ocorrem e o grau interligação de suas definições por referências cruzadas. O trabalho se situa na confluência entre estudos lingüísticos, lexicográficos e computacionais e explora a potencialidade da análise de corpora para a fundamentação e evidenciação de conceitos lingüísticos.

Palavras-chave: Substantivo-suporte; sintagma nominal; lexicografia computacional; lingüística computacional.

## ABSTRACT

*This article presents a lexicographic study of the support-nouns. These words have analogous properties to light verbs, or support verbs, and distinctive lexicographic features, such as a large number of dictionary senses, the fact that they occur in a large number of multi-word expressions and the degree with which their definitions are inter-related, as indicated by cross-references between these definitions. In this work, linguistic, lexicographic and computational studies converge, and the analysis of corpora provides a more objective way of establishing concepts which underlie linguistic analysis.*

*Keywords:* support noun; noun phrase; computational lexicography; electronic dictionary

## Introdução:

Este trabalho apresenta um estudo lexicográfico dos substantivos-suporte. Esses nominais têm como principal característica a possibilidade de ocorrerem em combinações S+Adj ou S+Prep+S, 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 explora a potencialidade da análise de corpora para a fundamentação e evidenciação de conceitos lingüísticos.

O objetivo do trabalho é descrever o conjunto de substantivos-suporte do Português do Brasil, utilizando suas características lexicográficas extraídas de um corpus e de um dicionário brasileiro padrão. O conceito de substantivo-suporte em que se baseia o trabalho (RefAutores 00) pressupõe um modelo de léxico que dê conta de construções regulares e semi-regulares, tais como as construções verbais com verbo-suporte. Os modelos lexicais computacionais inspirados em teorias lingüísticas de orientação funcionalista oferecem um quadro adequado para a utilização dos resultados desta pesquisa em sistemas computacionais.

As propriedades lexicográficas investigadas incluem: o elevado número de acepções em suas entradas, indicando que são palavras polissêmicas do ponto de

\* Instituto Militar de Engenharia Departamento de Engenharia de Computação

\*\* PUC-Rio, Departamento de Letras.

vista lexicográfico; o elevado número de expressões multi-vocabulares em que ocorrem; e o grau interligação de suas definições por referências cruzadas. Com relação à frequência, substantivos-suporte são abundantes tanto no discurso formal quanto no informal, o que justifica o presente estudo. Dentro de um projeto mais amplo de identificação de expressões denotativas úteis para um léxico computacional, a classe dos sintagmas nominais com substantivos-suporte apresentam muitas peculiaridades que podem ser exploradas.

A artigo está assim organizado: na Seção 2, uma caracterização do substantivo-suporte é apresentada, levando em consideração, por um lado, a dificuldade na definição formal dessa classe de palavras e, por outro lado, a facilidade com que o falante identifica um exemplo de substantivo-suporte por sua vagueza. Outras classes de substantivos relacionadas ao substantivo-suporte são descritas com o propósito de ressaltar algumas características importantes da classe em foco. A análise do substantivos-suporte é feita posicionando-o no contexto de um adjetivo denominal, já que essa combinação representa uma perfeita simbiose. A partir de uma lista de substantivos que, de acordo com uma pesquisa de corpus preliminar, desempenham função de suporte, na Seção 3 analisamos as entradas desses substantivos-suporte potenciais em um dicionário tradicional. Das características observadas, uma classificação é proposta tendo em vista a função do substantivo no sintagma. A Seção 4 apresenta um resumo dos dados obtidos e uma análise de um grupo de substantivos-suporte com respeito a seus contextos e frequências de ocorrências. Na Seção 5 são apresentadas as conclusões do trabalho.

## 2. A noção de substantivo-suporte

Os substantivos-suporte constituem uma classe não muito extensa de substantivos abstratos do português que se caracterizam por sua generalidade semântica. Enquadram-se no âmbito de um fenômeno mais amplo da linguagem, a função de suporte, amplamente estudada com respeito aos verbos. Em português, (Neves 1996) e (Neves 1999) discutem o **verbo-suporte**; (Scher 2003) utiliza a denominação **verbo leve**.

Os substantivos-suporte, segundo (RefAutores 00), possuem as seguintes características que se apresentam analogamente nos verbos-suporte:

1. O substantivo-suporte é semanticamente vago. Identificando como vagueza o esvaziamento lexical, observado por vários autores a respeito dos verbos-suporte, pode-se demonstrar o quanto ausência de especificação do substantivo-suporte torna o enunciado problemático. Em b), o apagamento do adjetivo denominal *POLÍTICA* torna o enunciado anômalo, pois *ESFERA*, como substantivo-suporte, é demasiadamente vago; em contraste, em d) o apagamento do adjetivo não provoca o mesmo efeito.

- a) Na *ESFERA* política, os manda-chuvas aprenderam a tratar os conflitos intestinos com maior flexibilidade e eficácia.
- b) ? Na *ESFERA*, os manda-chuvas aprenderam a tratar os conflitos intestinos com maior flexibilidade e eficácia.
- c) No *PARTIDO* político, os manda-chuvas aprenderam a tratar os conflitos intestinos com maior flexibilidade e eficácia.
- d) No *PARTIDO*, os manda-chuvas aprenderam a tratar os conflitos intestinos com maior flexibilidade e eficácia.

2. O complemento do substantivo-suporte, frequentemente um adjetivo denominal, tem como núcleo um nome, que veicula a parte mais relevante da denominação.

e) Os treinadores, do mesmo jeito, tratam-se de uma FORMA *diplomática*, mas não levam uma conversa por mais de 10 minutos.

3. A construção admite paráfrase, quer em termos da substituição da construção *S+Adj* por uma expressão cujo núcleo é o substantivo base do adjetivo (f), quer em termos da substituição do substantivo-suporte, mantendo-se o adjetivo (g).

f) Os treinadores, do mesmo jeito, tratam-se com DIPLOMACIA, mas não levam uma conversa por mais de 10 minutos.

g) Apesar do sucesso das modinhas no(a) {ÂMBITO, ÁREA, CAMPO, ESFERA} doméstico(a), o cenário musical durante o Império foi dominado pelo canto lírico.

## 2.1. Classes Relacionadas

Na lingüística, o termo **substantivo vazio** é utilizado para fazer referência a substantivos que não denotam conceitos, ou substantivos com um conteúdo semântico mínimo, identificados, na análise sintática, com elementos sem expressão fonológica que atuam como núcleos de sintagmas nominais. De acordo com Panagiotidis (2003), no léxico de qualquer língua existe um número limitado de substantivos vazios, que formam uma classe fechada. Eles são considerados palavras gramaticais, não lexicais, com forte característica pronominal. Tendo em vista sua reduzida capacidade de denotação, substantivos vazios distinguem-se entre si por meio de seus traços fonológicos e morfo-semânticos, tais como o gênero. O substantivo vazio típico do Inglês seria *ONE*, como em *a big ONE*.

Na teoria lexical, tem havido muito pouco interesse no fenômeno, com algumas exceções como (Schmid, 2000), que apresenta um estudo sobre substantivos abstratos. De acordo com Schmid, há um subconjunto dos substantivos abstratos em inglês que atuam como conchas conceituais, que realizam seu potencial semântico em associação com outros itens lexicais. Palavras como *THING, FACT, CASE, POINT, IDEA, REASON, PROBLEM* e *QUESTION*, entre outros substantivos de alta frequência no inglês, são exemplos desse classe.

Estudos lexicais aplicados ao processamento automático de textos ampliaram essa noção, incluindo palavras que podem ser apagadas ou ignoradas para o propósito de sumarização, classificação, agrupamento e outras operações computacionais sobre textos. Substantivos ocorrendo como quantificadores em expressões multi-vocabulares, tais como *GROUP, BUNCH* e *LOTS*, em *GROUP of students, BUNCH of students, LOTS of students*, são exemplos especificamente mencionados em (Muresan et al, 2001) para o inglês. Na construção de listas de palavras vazias, tradicionalmente, incluem-se também as palavras funcionais e os verbos auxiliares. Substantivos-suporte não são palavras vazias, a julgar pelo critério usual.

## 2.2. O Substantivo-suporte em Contexto

Nossos primeiros questionamentos sobre por este tipo de palavra surgiram

a partir de uma pesquisa de corpus cujo foco eram os sintagmas nominais *S+Adj*, onde *Adj* é um adjetivo denominal. O resultado da pesquisa mostrou que, combinados aos mais variados adjetivos, sempre havia termos, como *FATOR*, *PERSPECTIVA* e *ASPECTO*, que são aparentemente intercambiáveis, apesar de terem significados literais bastante distintos. Os dados mostram que a contribuição desses termos para o significado geral da expressão é bastante reduzida. Dado que as unidades do processamento de linguagem natural são a palavra e o sintagma e que o substantivo é o núcleo do sintagma nominal, é comum esperar que o substantivo seja o centro do significado do sintagma. No caso de adjetivos predicativos, como em *céu aberto*, a idéia central está no substantivo; no caso de adjetivos relacionais, como em *lesão cerebral*, o substantivo e o adjetivo se fundem em uma única unidade denotativa. Em contraste, em um sintagma contendo um substantivo-suporte, por exemplo em *ASPECTO comercial*, fica claro que 'comércio' é a idéia central, ainda que o substantivo núcleo seja *ASPECTO*.

Assim, no contexto de processamento automático e interpretação de textos, a possibilidade de identificar esses substantivos é de grande importância. Estamos buscando os casos em que a ocorrência do substantivo não se corresponda a um conceito, deixando de cumprir seu papel denominativo. Como exemplo, consideramos o substantivo *ASPECTO* na seguinte sentença do corpus:

a) Examinava o *ASPECTO constitucional* que lhe garantia o direito do voto.

O apagamento do adjetivo sublinhado enfraquece consideravelmente o conteúdo do enunciado, tornando o significado do sintagma amplo demais; o adjetivo especifica completamente o objeto. Esse é um exemplo de um substantivo-suporte, um substantivo cujo complemento é um adjetivo ou sintagma preposicional majoritariamente responsável pelo significado da expressão como um todo. Mesmo que sintaticamente o substantivo seja o núcleo do sintagma, semanticamente ele ocorre sem conteúdo específico.

Sendo assim, uma ótima pista para o reconhecimento de um substantivo-suporte está em seus complementos. Investigamos os tipos de adjetivos que ocorrem como complementos destes substantivos no corpus e observamos que eles são em grande parte denominais, ainda que possam aparecer em função **relacional** ou **qualitativa**. O conceito de adjetivo relacional, em oposição a adjetivo qualitativo, está bem discutido por Nirenburg & Raskin (2004): qualitativo refere-se a adjetivos predicadores e relacional refere-se a não-predicadores. Em português são mais frequentes os termos **predicativo** e **denotativo**, respectivamente.

Basilio & Gamarski (1995) discutem os sintagmas nominais *S+Adj* do ponto de vista do adjetivo, observando como opera a função denotativa em combinação com substantivos "de cunho muito geral que necessitam de um preenchimento semântico a ser fornecido via de regra por um adjetivo denominal, substantivo ou sintagma preposicionado". Exemplificando a partir do corpus do NURC (Projeto de Estudo Conjunto e Coordenado da Norma Urbana Oral e Culta), de discurso oral, as autoras selecionaram seis enunciados, quatro dos quais construídos com substantivo-suporte listados na Tabela 1 (*FATOR*, *PARTE*, *ASPECTO* e *COISA*).

Uma análise inicial, de exemplos coletados assistematicamente, motiva os principais questionamentos do nosso trabalho. Por exemplo, em:

b) ... estudantes talentosos, atraídos para os campos cognitivistas, ditavam

o TOM intelectual do momento ...

o substantivo TOM atua como um focalizador da essência da denotação do adjetivo *intelectual*. Em contraste, o substantivo ASPECTO em

c) Os temas incluem o ASPECTO comercial da moda, as tendências internacionais ...

secciona o escopo do adjetivo relacional *comercial* em sub-divisões de um todo.

d) Enfrentamos, portanto, um segundo dilema ... que é de NATUREZA ambiental.

Em d) o substantivo NATUREZA pode ser apagado pois possui um papel apenas textual. Algumas vezes o apagamento do substantivo-suporte não provoca mudança significativas, como no seguinte exemplo:

e) A decisão foi de CUNHO pessoal. → A decisão foi pessoal.

Em outras ocasiões, é necessário ajustar o enunciado, sintaticamente ou morfologicamente, como em:

f) Agiu de FORMA cerebral. → Agiu cerebralmente.

g) Atuava no ÂMBITO municipal. → Atuava no município.

A partir desses fatos podemos concluir que um critério operacional promissor para a caracterização do substantivo-suporte é a extração de sintagmas nominais *S+Adj*, onde *Adj* é um adjetivo denominal. A lista de substantivos resultantes podem ser analisados de acordo com sua intercambialidade, ou seja, sua co-ocorrência com o mesmo grupo de adjetivos, como no seguinte exemplo, que mostra três destes substantivos que ocorreram combinatoriamente com três diferentes adjetivos denominais:

<i>fator</i>		<i>fator</i>		<i>fator</i>		
<i>perspectiva</i>	> <i>racial</i>	<i>perspectiva</i>	> <i>fiscal</i>	<i>perspectiva</i>	> <i>ambiental</i>	
<i>aspecto</i>		<i>aspecto</i>		<i>aspecto</i>		

### 3. A Lexicografia do Substantivo-suporte

#### 3.1 O dicionário e a questão da multiplicidade de sentidos

O falante comum admite de bom grado que as palavras possuam múltiplos significados. Fora dessa hipótese, como explicar, sem recorrer a uma teoria lexicológica sofisticada, as possibilidades abaixo?

a) Raimundo sorriu, amarrotou a folha de PAPEL e lançou-a no chão.

b) O flúor desempenha um PAPEL importante na resistência dos dentes.

c) O PAPEL subiu 20%, pulando dos R\$20 o lote de mil ações para R\$24...

d) Ele rasga uma cruz que havia sobre sua cama, joga santinhos no vaso sanitário e usa as vestimentas de primeira comunhão como PAPEL higiênico.

É para esse usuário que os dicionários separam entradas diferentes para cada significado percebido pelo lexicógrafo, procurando, de modo sistemático e consistente, elucidar as diferenças entre os significados por meio de definições e exemplos.

Por outro lado, é cada vez mais claro, como afirma (Marcuschi 2004), que as expressões e seus sentidos, conteúdos, referentes, etc. não são dois lados da mesma moeda. Não se podem arranjar os sentidos de acordo com uma categorização da polissemia pois o problema é uma consequência do “**princípio de simbolização**, que resulta num sistema categorial fluido”.

Apesar de todas as restrições aos significados prontos, ao modelo de léxico como conjunto de etiquetas que se atribuem a elementos de um mundo discretizado, à semântica imanente das palavras, ainda assim, é necessário operar com acepções dicionarizadas dos substantivos-suporte. Para a operacionalização desse trabalho, a categorização lexicográfica é de grande relevância pelo fato de que o dicionário é uma base de dados, um repositório do conhecimento tradicional e consagrado sobre o vocabulário da língua portuguesa.

### 3.2 Freqüência na língua

Freqüência de ocorrência é um dado lingüístico que apenas recentemente vem ganhando a relevância devida na lexicografia, com o desenvolvimento de dicionários baseados em corpus, como (Sinclair 2001).

A freqüência pode ser um fator de identificação de substantivos-suporte. MURESAN et al (2001) identificam 141 **substantivos vazios** em um corpus do inglês, uma classe com muitas interseções com a dos substantivos-suporte, baseando-se apenas em contagem de freqüência. Com relação às conchas nominais, Schmid (2000) relata que, em um corpus de 225 milhões de palavras correntes do inglês britânico, as formas singulares de CASE, FACT, IDEA, NEWS, POINT, PROBLEM, REPORT e THING estão entre os cem substantivos mais freqüentes. O substantivo THING, por exemplo, tem uma freqüência de 256 ocorrências por milhão, sendo que a vasta maioria dos itens lexicais do inglês possuem freqüência menor que 20 por milhão.

Em português, Marques (1995) descreve um estudo do léxico de alta freqüência numa parte do corpus do projeto NURC, um conjunto de textos orais, provenientes de entrevistas realizadas na cidade do Rio de Janeiro. Das 506.108 unidades textuais processadas, 370.777 (39% das unidades) eram da classe dos substantivos, correspondendo a ocorrências distintas de 9.539 lexemas (51,5% dos lexemas).

No corpus do NURC, os substantivos de alta freqüência (até 100 ocorrências) totalizam 488 lexemas. Marques considera que 75 deles são de “sentido geral”, ou seja, não dependente de contexto temático, e o restante é classificado por campo temático, como: ‘tempo’, ‘indivíduos’, ‘corpo humano’, ‘vestuário’, ‘casa’, ‘alimentação’, ‘família’, etc. Ao analisarmos os quadros de substantivos de sentido geral, encontramos 15 dos 35 substantivos-suporte preliminarmente selecionados no corpus (Tabela1).

### 3.3 O que diz o dicionário

Para dar partida na investigação dos substantivos-suporte e suas combinações, extraímos de um dicionário do português brasileiro (Houaiss, 2001) um

grupo de 166 adjetivos denominais regulares, ou seja, adjetivos cujo significado é um produto motivado pela atuação do processo derivacional sobre o significado do substantivo básico. Não se pode garantir *a priori* se uma dada ocorrência de um adjetivo denominar será ou não denotativa, portanto não utilizamos esse critério, que no entanto geraria maior precisão dos dados.

<i>âmbito</i>	<i>cunho</i>	<i>matéria</i>	<i>perspectiva</i>
<i>área</i>	<i>dimensão</i>	<i>modo</i>	<i>plano</i>
<i>aspecto</i>	<i>elemento</i>	<i>natureza</i>	<i>ponto</i>
<i>base</i>	<i>esfera</i>	<i>nível</i>	<i>quadro</i>
<i>campo</i>	<i>fator</i>	<i>ordem</i>	<i>questão</i>
<i>caráter</i>	<i>forma</i>	<i>panorama</i>	<i>sentido</i>
<i>coisa</i>	<i>lado</i>	<i>papel</i>	<i>tipo</i>
<i>componente</i>	<i>maneira</i>	<i>parte</i>	<i>tom</i>

TABELA 1: Lista de substantivos-suporte do Português

Usando o critério de intercambialidade, selecionamos de um corpus uma lista preliminar de substantivos-suporte, colocados na Tabela 1. A partir dessa lista, procuramos inferir algumas propriedades lexicográficas importantes. Ainda que tenhamos encontrado um número reduzido de substantivos-suporte no corpus, classificá-los não é tarefa fácil, dada o seu alto grau de polissemia e a variedade de entradas que eles geram em dicionários convencionais. Além disso, muitas dessas entradas não correspondem em sentido ao substantivo-suporte.

Dos 166 adjetivos escolhidos, apenas 56 não ocorreram em associação com o substantivo-suporte. Esses adjetivos eram pouco freqüentes no corpus, como por exemplo *DORSAL* e *CONDOMINIAL*. A tabela 2 mostra os substantivos que se combinam com maior freqüência a diferentes adjetivos.

<i>Substantivo</i>	<i>Sentidos</i>	<i>Substantivo</i>	<i>Sentidos</i>
<i>caráter</i>	31	<i>plano</i>	17
<i>aspecto</i>	24	<i>área</i>	16
<i>parte</i>	23	<i>questão</i>	15
<i>natureza</i>	21	<i>tipo</i>	15
<i>forma</i>	20	<i>campo</i>	14
<i>problema</i>	20	<i>papel</i>	13
<i>ordem</i>	19	<i>base</i>	12
<i>lado</i>	17	<i>sentido</i>	12
<i>nível</i>	17	<i>âmbito</i>	10

TABELA 2: Número de adjetivos distintos em sintagmas *S+Adj* com substantivos-suporte

Ao procurar a lista preliminar de substantivos-suporte em um dicionário padrão do português brasileiro (Houaiss, 2001), percebemos que algumas peculiaridades se destacam. Em primeiro lugar, o número de acepções de um substantivo-suporte é, em geral, maior que de um substantivo qualquer. Observa-se que os substantivos-suporte apresentam um número elevado de acepções, muitas delas marcadas como derivações semânticas, como metonímia, metáfora e analogia. A

Tabela 3 mostra os substantivos-suporte mais polissêmicos da lista da Tabela 1.

<b>Substantivo</b>	<b>Sentidos</b>	<b>Substantivo</b>	<b>Sentidos</b>
<i>ponto</i>	57	<i>natureza</i>	18
<i>base</i>	30	<i>matéria</i>	17
<i>ordem</i>	30	<i>plano</i>	17
<i>forma</i>	27	<i>modo</i>	16
<i>quadro</i>	27	<i>sentido</i>	14
<i>coisa</i>	21	<i>área</i>	13
<i>tipo</i>	20	<i>caráter</i>	13
<i>tom</i>	19	<i>maneira</i>	12
<i>campo</i>	18	<i>parte</i>	12

TABELA 3: Número de sentidos dicionarizados

Esta flutuação em significado mostra que o substantivo-suporte é semanticamente tão vago que parece exigir complementação. Por exemplo, o substantivo *COISA* aparece em (Houaiss, 2001) com 21 sentidos distintos, como mostra a Tabela 4. O substantivo *COISA* funciona como um coringa, cujo significado só pode ser compreendido no contexto de um enunciado. Na verdade, várias línguas Indo-Européias, como espanhol, alemão, inglês e francês possuem um substantivo-suporte que pode ser traduzido como *COISA*, que pode ser considerado como a manifestação ostensiva de um esquema cognitivo de substantivo e pode substituir quase todos os substantivos (Mihatsch, 2003).

Substantivos-suporte também apresentam muitas definições qualificadas em campos semânticos de especialidade. O substantivo *BASE*, por exemplo, possui 30 sentidos dicionarizados, 23 dos quais são em rubricas especializadas.

1. tudo quanto existe ou possa existir, de natureza corpórea ou incorpórea;	12. negócio, transação;
2. qualquer ser inanimado;	13. algo que provoque estímulo, que entusiasme; motivo, incentivo, compensação;
3. realidade, fato concreto, em relação ao que é abstrato ou assim considerado;	14. o que não se sabe; mistério, enigma;
4. algo que não se quer ou não se pode nomear;	15. mal-estar ou indisposição súbita; ataque, perda dos sentidos;
5. aquilo de que se está tratando ou falando;	16. cigarro de maconha; baseado;
6. aquilo que se pensa; pensamento, idéia;	17. órgão genital do homem ou da mulher;
7. relação, ligação, vínculo;	18. algo imprestável, velho ou maltratado; traste, troço, bagulho;
8. interesse próprio, negócio, ocupação;	19. o diabo;
9. ato, empreendimento, empresa;	20. bens, propriedades, valores;
10. o que acontece; ocorrência, evento, caso;	21. negócios, interesses, ocupações.
11. assunto, tema, matéria;	

TABELA 4: Entrada do Houaiss para *COISA*

<b>Subst.</b>	<b>EMVs</b>	<b>Exemplos</b>
<i>papel</i>	113	<i>p. alçaço, confiar ao p., passar p. com, ficar no p.</i>
<i>ponto</i>	110	<i>p. alto, p. cardeal, p. crítico, p. culminante, p. cego</i>
<i>campo</i>	49	<i>c. conceitual, c. de concentração, c. de força, c. visual</i>
<i>base</i>	42	<i>b. aérea, b. monetária, b. vetorial, à b. de, b. de dados</i>
<i>plano</i>	38	<i>p. geral, p. de saúde, p. inclinado, primeiro p.</i>
<i>fator</i>	31	<i>f. de correção, f. Rh, f. abiótico, f. de crescimento</i>
<i>ordem</i>	30	<i>o. civil, o. cronológica, o. de grandeza, o. do dia</i>
<i>forma</i>	29	<i>f. canônica, f. livre, de certa f., de f. alguma, de f. que</i>
<i>sentido</i>	26	<i>s. anti-horário, s. estrito, s. figurado, duplo s., fazer s.</i>
<i>parte</i>	25	<i>p. do discurso, p. ideal, p. íntimas, à p., a p. do leão</i>
<i>área</i>	24	<i>á. de livre comércio, á. de transferência, á. de serviço</i>
<i>modo</i>	22	<i>m. de ser, m. de ação, m. maior, de m. a, de todo m.</i>
<i>coisa</i>	20	<i>c. de, c. pública, cheio de c., não dizer c. com c.</i>
<i>nível</i>	19	<i>n. de energia, n. de vida, alto n., ao n. de, baixo n.</i>
<i>elemento</i>	17	<i>e. neutro, e. mórfico, e. de composição, estar no seu e.</i>
<i>questão</i>	16	<i>q. aberta, q. de ordem, q. de tempo, q. fechada</i>
<i>lado</i>	15	<i>l. a l., ao l. de, de l., de um l. para outro, pôr de l.</i>
<i>caráter</i>	13	<i>c. hereditário, c. tipográfico, a c., de c.</i>
<i>matéria</i>	13	<i>m. em questão, m. pré-estrelar, m. processual</i>
<i>quadro</i>	11	<i>q. clínico, q. mural, q. de horário, q. vivo</i>
<i>tipo</i>	11	<i>t. comum, t. de caixa, t. ideal, fazer t.</i>

TABELA 5: Número de expressões multi-vocabulares dicionarizadas

A segunda característica lexicográfica mais marcante dos substantivos-suporte é o número de expressões multi-vocabulares dicionarizadas em que ocorrem. Por exemplo, se analisarmos a entrada de PAPEL, observamos que 100 das 113 expressões dicionarizadas são tipos especiais de papel, tais como PAPEL CELOFANE e PAPEL CREPOM, sob a rubrica 'indústria de papel'. Portanto não podem ser consideradas ocorrências do sentido de suporte de PAPEL. Já na entrada de PONTO, as 110 expressões listadas são de grande diversidade em sentido e campo semântico. Apesar dessas considerações, é justo afirmar que o substantivo-suporte tende a formar expressões cristalizadas, como ilustra a tabela 5.

Quando analisamos as entradas do sentido de suporte dos substantivos, encontramos uma grande quantidade de referências cruzadas, revelando uma rede muito interessante de relacionamentos, que pode ser utilizada para classificar os grupos de substantivos-suporte por função semântica.

<i>Substantivo-suporte</i>	<i>Referências cruzadas</i>	<i>Substantivo-suporte</i>	<i>Referências cruzadas</i>
<i>âmbito</i>	<i>esfera, campo</i>	<i>lado</i>	<i>maneira, aspecto</i>
<i>área</i>	<i>campo</i>	<i>maneira</i>	<i>modo, forma, processo</i>
<i>aspecto</i>	<i>maneira, lado</i>	<i>modo</i>	<i>forma</i>
<i>base</i>	<i>parte, aspecto</i>	<i>natureza</i>	<i>caráter, tipo</i>
<i>campo</i>	<i>área, esfera, âmbito</i>	<i>panorama</i>	<i>matéria</i>
<i>caráter</i>	<i>cunho</i>	<i>parte</i>	<i>matéria, papel, área, lado</i>
<i>cunho</i>	<i>caráter</i>	<i>perspectiva</i>	<i>forma</i>
<i>dimensão</i>	<i>aspecto</i>	<i>plano</i>	<i>nível</i>
<i>elemento</i>	<i>parte</i>	<i>ponto</i>	<i>parte, aspecto</i>
<i>esfera</i>	<i>área</i>	<i>questão</i>	<i>matéria, ponto</i>
<i>fator</i>	<i>elemento</i>	<i>sentido</i>	<i>modo, aspecto</i>
<i>forma</i>	<i>modo, maneira, tipo</i>	<i>tom</i>	<i>modo, caráter</i>

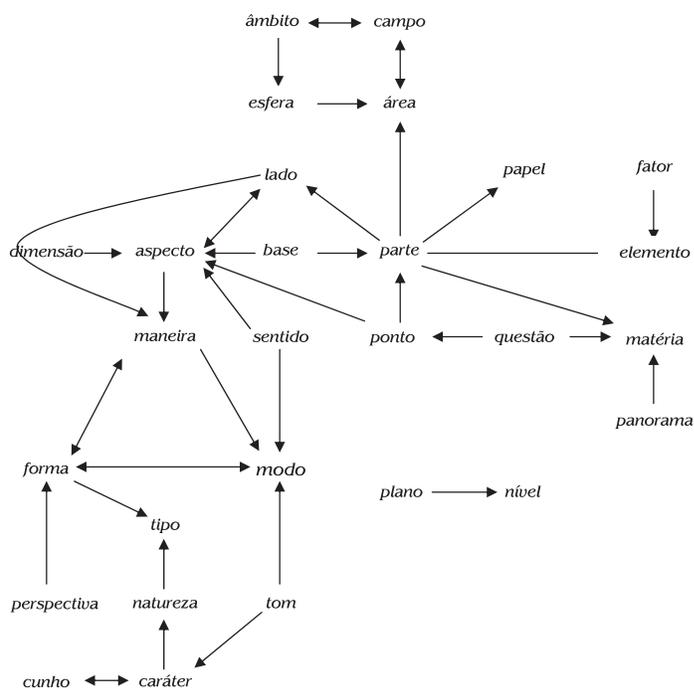


TABELA 6: Referências cruzadas em entradas do dicionário para substantivos-suporte

Os dados da Tabela 6 foram utilizados para construir um grafo de sentidos conectados da Figura 1. As definições dos substantivos na primeira coluna fazem referência aos substantivos na segunda coluna. Por exemplo, na definição de "âmbito" há referências a "campo" e "esfera".

Âmbito - n substantivo masculino

1. espaço que circunda, rodeia, envolve; periferia
2. espaço físico compreendido dentro de determinados limites; recinto, ambitude. Ex.: o â. do plenário. Área que os antigos romanos eram obrigados a deixar em volta de suas casas
3. Derivação: sentido figurado. Esfera de ação ou de pensamento; campo ou espaço em que ocorre ou se exerce alguma atividade. Ex.: era um assunto do â. da psicologia.
4. Derivação: sentido figurado. Núcleo central ou predominante dessa esfera. Ex.: aquela mulher não estava no â. de suas cogitações.
5. Rubrica: música. Intervalo entre a nota mais grave e a mais aguda de um trecho musical considerado.

FIGURA 1: Diagrama das referências cruzadas entre definições de substantivos-suporte

Analisando o grafo da Figura 1, percebe-se o quanto ele parece induzir uma classificação, por sua configuração agrupada. Os quatro grupos de substantivos-suporte resultantes da interpretação da figura podem ser descritos de acordo com as seguintes funções:

1. **função de delimitação**, que especifica o objeto do enunciado numa estrutura abstrata de campos de informação, onde são classificados ÂMBITO, CAMPO, ESFERA, QUADRO e ÁREA; essa função ficou concentrada na região superior do grafo;
2. **função de sub-divisão**, que determina os níveis ou partes dentro de um todo, onde são classificados LADO, ASPECTO, PARTE, PONTO, FATOR ELEMENTO, MATÉRIA, QUESTÃO, COMPONENTE e NÍVEL;
3. **função de enfoque**, que especifica o modo como o objeto é abordado, onde são classificados SENTIDO, MODO, MANEIRA, FORMA, DIMENSÃO, PERSPECTIVA e PLANO;
4. **função de essência**, que define o objeto do enunciado total ou basicamente, onde são classificados CUNHO, CARÁTER, ORDEM, NATUREZA, BASE, TIPO, ORDEM e TOM. Essa função ficou concentrada na região inferior do grafo.

#### 4. O Que Diz o Corpus

A utilização de análises lingüísticas baseadas em corpus vem ampliando as possibilidades de detecção de padrões construcionais das línguas. O apoio computacional disponível viabiliza o trabalho de pesquisa em corpora gigantescos, de grande cobertura e variabilidade textual. Do ponto de vista operacional, construções complexas, restritas por condições gramaticais de diversas naturezas, podem ser codificadas para que sejam identificadas em contexto, liberando o lingüista da tarefa enfadonha da busca manual de exemplos. O resultado da busca automática é mais consistente, pois se dá de maneira homogênea ao longo do tempo, o que é extremamente difícil para o lingüista individualmente e, ainda mais, em trabalhos cooperativos.

##### 4.1. A Descrição do Corpus

Como já dito previamente, este estudo do substantivo-suporte originou-se de uma pesquisa em corpus cujo foco era o sintagma N+Adj, onde Adj é um adjetivo nominal resultante de um processo de derivação.

O corpus do português do Brasil utilizado foi compilado pelo MILC (Núcleo Interinstitucional de Lingüística Computacional), e está descrito em (Aires & Aluísio, 2001). As consultas, realizadas por meio de uma linguagem especializada de especificação de padrões lingüísticos, resultam em um conjunto de segmentos textuais que, de maneira geral, correspondem a parágrafos dos textos da coleção. O corpus contém cerca de 37 milhões de palavras em textos em prosa, divididos em textos corrigidos, textos não corrigidos e textos semi-corrigidos. Os textos clas-

sificados como corrigidos, totalizando 33.081.000 palavras, são aqueles publicados para grande número de leitores (livros, jornais, revistas, etc.), que são, portanto, supostamente corrigidos por especialistas em revisão de textos. Há cerca de 4.490 textos de diversos gêneros: livros (de literatura brasileira; didáticos - biologia, química, física, história, geografia; enciclopédias; temáticos - arte, ciências, etc.); revistas; constituição brasileira e textos jurídicos; jornais. Os textos não corrigidos, totalizando 736.000 palavras, são textos autênticos, escritos por pessoas de nível médio de escolaridade (segundo grau) e universitários. Há 2.430 textos que incluem redações, monografias e textos de publicidade, por exemplo. Finalmente, os textos semi-corrigidos, em número de 340, com aproximadamente 1.390.000 palavras, são textos publicados para um pequeno número de leitores, ou não publicados, que são corrigidos, mas, geralmente, não por especialistas em revisão de textos. Estão entre estes, contratos, relatórios, dissertações acadêmicas, etc.

#### 4.2. Uma Análise dos Dados

A análise detalhada de todos os substantivos-suporte em sintagmas S+Adj é uma necessidade computacional; esses itens devem ser observados individualmente, já que se trata de uma classe provavelmente fechada, de substantivos muito peculiares.

Para demonstrar a viabilidade, segurança e produtividade da estratégia utilizada, apresentamos aqui a análise de um subgrupo dos substantivos-suporte com função de essência: TOM, CARÁTER, CUNHO e NATUREZA. Os seguintes padrões foram utilizados para concordância no corpus: "cunho" [pos="ADJ"], "natureza" [pos="ADJ"], "caráter" [pos="ADJ"] e "tom" [pos="ADJ"], onde "pos" abrevia **classe de palavras**<sup>5</sup>. A Tabela 7 resume os resultados.

Padrão	Ocorrências Totais	Ocorrências Distintas	S é de suporte
"cunho" [pos="ADJ"]	89	49	100,00%
"natureza" [pos="ADJ"]	230	176	93,00%
"caráter" [pos="ADJ"]	730	427	97,00%
"tom" [pos="ADJ"]	478	113	72,00%

TABELA 7: Resultados de concordância

Analisamos os dados referentes às ocorrências de substantivos-suporte seguidos de adjetivos denominais. O substantivo CUNHO apareceu com sentido de suporte em 100% dos casos, no contexto S+Adj; nessas ocorrências 55% dos adjetivos eram denominais.

O substantivo NATUREZA apareceu como substantivo-suporte em 93% dos casos, 56% dos quais foram seguidos de adjetivos denominais. Nos casos em que o sentido não foi considerado de suporte (7% - 17 ocorrências), encontramos 4 adjetivos de origem: NATUREZA amazônica, NATUREZA californiana, NATUREZA brasileira e NATUREZA americana; dos 13 restantes apenas 4 eram seguidos de adjetivos denominais.

<sup>5</sup>POS é a abreviatura de part of speech, um rótulo consagrado mesmo embases de dados do português.

O substantivo *CARÁTER* apareceu como substantivo-suporte em 97% dos casos, apenas 36% dos quais eram seguidos de adjetivos denominais. Entre os sentidos plenos, 3 eram em línguas de especialidade (biologia); dos restantes apenas 3 eram seguidos de adjetivos denominais.

O substantivo *TOM* não apresentou a maioria esmagadora de ocorrências de suporte - 72% dos casos analisados - por aparecer freqüentemente (37 em 113) no sentido de tom de voz, que não consideramos de suporte. Nesses casos, *TOM* ocorre como um complemento de um verbo *dicendi* ou outra palavra dessa classe semântica, tais como 'discurso' ou 'saudação'. Como exemplos, apresentamos *disse em TOM malicioso*, *murmurou em TOM queixoso*, *falou em TOM grave* e *um discurso em TOM contemporalizador*. Outros casos do sentido pleno do substantivo *TOM* inclui tons de cor, tais como *TOM vermelho*, *TOM laranja* e *TONS musicais*, tais como *TOM menor*. Nos casos de suporte, o adjetivo que segue o substantivo é denominal em 46% da ocorrência; nos casos plenos 12% são adjetivos denominais.

Finalmente, observamos a intercambialidade entre *CUNHO* e *NATUREZA*, com respeito aos seus sentidos de suporte. Todas as ocorrências de sintagmas nominais com *CUNHO* foram encontrados também com *NATUREZA*, como os pares.

cunho acadêmico	->	natureza acadêmica
cunho social	->	natureza social
cunho jornalismo	->	natureza jornalística

O corpus demonstra que o sub-grupo dos substantivos-suporte com função de essência, de uma maneira geral, comportam-se com grande regularidade em relação às propriedades funcionais exploradas na Seção 2: vagueza semântica, tipo de complementação e possibilidade de paráfrase. Os outros sub-grupos ainda estão sendo estudados.

## 5. Conclusões

Neste artigo caracterizamos os substantivos-suporte do português utilizando um conjunto de propriedades lexicográficas desta classe que podem ser utilizadas para reconhecê-la e analisá-la. O dicionário convencional utilizado deu suporte aos fatos lingüísticos encontrados no corpus, no que tange à complementação, freqüência, número de sentidos, número de expressões multi-vocabulares e intercambialidade. Essa caracterização pode ser utilizada como base para o tratamento automático de um subconjunto muito importante de sintagmas nominais.

O artigo apresenta a análise detalhada de 4 dos 35 substantivos-suporte preliminarmente extraídos do corpus. Propomos que tal estudo seja estendido para cobrir toda a classe, individualmente, para uma dada língua. Esta empreitada é viável dado o limitado tamanho dos grupos de substantivos envolvidos.

Devemos mencionar as dificuldades encontradas na obtenção de referências corroborativas para esse tópico de estudo, já que a própria denominação **substantivo-suporte** ainda não se encontra totalmente estabelecida na literatura relacionada.

### Referências

1. Aires, R., Aluísio, S. Criação de um corpus com 1,000,000 de palavras etiquetado morfossintaticamente. **Relatório Técnico n. MILC-TR-01-8**, Unicamp, Campinas, SP. 2001.
2. Basilio, M. Teoria Lexical. **Editora Ática**, Brasil. 1999.
3. Basilio, M., Ęamarski, L. Adjetivos denominais no português falado. In: A. Castilho (org.), **Gramática do Português Falado vol. V.**, Editora Unicamp, Brasil. 1995.
4. **Dicionário Eletrônico Houaiss da Língua Portuguesa**. Editora Objetiva, Brasil. 2001.
5. Lobato, L. Adjetivo: Tipologia e Interpretação Semântica. **Boletim da ABRALIN 14**. 1993.
6. Marques, M. H. D. Léxico de alta frequência na língua portuguesa. In **Flores verbais**, Editora 34. 1995.
7. Muresan, S., Tzoukermann E., Klavans, J., Combining Linguistic and Machine Learning Techniques for Email Summarization. In: **Proceedings of CoNLL-2001: The Fifth Workshop on Computational Language Learning**. 2001.
8. Mihatsch, W. Nouns are THINGS - new evidence for a grammatical metaphor. In: **8th International Linguistics Conference**, La Rioja. 2003.
9. Neves, M. H. M. Estudo das construções com verbo-suporte em português, In Koch, I. V., ed., **Gramática do Português Falado, vol VI: Desenvolvimentos**, Editora Unicamp, Brasil. 1996.
10. Neves, M. H. M. A delimitação das unidades lexicais: **o caso das construções com verbo-suporte**. *Palavra* 5. 1999.
11. Nirenburg, S., Raskin, V., **Ontological Semantics, MIT Press**, Cambridge, MA. 2004.
12. Panagiotidis, P., Empty Nouns. **Natural Language and Linguistic Theory 21(2)**: 381-432.2003.
13. RefAutores 00. Referência a trabalho dos próprios autores.
14. Sinclair, J. ed. COBUILD English Dictionary for Advanced Learners. **HarperCollins Publishers, 3a. Edição**. 2001.
15. Scher, A. P. Quais são as propriedades lexicais de uma construção com verbo leve? In Müller, A. L., Negrão, E. V. e Foltran M. J., eds., *Semântica Formal, Contexto*, São Paulo. 205-219. 2003.
16. Schmid, H., English Abstratct Nouns As Conceptual Shells: **From Corpus to Cognition**, Mouton de Ęruyter. 2000.