

Maria Carolina Ávila

**Propriedades semânticas e alternâncias
sintáticas do verbo: um exercício
exploratório de delimitação do
significado**

**Araraquara
2006**

Maria Carolina Ávila

Propriedades semânticas e alternâncias
sintáticas do verbo: um exercício
exploratório de delimitação do significado

Dissertação apresentada à Faculdade de Ciências e Letras, Universidade Estadual Paulista – Câmpus de Araraquara, como parte dos requisitos para a obtenção do título de Mestre em Letras (Área de Concentração: Lingüística e Língua Portuguesa).

Orientador: *Prof. Dr. Bento Carlos Dias da Silva*

Araraquara

2006

Ávila, Maria Carolina

Propriedades semânticas e alternâncias sintáticas do verbo,
um exercício exploratório de delimitação de significados / Maria
Carolina Ávila – 2006

114 f. ; 30 cm

Dissertação (Mestrado em Lingüística e Língua Portuguesa)
– Universidade Estadual Paulista, Faculdade de Ciências e
Letras, Campus de Araraquara.

Orientador: Bento Carlos Dias-Da-Silva

1. Verbos
2. Estrutura de Argumentos
3. Estrutura Conceitual
4. Wordnets

Maria Carolina Ávila

Propriedades semânticas e alternâncias
sintáticas do verbo: um exercício
exploratório de delimitação do significado

Data de aprovação: 20 de março de 2006

Membros da comissão examinadora:

Prof. Dr. Bento Carlos Dias da Silva

Universidade Estadual Paulista Júlio de Mesquita Filho/UNESP - Araraquara

Profª. Drª. Maria Helena de Moura Neves

Universidade Estadual Paulista Júlio de Mesquita Filho/UNESP - Araraquara

Profª. Drª. Rove Luiza de Oliveira Chishman

Universidade do Vale do Rio dos Sinos/Unisinos- São Leopoldo

Aos meus pais, Sônia e Antônio Carlos,
Aos meus avós, Conceição e João,
por tornarem tudo possível...

Agradecimentos

A UNESP/Ar, pelo apoio institucional.

A FAPESP, pelo apoio financeiro.

Ao CELiC-FCL, por proporcionar a necessária infraestrutura de trabalho.

Ao meu orientador Bento Carlos Dias da Silva, pela dedicação, pela amizade e por, tantas vezes, fazer tudo ficar mais simples em poucas palavras.

Aos meus pais, pelo suporte de toda sorte e por acreditarem em mim incondicionalmente.

Ao Luís, pela paciência e por proporcionar momentos tão agradáveis; meu descanso.

Às amigas de sempre e aos meus novos amigos, em especial à Andresa, à Tânia, à Mirna, ao Hélio e à Ana Eliza.

A Deus, sempre em meu caminho.

Resumo

A partir da hipótese de que a estrutura de argumentos projetada pelo verbo reflete aspectos da sua estrutura conceitual, esta dissertação investiga um conjunto de teorias que analisam essa interface sintaxe-semântica com o objetivo de recortar uma representação das dimensões sintática e semântica para essa classe lexical. Do ponto de vista lingüístico, analisam-se as propriedades léxico-semânticas e léxico-sintáticas de um conjunto de verbos do português do Brasil, extraído da base de verbos da rede WordNet.Br e nocionalmente correspondente à classe semântica dos verbos do inglês “Verbos de Destituição de Posse-Verbos do Tipo ‘Roubar’”, isolada por Levin (1993). As representações léxico-semântica e léxico-sintática fundamentam-se, respectivamente, na teoria sobre as *Estruturas Conceituais* de Jackendoff (1990, 2002) e na teoria sobre a *Estrutura de Argumentos* de Hale e Keyser (2002). Do ponto de vista lingüístico-computacional, desenvolvem-se uma estratégia de construção e refinamento dos synsets de verbos da rede WordNet.Br e uma proposta de representação formal das dimensões sintática e conceitual para os verbos.

Abstract

This thesis presents an inquiry on the lexical-syntactic and the lexical-semantic representations of verbs from the perspective that aspects of verb's argument structure reflect its conceptual structure. In the linguistic domain, the thesis investigates both the lexical-semantic and lexical-syntactic properties of the *synset* of Brazilian Portuguese verbs extracted from the WordNet.Br lexical database that corresponds to Levin's (1993) class of *Verbs of Possessional Deprivation-Steal/Rob Verbs*. The lexical-semantic and lexical-syntactic representations are grounded in Jackendoff's (1990, 2002) *Semantic Structures Theory* and Hale and Keyser's (2002) *Argument Structure Theory*, respectively. In the computational-linguistic domain, it presents both a strategy for constructing and refining the WordNet.Br verb synsets and a formal representation for describing the syntactic and conceptual dimensions of verbs.

Lista de figuras

Figura 1a. Estrutura-P dos verbos *inergativos*.

Figura 1b. Estrutura-P dos verbos *inacusativos*.

Figura 2. Estrutura de argumentos diádica projetada pelo verbo “quebrar”.

Figura 3. Estrutura de argumentos monádica projetada pelo verbo “engasgar”.

Figura 4. Estrutura diádica projetada pelo verbo “quebrar” na forma transitiva.

Figura 5. Estrutura diádica projetada pelo verbo "selar" na voz média.

Figura 6. Estrutura monádica projetada pelo verbo “fazer”.

Figura 7a. Estrutura diádica projetada pela preposição “em”.

Figura 7b. Projeção-P encaixada em uma Projeção-V.

Figura 8a. Estrutura projetada pelo verbo "tornar" na variante intransitiva.

Figura 8b. Estrutura projetada pelo verbo “tornar” na variante transitiva.

Figura 9 a-d. Tipos de estrutura de argumento lexical.

Figura 10a. Estágio anterior às Fusões.

Figura 10b. Fusões do adjetivo “vermelho” com os afixos **-a** e **-ar** e do verbo derivado V1 com a matriz verbal vazia V2, compondo a variante transitiva do verbo “avermelhar”.

Figura 11a. Estágio anterior às Fusões.

Figura 11b. Fusão do sufixo “en-“, em P, com N “garrafa” e posterior Fusão de “engarrafa”, em P, com o sufixo verbal “-ar”, originando o verbo “engarrafar”.

Figura 12. Ilustração da impossibilidade da Fusão de N (gado) com V (ar-).

Figura 13. Ilustração do vestígio deixado pela Fusão de P (en-) com V (-ar), originando o verbo “encurralar.”

Figura 14a. Cópia da assinatura-p do complemento no núcleo.

Figura 14b. Apagamento da assinatura-p do complemento.

Figura 15a. Combinação do rótulo do núcleo com o rótulo do complemento.

Figura 15b. Cópia simultânea da assinatura-p do complemento sobre a assinatura-p defectiva do núcleo.

Figura 16a. Estrutura da variante incoativa do verbo “engrossar”.

Figura 16b. Estrutura da variante causativa do verbo “engrossar”.

Figura 17. Representação da operação de Fusão nos verbos com argumento cognato.

Figura 18a. Ilustração do processo de Inserção Lexical no SV “tocar um samba”.

Figura 18b. Ilustração do processo de Inserção Lexical no SV “dançar um samba”

Figura 19. Representação da relação seletional entre o verbo e o seu complemento, indicada pelos índices {dança}.

Figura 20. O processo de Coindexação entre o verbo “embolsar” e o complemento nominal da preposição.

Figura 21. O processo de Incorporação no verbo ‘deadjetival’ “alargar”.

Figura 22a. Estrutura projetada por um verbo ‘denominal’ com argumentos dos tipos Local e Localizado.

Figura 22b. Estrutura projetada por um verbo ‘deadjetival’.

Figura 23 a. Variante intransitiva projetada pelo verbo “respingar”.

Figura 23b. Variante transitiva projetada pelo verbo “respingar”.

Figura 24. Variante intransitiva projetada pelo verbo “lambuzar”.

Figura 25a. Variante intransitiva projetada por um verbo do tipo “modo-paciente”.

Figura 25b. Variante transitiva projetada por um verbo do tipo “modo-agente” (com a indicação do índice de associação entre o item lexical e o traço semântico).

Figura 26. Representação da correspondência entre as estruturas conceitual e sintática da frase *John went into the room*.

Figura 27. Representação da estrutura conceitual da frase *Paulo embolsou o dinheiro*.

Figura 28. Estrutura de argumentos do verbo “roubar” do Synset_3742a.

Figura 29. Estrutura de argumentos do verbo “roubar” do Synset_3742b

Figura 30. Estrutura de argumentos do verbo “roubar” do Synset_3742c.

Figura 31. Estrutura de argumentos do verbo “roubar” do Synset_3742d.

Lista de quadros

Quadro1. Estatísticas da base da rede WordNet.Br

Sumário

Agradecimentos.....	5
Resumo	6
Abstract.....	7
Lista de figuras	8
Lista de quadros	9
Sumário	10
Seção 1 Introdução.....	11
1.1 <i>As redes wordnets: uma motivação.....</i>	<i>11</i>
1.2 <i>Hipótese e objetivos da dissertação</i>	<i>14</i>
1.3 <i>Os domínios e o alcance da pesquisa.....</i>	<i>16</i>
1.4 <i>Estrutura da dissertação</i>	<i>18</i>
Seção 2 A interface sintaxe-semântica.....	20
2.1 <i>Introdução ao estudo das relações entre a sintaxe e a semântica.....</i>	<i>20</i>
2.2 <i>A Inacusatividade: um pouco de teoria.....</i>	<i>32</i>
2.3 <i>Abordagens da Inacusatividade.....</i>	<i>38</i>
Seção 3 A representação lexical dos verbos.....	45
3.1 <i>Questões de representação: as dimensões léxico-semântica e léxico-sintática</i>	<i>45</i>
3.2 <i>A representação léxico-sintática.....</i>	<i>52</i>
3.3 <i>As operações de Fusão, Combinação, Incorporação e Inserção Lexical.....</i>	<i>61</i>
3.4 <i>A interação entre a estrutura de argumentos e a estrutura conceitual</i>	<i>76</i>
3.5 <i>A representação léxico-conceitual dos verbos.</i>	<i>81</i>
3.6 <i>A interdependência sintaxe-semântica: uma advertência</i>	<i>90</i>
Seção 4 A sintaxe e semântica dos verbos do tipo “roubar” da WordNet.Br	95
4.1 <i>Análise preliminar Synset 3742</i>	<i>95</i>
4.2 <i>Proposta de descrição da dimensão léxico-sintática.....</i>	<i>99</i>
4.3 <i>Proposta de descrição da dimensão léxico-semântica.....</i>	<i>103</i>
Seção 5 Considerações finais.....	106
Referências.....	109

Seção 1 Introdução

1.1 As redes wordnets: uma motivação

O formato das redes *wordnets* foi estabelecido com a construção da *WordNet de Princeton*, para o inglês norte-americano, que se iniciou em 1985. As wordnets são “redes de palavras”, computacionalmente implementadas, que armazenam unidades lexicais e expressões de uma língua natural (MILLER et al, 1993); são bases relacionais de dados lexicais que objetivam sistematizar as categorias lexicais (substantivos, verbos, adjetivos e advérbios) em função de quatro relações de sentido fundamentais: sinonímia, antonímia, meronímia e hiponímia. Dessa forma, buscam modelar o conhecimento léxico-conceitual que os falantes possuem de sua língua.

Do ponto de vista formal, uma wordnet estrutura-se em termos de *synsets* (*synonym sets*), isto é, de conjunto de sinônimos, formados por unidades lexicais pertencentes a uma mesma categoria sintática e que se inter-relacionam por manterem entre si uma semelhança de sentido. É importante ressaltar que a relação de sinonímia que une as unidades constitutivas do synset não é a de sinonímia absoluta, mas a de sinonímia relativa a um contexto, assim definida: duas expressões, A e B, são sinônimas se, e somente se, a substituição de A por B e de B por A em um contexto C não altera o valor de verdade da frase em que a substituição se verifica (MILLER; FELLBAUM, 1991). Em outras palavras, dois termos são sinônimos se existir pelo menos um contexto em que ambos puderem ser intersubstituíveis sem que haja alteração substancial do significado (CRUSE, 1986).

Conseqüentemente, posto que as unidades de um synset se agrupam por sinonímia, cada synset deve apontar, com maior ou menor precisão, para um único conceito (o conceito que é lexicalizado pelas unidades do synset). Além da relação léxico-semântica de sinonímia,

que se estabelece entre as unidades de um único synset, uma wordnet codifica também outras relações que se estabelecem entre synsets. Destacam-se a antonímia, outra relação léxico-semântica, e as relações de meronímia, hiponímia, acarretamento, causa e troponímia, que são relações de natureza lógico-conceitual.

Em termos gráficos, a relação de sinonímia, constitutiva do synset, é representada pelos nós na rede. As outras relações são representadas por arcos, que interligam esses nós. Dessa forma, o sentido de cada unidade lexical não é dado por definições, como em dicionários convencionais, mas ele se manifesta por meio das relações expressas pela rede. Precisamente por codificarem essas relações lexicais dessa maneira específica é que as wordnets se diferenciam dos dicionários eletrônicos convencionais (JURAFSKY; MARTIN, 2000).

As redes wordnets codificam ainda outros expedientes que facilitam a apreensão do sentido lexicalizado pelo synset: glosas e exemplos. As glosas são definições informais que explicitam o conceito implicitamente codificado no synset. Os exemplos, ou melhor dizendo, as frases-exemplo, são frases extraídas de córpus ou de dicionários e fornecem, portanto, os co-textos mínimos que exemplificam e atestam o uso de cada unidade do synset.

O modelo subjacente às wordnets apresenta três características marcantes. A primeira delas é a adoção do método diferencial de construção de léxicos. Esse método, que pressupõe o princípio de ativação de conceitos na mente do falante por meio de um conjunto de formas lexicais relacionadas pela sinonímia, possibilita a descrição de parte do significado das unidades lexicais por meio da especificação de relações de sentido entre elas, e não por meio de definições clássicas encontradas nos dicionários. A segunda traduz-se pelo fato de o modelo codificar computacionalmente a semelhança de sentido entre formas lexicais da língua por meio de conjuntos. A terceira característica do modelo é valer-se da noção de matriz lexical para representar a correspondência biunívoca que se estabelece entre o

significado das unidades lexicais (ou conceito lexicalizado), representado pelo synset, e as formas, isto é, as próprias unidades lexicais. Essa matriz lexical (FELLBAUM, 1998) pode ser graficamente representada em um plano cartesiano no qual, no eixo das abscissas, estão as formas lexicais (F1...Fn) e, no eixo das ordenadas, os conceitos (C1...Cn), representados pelos synsets.

Para o português do Brasil, a base da WordNet.Br está em fase de construção, empreendimento que tomou forma a partir da base do Thesaurus Eletrônico para o Português do Brasil (DIAS-DA-SILVA et al, 2000). Essa base contém, no estágio atual, cerca de 40 mil unidades lexicais do português, distribuídas entre aproximadamente 18 mil synsets (DIAS-DA-SILVA, OLIVEIRA, MORAES, 2002; DIAS-DA-SILVA, 2003), conforme registra o Quadro 1.

Categoria	Nº de unidades lexicais	Nº de synsets
Verbos	11.000	4.000
Nomes	17.000	8.000
Adjetivos	15.000	6.000
Advérbios	1.000	500
Total	44.000	18.500

Quadro1. Estatísticas da base da rede WordNet.Br

No processo de coleta e seleção das frases-exemplo, depara-se com problemas variados, como a dificuldade de identificação do conceito lexicalizado pelas unidades do synset; a constatação de synsets mal-formados, devido ao fato de as unidades não se conformarem ao significado dominante do synset e a verificação de synsets cujas unidades não projetam a mesma configuração sintática.

A reflexão sobre essas dificuldades, decorrentes da carência de técnicas precisas de análise lexical, de análises lexicais insuficientes ou equivocadas ou, até mesmo, de erros de

impressão ou de digitação que se detectaram nos dicionários dos quais muitas unidades foram extraídas, entre outros fatores, são o que motivam este trabalho.

Em particular, a variação na realização sintática dos argumentos dos verbos agrupados em um mesmo synset apontou para a investigação da interação entre as dimensões gramatical e conceitual dos verbos. No estudo desenvolvido por Levin e Rappaport-Hovav (1991) é que se buscam os subsídios para a estratégia de análise léxico-semântica explicitada na seção seguinte que serve de norte para a investigação mais aprofundada do aspecto da interface entre léxico e gramática.

Assim, o estudo relativo ao conhecimento envolvido na construção de uma wordnet, em particular da rede WordNet.Br, foi essencial para o desenvolvimento do trabalho que aqui se apresenta pois, em primeiro lugar, foi a atividade de análise de synsets e coleta de frases-exemplo, desenvolvida em estágio preliminar (ÁVILA, 2003; OLIVEIRA, ÁVILA, 2003a, 2003b), que motivou a investigação de divergências sintáticas de que os verbos podem participar; em segundo lugar, essa rede fornece milhares de synsets de verbos do português do Brasil disponíveis para análise fornece dados para análise (e que serão complementados com outros selecionados no cópús de referência da pesquisa), parte dos quais será selecionada conforme sua correspondência semântica com os verbos do inglês compilados por Levin (1993); em terceiro lugar, a base de verbos da WordNet.Br tem sido alvo de aplicação e verificação das teorias de Levin e Rappaport-Hovav (1991; 1995), de Levin (1993), de Hale e Keyser (2002) e de Jackendoff (1990; 2002), conforme será demonstrado ao longo desta dissertação.

1.2 Hipótese e objetivos da dissertação

O objetivo da dissertação é investigar um conjunto de propostas de análise da interface entre o léxico e a sintaxe de uma língua natural, a partir de estudos que consideram a **hipótese** de que aspectos da estrutura sintática dos argumentos do verbo (GRIMSHAW, 1994) são dados em função do seu significado (LEVIN; RAPPAPORT-HOVAV, 1991, 1995).

Assim, a partir do estudo realizado para os verbos do inglês (HALE; KEYSER, 2002; LEVIN; RAPPAPORT-HOVAV, 1991, 1995; JACKENDOFF, 1990, 2002), pretende-se investigar também as propriedades léxico-semânticas e sintáticas de um pequeno conjunto de verbos do português, o conjunto dos verbos intuitivamente relacionados ao verbo “roubar”, integrantes do Synset 3742 da WordNet.Br. Esse synset foi selecionado da base da rede WordNet.Br por conter um número significativo de verbos, o que exemplifica a dificuldade de apreensão do conceito representado pelo synset, por apresentar uma grande variação no comportamento sintático exibido pelos verbos e por se assemelhar semanticamente a uma classe discriminada por Levin (1993).

Levin analisa as alternâncias sintáticas de que o verbo (do inglês) participa, isto é, o modo alternativo de expressão dos seus argumentos, e, assim, o classifica em determinadas classes semanticamente coerentes. Do mesmo modo, para o português do Brasil, propõem-se a análise de um conjunto de alternâncias sintáticas de que os verbos, classificados como sinônimos de “roubar”, participam, a partir da coleta de ocorrências no corpúsculo de referência da dissertação, composto por três fontes digitais de informação lexical: o Corpus do Nilc¹, textos do português do Brasil disponíveis na Web e dela selecionados por meio do motor de busca Google e abonações registradas nos dicionários eletrônicos Michaelis (WEISZFLOG, 1998), Aurélio (FERREIRA, 1999), Houaiss (HOUAISS, 2001) e no DUP (BORBA, 2002).

Paralelamente aos estudos de Levin (1993) e Levin e Rappaport-Hovav (1991, 1995), este trabalho vale-se também dos estudos de Hale e Keyser (2002), que teorizam sobre a

¹ O Corpúsculo do Nilc, apresenta cerca de 40 milhões de ocorrências e é composto por textos escritos em português do Brasil, nos registros jornalísticos, didáticos e epistolar; <http://acdc.linguatca.pt/acesso/>

estrutura de argumentos dos itens lexicais, discutindo a descrição das configurações sintáticas projetadas pelos verbos, que são consideradas, nesta proposta, como complementar à teoria de Levin e Rappaport-Hovav.

Jackendoff (1990, 2002), por sua vez, discute a especificação formal dos componentes de significado relevantes para a projeção sintática, contribuindo para a proposta de representação da dimensão léxico-semântica dos verbos.

Além de investigar os aspectos da interface entre léxico, sintaxe e semântica, pretende-se também desenvolver uma estratégia de construção e refinamento das bases de verbos de redes wordnets, em particular, da rede WordNet.Br (DIAS-DA-SILVA, 2003a; DIAS-DA-SILVA et al, 2002). Nesse contexto, “refinar” significa estabelecer ou verificar a correspondência biunívoca que deve ser especificada entre o synset e o conceito por ele lexicalizado, relação essencial e estruturadora das wordnets.

Assim, esta pesquisa é relevante tanto do ponto de vista lingüístico, ao investigar as propriedades sintáticas e semânticas dos verbos, de modo a contribuir para a descrição dessa classe sintática, quanto do ponto de vista lingüístico-computacional, ao fornecer uma técnica de construção e refinamento da base de verbos de wordnets e, por extensão, de sistemas de PLN, conforme destacado na próxima subseção.

1.3 Os domínios e o alcance da pesquisa

Há que se destacar que este trabalho, do ponto de vista da metodologia global, desenvolve-se em um contexto interdisciplinar de pesquisa que busca relacionar conhecimentos construídos no âmbito da Lingüística e conhecimentos construídos no âmbito do PLN (Processamento Automático de Línguas Naturais). Adota-se, desse modo, a metodologia de pesquisa em lingüística computacional de Dias-da-Silva (1996, 1998, 2003a),

que fatora a investigação em dois **Domínios** de investigação distintos, mas complementares: o **Lingüístico** e o **Lingüístico-Computacional**, sendo que cada um desses domínios envolve um conjunto de atividades e competências específico à cada especialidade.

De acordo com essa metodologia, esta pesquisa, no Domínio Lingüístico, busca a sistematização do conhecimento lingüístico envolvido no processo de projeção sintática de aspectos da estrutura lexical e da estrutura conceitual dos verbos, isto é, dos componentes sintáticos e semântico-conceituais que integram a especificação dos verbos e as configurações sintáticas por eles projetadas (GRIMSHAW, 1994; LEVIN, 1993; LEVIN, RAPPAPORT-HOVAV, 1995; BAKER, KEYSER, 2002; TSUNODA, 1999; BROWN, MILLER, 1996). No Domínio Lingüístico-Computacional, por sua vez, o trabalho volta-se para a investigação das questões referentes à representação formal dos elementos trabalhados no Domínio Lingüístico, isto é, aborda-se a questão da representação formal das alternâncias sintáticas, da estrutura de argumentos e da estrutura conceitual dos verbos e da aplicação da análise das alternâncias no desenvolvimento e aprimoramento da base de verbos das wordnets, sobretudo porque essas bases podem ser integradas a bases lexicais de diversos tipos de sistemas de PLN como, por exemplo, sistemas de tradução automática, sumarizadores, sistemas de recuperação de informação e de dados, entre outros (VOSSEN, 1998; FELLBAUM, 1998; DIAS-DA-SILVA, OLIVEIRA, MORAES, 2002; DIAS-DA-SILVA, 2003; JURAFSKY, MARTIN, 2000; KORHONEN, 2002).

Assim, com o viés lingüístico, enfatiza-se que esta dissertação contribui para a sistematização de uma metodologia de análise léxico-conceitual e gramatical dos verbos do português, especialmente no que tange à questão de como isolar e descrever aspectos do significado dos predicadores (sobretudo verbos) que contribuem para a projeção das suas configurações sintáticas. Com o viés lingüístico-computacional, contribui para o desenvolvimento da pesquisa em PLN, visto que estuda o formato da base de verbos de redes

wordnets e, em especial, para o refinamento e o aprimoramento da base de verbos da rede WordNet.Br e da sua futura conexão com a rede *EuroWordNet* (VOSSSEN, 1998), projeto que inter-relaciona, por meio de um indexador particular, *ILLI*, *Inter-Lingual-Index* (um índice interlingüal que serve de elo de conexão entre as unidades lexicais de línguas diferentes que lexicalizam um mesmo conceito), as wordnets em construção para as línguas da União Européia e a *WordNet de Princeton*.

1.4 Estrutura da dissertação

Esta dissertação, além desta seção introdutória, em que se apresentaram a motivação do estudo (1.1), os objetivos e as hipóteses da investigação (1.2) e os domínios e o alcance da pesquisa (1.3), compõe-se de mais quatro seções.

A Seção 2, que trata da associação entre semântica e sintaxe, subdivide-se em três subseções: a subseção 2.1, que introduz a discussão particular das relações entre semântica e sintaxe baseada em Levin e Rappaport-Hovav (1991, 1993); a subseção 2.2, que focaliza a *Inacusatividade*, fenômeno cuja investigação, como se verá, possibilita a delimitação de componentes do significado sintaticamente relevantes para a sintaxe das unidades lexicais; e a subseção 2.3, que descreverá as abordagens teóricas para esse fenômeno.

A Seção 3, que trata da representação lexical dos verbos, subdivide-se em seis subseções: a subseção 3.1, que descreve um tipo de representação formal para os verbos de uma língua natural que contempla as dimensões léxico-semântica e léxico-sintática da informação lexical; a subseção 3.2, que sumariza uma proposta de representação léxico-gramatical para as categorias lexicais com base em Hale e Keyser (2002); a subseção 3.3, que tratará de operações que atuam na caracterização dessas representações sintáticas; a subseção 3.4, que demonstrará a relação entre a estrutura de argumentos e os componentes de

significado dos verbos; a subseção 3.5, que sumariza a proposta de representação léxico-conceitual para verbos com base em Jackendoff (1990, 2002); e a subseção 3.6, também embasada nesses trabalhos de Jackendoff, que faz uma breve avaliação sobre até que ponto a estrutura sintática de um verbo é determinada por seu significado.

A Seção 4, subdividida em três subseções, propõe um exercício exploratório de aplicação das três investigações teóricas descritas nas seções anteriores: na subseção 4.1, aplica-se a proposta de Levin e Rappaport-Hovav (1991) na análise do Synset 3742 da WordNet.Br e, nas subseções 4.2 e 4.3, exemplificam-se, respectivamente, as representações léxico-sintática e léxico-semântica para o verbo “roubar”.

Finalmente, a Seção 5 tece as considerações finais da dissertação. Nela, sumariza-se a contribuição das teorias estudadas para investigação exploratória da delimitação do significado e da projeção sintática de verbos e da aplicação dessa investigação na descrição lingüístico-computacional dos verbos de uma wordnet.

Seção 2 A interface sintaxe-semântica

2.1 Introdução ao estudo das relações entre a sintaxe e a semântica

Conforme já salientado na seção anterior, o estudo das relações entre semântica e sintaxe originou-se ao se detectar que, em alguns synsets da base da WordNet.Br, os verbos constituintes não projetavam a mesma configuração sintática. Desse modo, partiu-se para a investigação teórica sobre a correlação entre a estrutura conceitual e a configuração sintática da estrutura de argumentos dos verbos.

Orientando-se pela hipótese de que a configuração sintática do verbo reflete aspectos da sua estrutura conceitual, Levin (1993) propõe que verbos que se comportam de modo semelhante em função das alternâncias de diátese de que participam, isto é, que apresentam os mesmos modos alternativos de expressão sintática dos seus argumentos, às vezes acompanhados de mudanças sutis de significado, compartilham, também, algum aspecto de significado e formam uma classe semanticamente coerente.

Essa hipótese, no contexto das wordnets, implica supor que verbos de um mesmo synset, que estão agrupados em um mesmo conjunto por serem sinônimos, deverão, em princípio, projetar configurações sintáticas semelhantes. Não se descarta, entretanto, que existam outros fatores que interfiram na projeção sintática da estrutura de argumentos do verbo, como a interação da semântica do verbo com a semântica do próprio complemento (PUSTEJOVSKY, 1996), como será demonstrado mais adiante nesta seção. Ainda assim, entende-se, nesta dissertação, que as alternâncias de diátese constituem um caminho capaz de elucidar o significado dos verbos, apontando para diferenças e similaridades sutis que possam existir entre eles e que, de outra maneira, não seriam identificáveis.

Com o objetivo de verificar até que ponto o comportamento sintático do verbo é função do seu significado, examinam-se, de modo exploratório, os verbos do inglês *break*, *cut*, *hit* e *touch*.

De acordo com o que se propõe neste trabalho, a investigação sobre a correlação entre sintaxe e semântica parte de análises propostas para o inglês, mas sem se descuidar de análises paralelas que contribuam também para a descrição dos verbos do português. Desse modo, ao serem selecionados verbos do inglês, como os citados, por exemplo, analisam-se também os verbos do português “quebrar”, “cortar”, “golpear”/“espancar”/“bater” e “tocar”, correspondentes, respectivamente, a *break*, *cut*, *hit* e *touch*.²

Conforme mostra (1), tais verbos são transitivos, apresentando, portanto, dois argumentos sintáticos: o SN o sujeito e o SN objeto.

(1)

- a. *Margaret cut the bread.* (“Margaret cortou o pão.”)
- b. *Margaret broke the vase.* (“Margaret quebrou o vaso.”)
- c. *Terry touched the cat.* (“Terry tocou o gato.”)
- d. *Carla hit the door frame.* (“Carla golpeou o batente da porta (com algo).”)³

Esses verbos, porém, diferem quanto à participação da **alternância de voz média**. Como demonstra (2), *cut* e *break*, e os seus correspondentes em português, participam dessa alternância, que se caracteriza por denotar uma referência temporal genérica e um Agente não expresso, mas subentendido.

(2)

- a. *The bread cuts easily.* (“O pão corta facilmente.”)
- b. *Crystal vases break easily* (“Vasos de cristal quebram facilmente.”)
- c. **Cats touch easily.* (* “Gatos tocam facilmente.”)⁴

² As fontes de busca dos verbos correspondentes no português são os dicionários: The Random House (1996), Webster’s (1982), Borba (2002), Ferreira (1999), Houaiss (2001) e Weiszflog (1998).

³ O verbo “bater”, em “Carla bateu no batente da porta (com algo)”, paráfrase de (2b) não projeta a mesma estrutura sintática, pois o objeto é um SP_{EM} (sintagma preposicional).

⁴ O símbolo “*” indica agramaticalidade: as frases *Cats touch easily* e “Gatos tocam facilmente” são agramaticais no significado aqui pretendido (gatos são fáceis de serem tocados).

- d. **Door frames hit easily.* (* “Batentes de porta golpeiam facilmente.”)

Cut e *hit*, por sua vez, participam da **alternância conativa**, que se caracteriza por ser também uma alternância transitiva na qual o sujeito da frase transitiva e o sujeito da frase intransitiva, respectivamente, estabelecem a mesma relação agentiva com o verbo. O que se diferencia nessa alternância é a expressão do argumento objeto, que é um sintagma preposicional nucleado pela preposição do inglês *at*, conforme mostra os exemplos em (3). Nessa alternância, o verbo denota que a ação não foi finalizada.

(3)

- a. *Margaret cut at the bread.* (“Margaret quase cortou o pão.”)⁵
 b. **Margaret broke at the vase.*⁶ (“Margaret quase quebrou o vaso.”)
 c. *Terry touched at the cat.* (“Terry quase tocou o gato.”)
 d. *Carla hit at the door.* (“Carla quase bateu a porta.”)

A alternância examinada a seguir, que também pode ser observada para os verbos do português correspondentes, é a **alternância de ascensão do possuidor de partes do corpo** (*body-part possessor ascension*). Essa alternância, que se caracteriza pela mudança da expressão do argumento que expressa a parte do corpo, ocorre com os verbos *cut*, *hit* e *touch*, mas não com o verbo *break*, conforme ilustram os exemplos de (4), (5), (6) e (7).

(4)

- a. *Margaret cut Bill’s arm.* (“Margaret cortou o braço de Bill.”)⁷
 b. *Margaret cut Bill on the arm.* (“Margaret cortou Bill no braço.”)

(5)

- a. *Margaret broke Bill’s finger.* (“Margaret quebrou o dedo de Bill.”)
 b. **Margaret broke Bill on the finger.* (* “Margaret quebrou Bill no dedo.”)

(6)

- a. *Terry touched Bill’s shoulder.* (“Terry tocou o ombro de Bill.”)

⁵ Observe-se que essa alternância não se verifica no português. Do ponto de vista semântico, entretanto, essa frase corresponde às frases “Margaret tentou cortar o pão” e “Margaret quase cortou o pão”, conforme ilustram as traduções.

⁶ Essa frase, se fosse gramatical em inglês, denotaria a situação em que Margaret tentou quebrar o vaso.

⁷ (4a) pode também significar que Margaret amputou o braço de Bill.

b. *Terry touched Bill on the shoulder.* (“Terry tocou Bill no ombro.”)

(7)

a. *Carla hit Bill’s back.* (“Carla golpeou as costas de Bill.”)

b. *Carla hit Bill on the back.* (“Carla golpeou Bill nas costas.”)

Os tipos de comportamento exibido por esses verbos podem também ser observados em outros verbos semanticamente a eles relacionados, permitindo assim o estabelecimento de quatro classes: a classe dos verbos do tipo “*break*”, incluindo *crack* (“rachar”), *rip* (“rasgar”), *shatter* (“despedaçar”), *snap* (“trincar”), entre outros; a classe dos verbos do tipo “*cut*”, que incluem *hack* (“picar”), *saw* (“serrar”), *scratch* (“rasgar”), etc; a classe dos verbos do tipo “*touch*”, que são *pat* (“afagar”), *stroke* (“rebater”), *tickle* (“fazer cócegas”); a classe dos verbos do tipo “*hit*”: *bash* (“surrar”), *kick* (“chutar”), *pound* (“esmurrar”), *tap* (“bater de leve”), etc.

Cada classe de verbos caracteriza-se pelo fato de seus membros participarem das mesmas alternâncias de diátese, decorrendo assim, por hipótese, que os verbos de cada classe também compartilham aspectos do significado.

Ao analisar os componentes de significado dos verbos dessas classes, detecta-se, em primeiro lugar, que os verbos *cut*, *touch* e *hit*, que participam da **alternância de ascensão do possuidor de partes do corpo**, denotam contato, enquanto que o verbo *break*, do ponto de vista de sua manifestação lingüística, não necessariamente envolve essa noção. Como essa alternância está inerentemente relacionada a essa noção de ‘contato’, deve-se supor que os verbos que dela participam também devam possuir esse componente de significado.

Não obstante, deve haver outros componentes de significado relevantes para a descrição do significado dos verbos *touch*, *hit* e *cut*, visto que o primeiro deles não participa da **alternância conativa**, cf. o exemplo (3), como os outros dois. Sugere-se, então, que os verbos que participam dessa alternância exibem componentes de significado relacionados a

‘movimento’ e ‘contato’. *Touch* não possui o componente de ‘movimento’, apenas o de ‘contato’, *cut* e *hit* apresentam ambos e *break* não apresenta nenhum deles.

A **alternância de voz média** de que participam os verbos *cut* e *break*, cf. o exemplo (2), é característica dos verbos de mudança de estado. Decorre que *hit* e *touch*, por não participarem dessa alternância, não acarretam mudança de estado. *Cut* e *break*, por sua vez, compartilham outra propriedade: ambos estão associados a substantivos que denotam o resultado final do processo: *a break* (“uma quebra”) e *a cut* (“um corte”), respectivamente.

Por outro lado, os verbos *hit* e *touch* estão relacionados a substantivos que denotam ação apenas: *a hit* (“uma pancada”) e *a touch* (“um toque”), respectivamente. Essa característica reforça a hipótese de que *cut* e *break* são verbos de mudança de estado, pois o substantivo deles derivado denota o resultado da mudança de estado.

Já *cut* e *break* distinguem-se pelo fato de que o significado de *cut* envolve as noções de ‘contato’ e ‘movimento’, conforme já destacado nos parágrafos anteriores, e descreve mudança de estado por meio de movimento. Observe-se a discussão, agora transposta para os verbos correspondentes no português: o verbo “cortar” (*cut*) envolve o posicionamento de um objeto cortante em contato com uma superfície, promovendo uma mudança na integridade dessa superfície. O verbo “quebrar” (*break*), por sua vez, denota apenas mudança de estado nos usos intransitivo e transitivo, sem especificar como essa mudança ocorre. Por exemplo, há várias maneiras de se quebrar algo, deixando-o cair, atirando-se nele uma pedra, pisando-se sobre ele, entre outras.

Os verbos que estritamente denotam mudança de estado participam da **alternância causativo/incoativo**. De acordo com (8), o verbo *break* e o seu correspondente “quebrar” no português participam dessa alternância, demonstrando que esses verbos pertencem à classe dos verbos de mudança de estado. Por outro lado, *cut* e o seu correspondente “cortar” no português não participam dessa alternância, como mostra (9). Esse fato sugere que não se

trata de verbos de “mudança de estado pura”, mas de verbos que denotam “mudança de estado mediada por contato/movimento”.

(8)

- a. *The window broke.* (“A janela quebrou.”)
- b. *The little boy broke the window.* (“O garotinho quebrou a janela.”)

(9)

- a. *Margareth cut the string.* (“Margareth cortou o fio.”)
- b. **The string cut.* (* “O fio cortou.”)

Um verbo de “mudança de estado pura”, como *break*, é um verbo monoargumental que denota a entidade que sofre a mudança, conforme ilustra a forma incoativa (8a). A forma causativa (8b), com dois argumentos, realiza-se quando se introduz a noção de causa. Contrariamente, o *cut* inerentemente necessita da expressão do Instrumento e do Agente que o manipula. Trata-se, então, de um verbo que projeta dois argumentos e que não participa da **alternância causativo/incoativo**. Esses dois verbos, entretanto, participam da **alternância de voz média**, como verificado em (2), porque ela está relacionada tanto aos verbos de “mudança de estado pura” quanto aos verbos cujos significados especificam “mudança de estado mediada por contato/movimento”, como *cut*.

A partir dessa análise, conclui-se o seguinte: o verbo *touch* (“tocar”) é um verbo “estritamente de contato”; o verbo *hit* (“golpear”) é um verbo de “contato/movimento”; o verbo *cut* (“cortar”) é um verbo de “mudança de estado mediada por contato/ movimento”; e o verbo *break* (“quebrar”) é um verbo de “mudança de estado pura”.

Note-se que essa análise exploratória não isola todos os componentes de significado dos verbos examinados, somente aqueles que provavelmente são responsáveis pela participação dos verbos das alternâncias descritas. Assim, pode-se dizer que a **alternância de ascensão do possuidor de partes do corpo** é sensível ao componente ‘contato’; a **alternância conativa** é sensível aos componentes ‘contato’ e ‘movimento’; a **alternância**

causativo/incoativo é sensível aos verbos de ‘mudança de estado’; **alternância de voz média** é sensível tanto aos verbos de mudança de estado pura, quanto aos verbos de mudança de estado mediada por contato/movimento.

A seguir, apresenta-se mais um exercício exploratório de discriminação de classes de verbos que, do ponto de vista intuitivo e nocional, parecem constituir uma única classe. Trata-se do conjunto de verbos nocionalmente rotulados “verbos de remoção”, formado pelos verbos *clear* (“limpar /tirar /desobstruir /desatracar”), *wipe* (“limpar /esfregar”) e *remove* (“remover /tirar”) (LEVIN, RAPPAPORT-HOVAV, 1991).⁸

A avaliação intuitiva desses verbos do inglês leva à conclusão de que todos eles projetam a mesma configuração sintática: o argumento A1, tematicamente o Agente, projeta-se como Sujeito (Suj); o argumento A2, tematicamente o Localizado (ou o Tema), projeta-se como Objeto (Obj) e o argumento A3, tematicamente o Local (ou a Origem), como Complemento Obl_{FROM} (Oblíquo). Em outras palavras, essa avaliação intuitiva leva à conclusão de que esses verbos descrevem a remoção de um objeto físico, o Localizado (A2), de um determinado lugar, o Local (A3), sob o agenciamento de um Agente (A1), conforme ilustram os exemplos (10), (11) e (12).

(10) *Doug* (A1=Agente-Suj) *cleared the dishes* (A2=Localizado-Obj) *from the table* (A3=Local-Obl_{FROM}).

(“Doug *limpou/tirou/*desobstruiu/*desatracou os pratos da mesa.”)⁹

(11) *Kay* (A1=Agente-Suj) *wiped the fingerprints* (A2=Localizado-Obj) *from the counter* (A3=Local- Obl_{FROM}).

(“Kay limpou/*esfregou as impressões digitais do balcão.”)

⁸ Cumpre, neste ponto, advertir o leitor sobre o exercício: a transposição desse exercício para a investigação dos verbos do português potencialmente correspondentes a esses verbos será apenas parcialmente observada, posto que os verbos do inglês objetos da análise parecem lexicalizar certos componentes de significado (como ‘modo’ e ‘instrumento’, por exemplo) não lexicalizados nos verbos do português aos quais parcialmente correspondem. Assim, mesmo que a análise dos verbos do português não seja aqui explorada, tarefa complexa que merecerá atenção em trabalhos futuros, o que se pretende é subsidiar o leitor com mais evidências que motivem a busca de componentes de significado do verbo que sejam sintaticamente relevantes.

⁹ Neste exercício exploratório, parcialmente por razões apontadas na nota anterior, é interessante observar a frágil correspondência que se pode estabelecer entre os verbos das duas línguas em questão. O asterisco ‘*’ sinaliza as formas verbais do português que tornam a frase agramatical. As formas verbais do português grifadas participam da variante sob análise.

(12) *Monica* (A1=Agente-Suj) *removed the groceries* (A2=Localizado-Obj) *from the bag* (A3=Local-Obl_{FROM}).
 (“Mônica removeu/tirou os mantimentos da bolsa.”)

No entanto, a investigação mais acurada revela que esses verbos podem participar de diferentes alternâncias de diátese, cuja observação auxilia o analista na discriminação de variações sutis do significado lexical.

Nos exemplos (13), (14) e (15), a seguir, verifica-se que os verbos *clear* e *wipe*, mas não o verbo *remove*, participam da **variante local como objeto** (*location-as-object-variant*), isto é, o papel temático Local realiza-se na função de Obj. Nessa variante, o argumento A2 dos verbos *clear* e *wipe* não se realiza sintaticamente e o argumento A3 é deslocado da função de Obl_{FROM} para a função de Obj.¹⁰

(13) *Doug* (A1=Agente-Suj) *cleared the table* (A3=Local-Obj).
 (“Doug limpou/tirou/desobstruiu/desatravancou a mesa.”)

(14) *Kay* (A1=Agente-Suj) *wiped the counter* (A3=Local-Obj).
 (“Kay limpou/esfregou o balcão.”)

(15) **Monica* (A1=Agente-Suj) *removed the bag* (A3=Local-Obj).
 (“Mônica *removeu/*tirou a bolsa.”)¹¹

A análise análoga dos verbos *clear* e *wipe* também não se estende muito além no que se refere à alternância sintática de seus argumentos. Os exemplos (16) e (17) demonstram que o *clear*, mas não *wipe*, participa de um outro tipo de variante, a **variante of** (*of-variant*), em que o argumento A3, de modo análogo ao da alternância descrita no parágrafo anterior, é deslocado da função Obl_{OF} para a função Obj. Já o argumento A2 é deslocado da função Obj para a função Obl_{OF}, marcada pela preposição do inglês *of*.

¹⁰ É nessa variante que as autoras afirmam que os verbos parecem exibir o seu significado primário.

¹¹ O exemplo (15) é agramatical porque *remove* significa "tirar algo de algum local", e não "esvaziar, limpar um local", denotado pelos outros dois verbos, *clear* e *wipe*.

(16) *Doug* (A1=Agente-Suj) *cleared the table* (A3=Local-Obj) *of dishes* (A2=Localizado-Obl_{OF}).
 (“Doug *limpou/*tirou/*desobstruiu/*desatravancou a mesa de pratos.”)

(17) **Kay* (A1=Agente-Suj) *wiped the counter* (A3=Local-Obj) *of fingerprints* (A2=Localizado- Obl_{OF}).
 (“Kay *limpou/*esfregou o balcão de impressões digitais.”)

O fato de os argumentos dos verbos *clear*, *wipe* e *remove* participarem de variantes sintáticas distintas sugere que devam pertencer a classes semânticas também distintas, mesmo que eles compartilhem a noção geral de "remoção".

Assim, delimitam-se três classes: os verbos do tipo “*remove*” (*remove*-“remover /tirar”, *dislodge*-“desalojar”, *draw*-“sacar, arrancar, tirar”, *extract*-“extrair” e *withdraw*-“retirar”), que lexicalizam o processo de remoção; os verbos do tipo “*clear*” (*clear*-“limpar /tirar /desobstruir/desatracar”), *clean*-“limpar”; *empty*-“esvaziar”), que lexicalizam o estado resultante do processo de remoção, mas não como esse estado final é alcançado; os verbos do tipo “*wipe*” (*wipe*-“limpar /esfregar”, *buff*, *rush*, *erase*, *file*, *rub*, *scrape*, *sweep*, *vacuum*), que lexicalizam o modo ou o instrumento de remoção, caracterizando duas subclasses: verbos de ‘modo de remoção’ (*erase*- “limpar raspando ou esfregando”, *shave*-“limpar cortando”, *rub*-“limpar esfregando” e *scrape*-“limpar raspando”) e verbos de ‘instrumento de remoção’ (*brush*-“limpar com uma escova”, *mop*-“limpar com um esfregão”, *rake*-“limpar com um ancinho”, *vacuum*-“limpar com um aspirador de pó” e *buff*-“limpar com uma camurça”).

Quando *clear* e *wipe* participam da **variante local como objeto**, cf. exemplos (13) e (14), eles não são verdadeiramente verbos de remoção, mas verbos de mudança de estado: o Local é o argumento que muda de estado. A **alternância causativo/incoativo** de que *clear* participa, mas não *wipe*, e a **alternância conativa** de que *wipe* participa, mas não *clear*, sugerem que esses verbos devem pertencer a duas classes distintas: respectivamente, à classe

de “verbos de realização”¹², cf. o exemplo (18), e à classe de “verbos de atividade”, cf. exemplo (19).

- (18) *The strong winds cleared the skies.*
 “O vento forte limpou/*tirou!/desobstruiu!/desatravancou o céu.”¹³
 / *The skies cleared.* (Alternância causativo/incoativo)
 “O céu limpou/*tirou!/desobstruiu!/desatravancou.”
 / **The strong winds cleared at the skies.* (Alternância conativa)
 “O vento forte quase limpou/*tirou/desobstruiu/desatravancou o céu.”
- (19) *Kay wiped the counter.*
 “Kay limpou/esfregou o balcão.”
 / *Kay wiped at the counter.* (Alternância conativa)
 “Kay quase limpou/esfregou o balcão.”
 / **The counter wiped.* (Alternância causativo/incoativo)
 “O balcão *limpou/*esfregou”.

A partir dessa discussão, conclui-se que o que parece ser uma propriedade idiossincrática de um verbo, reflete, na verdade, sutis diferenças do seu significado. Isso pode explicar o fato de *clear* não ser propriamente um verbo de remoção, mas um verbo de mudança de estado. Outro fator importante é notar que o conhecimento de mundo, ou enciclopédico (HAIMAN, 1980), associado aos eventos denotados pelos verbos, não tem necessariamente que estar refletido na sua caracterização lingüística; é o que ocorre, por exemplo, com *break*, no conjunto dos verbos analisados, que não pressupõe, necessariamente, a noção de contato.

As correspondências analíticas parciais que se depreendem entre os exemplos do inglês e do português demonstram que o trabalho de Levin e Rappaport-Hovav pode ser estendido para a análise dos verbos de outras línguas. É também importante observar que, mesmo que as línguas não exibam o mesmo inventário de alternâncias sintáticas e de classes

¹² Esses termos foram propostos por Vendler (1957 apud LEVIN 1993), que classifica os verbos em quatro classes aspectuais: verbos de estado (por exemplo “Ela está calma”); verbos de atividade (por exemplo, “Ela corre todos os dias”), que expressam eventos dinâmicos; verbos de realização (*achievement*) (por exemplo, “As folhas avermelharam”), que denotam uma mudança de estado; verbos de finalização (*accomplishment*), que envolvem noção de causa e fim (por exemplo “O vaso quebrou”).

¹³ O símbolo ‘!’ indica estranheza.

verbais, pode-se verificar que há certa semelhança no que tange aos componentes do significado do verbo que são sintaticamente relevantes para sua realização sintática.

As variações entre as línguas, no que diz respeito ao mapeamento entre semântica e sintaxe, conforme já antecipado na nota 8, podem também, segundo Talmy (1985), ser atribuídas às diferenças nos padrões de lexicalização, definidos como generalizações sobre os tipos de significados associados aos verbos de uma língua. Diferentes verbos, freqüentemente, parecem compartilhar componentes de significado em um certo nível de abstração, diferindo um dos outros no que diz respeito a componentes de significado específicos. Assim, grande parcela da variação entre as línguas se deve à diferença nesses padrões de lexicalização e não aos conjuntos de componentes do significado que desempenham um papel na sintaxe. Por exemplo, itens lexicais que parecem ser equivalentes na tradução de duas línguas podem se diferenciar na expressão sintática dos seus argumentos devido ao fato de que os componentes de significado lexicalizados no significado central são divergentes nas duas línguas. A variação na associação entre semântica e sintaxe também pode ocorrer em função do grau de proeminência que as línguas atribuem a determinados componentes do significado no processo de expressão sintática dos argumentos dos verbos; ou seja, cada componente de significado de um determinado argumento pode não ter a mesma expressão sintática em diferentes línguas.

A argumentação desta subseção teve como meta dar contornos mais claros à hipótese de que o comportamento estrutural do verbo, particularmente no que diz respeito à expressão sintática dos seus argumentos é função do seu significado. Há, entretanto, uma breve consideração que deve ser feita para finalizá-la.

De acordo com Pustejovsky (1995), na investigação do significado lexical, além da semântica do verbo, a semântica dos seus complementos também deve ser considerada. O estudo da **alternância de apagamento de SN**, ou como nomeia Levin (1993), a **alternância**

de objeto não especificado, encaminha essa argumentação, conforme o exame dos exemplos (20)-(27) do português e os seus correspondentes no inglês.

(20)

- a. “A mulher comeu a refeição rapidamente.” (*The woman ate her meal quickly.*)
- b. “A mulher comeu rapidamente.” (*The woman ate quickly.*)

(21)

- a. “O cachorro devorou o biscoito.” (*The dog devoured the cookie.*)
- b. *“(O cachorro devorou.” (**The dog devoured.*)

(22)

- a. “John bebeu a cerveja rapidamente.” (*John drank his beer feverishly.*)
- b. “John bebeu rapidamente.” (*John drank feverishly.*)

(23)

- a. “John engoliu a cerveja rapidamente.” (*John gulped his beer feverishly.*)
- b. *“(John engoliu rapidamente.” (**John gulped feverishly.*)

O apagamento do objeto, permitido para os verbos “comer” e “beber” (*eat* e *drink*), pode ser explicado pelo fato de “comer” e “beber” denotarem uma atividade sem duração definida e “devorar” e “engolir” (*devour* e *gulp*) denotarem uma transição que implica complementação, ou seja são verbos ‘télicos’. No entanto, como podem ser explicados as ocorrências em que um mesmo verbo se comporta de maneira diferente em relação ao licenciamento do apagamento do complemento, como se verifica em (24)-(27)?

(24)

- a. “João deu um livro para Maria.” (*John gave a book to Mary.*)
- b. *“(John deu um livro.” (**John gave a book.*)

(25)

- a. “João deu uma palestra para a academia.” (*John gave a lecture to the academy.*)
- b. “João deu uma palestra.” (*John gave a lecture.*)

(26)

- a. “John mostrou um livro para Maria.” (*John showed a book to Mary.*)
- b. * “John mostrou um livro.” (**John showed a book.*)

(27)

- a. “Bill mostrou um filme para a platéia.” (*Bill showed a movie to the audience.*)
- b. “Bill mostrou um filme.” (*Bill showed a movie.*)

Segundo Pustejovsky (1995), o que torna possível o apagamento do argumento Meta em (25b) e em (27b) é a relação entre a semântica do verbo e a semântica do seu complemento, ou seja, a configuração sintática não é independente da informação trazida pelos complementos. Essas considerações são, portanto, um alerta para que se considere também a semântica do complemento na investigação das alternâncias sintáticas.

2.2 A Inacusatividade: um pouco de teoria

Com o intuito de embasar, do ponto de vista teórico, os exercícios da seção anterior e a argumentação que oferece suporte à hipótese de que o comportamento sintático exibido na estrutura de argumentos do verbo é, em certa medida, reflexo da estrutura conceitual, parte-se, nesta subseção, para o estudo da *Inacusatividade*, lançado pela *Hipótese da Inacusatividade* de Permuter (1978 apud LEVIN; RAPPAPORT-HOVAV, 1995) e retomado por Levin e Rappaport-Hovav (1995). Por ora, basta dizer que essa hipótese estipula que a classe dos verbos intransitivos deve ser subdividida e que a cada subclasse, como se discutirá mais adiante, associa-se a uma representação sintática particular. Desse modo, o estudo da *Inacusatividade* pode ser profícuo na busca dos componentes de significado que são sintaticamente relevantes.

Há que se recordar, neste ponto, que a suposição de que existem propriedades sintáticas do verbo que são determinadas pelo seu significado preocupa gramáticos e lingüistas desde há muito tempo. Os gramáticos, desde Pānini, já observavam essa tendência dos argumentos dos predicadores de estarem associados a posições sintáticas particulares, tendência que está materializada na “regra de preenchimento do sujeito” (*subjectivalization*) de Fillmore (1968), que estabelece como padrão a seguinte ordem de precedência no preenchimento do sujeito sintático: o Agente > o Instrumento > o Tema, em que o símbolo ‘>’ indica precedência.

As regras que efetuam essas associações (*linking*) entre os argumentos do verbo e as posições sintáticas em eles se realizam denominam-se “regras de associação” (*linking rules*). Na medida em que o papel semântico de um argumento é determinado pelo significado do verbo (ou do verbo e do seu complemento, quando se tratar do argumento Sujeito) que o seleciona, a existência dessas regras de associação reforça a suposição de que o significado do verbo é um fator relevante para a determinação da estrutura sintática das frases de que são núcleos, além de sugerir que as similaridades nas regularidades de associação verificadas em diferentes línguas sejam fortes indícios de que essas regularidades sejam parte da arquitetura universal das línguas naturais.

Embora a existência dessas regularidades de associação seja reconhecida, há ainda muito a ser investigado para que se possa desenvolver uma teoria que explique a correspondência entre a semântica lexical e a sintaxe. É necessário, primeiramente, determinar até que ponto a expressão sintática dos argumentos é previsível a partir do significado e regular entre as línguas; também é importante conhecer a natureza da representação léxico-conceitual, já que as regras de associação devem operar sobre tais representações, objetos de estudo das subseções 3.5 e 3.6

O encaminhamento de tais questões se torna cada vez mais urgente devido ao recente desenvolvimento de teorias sintáticas que utilizam as propriedades idiossincráticas dos verbos, particularmente os seus significados, para explicar as configurações sintáticas em que eles ocorrem. Merece destaque mencionar o *Princípio de Projeção*¹⁴ (CHOMSKY, 1981) e a “seleção-s”, isto é, a seleção semântica (*s-selection*), que determina a “seleção-c”, isto é, a seleção categorial (*c-selection*).

A seleção-s (CHOMSKY, 1986) é a propriedade seletional das entradas lexicais de categorias que podem ser núcleos lexicais, isto é, as categorias substantivo (N), verbo (V), adjetivo (A) e preposição (P). Essa propriedade é especificada na entrada lexical do verbo em termos de papéis temáticos: entrada do verbo "golpear", por exemplo, deve ter, como especificação seletional, que esse verbo seleciona um complemento Tema e um sujeito Agente. Já a seleção-c diz respeito às propriedades sintáticas das entradas lexicais de especificarem os tipos sintáticos dos seus argumentos (SN “sintagma nominal”, SD “sintagma determinante”,¹⁵ SP “sintagma preposicional” ou SV “sintagma verbal”, por exemplo). Chomsky (1986) sugere ainda que o papel temático apresenta realizações sintáticas canônicas: por exemplo, a realização sintática canônica do Tema é um SN na função de objeto direto.

Pode-se dizer que todos os modelos gerativistas dos últimos anos assumem que os contornos centrais da sintaxe das frases são projetados das propriedades lexicais dos verbos (e de outros predicadores); para que o *Princípio de Projeção* seja implementado, os verbos precisam ter representações lexicais cuja estrutura possa incorporar os aspectos estruturais potenciais que são relevantes para a projeção sintática da frase. Essas representações tomam a forma de uma estrutura de argumentos, que será estudada na subseções 3.2, 3.3 e 3.4. Daí o

¹⁴ O *Princípio de Projeção* estipula que a informação lexical deve estar sintaticamente representada. No contexto do Gerativismo Chomskiano, a estrutura lexical deve estar representada em todos os níveis de descrição sintática: Estrutura-P (representação das projeções sintáticas canônicas), Estrutura-S (representação das projeções sintáticas que sofreram “deslocamentos”), Forma Lógica (representação do significado nuclear da frase) e Forma Fonética (representação da realização fonética da frase).

¹⁵ SD é a abreviação de *Sintagma Determinante*, a projeção sintática de seqüências como [SD [D O] [N menino]].

desafio: compreender porque os verbos podem ocorrer em vários contextos sintáticos. Se esse tipo de variação se torna regra, ao invés de exceção, manter o *Princípio de Projeção* acarreta uma proliferação desordenada, e indesejada, de entradas lexicais para os verbos.

Há que se advertir, no entanto, que existem abordagens teóricas que rejeitam a existência de representações lexicais estruturadas que determinam a sintaxe da frase. Na abordagem construcional de Golberg (1996), por exemplo, o verbo é associado a algum conceito básico que não apresenta estrutura lingüística interna. Os argumentos são projetados livremente em estruturas sintáticas básicas disponíveis nas línguas; o significado exato do verbo em uma configuração sintática particular é dado pelo conceito ao qual ele está associado, em conjunção com o significado associado à estrutura sintática em que se encontra.

A *Hipótese da Inacusatividade*, como originariamente formulada por Perlmutter (1978 apud LEVIN; RAPPAPORT-HOVAV, 1995), como já se mencionou, é a hipótese sintática de que os verbos intransitivos subdividem-se em duas classes: os **verbos inacusativos** e os **verbos inergativos**. De acordo com Chomsky (1981), o verbo inergativo projeta, na Estrutura-P, ou seja, em um nível de representação sintática mais abstrato (“estrutura canônica”), apenas um argumento externo (o sujeito); já o verbo inacusativo projeta, nesse mesmo nível de representação, apenas um argumento interno (o objeto). Têm-se, assim, as configurações representadas em (28) para os dois tipos de verbos intransitivos.

(28).

Configuração sintática do verbo inergativo:	SN(Sujeito) [_{SV} V]
Configuração sintática do verbo inacusativo:	[_{SV} V SN (Objeto)]

Em outras palavras, assume-se que os verbos inergativos apresentam um sujeito subjacente, mas nenhum objeto; os verbos inacusativos, por outro lado, apresentam um objeto

subjacente, que se torna o sujeito da frase, mas nenhum sujeito subjacente (SPENCER, 1991), conforme ilustram as fig.1a e fig.1b, respectivamente.

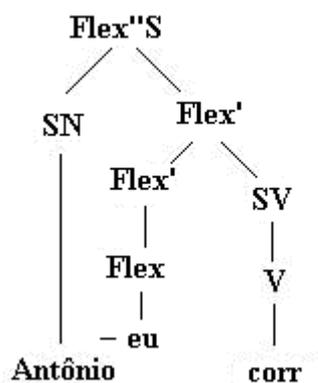


Figura 1a. Estrutura-P dos verbos *inergativos*.

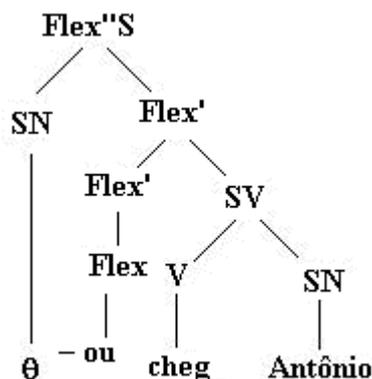


Figura 1b. Estrutura-P dos verbos *inacusativos*.

Os verbos inacusativos ainda têm a propriedade de não atribuir o caso Acusativo ao seu único argumento interno. Burzio (1986 apud LEVIN, 1995) estudou a correlação entre a habilidade de o verbo apresentar um argumento externo e a sua habilidade de atribuir caso acusativo, chegando à seguinte generalização, conhecida como *Generalização de Burzio*: se um verbo não atribui o caso Acusativo ao seu argumento interno (isto é, o seu complemento), ele também não atribui papel temático ao sujeito da frase, o que implica dizer que ele não projeta argumento externo algum. Em outras palavras, um verbo só atribui caso (estrutural) Acusativo quando projeta um argumento externo. Desse modo, o SN “Antônio”, na estrutura 1b, recebe o papel temático Tema, mas não recebe caso, a menos que se desloque para a posição de sujeito, onde receberá caso Nominativo da Flex (a flexão de tempo), condição essencial para a gramaticalidade da frase.¹⁶

Um verbo inacusativo é, portanto, aquele que não atribui papel temático ao sujeito da frase e nem caso ao seu objeto (argumento interno). Daí, de acordo com a definição original

¹⁶ A estrutura da frase “Chegou o Antônio”, segundo essa teoria, seria descrita como resultado de dois deslocamentos do SD “o Antônio”. O primeiro, descrito no texto, desloca esse SD para a posição φ da fig. 1b; o segundo desloca mais uma vez esse SP para a posição de adjunto de Flex’’=S.

de Perlmutter, um verbo inacusativo é aquele que apresenta apenas um argumento interno. Considerando essa definição, os verbos inacusativos são sintaticamente idênticos, na Estrutura-P, aos verbos na voz passiva, que apresentam apenas um argumento interno que precisa, em geral, deslocar-se para a posição de sujeito para receber o caso Nominativo de Flex, já que o verbo na voz passiva, como o inacusativo, possui a característica inerente de não atribuir caso algum ao seu objeto.

Já os verbos inergativos, como “correr” e “conversar”, são “responsáveis” pela denotação da ação, projetando apenas o argumento externo o Agente. Os inacusativos, contrariamente, apresentam sujeitos que não são Agentes, como “morrer”, “cair”, “chegar”, entre outros.

Com a introdução da *Hipótese da Inacusatividade*, muitos fenômenos em línguas diversas foram estudados com o propósito de distinguir verbos inacusativos de verbos inergativos. Assim, surgiram os *diagnósticos da inacusatividade*, isto é, testes para distinguir as duas classes de verbos. Deve-se ter cautela, no entanto, na eleição de um diagnóstico como discriminador da inacusatividade, pois nem todo teste que parece distinguir entre a inacusatividade e a inergatividade é um diagnóstico válido.

A pesquisa inicial sobre a inacusatividade concentrou-se no estabelecimento dos aspectos sintáticos da inacusatividade; pouca atenção foi dada à relação entre o significado dos verbos intransitivos e as classes semânticas as quais pertenceriam. Entretanto, há que se alertar que a *Hipótese da Inacusatividade* foi introduzida por Perlmutter no contexto da *Hipótese do Alinhamento Universal (Universal Alignment Hypothesis)*, para a qual a expressão sintática dos argumentos é passível de ser determinada pelo significado do verbo. De fato, a semelhança entre os verbos selecionados por diagnósticos inacusativos em várias línguas pode demonstrar que há traços semânticos importantes para a caracterização da inacusatividade.

2.3 Abordagens da *Inacusatividade*

Embora a Hipótese da Inacusatividade (LEVIN; RAPPAPORT-HOVAV, 1995) postule a existência de duas classes de verbos intransitivos, a realidade lingüística é mais complexa do que essa generalização. Há “divergências de inacusatividade” (*unnacusative mismatches*) entre os verbos que, segundo vários diagnósticos, seriam classificados ora como inacusativos ora como inergativos, considerando a semântica ou a sintaxe que exibem, e os verbos que realmente são isolados por esses diagnósticos.

Há dois casos de divergências de inacusatividade: uma que leva à abordagem sintática (ROSEN, 1984) e a outra que leva à abordagem semântica (VAN VALIN, 1990). Aquela não aceita que a inacusatividade seja semanticamente previsível e esta nega que a inacusatividade seja codificada sintaticamente. Adota-se, aqui, a posição de Levin e Rappaport-Hovav (1995), para quem a inacusatividade é sintaticamente codificada e semanticamente motivada.

Na abordagem sintática (ROSEN, 1984), os verbos inacusativos compartilham uma configuração **sintática** particular. Nessa abordagem, o importante é notar que não há uma única propriedade semântica comum a todos os verbos inacusativos selecionados por todos os diagnósticos inacusativos em todas as línguas, embora possa haver certa correspondência entre o significado do verbo e a sua classificação em inacusativo ou inergativo. Entretanto, dizer que essa classificação seja motivada pelo significado não implica dizer que todos os verbos inacusativos ou que todos os verbos inergativos representam uma classe semântica unificada.

Esse fato decorre de outro: a correspondência entre a semântica lexical e a sintaxe tem a característica de ser a de “muitos para um”, ou seja, vários significados convergem para uma mesma configuração sintática, o que conforma-se ao princípio da economia lingüística e da parcimônia teórica, de modo que não há razão para se afirmar que todos os verbos que

apresentam propriedades sintáticas atribuídas aos verbos inacusativos formarão uma classe semântica homogênea, como também não faz sentido afirmar que todos os verbos transitivos formam uma classe semântica única.

A favor da abordagem sintática há o fato de que verbos com significados semelhantes podem ser classificados diferentemente quanto à inacusatividade em línguas diversas. O verbo que corresponde a “morrer”, por exemplo, se comporta como inacusativo no português e como inergativo na língua choctaw e mesmo os verbos do português de “fenômenos físicos” divergem quanto aos seus comportamentos sintáticos: “roncar” apresenta características dos verbos inergativos e “enrubescer-se” comporta-se como inacusativo.

Essas divergências, porém, não são argumentos contra a *Hipótese da Inacusatividade*. O que deve ser levado em conta é o fato de alguns aspectos mais abstratos do significado do verbo serem relevantes para a caracterização semântica e outros aspectos não o serem. Assim, somente depois de isolarem-se esses aspectos relevantes do significado é que se pode avaliar a possibilidade de os verbos compartilharem a mesma classificação semântica.

Os verbos “enrubescer” e “roncar”, por exemplo, só representariam problema para a *Hipótese da Inacusatividade* se eles realmente fizessem parte da mesma classe semântica. No entanto, a noção de fenômeno físico pode não ser suficiente para definir uma classe: “roncar” pode ser descrito como um verbo de atividade, enquanto “enrubescer”, um verbo de atividade ou de mudança de estado (“tornar vermelho”). Atividade e mudança de estado são considerados aspectos do significado importantes na classificação dos verbos, como foi ilustrado na seção 2.1 para os verbos do tipo “*wipe*” e do tipo “*clear*”. Assim, para a devida classificação semântico-conceitual dos verbos, é imprescindível isolarem-se os aspectos do significado que têm relevância para a sintaxe e explicitar como esses componentes contribuem para a classificação (LEVIN; RAPPAPORT-HOVAV, 1995).

Um outro argumento levantado pelos defensores da abordagem sintática (ROSEN, 1984), que não se sustenta, é a existência de verbos que são classificados como inacusativos e inergativos pelo mesmo diagnóstico. Por exemplo, na expressão do passado, há verbos intransitivos do italiano que ocorrem com ambos os auxiliares *avere* (“haver”), sinal de inergatividade, e *essere* (“ser”), sinal de inacusatividade: *Mario ha continuato* (“Mario continuou”) e *Il dibattito è continuato* (“O debate continuou”). No entanto, a existência de verbos que selecionam auxiliares diferentes não é evidência de que a distinção entre as duas classes de verbos intransitivos não pode ser caracterizada pelo significado, pois tais verbos são apenas exemplos de verbos que apresentam uma classificação múltipla no que diz respeito à variedade de fenômenos sintáticos. Tais verbos são chamados de “verbos de comportamento variável” e cada um desses “comportamentos” está associado a um significado distinto. Conseqüentemente, a existência de verbos com classificação múltipla demonstra a importância da busca de componentes de significado sintaticamente relevantes; verbos de comportamento variado são evidências adicionais para o se estudar como aspectos do significado influenciam na representação sintática.

Em português, conforme salientam Campos e Xavier (1991), nem sempre são claros os testes de classificação dos verbos intransitivos em inergativos e inacusativos, pois nem sempre é possível isolarem-se fenômenos sintáticos bem definidos e que auxiliam o linguísta na distinção das duas classes, como há no italiano a seleção dos auxiliares *essere*, para os inacusativos, e *avere*, para os inergativos.

O que se observa no português, por exemplo, são os seguintes fatos sintáticos: (i) diferentemente dos verbos inacusativos, que ocorrem livremente em construções com o particípio (“falida a empresa” / “desmaiado o João” / “destruída a casa”), os verbos inergativos não ocorrem nesses contextos sintáticos (*“trabalhada a Maria” / * “corrido o João”); (ii) diferentemente dos verbos inacusativos, que ocorrem livremente em construções predicativas

(“A laranjeira está florida” / “A empresa está falida”), os verbos inergativos não ocorrem em tais configurações (*“O João está telefonado” / * “A Maria está cantada”).

Já a abordagem semântica (VAN VALIN, 1990), em contraposição à sintática, afirma que as duas classes de verbos intransitivos podem ser diferenciadas pelo significado e que a caracterização semântica das duas classes torna desnecessária a atribuição de representações sintáticas diferentes.

Como já se observou, os verbos inacusativos e os participípios passivos apresentam a mesma Estrutura-P e há vários fenômenos sintáticos e morfológicos que são comuns a esses verbos. Como exemplos, citam-se os sintagmas resultativos que, na Estrutura-S, podem ser predicados do sujeito da passiva dos verbos inacusativos, mas não dos inergativos ou dos transitivos; o participípio e o participípio passivo podem modificar a Estrutura-S do sujeito dos verbos na voz passiva (“uma carta mal escrita”) e dos inacusativos (“um livro recém lançado”), mas não podem modificar a estrutura dos inergativos (*“um advogado trabalhado duro”) nem a estrutura dos transitivos (*“um artista muito pintado”, no sentido de que o artista pinta muito).

A existência de tais fenômenos poderia ratificar a abordagem sintática, visto que os verbos inacusativos e os verbos na voz passiva aparecem na mesma configuração sintática, mas é difícil de encontrar uma propriedade semântica que seja compartilhada pelas duas classes de verbos.

Os proponentes da abordagem semântica (VAN VALIN, 1990), contudo, sugerem que o sujeito dos verbos inacusativos e o objeto dos verbos transitivos compartilham um mesmo componente semântico: o papel temático *Undergoer* (“aquele que sofre, ou que se submete a, a ação ou processo”), aquele que é o paciente, não existindo argumento que tenha o papel temático “Ator”. A noção de *undergoer* não equivale ao objeto direto, visto que o objeto de um verbo como “comer”, por exemplo, é um objeto direto, mas não um *undergoer*. No

entanto, a noção de *undergoer* não pode ser apenas semântica, pois a idéia de papel temático é sintático-semântica. A noção de *undergoer* pode ser caracterizada como uma generalização sobre um número específico de papéis semânticos que dão suporte a uma abordagem sintática.

A abordagem semântica para a inacusatividade considera também que o comportamento não homogêneo dos verbos intransitivos origina-se do fato de que certas construções são compatíveis com verbos com certos tipos de significados e outras são compatíveis com verbos com outro tipo de significado. A divergência na classe intransitiva, portanto, não se reduz a nenhum traço sintático dos verbos, mas é fruto da compatibilidade entre classes semânticas verbais com restrições semânticas que competem às próprias construções diagnósticas. Isso tornaria desnecessária a caracterização sintática. Cada construção é relacionada com sua própria restrição semântica, não havendo razão, assim, para esperar que todos os diagnósticos diferenciem os verbos intransitivos da mesma maneira. Por exemplo, uma construção pode distinguir verbos télicos de atélicos, outra pode distinguir entre verbos agentivos e não-agentivos, resultando no fato de que um verbo pode ser considerado inergativo por um diagnóstico e, inacusativo por outro. Dessa maneira, a abordagem semântica explica porque muitos diagnósticos identificam subclasses verbais semanticamente coerentes, enquanto permitem alguns tipos de divergências. Porém, o fato de que pode ser dada uma caracterização semântica para as classes de verbos não impede que sejam atribuídas propriedades sintáticas comuns aos verbos inacusativos.

Para os defensores da abordagem semântica (VAN VALIN, 1990), basta demonstrar a conexão entre a classificação semântica de um verbo e os diagnósticos, descartando a necessidade de se postular uma diferença sintática entre verbos inacusativos e inergativos. A hipótese central é que algumas construções, por virtude dos seus significados, selecionam verbos pertencentes a classes semânticas específicas.

Contudo, é necessário considerar que as propriedades semânticas de um verbo podem ser necessárias, mas não suficientes, para sinalizar a inacusatividade. Como exemplo, a escolha do auxiliar é um diagnóstico sensível à telicidade, que é uma propriedade semântica; os verbos télicos, aqueles que veiculam a idéia de terminalidade da ação ou processo e podem ser usados como indicadores de inacusatividade, selecionam o auxiliar *zijn*, no holandês, (equivalente ao “ser”, no português, e ao *be*, no inglês). No entanto, no holandês, uma frase que traz a noção de telicidade pode ser encontrada com o auxiliar *hebben* (equivalente ao “ter”, no português, e ao *have*, no inglês), indicador de inergatividade, se tal frase apresentar um verbo suporte ou uma expressão idiomática, como pode ser verificado nas frases do holandês *Het vliegtuig heeft een landing gemaakt.* (“*The plane has made a landing*”/ “O avião fez uma aterrissagem”).¹⁷ Desse modo, no holandês, é necessário que os verbos télicos selecionem o auxiliar *zijn* (“ser”, “be”), mas somente esse dado não é suficiente para determinar a escolha do auxiliar e sinalizar a inacusatividade.

Em resumo, a abordagem sintática (ROSEN, 1984) considera a inacusatividade um fenômeno unificado, ou seja, todos os verbos inacusativos, não importando a qual classe semântica pertençam, compartilham algumas propriedades sintáticas: a seleção de um argumento direto interno, a não seleção de um argumento externo e a incapacidade de atribuir caso Acusativo. Por outro lado, a abordagem semântica (VAN VALIN, 1990) considera a inacusatividade um fenômeno não-unificado, isto é, um verbo pode ser inacusativo de acordo com um diagnóstico e inergativo de acordo com outro.

A discussão sobre as abordagens semântica e sintática demonstra a importância de se estabelecerem considerações metodológicas ao se estudar a inacusatividade (LEVIN; RAPPAPORT-HOVAV, 1995). Esse fenômeno já é por si só complexo e há, ainda, questões

⁷ Em português, esse diagnóstico não é válido, pois não é selecionado nenhum verbo auxiliar.

sobre a semântica lexical e a sintaxe da língua sob investigação que precisam ser bem entendidas.

Em primeiro lugar, o diagnóstico usado para se estabelecer a inacusatividade ou a inergatividade de um verbo deve ser legítimo. Ou seja, é importante questionar e reavaliar os diagnósticos já aceitos. Um diagnóstico legítimo é aquele que testa uma propriedade sintática cuja explicação esteja ligada a uma configuração sintática inacusativa. A inacusatividade é, acima de tudo, uma propriedade sintática, mesmo que a consideremos semanticamente previsível; ela é usada para explorar a associação entre semântica lexical e sintaxe e por isso é necessário que se estabeleçam meios sintáticos para identificar verbos inacusativos, de modo que se possa também verificar a motivação semântica do fenômeno.

Outra consideração metodológica que deve ser feita para se encerrar esta seção é a respeito da escolha das línguas que estão sendo analisadas . Embora seja desejável a investigação da inacusatividade em línguas diferentes, é recomendável que se restrinja às línguas que sejam do domínio do pesquisador, devido à necessidade de julgamentos sutis quanto ao significado verbal que são necessários na identificação dos componentes do significado relevantes para a sintaxe, conhecimento que vai além daquilo que o pesquisador encontra em dicionários e gramáticas (LEVIN; RAPPAPORT-HOVAV, 1995).

Seção 3 A representação lexical dos verbos

3.1 Questões de representação: as dimensões léxico-semântica e léxico-sintática

O estudo dos trabalhos de Levin e Rappaport-Hovav (1991, 1993, 1995) sobre a natureza do conhecimento lexical permite que se conclua que a entrada lexical ideal deve minimizar a informação direcionada para a palavra, o que pode ser alcançado ao se fatorar a informação previsível, sendo descritas apenas as informações idiossincráticas. Se há propriedades sintáticas que derivam do significado, o significado do verbo, necessariamente idiossincrático, deve estar representado na entrada lexical. Porém, há características que são consideradas idiossincrasias e, por essa razão, são incluídas nas entradas lexicais das unidades do léxico sem a necessária avaliação da real necessidade de serem explicitadas, posto que essa avaliação pode sugerir que se trata de características que podem ser derivadas a partir do significado do verbo.

É nesse contexto de busca de estratégias de descrição das unidades do léxico que se exploram, nesta seção, algumas hipóteses sobre a estrutura do léxico e sobre a representação lexical, estudo importante para a identificação de aspectos do significado que influenciam na projeção sintática para o estabelecimento das regras de associação entre sintaxe e semântica e também para a investigação de duas propostas de representações: a proposta de representação sintática de Hale e Keyser (2002) e a proposta de representação semântica de Jackendoff (1990, 2002). Adota-se, assim, a descrição sugerida por Levin e Rappaport-Hovav (1995): a descrição da entrada lexical, além de outras, deve contemplar as dimensões léxico-sintática e léxico-semântica (ou léxico-conceitual).

Do ponto de vista sintático, a cada verbo está associada uma estrutura de argumentos que traz informações sobre suas propriedades de selecionar seus argumentos (LEVIN; RAPPAPORT-HOVAV, 1995). A representação léxico-sintática permite, então, que se distinga o argumento externo (o sujeito) dos argumentos internos (complementos), apontando se estes são diretos ou oblíquos (indiretos).

O argumento externo é representado na sintaxe externamente ao SV, que tem como núcleo o verbo. Os argumentos internos são projetados dentro do SV. O argumento interno direto é realizado como o argumento que é sintaticamente irmão do V (que é o núcleo do SV) e, portanto, regido por V. Neste trabalho, não se descrevem as estruturas de argumentos por meio de papéis temáticos, pois a estrutura de argumentos é aqui tomada como uma representação puramente sintática. A estrutura de argumentos do verbo “colocar”, por exemplo, é dada por: $e, x <y, P_{loc} z >$. Essa estrutura indica que o verbo “colocar” é triádico, isto é, apresenta valência três (x, y, z); as variáveis fora dos colchetes angulares $< >$ indicam, além do evento (e), o argumento externo (x); as variáveis dentro dos colchetes angulares, os argumentos internos (y, z), em que o primeiro é direto e o segundo é oblíquo, introduzido pela preposição locativa (P_{loc}).

A informação na estrutura de argumentos de um verbo, juntamente com o *Princípio de Projeção*, que estipula que as representações sintáticas são projetadas do léxico, como já mencionado, e o *Critério- θ* , que estipula que todos os papéis temáticos previstos na estrutura lexical devem estar projetados na sintaxe, determinam a configuração sintática do verbo. Desse modo, a relação entre a estrutura de argumentos e a representação sintática da Estrutura-P é trivial.

Como se parte da hipótese de que há aspectos do comportamento sintático dos verbos que são determinados por sua semântica, pelo fato de eles pertencerem a determinadas classes semânticas (LEVIN, 1993), suas propriedades sintáticas podem ser inferidas das propriedades

compartilhadas por todos os outros verbos da mesma classe semântica e não precisam ser aprendidas verbo a verbo. Qualquer representação léxico-semântica deve, pois, acomodar essas propriedades dos verbos, bem como as classes às quais eles pertencem.

Do ponto de vista da semântica, segundo Levin e Rappaport-Hovav (1995) a representação do significado do verbo pode essencialmente ser caracterizada de duas formas: uma em que são listados os papéis semânticos por ele projetados, em que o significado é reduzido a essa lista, e outra em que se estipula um conjunto fixo de predicados primitivos, que são depreendidos por meio da decomposição. Adota-se, aqui, a segunda forma, conforme será verificado na sessão 3.5.

Em uma primeira aproximação, a representação léxico-semântica em termos da decomposição semântica dos predicados envolve dois tipos básicos de elementos primitivos: um **conjunto fixo de predicados primitivos**, e um conjunto de **constantes**, que são escolhidas entre um número limitado de tipos semânticos. As constantes preenchem posições argumentais do predicado, podendo agir também como modificadores. Para a forma intransitiva (não causativa) do verbo "quebrar" ("A janela quebrou"), por exemplo, a decomposição de predicados resultaria em (29).

(29) quebr(ar) : [_Y TORNAR-SE QUEBRADO].

Nessa representação, TORNAR-SE é um predicado primitivo e QUEBRADO é a constante que representa o elemento do significado que discrimina o "estado de estar quebrado" de outros estados¹⁸.

Os predicados primitivos, que são empregados na descrição dos significados dos verbos, são selecionados de modo que os verbos pertencentes à mesma classe semântica

¹⁸ Se nos transportarmos à representação lexical proposta pelo modelo das wordnets, a constante poderia ser os próprios *LLIs*, *Inter-Lingual-Index*, que codificam os conceitos lexicalizados pelos synsets da rede.

tenham sub-estruturas conceituais comuns em suas decomposições; essa característica permite que os predicados primitivos sejam numericamente restritos, dado que o mesmo primitivo pode ser recorrente em estruturas diferentes. Por exemplo, todos os verbos de mudança de estado deveriam ter em comum, nesta discussão teórica preliminar, a sub-estrutura conceitual formada pelo predicado primitivo TORNAR-SE; a constante que preenche o seu segundo argumento especifica um estado, como em (30).

(30) Verbo não-causativo de mudança de estado: [y TORNAR-SE STATE]

Essa sub-estrutura é uma “estrutura semântico-lexical aberta” (*lexical semantic template*). Para verbos como “embolsar e “amanteigar”, a estrutura semântico-lexical aberta, seria a mesma e descreve todos os verbos do tipo de “colocar”, conforme ilustra (31).

(31) Verbos do tipo "colocar": [x CAUSA [y TORNAR-SE P_{LOC} z]]

Como ilustram (32) e (33), a diferença do significado desses dois verbos é descrita em termos das constantes e da posição que elas ocupam na estrutura conceitual.

(32) amanteig(ar): [x CAUSA [MANTEIGA TORNAR-SE P_{LOC} z]]

(33) embols(ar): [x CAUSA [y TORNAR-SE P_{LOC} BOLSO]]

Das estruturas (32) e (33) é possível extrair também as estruturas léxico-semânticas abertas (34) e (35), que podem ser características de outros verbos das mesmas classes de “amanteigar” e “embolsar”.

(34) [x CAUSA [COISA TORNAR-SE P_{LOC} z]]

(35) [x CAUSA [y TORNAR-SE P_{LOC} LOCAL]]

A constante, na representação léxico-semântica de um verbo pertencente a uma determinada classe semântica, é o que distingue esse verbo dos outros verbos da mesma classe. Essa constante também determina a subclasse a que o verbo pertence, no sentido de que certas constantes têm propriedades que as fazem compatíveis somente com certas posições em uma representação léxico-semântica aberta particular, devido à restrições seletivas das posições; tais posições, em (34) e (35), estão indicadas pelos rótulos COISA e LOCAL. Por exemplo, se a constante representa um objeto físico que serve como recipiente (um bolso) ela se encaixa na estrutura aberta (35). Por meio desse papel da constante, a forma fonológica que dá suporte ao significado do verbo pode ser prevista. Trata-se do “nome” ou “rótulo” do verbo, que é morfologicamente derivado, ou idêntico, ao “nome” da constante. Essa propriedade das constantes, de completar posições específicas na decomposição, é o que torna possível uma caracterização finita de possíveis significados dos verbos existentes em uma língua e, ao mesmo tempo, é o que permite a criação de novos verbos.

Estudar a maneira como a língua permite que as estruturas léxico-semânticas abertas sejam relacionadas às formas fonológicas é a chave para a compreensão dos verbos de comportamento variável. Às vezes, um significado está associado a várias formas fonológicas, dando origem à sinonímia; outras vezes, significados diferentes estão associados a uma única forma, originando a polissemia. A polissemia é de particular interesse para o estudo da inacusatividade, pois é a fonte dos verbos de comportamento variável, mas não será abordada nesta dissertação.

É relevante para os propósitos desta dissertação enfatizar-se que uma teoria sobre a representação léxico-semântica deve ser uma teoria gerativa que dê conta da caracterização dos significados possíveis. Os significados que são realizados precisam estar associados às formas fonológicas, isto é, aos “nomes” dos verbos, e essa associação é um processo arbitrário. No entanto, parece haver algumas restrições na determinação da escolha dos

“nomes” dos verbos quando duas representações léxico-semânticas compartilham o mesmo “nome”.

Como já foi dito, são as constantes que servem para identificar um membro particular de uma classe verbal. Muitas classes apresentam verbos que se diferem apenas na escolha das constantes e, nesses casos, a forma fonológica do verbo coincide com a constante. Isso, diferentemente do português, ocorre com muitos verbos do inglês, que apresentam derivação zero da forma do substantivo, ou seja, a forma verbal é a mesma forma do substantivo, sem nenhuma diferença morfológica, como *whistle* (“assobiar”/“assobio”)¹⁹ e *creak* (“chiar”/“chiado”); essas formas servem de constantes na estrutura léxico-semântica. Igualmente, muitos verbos deadjetivais de mudança de estado, como *dry* (“secar”), *warm* (“aquecer”) e *empty* (esvaziar) também não apresentam diferença da forma do adjetivo e é essa forma que nomeia o verbo e corresponde à constante na representação léxico-semântica. Assim, constroem-se as representações (36), (37) e (38), para os verbos não causativos de mudança de estado do inglês. As representações correspondentes para os verbos do português, que não apresentam derivação zero dos respectivos adjetivos, são dadas em (39), (40) e (41).

(36) *dry*: [y BECOME DRY]

(37) *empty*: [y BECOME EMPTY]

(38) *warm*; [y BECOME WARM]

(39) *sec(ar)*: [y TORNAR-SE SEC(O)]

(40) (es)*vazi(ar)*: [y TORNAR-SE VAZI(O)]

(41) (a)*quec(er)*; [y TORNAR-SE QUE(NTE)]

No português, há marcadores morfológicos explícitos que estão envolvidos na criação dos nomes dos verbos, como, por exemplo, em (40), em que os morfema **es-** e **-ar** combinam-

¹⁹ Note que em português a relação entre a forma do verbo e a forma do substantivo é sempre marcada por sufixação.

se com a constante VAZI(O) para formar o verbo “esvaziar”. Em inglês, naturalmente, há também esse tipo de derivação, como pode ser verificado em *defuzz a sweater* (“desfiar um casaco”), em que o morfema *de-* opera na formação do nome do verbo, cujo significado especifica a remoção de uma entidade de um local. É importante, contudo, notar que nem todos os verbos desse tipo são formados com o morfema *de-* no inglês e com o prefixo *des-* no português. Pode haver conversão de certos nomes, quando há uma relação inalienável de posse entre a entidade que denota local e a que denota a entidade localizada, como em *bone* (“desossar”) ou *core* (tirar o caroço de uma fruta). A variante “cascar” do português pode também ilustrar esse tipo de derivação.

Uma única constante pode estar relacionada a mais de uma estrutura léxico-semântica aberta e, visto que os verbos têm seus nomes derivados dessas constantes, mais de uma representação léxico-semântica terá o mesmo nome. Essa associação de um único nome a várias representações é a origem de verbos polissêmicos (os verbos de comportamento variável). O verbo denominal *shovel* (“trabalhar com a pá”) serve como ilustração desse comportamento: ele pode denotar tanto a ação de adicionar uma entidade a um local com uma pá ou a ação de remover uma entidade de um local com uma pá (*shovel the snow off the walk* /“retirar a neve da passagem com uma pá”; *shovel the sand into the truck* /“colocar a areia dentro do caminhão com uma pá”). Conseqüentemente, esse verbo pode fazer parte de duas classes semânticas opostas.

São os verbos que compartilham a mesma estrutura léxico-semântica aberta, mas cujas constantes se diferem, que formam classes cujos membros apresentam a mesma expressão dos argumentos: todos os verbos do tipo “*shovel*”, quando denotam remoção, devem compartilhar a mesma estrutura léxico-semântica aberta. Por outro lado, verbos que apresentam a mesma representação léxico-semântica aberta e que compartilham a mesma constante deverão estar associados à configurações de argumentos distintas, que refletem os diferentes significados do

verbo. Por exemplo, *shovel* (“rastelar”) poderá ter uma estrutura de argumentos quando é verbo de preenchimento e outra diferente quando é verbo de remoção.

Essas considerações preliminares sobre a representação lexical dos verbos conduzem à investigação dos aspectos sintáticos e à investigação dos aspectos semânticos. A primeira, que estuda a codificação das propriedades estruturais das unidades lexicais que são relevantes para a seleção e para a projeção sintática dos argumentos, será discutida, nas subseções 3.2, 3.3 e 3.4, em termos da estrutura de argumentos essencialmente proposta por Hale e Keyser (2002). A segunda, que estuda a codificação dos aspectos do significado da unidade lexical que refletem na sintaxe, será abordada, nas subseções 3.5 e 3.6, em termos da estrutura conceitual proposta por Jackendoff (1990, 2002).

3.2 A representação léxico-sintática

A hipótese analisada no trabalho de Hale e Keyser (2002) é a de que a estrutura sintática dos itens lexicais pode ser descrita em termos de em um pequeno número de constructos e princípios, a saber: um item lexical projeta uma configuração estrutural binária definida por apenas duas relações: relação de complementação e relação de especificação.

O termo “estrutura de argumentos” é definido como a configuração estrutural projetada pelo item lexical. É, portanto, o sistema engendrado pelos dois tipos de relação estrutural nomeados no parágrafo anterior, relações que se estabelecem entre o predicador núcleo e seus argumentos.

Ao se analisarem as frases (i) “O vaso quebrou”, que corresponde à frase do inglês *The vase broke*, e (ii) “O motor engasgou”, que corresponde à frase do inglês *The engine caughed*, observa-se que o verbo da primeira participa da **alternância causativa/incoativa**

(descrita na subseção 2.1), enquanto que o verbo da segunda frase, não: “Eu quebrei o vaso” (*I broke the vase*)/ * “Eu engasguei o motor” (**I coughed the engine*). As frases (i) e (ii) estruturam-se em termos do *frame* canônico dos verbos intransitivos (SD V), mas elas não são isomorfas. Essa diferença decorre, dentre outros, de fatores estruturais. A estrutura de argumentos do verbo “quebrar”, ilustrada esquematicamente na fig.2.,²⁰ consiste de dois componentes: um componente raiz (R), que contém os traços semânticos e fonológicos associados ao verbo "quebrar", e um componente verbal (V). Este componente, por razões categoriais, exige um complemento, que é o elemento R. O componente R deste verbo, por razões léxico-gramaticais, exige a projeção de um especificador, o SD.

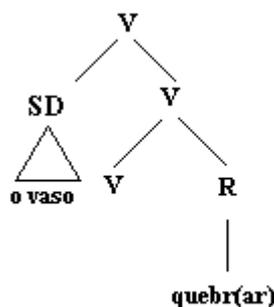


Figura 2. Estrutura de argumentos diádica projetada pelo verbo “quebrar”.

A estrutura de argumentos do verbo “engasgar”, ilustrada na fig.3, também apresenta dois componentes, R e V. Entretanto, o componente R veta a projeção de um especificador. Assim, * “Eu engasguei o motor” é uma frase agramatical porque o verbo “engasgar” não projeta uma estrutura sintática transitiva, pois, na estrutura de argumentos do verbo "engasgar", o núcleo verbal, como já foi sinalizado, por razões categoriais, não projeta um especificador e a natureza do complemento R também veta a projeção de um especificador. Essas restrições impedem a transitivização do verbo e, conseqüentemente, inviabilizam a frase

²⁰ Embora os autores utilizem o mesmo tipo de representação arbórea empregada para a descrição da dimensão sintagmática da frase, é essencial esclarecer ao leitor que as “árvores” não devem ser entendidas como projeções sintáticas, mas como esquemas que representam/descrevem, de modo gráfico, as relações ou propriedades abstratas que caracterizam a estrutura de argumentos dos itens lexicais.

em questão. A propriedade específica dessa raiz R é decorrente de ela ser um núcleo lexical de uma projeção nominal ("engasgo") que, por razões categoriais, não projeta um especificador. Já um núcleo lexical de uma projeção de adjetivos, como se mostrará mais adiante, exigiria a projeção de um especificador.

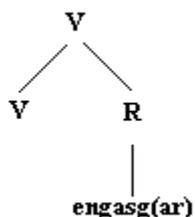


Figura 3. Estrutura de argumentos monádica projetada pelo verbo “engasgar”.

A **alternância causativa** (“Eu quebrei o vaso”), da qual “quebrar” participa, isto é, a transitivização do verbo "quebrar", é resultado da operação de Combinação (*Merge*)²¹ (RADFORD, 1997), que combina a estrutura representada na fig.3 com um núcleo verbal V1, como ilustra a fig.4.

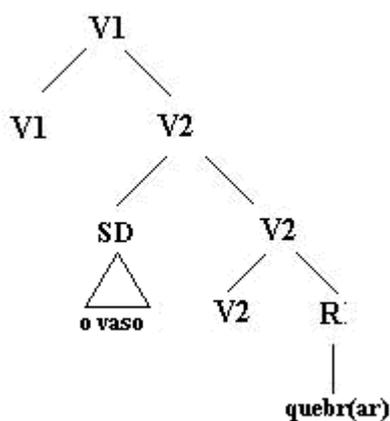


Figura 4. Estrutura diádica projetada por “quebrar” na forma transitiva.

²¹ Essa operação será explicada na subseção seguinte.

A frase (iii) “Ele selou esse cavalo”, correspondente à frase do inglês *He saddled this horse* e a frase (iv) “Ele fez gracinha”, correspondente a *He made a fuss*, também parecem possuir a mesma estrutura sintática. Entretanto, elas divergem quanto à participação na **alternância de voz média**, apresentada na subseção 2.1, pois a frase “Esse cavalo sela facilmente” (*This horse saddles easily*) é gramatical, mas * “Gracinha faz facilmente” (**A fuss makes easily*) é agramatical. A propriedade gramatical da construção média é cancelar a propriedade do verbo de atribuir caso Acusativo ao seu complemento, forçando-o a se deslocar para a posição do sujeito sintático, para que receba o caso Nominativo de Flex. É esse deslocamento que impede que haja um argumento sintático externo ao SV.

De acordo com Hale e Keyser (2002, p.4), a participação dos verbos na **alternância de voz média** também se deve a fatores estruturais: um verbo pode participar da voz média quando seu complemento exigir um especificador, formando uma projeção diádica, ou seja, uma projeção que exiba complemento e especificador, como demonstrado na fig.5, em que o complemento N exige um especificador SD.

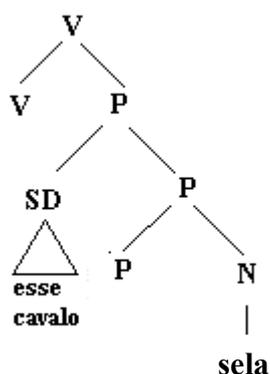


Figura 5. Estrutura diádica projetada pelo verbo "selar" na voz média.

Como mostra a fig.5, o núcleo dessa construção é uma preposição, diferentemente da estrutura da fig.4, cujo núcleo é um verbo, pois o verbo “selar” incorpora a preposição “em”:

“pôr sela em”. A frase (iv), por sua vez, não participa da **alternância de voz média** porque a estrutura de argumentos do verbo “fazer” não projeta um especificador, conforme ilustra a fig.6.

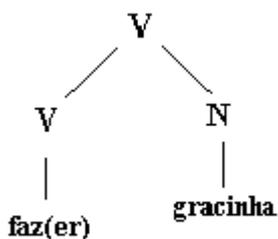


Figura 6. Estrutura monádica projetada pelo verbo “fazer”.

A projeção do verbo “fazer”, na fig.6, é uma projeção lexical do tipo *monádico*, em que o verbo projeta apenas o complemento N. O complemento (“gracinha”) é estruturalmente o único nó irmão do núcleo. O termo *monádico* diz respeito aos argumentos que aparecem internamente à configuração do item lexical. O sujeito frasal, que é o argumento externo à essa configuração, não é considerado na nomeação das estruturas em *monádica*, *diádica* ou *ternária*.

A projeção da estrutura de argumentos do verbo “fazer” diferencia-se da projeção de estruturas cujos núcleos são preposições, como “pôr os livros na estante”. Como ilustra a fig. 7a, a estrutura de argumentos do exemplo tem como núcleo a preposição “em”, e exige os dois componentes previstos na teoria: o complemento SD (“a estante”) e o especificador SD (“os livros”).

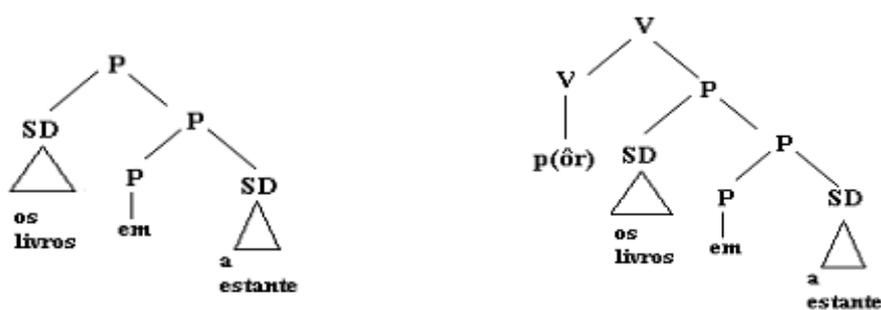


Figura 7a. Estrutura diádica projetada por “em”. **Figura 7b.** Projeção-P encaixada em uma Projeção-V.

A configuração na fig.(7a) é diádica, ou seja, a estrutura de argumentos definida pelo núcleo P projeta duas posições de argumentos que são internas: o complemento é irmão do núcleo, e o especificador é o único irmão da primeira projeção do núcleo.

A estrutura argumental de “pôr”, na fig.7b, que é monádica, exige como seu único argumento interno uma Projeção-P, diádica, encaixada como complemento dentro da Projeção-V. O especificador da Projeção-P projeta-se, na voz ativa, como objeto gramatical na sintaxe.

Assim, a propriedade categorial das preposições é projetar uma estrutura que contém um complemento e um especificador; hipótese que é reforçada pelo fato de as preposições especificarem uma relação (temporal, espacial ou outra qualquer) entre duas entidades, dois eventos ou duas circunstâncias. O núcleo P e o seu complemento SD/SN se combinam para formar uma espécie de predicação, que exige um "sujeito" sobre o qual essa predicação deve ser feita. Assim, a existência do especificador e do complemento é consequência da natureza da preposição. A estrutura diádica nesses casos é denominada “estrutura diádica básica”.

Há, ainda, outro tipo de estrutura de argumentos que também apresenta um argumento especificador interno, mas que se diferencia do tipo apresentado nas figs.7a e 7b. No português e inglês, há construções como a listada no exemplo (42).

(42)

(a) “As folhas tornaram-se vermelhas.”

(b) “O frio tornou as folhas vermelhas.”

O núcleo verbal de frases como as exemplificadas em (44) também requer um especificador (“as folhas”) e um complemento (“vermelhas”), conforme ilustração na fig. 8a.

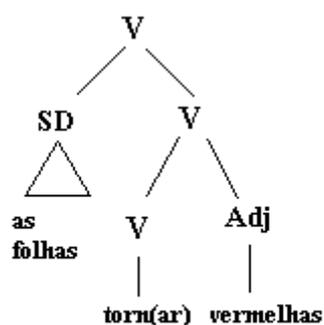


Figura 8a. Estrutura projetada por “tornar” na variante intransitiva.

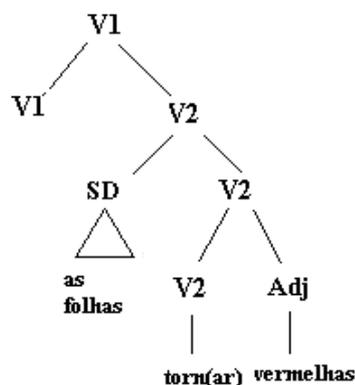


Figura 8b. Estrutura projetada por “tornar” na variante transitiva.

A variante intransitiva (42a) realiza-se na estrutura da fig. 8a e a variante transitiva (42b) na estrutura da fig. 8b²². Na estrutura da fig. 6b, o V1 é um núcleo monádico que exige V2 como seu complemento; este, por sua vez, corresponde à estrutura diádica da fig. 8a.

Nas estruturas correspondentes às figs. 8a e 8b, no entanto, não é o núcleo verbal V (“tornar”), por si mesmo, que projeta o especificador SD (“as folhas”), mas é o complemento Adj (o adjetivo “vermelh(as)”), que exige essa projeção. Isso, como já se mencionou, é

²² Pode ser traçado um paralelo entre as estruturas transitivas dos verbos “pôr” e “tornar” (“pôr os livros na estante” / “o frio torna as folhas vermelhas”). Note-se que, na estrutura de “pôr” (fig. 7b), o V1 está preenchido, enquanto que na estrutura de “tornar” (fig. 8b), o V1 é uma matriz vazia. Para demonstrar como esse V1 vazio é preenchido, é necessário recorrer às operações de Fusão e Combinação, que serão explicadas na sub-subseção 3.21.

característica dos adjetivos e, em construções como essas, o núcleo verbal fornece a estrutura para alojar o especificador, conforme a exigência do complemento adjetivo. Essa estrutura de argumentos é formada por duas estruturas monádicas: uma formada pela projeção do complemento pelo núcleo verbal e a outra formada pela projeção "forçada" do especificador pelo complemento adjetivo. Por esse motivo, são denominadas "estruturas binárias compostas", nas quais o adjetivo satisfaz a necessidade de complementação do verbo e o verbo, por sua vez, disponibiliza a posição de especificador exigida pelo adjetivo.

A partir dessa discussão, observa-se que as duas relações fundamentais que caracterizam a estrutura de argumentos dos itens lexicais são formalmente definidas em (43a) e (43b).

(43)

- a. *Relação 1*: Relação "Núcleo-Complemento"; se X é complemento de um núcleo Nuc, então X é o único irmão de Nuc.
- b. *Relação 2*: Relação "Especificador-Núcleo"; se X é o especificador de um núcleo Nuc, e se P1 é a primeira projeção de Nuc, então X é o único irmão de P1.

Com essas duas relações, quatro tipos fundamentais de configurações de estrutura de argumentos podem ser logicamente previstos. No primeiro tipo, fig. 9a, o núcleo não projeta nem complemento e nem especificador, dando origem ao tipo mais simples de estrutura; no segundo tipo, fig. 9b, o núcleo projeta um complemento, mas não um especificador, dando origem à estrutura monádica. No terceiro tipo, fig. 9c, o núcleo projeta tanto um complemento, quanto um especificador na estrutura binária básica. Finalmente, no quarto tipo, fig. 9d, o núcleo exige um especificador, mas exclui o complemento. Essa estrutura só pode ser gerada por composição: o núcleo que possui essa propriedade deve aparecer como complemento de outro núcleo, Núcleo*.

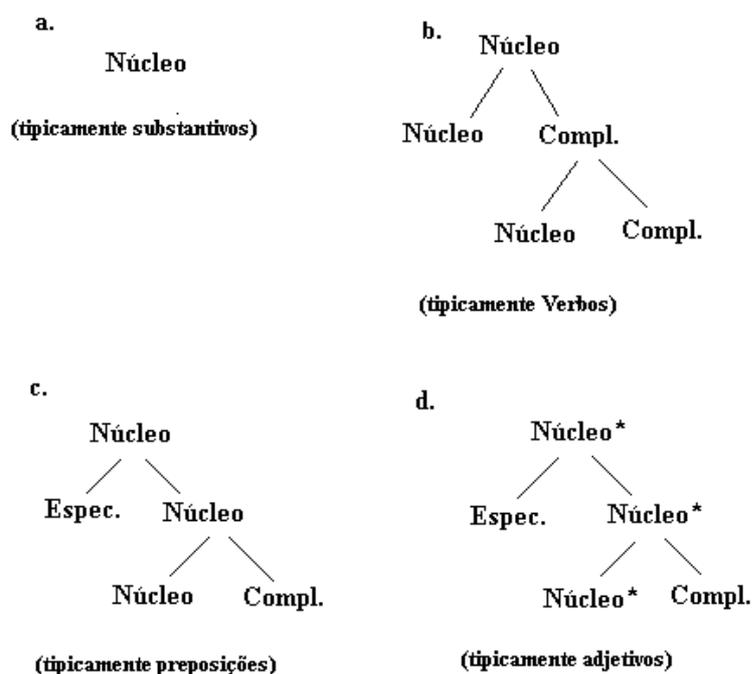


Figura 9 a-d. Tipos de estrutura de argumento lexical.

No inglês e no português, a realização típica da categoria morfossintática dos núcleos das estruturas ilustradas nas figs. (9a), (9b), (9c) e (9d) é, respectivamente, substantivo, verbo, preposição e adjetivo. Porém isso não pode ser afirmado para todas as línguas. É importante reafirmar que os esquemas ilustrados acima não são projeções sintáticas, mas projeções da estrutura de argumentos no nível lexical.

Três operações estão envolvidas na configuração das estruturas de argumentos dos verbos: a Combinação (*Merge*) (RADFORD, 1997), a Fusão (*Conflation*) (HALE; KEYSER, 2002) e a Incorporação (*Incorporation*) (BAKER, 1988). Na subseção que se segue, detalham-se essas operações e focaliza-se a relevância delas na descrição da dimensão léxico-sintática dos verbos.

3.3 As operações de Fusão, Combinação, Incorporação e Inserção Lexical

A **Fusão** parece ser, até o momento, uma operação importante para o estudo das estruturas de argumentos. Ela consiste na junção de núcleos sintáticos, em que a matriz fonológica do núcleo de um complemento é inserida no núcleo, vazio ou afixional, que o rege, dando origem a uma palavra única.

A Fusão opera, por exemplo, na formação dos verbos deadjetivais que apresentam um componente morfológico explícito, tipicamente o sufixo **-ar**, no português. Na estrutura da fig. 10a., representa-se o estágio anterior à fusão do adjetivo “vermelh(o)” com a matriz afixal de V1 e, na fig. 10b, representa-se a forma transitiva de “avermelhar”, que resulta de duas fusões: a primeira, do adjetivo “vermelh(o)” com a matriz afixal **-ar**, formando “avermelhar” (V1), e a segunda, referente à Fusão do primeiro verbo derivado, V1, com a matriz verbal vazia, V2, gerando o verbo da frase “O pôr do sol avermelhou o céu”.

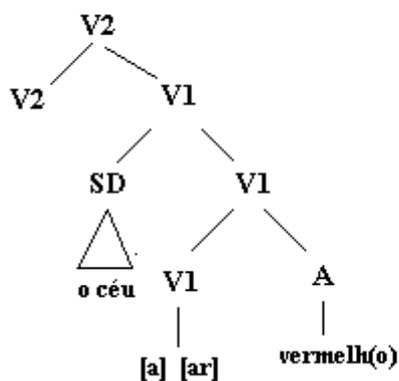


Fig 10a. Estágio anterior às Fusões.

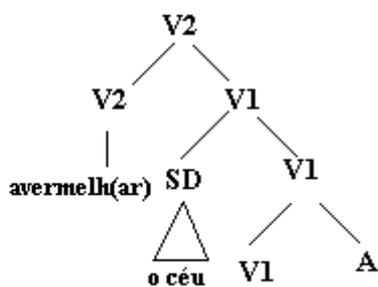


Figura 10b. Fusões do adjetivo “vermelho” com os afixos **a-** e **-ar** e do verbo derivado V1 com a matriz verbal vazia V2, compondo a estrutura transitiva do verbo “avermelhar”.²³

²³ A primeira fusão [V1+A] produz a forma intransitiva do verbo “tornar” “As folhas tornaram(se) vermelhas”.

A operação de Fusão também se aplica aos verbos que tem os argumentos Local e Localizado (LEVIN; RAPAPPORT-HOVAV, 1991). Esses verbos, como os denominais “embolsar”, “engarrafar”, “encaixotar”, “aprisionar”, “amanteigar”, “salgar”, “aguar”, “selar”, entre outros, são "hipônimos" do verbo “colocar”/“pôr”. Portanto, apresentam estrutura semelhante à ilustrada nas figuras 7a e 7b da subseção anterior, em que o núcleo interno, a preposição P, tem a propriedade léxico-gramatical de projetar um complemento e um especificador (de acordo com a estrutura de argumentos exposta na fig. 9c).

Assim, para formar o verbo “engarrafar”, em “Ele engarrafou o vinho”, o núcleo P funde-se com o seu complemento N. O verbo, V, que apresenta uma matriz afixional **-ar**, funde-se com o núcleo de seu complemento, P. A fig. 11a ilustra a estrutura de argumentos do verbo denominal “engarrafar” antes da dupla aplicação da operação de Fusão, e a fig. 11b, ilustra esse verbo depois das Fusões.

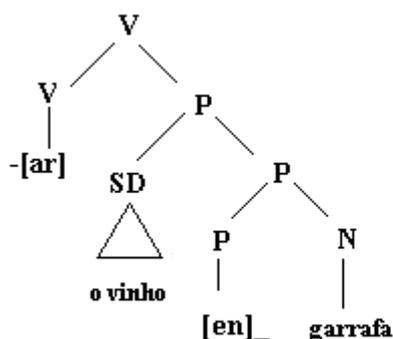


Figura.11a Estágio anterior às Fusões.
e

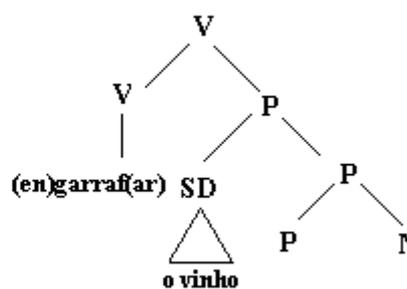


Figura.11b. Fusão do sufixo “en-“, em P, com N “garrafa”
posterior Fusão de “engarrafa”, em P, com o sufixo verbal
“-ar”, originando o verbo “engarrafar”

A Fusão e a **Incorporação** (BAKER, 1988) são processos sintáticos que estão de acordo com a Restrição ao Movimento do Núcleo (*Head Movement Constraint*) e com o

Princípio das Categorias Vazias (*Empty Category Principle*) (CHOMSKY, 1981). A Restrição ao Movimento do Núcleo preconiza que um item lexical, como um verbo, somente pode se incorporar ou se fundir com as palavras por ele propriamente regidas. O Princípio das Categorias Vazias também está relacionado à relação de regência²⁴ (*government*) e se refere à condição de que os vestígios deixados em processos que envolvem deslocamento devem ser propriamente regidos, ou seja, A rege propriamente B se A rege B e A e B estão coindexados.

As duas operações, Fusão e Incorporação, formam uma palavra por anexação do núcleo de um complemento ao núcleo daquele que o governa sintaticamente. Porém, a Incorporação é diferente da Fusão no que se refere à questão da regência, pois a operação de Incorporação ocorre em qualquer posição, desde que o núcleo a ser incorporado seja regido pelo núcleo que o incorpora. Por outro lado, a operação de Fusão exige que o núcleo a ser incorporado não esteja na posição de especificador, conforme pode ser entendido a partir da fig 12, que representa a construção “encurrular o gado”. A operação de Incorporação do especificador da projeção-P, o N (gado), pelo núcleo V (ar-), produziria verbos como “gadar”, e licenciaria frases agramaticais como *“João gadou no curral.”

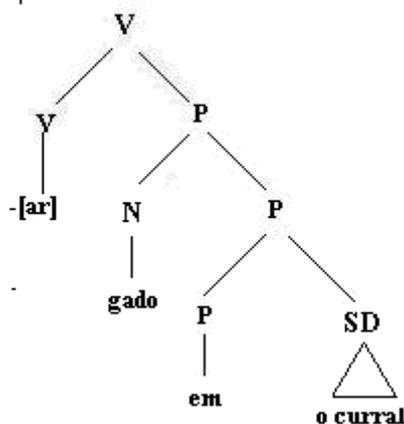


Figura 12. Ilustração da impossibilidade da Fusão de N (gado) com V (-ar).

²⁴ A Teoria da Regência estipula que A rege B se A c-comanda B e não há nenhuma categoria C que seja barreira entre A e B.

A Fusão parece deixar um vestígio, suposição que será discutida ainda nesta subseção. Por exemplo, o verbo “encurralar” é formado, pela fusão de N (curral), complemento de P, com o P (en-) e pela fusão de P + N (encurrar) com V (-ar). Essa última fusão é que deixaria um vestígio (v) na estrutura dominada por P, conforme demonstra a fig. 13.

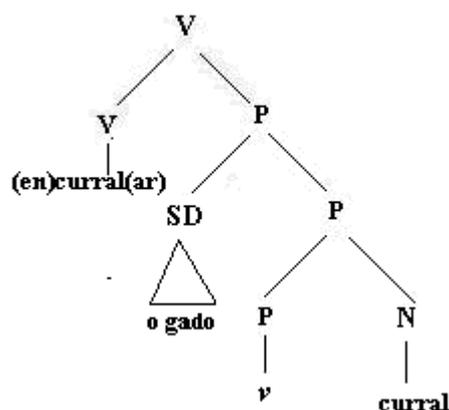


Figura 13. Ilustração do vestígio deixado pela Fusão de P (en-) com V (-ar), originando o verbo “encurralar.”

Na Fusão, a relação sintática entre os núcleos envolvidos é a de *complementação estrita*, assim é definida: *um núcleo X é o complemento estrito de um núcleo Y se Y está na relação de c-comando com a projeção categorial máxima de X*. Na fig. 13, a Fusão está de acordo com a complementação estrita, visto que a projeção máxima de P (o nó P irmão de V) c-comanda N. Essa exigência de complementação estrita também explica o fato de não se poder fundir especificadores de núcleos, pois um especificador nunca será irmão de V.

Sugere-se, ainda, que a Fusão deve ocorrer concomitantemente à operação de **Combinação** (*Merge*) (RADFORD, 1997), responsável pela construção dos sintagmas e, conseqüentemente, da frase. Essa é operação sintática fundamental responsável pela projeção sintática das categorias lexicais. De forma simplificada, a projeção sintagmática “causar problema”, por exemplo, é formada pela seleção dos itens [caus(ar)] e [problema] e pela

combinação dos dois por meio da operação de Combinação, conforme ilustrado nos passos (44a) (44b) e (44c).

(44)

- a. Seleção de [caus(ar)]
- b. Seleção de [problema]
- c. Combinação ([caus(ar)], [problema]) = {[caus(ar)] [problema]}²⁵

Essa combinação gera uma configuração sintática na qual os dois itens são irmãos; os itens são representados pelos rótulos categoriais convencionais, V para “caus(ar)” e N para “problema”. No entanto, não são apenas os nós terminais que apresentam rótulos, os sintagmas também o apresentam. Por exemplo, {[caus(ar)] [problema]} pode ser abreviado por {V, N} e deve estar associado ao rótulo SV. O rótulo do sintagma é determinado pelo constituinte núcleo (Nuc) da combinação.

Desse modo, o rótulo de uma configuração sintática é a cópia dos traços do seu núcleo. Em outras palavras, o rótulo de um objeto sintático X é o conjunto de traços [T, Nuc], onde [T, Nuc] são os traços (T) morfológicos, fonológicos, sintáticos e semânticos de Nuc, o núcleo de X.

Os traços que estão registrados nos rótulos são chamados de **assinatura-p** (*p-signature*). Esse constructo representa o conjunto de traços fonológicos presentes nos rótulos. A assinatura-p é bastante relevante para a teoria da Fusão e será simbolizada pela ortografia convencional do item. A assinatura-p de gado, por exemplo, é, de forma simplificada, a forma vocabular “gado”.

Os autores propõem que a Fusão opere somente sobre os rótulos, tendo acesso aos mesmos elementos lingüísticos que a operação de Combinação, já que ocorrem concomitantemente. Assim a Fusão é definida do seguinte modo: *é o processo de cópia da*

²⁵ As chaves "{}" sinalizam o resultado da operação de Combinação.

*assinatura-p do complemento na assinatura-p do núcleo, quando esta é defectiva. Uma assinatura-p é defectiva quando for vazia, isto é, não contém traços fonológicos, ou quando for afixal. O exemplo (45) ilustra um verbo denominal do inglês que possui uma assinatura-p vazia. Sua matriz fonética procede do complemento.*²⁶

(45)

núcleo	complemento
{V, [∅]}	{N, [laugh]}

Depois da aplicação das operações de seleção e de combinação, o rótulo do Núcleo passa a ser {V, [laugh]} e é projetado para definir os traços da construção como um todo; a operação de fusão (de cópia) ocorre concomitantemente com a operação de combinação sintática, como mostra a fig. 14a. Depois da operação de cópia, a assinatura-p do complemento é apagada, como mostra a fig. 14b.

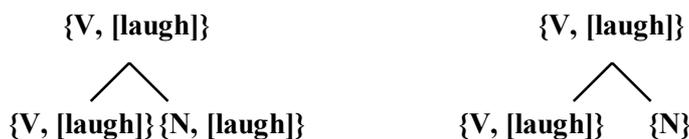


Figura 14a. Cópia da assinatura-p do complemento no núcleo.

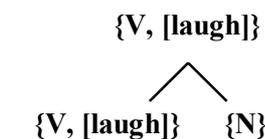


Figura 14b. Apagamento da assinatura-p do complemento.

Os verbos deadjetivais resultam da aplicação de operações análogas. O verbo “engrossar”, por exemplo, é resultado da Fusão e da Combinação da assinatura-p da matriz verbal afixal [en][∅][ar] com a assinatura-p da matriz nominal plena [gross(o)], conforme ilustram as figs. 15a e 15b.

²⁶ Os símbolos [∅] e [laugh] representam as assinaturas-p dos itens vazio e pleno, respectivamente.

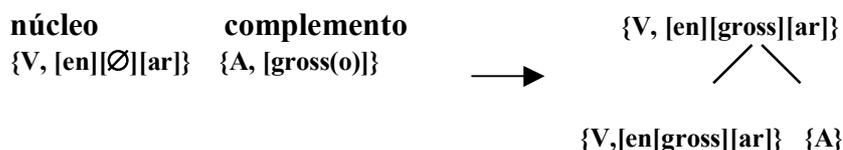


Figura 15a. Combinação do rótulo do núcleo com o rótulo do complemento.

Figura 15b. Cópia simultânea da assinatura-p do complemento sobre a assinatura-p deficiente do núcleo.

Os verbos deadjetivais projetam um especificador que corresponde ao sujeito, na variante intransitiva, e ao objeto, na variante transitiva; assim, adicionando-se o especificador “o caldo” por meio da operação de Combinação chega-se a uma estrutura na qual a assinatura-p do núcleo determina o rótulo da construção: a projeção verbal, que tem como núcleo o verbo “engrossar”. A fig.16a. representa a variante intransitiva inacusativa (incoativa) da **alternância causativo/incoativo**. A fig.16b representa a variante transitiva/causativa de “engrossar”, formada pela combinação da estrutura da fig 16a com um verbo de assinatura-p vazia, que se funde com a assinatura-p do núcleo do seu complemento (o {V} da estrutura da fig. 16b).

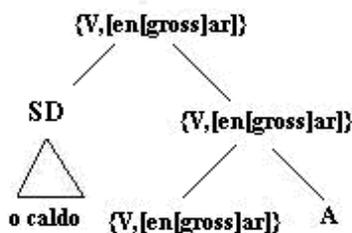


Figura 16a. Estrutura da variante incoativa do verbo “engrossar”.

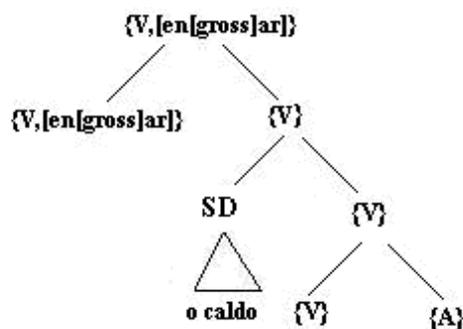


Figura 16b. Estrutura da variante causativa do verbo “engrossar”.

Considerando que a Fusão opera sobre as assinaturas-p dos rótulos, é importante ressaltar que ela não tem efeito sobre o que se pode chamar de assinatura-s, isto é, a assinatura que se refere aos traços semânticos e sintáticos de um nó sintático; as assinaturas-p de todos os nós são irrelevantes e invisíveis para a sintaxe.

A conjectura sobre a possibilidade da Fusão deixar um vestígio, conforme mostrado na fig 13, deve ser melhor avaliada.

Por um lado, pode-se pensar que a Fusão deixa um vestígio, pois as estruturas semântica e sintática permanecem intactas. Por outro lado, pode-se pensar também que a Fusão não deixa vestígio algum, já que ela não é uma operação de movimento, mas um processo de “substantivação” de uma assinatura-p defectiva.

Em decorrência dessas suposições, seja para validação seja para rejeição da necessidade de se postular uma operação de Fusão, a análise de construções com argumento cognato e com argumento hipônimo trazem alguma luz sobre o problema.

Há diferença entre as construções com objeto cognato, como em “Ele sorriu [seu último sorriso]_{ObjCog}”, “Ela sonhou [um sonho horrível]_{ObjCog}” ou “Choveu uma [chuva fina]_{ObjCog}”, e construções com objeto hipônimo como “Ela dançou [balé]_{ObjHip}”, em que o objeto é hipônimo de dança.

Nessas construções, o objeto cognato possui como núcleo um substantivo com raiz idêntica ao verbo. Assim, elas são problemáticas para a operação de Fusão, pois o verbo denominal é produto de uma fusão do N (sonh(o)) com a matriz afixional V (ar), de modo que o objeto sintático “sonho” acaba por co-ocorrer com o N (sonh(o)) fundido ao V. Conseqüentemente, a construção com objeto cognato não pode ser entendida por meio da operação de Fusão canônica, na qual o elemento fundido é o núcleo da projeção lexical que é complemento (é o nó irmão) do verbo alvo. Em decorrência dessa impossibilidade, a Fusão que opera nas construções com objeto cognato deve ocorrer entre o verbo e uma projeção

irmã “estendida”. Essa extensão consiste na cópia da assinatura-p do N na assinatura-p do V, levando-se em conta que alguns traços do rótulo de um item lexical projetam-se até a projeção máxima de uma projeção estendida do núcleo lexical, não apenas até a projeção máxima de um núcleo lexical. Assim, conforme esquema da fig.17, o verbo defectivo é irmão de D(Determinante), que é uma projeção estendida, e carrega consigo a assinatura-p do núcleo lexical, passando-a para o rótulo de V.

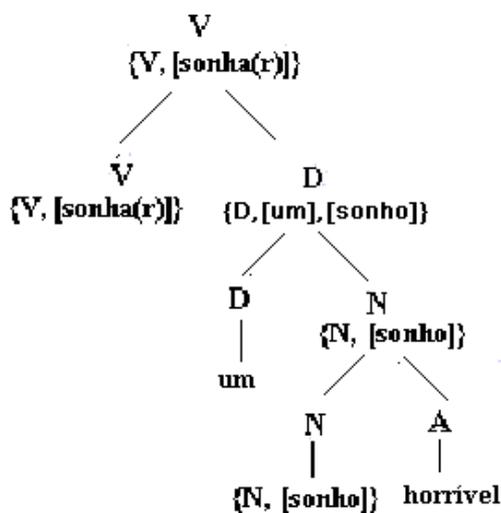


Figura 17. Representação da operação de Fusão com verbos com argumento cognato.

Na fig. 17, N e V fundem-se porque D, a projeção estendida de N, é complemento (irmã de V), o verbo alvo, e a Fusão opera em longa distância, no sentido de que atravessa uma projeção estendida, mas obedece a relação entre núcleo e complemento.

Constata-se, assim, que a Fusão é um processo de cópia. Quando a posição de origem, isto é, o N, é dominada por uma categoria funcional (por exemplo, D) a cópia é explícita, conforme mostra a fig. 17; quando, por outro lado, a posição de origem é dominada por categoria lexical, como ocorre na Fusão canônica, a cópia é apagada na posição de origem, como ilustrado na fig. 16b, em que as posições de origem são V e A.

As construções com argumentos hipônimos, como “O casal dançou samba” podem ser tratadas de modo análogo ao tratamento sintático dado à frase "O casal tocou um samba".

Nessa frase, em que não se recorre à operação de Fusão para explicar sua formação, há somente o que se pode denominar **Inserção Lexical**, conforme ilustram as figs. 18a e 18b.

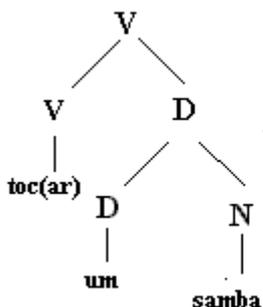


Figura 18a: Ilustração do processo de Inserção Lexical no SV em "tocar um samba".

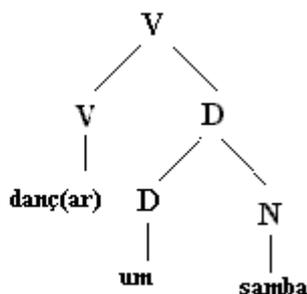


Figura 18b. Ilustração do processo de Inserção Lexical no SV em "dançar um samba".

Conseqüentemente, é possível pensar que, em "dançar um samba", as duas operações entrem em ação: Combinação de V com [_D D N] e Inserção Lexical. Como resultado, indicado na fig.18b, "danç(ar)" já apresenta uma matriz fonológica integrante à sua entrada lexical, o que desqualifica a noção de assinatura-p e também a noção de Fusão, pois, se o léxico possui o verbo "danç(ar)", com matriz fonológica completa, ele pode ser inserido diretamente na posição V, sem que a assinatura-p de N tenha que ser copiada para V.

No entanto, se o verbo pode ser licenciado por Inserção Vocabular, é necessário explicar o licenciamento de um N vazio. Alguns verbos não admitem que o objeto seja apagado, é o caso dos verbos suporte, como *"Ela deu" ("Ela deu um beijo"), nem os verbos na variante transitiva, como *"Ele limpou" ("Ele limpou a mesa") ou *"Ela apertou" ("Ela apertou o cinto"). Nesses casos, o verbo não fornece informação suficiente para "dispensar" o objeto.

Por outro lado, em casos em que o verbo é inergativo, como em "O casal dançou" ou "A garota espirrou", ele oferece informação suficiente para a interpretação do complemento,

mesmo quando este não está explícito. Assim, pode ser entendido que o componente nominal desses verbos está realmente presente e serve para licenciar o complemento implícito.

Um fator que é relevante para se avaliar a possibilidade de o núcleo apresentar um complemento implícito que seja identificado pelo caráter do verbo é a relação classificatória entre certos traços semânticos do núcleo e do complemento. No caso das construções com objeto hipônimo, entende-se “samba” como um tipo de dança e, nos casos dos inergativos, entendem-se os complementos implícitos como os eventos ou as entidades que correspondem aos substantivos “dança” e “espirro”. Essa seleção classificatória, como mostra a fig.19, pode ser representada por índices, no exemplo, os índices {dança}, que ligam os verbos aos seus complementos.

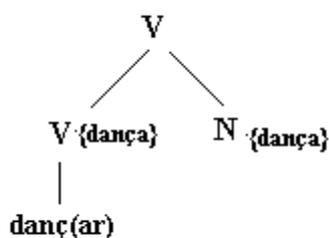


Figura 19. Representação da relação seletiva entre o verbo e seu complemento, indicada pelos índices {dança}.

Assim, os verbos suporte não podem licenciar um complemento implícito, pois não têm um componente semântico para participar de uma relação classificatória; a diferença com os inergativos reside, portanto, no componente semântico lexical presente nestes e ausente naqueles.

Outro fator importante no licenciamento do complemento implícito são as relações expressas na configuração da estrutura de argumentos. A relação expressa na fig. 19 estabelece-se somente entre o núcleo e o argumento selecionado, isto é, o complemento. Os verbos “limpar” e “apertar” nas frases citadas não licenciam um complemento implícito, porque o complemento não entra na relação estrutural com o verbo, ou seja, o argumento

deveria estar na posição de complemento, mas está na posição de especificador de V, não mantendo relação argumental direta com o verbo em posição hierarquicamente superior da estrutura. Como o complemento não pode ser indexado, ele deve ser explícito.

Como conclusão interessante desta discussão, no que se refere aos verbos denominais, não se deveria postular substantivos intimamente relacionados a verbos, mas sim postular “raízes lexicais” que não são, inerentemente, nem verbos e nem substantivos. A categoria dessa raiz dependerá do contexto, ou seja, será V se ocorrer em uma projeção verbal estendida e, assim, é será apropriadamente flexionada como verbo, com tempo, modo, aspecto e todas as outras categoria que lhe são próprias. Analogamente, será N se ocorrer em uma projeção nominal estendida, assumindo os traços próprios da categoria dos substantivos. Decorre, então, não ser desejável afirmar que *dance* do inglês e a raiz “danç-” do português, em *He danced samba* / “Ele dançou um samba”, sejam verbos denominais; sugere-se, de outro modo, que se trata é um item de categoria indeterminada cujas características fonológicas, por meio da Inserção Vocabular, assume a posição correspondente ao núcleo da projeção verbal estendida.

Resumindo, não há mais razão para se pautar na Fusão para explicar os verbos denominais; a relação que existe entre o verbo “dançar” e seu complemento explícito N é uma relação de seleção semântica, não constituindo, portanto, evidência para tal operação. Além disso, a relação entre “danç_” (V) e “danç_” (N) é simplesmente a de indeterminação categorial. Constata-se, conseqüentemente, que a Fusão não deverá fazer parte da descrição da estrutura de argumentos.

Paralelamente, certos verbos cujos argumentos são do tipo Local e Localizado podem apresentar complemento implícito, enquanto outros, não. No par * “Ricardo pôs as balas” (cf. “Ricardo pôs as balas no bolso”) e “Ricardo embolsou as balas”, o verbo “pôr” é também fraco para licenciar um argumento não explícito. O significado associado a esse verbo é o de

uma relação entre duas entidades: o especificador interno funcionando como “Localizado” e o complemento de P, como “Local”. O esquema de subcategorização de “pôr” exige um complemento direto SD e um complemento oblíquo SP, o que não é satisfeito por “*Ricardo pôs as balas”, em que o argumento Local (o complemento oblíquo) está implícito.

Entretanto, verbos como “embolsar”, “engarrafar” e “selar”, entre outros do tipo Local e Localizado, permitem o apagamento do sintagma preposicional. Aliás, é o que comumente ocorre, não se descartando, porém, a possibilidade de que o argumento que indica lugar explícito, como em “Ricardo embolsou as balas no **bolso furado**”, ocorra, pois “embolsar” possui a necessária semântica lexical para permitir um complemento explícito. Assim, o verbo “embolsar” deve ser coindexado, juntamente com o complemento nominal da preposição, conforme mostra a fig.20.

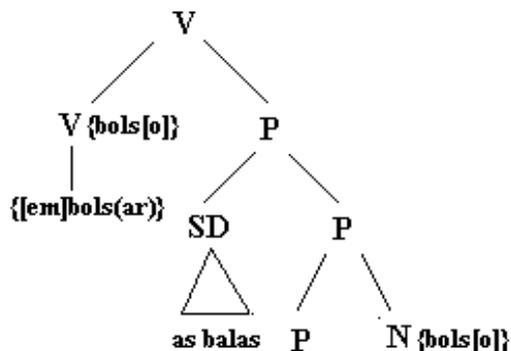


Figura 20. O processo de Coindexação entre o verbo “embolsar” e o complemento nominal da preposição.

Na fig. 20, V estruturalmente não mantém uma relação direta com N (dentro do SD); P está entre V e esse N e é responsável pela seleção do N “bolso”. Porém, é o verbo que semanticamente e estruturalmente seleciona a preposição, estabelecendo uma cadeia de seleção que se estende do verbo até o objeto da preposição. Desse modo, a seguinte generalização é possível: os verbos que apresentam argumentos do tipo Local e Localizado

mantêm uma relação seletional com o objeto nominal da preposição. Essa generalização pode ser verificada em frases como “Encurrelei o gado no cercado”, em que “cercado” qualifica algo que pode ser usado para prender o gado.

No entanto, se o verbo pode semanticamente selecionar o objeto do seu complemento preposicional, sendo essa propriedade suficiente para licenciar um complemento explícito, então frases como “*Ricardo embolsou as balas no” ou “O vaqueiro encurralou as vacas dentro” seriam gramaticais. A explicação da agramaticalidade de frases como essas é a seguinte: a relação entre o verbo e o objeto da preposição não é local, ou seja, não ocorre em um domínio estritamente restrito, pois P intervém entre as categorias V e N relevantes, bloqueando a ligação entre ambas.

Como pode ser visualizado em na fig. 20, N deve ser governado por V em uma relação de antecedência (*antecedent-binding relation*); se a preposição for vazia, conseqüentemente, ela não constitui um impedimento para essa relação e o complemento pode ser não explícito.

Diante do que foi exposto, o que se chamou Fusão é, de outro modo, a situação em que um verbo, com conteúdo semântico, é coindexado com um argumento nominal que se posiciona em um local da estrutura sintática em que pode ser governado por V; N pode, assim, ser implícito, desde que a preposição seja nula.

Porém, é necessário que se considerem, ainda, os verbos ‘deadjetivais’ como “alargar”, que participa da alternância causativa/incoativo, como em “O buraco alargou” / “A erosão alargou o buraco”. Uma das questões que se coloca é se esses verbos podem ser analisados a partir da seleção semântica e da coindexação, como os verbos ‘denominais’. O que faz pensar que essa não é a abordagem adequada para esse tipo de verbos é que eles não apresentam construções produtivas com argumentos cognatos e hipônimos.

O que explica a formação dos verbos ‘deadjetivais’, por outro lado, é a operação de Incorporação, que estipula que o complemento adjetival deve deslocar-se de sua posição de

base para a posição do verbo, de acordo com o Princípio de Restrição ao Núcleo, criando uma cadeia entre a posição de base (a do adjetivo) e a posição final (a do verbo). Esse deslocamento é o que impede a formação de objetos cognatos, visto que o vestígio bloqueia a inserção de um item lexical.

Da mesma forma, é a Incorporação que explica a composição morfológica característica dos verbos ‘deadjetivais’: “alargar”, “clarear”, “alongar”, “escurecer”, etc; afixos que encontram paralelo no afixo *-en* do inglês: *redden*, *widen*, *lengthen*, *darken*, etc. De acordo com a fig.21, o complemento deadjetival se desloca até V, posicionando-se à sua esquerda; A é o vestígio do movimento; “larg(o)”, em V, governa por antecedência o A vazio, sendo o núcleo da cadeia de Incorporação.

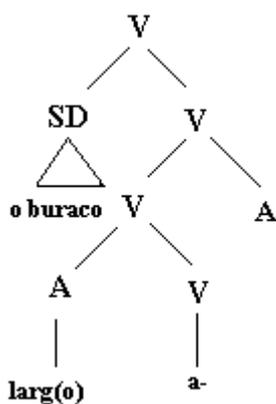


Figura 21. O processo de Incorporação do verbo ‘deadjetival’ ‘alargar’.

O que se pode concluir sobre a discussão desta subseção, é que a Fusão não uma operação de deslocamento, mas é um processo de ligação entre os traços semânticos do verbo e os traços semânticos do núcleo nominal do seu complemento. Essa relação de ligação, por sua vez, é resultado da relação de seleção entre o verbo e o seu complemento, não atuando, portanto, entre o verbo e o seu especificador.

3.4 A interação entre a estrutura de argumentos e a estrutura conceitual

Os tipos de estrutura de argumentos delineados na subseção 3.2 interagem com as operações de Combinação e de Incorporação (subseção 3.3), conforme demonstrado na subseção anterior. Contudo, o que está em foco nesta subseção é a interação entre a estrutura de argumentos e a semântica, mais especificamente, analisam-se os traços semânticos associados aos itens lexicais para explicar o comportamento da estrutura de argumentos, que não pode ser explicado apenas pelas relações de complementação e especificação.

Os verbos ‘denominais’ do tipo Local e Localizado não participam da **alternância causativo/incoativo**. Esses verbos apresentam somente a forma transitiva/causativa, como em “João engarrafou o vinho” /*“O vinho engarrafou”. Os verbos ‘deadjetivais’, por outro lado, participam dessa alternância, como em “O céu clareou” /“O sol clareou o céu depois da chuva”. Isso ocorre porque, ao se retirar o V da estrutura projetada pelos verbos ‘denominais’ de Local e Localizado, resta uma projeção preposicional, isto é, uma projeção-P, não uma projeção-V, conforme mostra a fig.11a, repetida na fig 22a. Por outro lado, os verbos ‘deadjetivais’, como mostra a fig.22b, projetam uma estrutura verbal intransitiva.

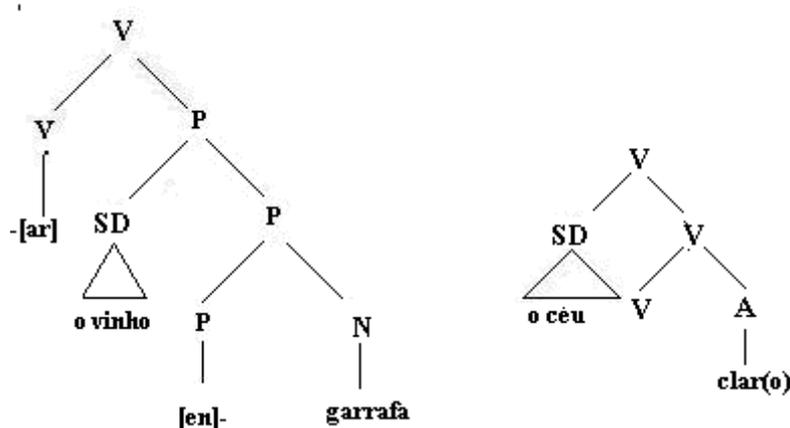


Figura 22a. Estrutura projetada por um verbo ‘denominal’ com argumentos do tipo Local e Localizado.

Figura 22b. Estrutura projetada por um verbo ‘deadjetival’.

A descrição da realização sintática dos verbos “lambuzar” (no sentido de “passar X sobre uma superfície Y, em que X tem características gordurosas”) e “respingar” (no sentido de “respingar líquido em Z”), diverge, entretanto, desse padrão, pois o verbo “respingar” participa da **alternância causativo/incoativo**, mesmo tendo uma projeção-P como complemento (“em” e “com”), como demonstram (46a) e (46b).

(46)

- a. “Os porcos respingaram lama na parede.”
- b. “A lama respingou na parede.”

O verbo lambuzar, adequando-se ao padrão dos verbos que apresentam uma projeção-P de não participarem da **alternância causativo/incoativo**, não possui a forma incoativa, conforme mostra os exemplos (47a) e (47b).

(47)

- a. “O garoto lambuzou as mãos com chocolate.”
- b. * “O chocolate lambuzou”/ * “O garoto lambuzou.”

A estrutura projetada por “respingar” exhibe um especificador, que é exigido pela projeção-P; mas é o verbo que projeta o especificador, e a preposição preenche o complemento do verbo; como ocorre com os verbos ‘deadjetivais’, exemplificado na estrutura da fig. 9d, subseção 3.2. As figuras 23a e 23b ilustram, respectivamente, as duas configurações sintáticas intransitiva e transitiva projetadas pelo verbo “respingar”.

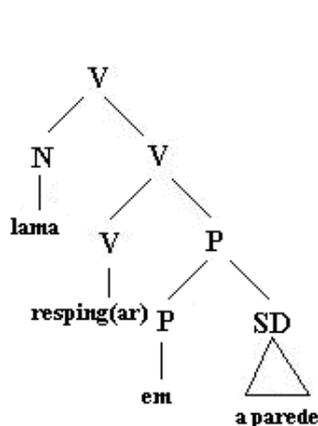


Figura 23a. Variante intransitiva projetada pelo verbo “respingar”.

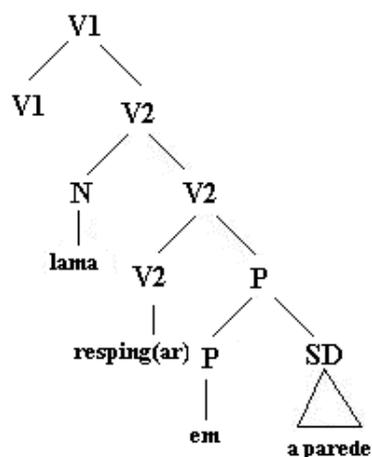


Figura 23b. Variante transitiva projetada pelo verbo “respingar”.

O verbo “lambuzar”, contudo, apresenta apenas a forma transitiva. Há a possibilidade de se considerar que a estrutura atribuída a ele seja a ilustrada na fig. 24, em que o verbo entra em construção direta com a projeção máxima de P, e é P que exige o especificador, não o verbo como na estrutura de “respingar”. A estrutura da fig. 24 é a estrutura canônica esperada para verbos com projeção-P, sendo a estrutura das figs. 23a e 23b, a estrutura marcada.

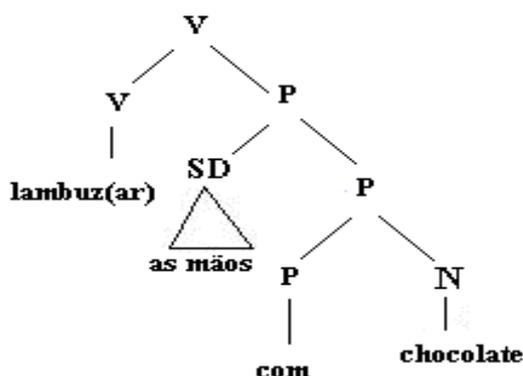


Figura 24. Variante intransitiva projetada pelo verbo “lambuzar”.

Por outro lado, uma forma de explicar essa divergência de comportamento nos verbos que apresentam como complemento uma projeção-P é considerar a relação entre as estruturas argumental e semântica, ou seja, considerar os aspectos do significado lexical que são relevantes para a estruturação sintática.

Verbos do tipo “respingar”, que incluem “gotejar” (“gotejar lama no carro”) “emplastar” (“emplastar lama no carro”), “pingar” (“pingar lama no carro”) e “derramar” (“derramar lama no carro”), são definidos como verbos “modo-paciente” porque incluem, na sua entrada léxico-enciclopédica, um traço de natureza adverbial que identifica o movimento físico, a distribuição e a dispersão da entidade denotada pelo argumento (interno) semanticamente Paciente, nesse caso a lama. É importante ressaltar que o traço semântico adverbial “modo-paciente” está associado ao argumento interno.

Verbos do tipo de “lambuzar”, que compreendem “passar” (“X passa reboco na parede”); “colocar” (“X coloca reboco no chão”), são definidos como verbos “modo-agente” por incluírem, em sua entrada léxico-enciclopédica, um traço de natureza adverbial que descreve a ação (os gestos e movimentos) da entidade denotada pelo argumento externo (sujeito sintático), semanticamente Agente. É necessário um Agente X que execute a ação

denotada por esses verbos; por isso é que se espera que um traço semântico de natureza adverbial deva estar associado ao argumento externo.

A estrutura intransitiva impede que o traço semântico “modo-agente” esteja ligado ao argumento externo, pois não há argumento externo em tal configuração. Por isso, verbos do tipo “lambuzar” não participam dessa alternância. Com o objetivo de relacionar o traço de natureza adverbial “modo-agente” ao argumento externo, propõem-se índices associados a itens lexicais ($\{i\}$ na fig. 25b). O índice deve estar ligado a um argumento externo disponível e que tenha o mesmo índice em um contexto sintático. Esse argumento estará disponível somente na construção intransitiva, representada na fig. 25b.

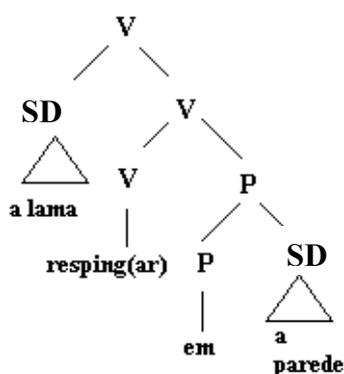


Figura 25a. Variante intransitiva projetada por um verbo do tipo “modo-paciente”.

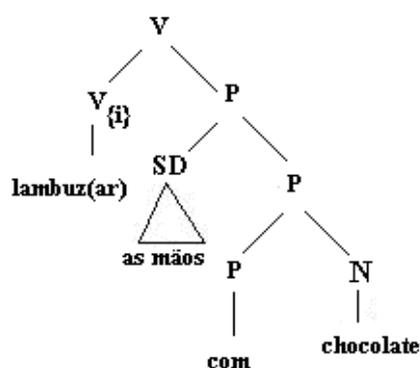


Figura 25b. Variante transitiva projetada por um verbo do tipo “modo-agente” (com a indicação do índice de associação entre o item lexical e o traço semântico).

Em suma, nesta subseção teceram-se considerações sobre a subordinação da estrutura de argumentos dos verbos aos traços semânticos a eles associados, demonstrando que os aspectos semânticos dos itens lexicais têm papel de destaque na teoria de estruturas sintáticas projetadas pelas estruturas de argumentos dos itens lexicais de Hale e Keyser (2002). Desse modo, entende-se que a teoria desses autores vem auxiliar sobremaneira na investigação da hipótese de que o comportamento sintático dos verbos é função de sua estrutura conceitual.

3.5 A representação léxico-conceitual dos verbos.

Jackendoff (1990; 2002) propõe um conjunto de funções e constantes semânticas primitivas a partir das quais a estrutura conceitual²⁷ do verbo pode ser representada. Nesta subseção, utilizam-se essas funções e constantes para a proposição da representação léxico-semântica compatível com o que se discutiu na subseção 3.1

A idéia de decomposição semântica do significado do verbo apresentada por Jackendoff deriva do estudo de Gruber (1965), que demonstra que os modelos de campos semânticos que descrevem objetos físicos no espaço podem ser utilizados também na descrição de modelos de campos semânticos não espaciais. Isso pode ser demonstrado por meio de ocorrências de verbos e preposições cujos significados transcendem o domínio espacial, como ilustram os exemplos (48), (49), (50) e (51).

(48) *Localização espacial e movimento*

- a. O pássaro **foi da grama para** a árvore
- b. O pássaro **está na** árvore.
- c. Pedro **manteve** o pássaro **na** gaiola.

(49) *Posse*

- a. A herança **foi de Paulo para** Pedro.
- b. O dinheiro **está com** Pedro.
- c. Ana **manteve** o dinheiro **na** sua carteira.

(50) *Atribuição de propriedades*

- a. A luz **mudou de** verde **para** vermelho
- b. A luz **é** vermelha
- c. O policial **manteve** a luz vermelha.

(51) *Planejamento de atividades*

- a. A reunião foi **mudada de** terça-feira **para** segunda-feira.
- b. A reunião **é na** segunda-feira.
- c. O presidente **manteve** a reunião **na** segunda-feira.

²⁷ A Estrutura Conceitual (*Conceptual Structure*) é o nome que Jackendoff dá a uma das divisões da estrutura do significado; é a contrapartida da Estrutura Espacial (*Spacial Structure*); trata-se de um arranjo hierárquico, constituído de traços discretos e funções. A Estrutura Conceitual de um item lexical codifica esses traços discretos e as funções por meio da relação de pertença a uma classe (taxonomia) e por meio da estrutura de argumentos.

Cada grupo apresenta uma frase com os verbos “ser”, “estar”, “ir”, “mudar” e “manter” e com as preposições “de” e “para” e outra com o verbo “manter”. As frases com os verbos “ser” e “estar” descrevem um estado, como por exemplo, estar em um lugar, pertencer a alguém, ser de determinada cor ou acontecer em determinado tempo; as frases com os verbos “ir” e “mudar” descrevem uma mudança de estado; as frases com o verbo “manter” denotam um Agente/Causativo que faz com que um objeto tenha determinadas características por um determinado tempo.

As frases dos exemplos acima refletem um conjunto de esquemas conceituais que podem ser aplicados a vários campos semânticos. Desse modo, Jackendoff (1990; 2002) propõe que há **funções** que sejam recorrentes na decomposição dos verbos e só é preciso especificar o campo semântico no qual um *Estado (State)* ou *Evento (Event)* é definido. Assim, postulam-se, por exemplo, a função GO (IR)²⁸, que é usada para descrever frases com verbos do tipo “ir” ou “mudar”; a função BE (SER/ESTAR), que descreve a estrutura conceitual frases que contém os verbos de estado do tipo “ser/estar”; a função STAY (Ficar), que descreve a estrutura conceitual de frases com o verbos permansivos, como “manter”. Essas são algumas das funções que descrevem os componentes conceituais primitivos de predicadores e se organizam em duas categorias semânticas ontológicas: *Event* (Evento) e *State* (Estado).

A função BE(X,Y), ilustrada em (52a), é uma função de dois lugares da categoria ontológica *State*, isto é, possui dois argumentos, representados por (X,Y), e é modificada por um traço de campo semântico, que determina o caráter dos argumentos de BE. Assim, quando

²⁸ Dado que o modelo de Jackendoff é universalmente conhecido com todos os seus constructos rotulados em inglês, nesta dissertação, seguindo essa tradição, os nomes das funções conceituais serão escritos em letras maiúsculas e em inglês e os nomes das categorias ontológicas (as constantes), também em inglês, serão subscritos adjacentes às funções. Mas é essencial destacar que esses rótulos não são palavras do inglês, mas mnemônicos que identificam funções e conceitos que, para Jackendoff, não dependem de língua. Por exemplo, os diferentes eventos que envolvem deslocamento no espaço conterão, na sua estrutura conceitual a função GO. Não se perderia nada, do ponto de vista teórico, se essa função fosse denominada IR.

o campo semântico for *Spacial* (Espacial), X é um Objeto e Y, a Localização de X; quando o campo for *Possession* (Posse), X é um Objeto e Y é o Possuidor de X; quando o campo for *Ascription* (Atribuição), X é uma Entidade e Y é uma Propriedade ou um Tipo de X; quando o campo for um *Scheduling* (Planejamento), X é um Evento e Y é um Período de Tempo.

(52)

- a. $[\text{State BE } ([\text{Thing } X], [\text{Place } Y])]$
- b. $[\text{Event STAY } ([\text{Thing } X], [\text{Place } Y])]$
- c. $[\text{Event GO } ([\text{Thing } X], \left[\begin{array}{l} \text{FROM } ([W]) \\ \text{TO}([Z]) \\ \text{Path} \end{array} \right])]$

A função $\text{STAY}(X,Y)$ pertence à categoria ontológica *Event*. Essa função está subjacente à descrição das estruturas conceituais de verbos como “manter”, como já foi mostrado, “permanecer” e “continuar”, e é a conceitualização de uma configuração estática que se estende por um período de tempo, como ilustra o exemplo (52b).

A função $\text{GO}(X,Y)$ também pertencente à categoria *Event*, é uma função de dois argumentos em que o argumento Y é um *Path* (Trajetória), conceitualizando o evento de X percorrer Y. Uma trajetória pode designar um ponto de início, o *Source* (Fonte), marcado pela função FROM (DE), e um ponto terminal, o *Goal* (Meta), marcado pela função TO (PARA), conforme demonstra os exemplos (48), (49), (50) e (51).

Há, ainda, outras funções que designam, por exemplo, uma direção (“O avião voou para o Norte”) ou o formato de uma trajetória (“O avião voou em círculos”). Quando a trajetória contém os componentes *Source-Goal*, ela pode ser representada conforme (52c), em que o segundo argumento de GO tem sua própria estrutura de argumentos.

Além dessas, postulam-se também as funções EXT (EXTENSÃO) e ORIENT (ORIENTAÇÃO). Essas funções podem ter *Path* (Trajetória) como argumento. Assim, em “A rodovia acompanha o rio”, a trajetória não implica a noção de temporalidade que é expressa no contexto de “mudança ao longo de um caminho”, mas representa uma extensão não temporal. A mesma análise aplica-se a “A seta aponta para o norte”, em que também não há temporalidade e *Path* é argumento da função ORIENT (X, Path).

Há também as funções semânticas que atuam na dimensão aspectual. A função INCH (abreviação de INCHOATIVE, isto é, INCOATIVO) é uma delas. Essa função é de um lugar e seu argumento codifica o estado em que o evento culmina. Por exemplo, a relação entre o significado do adjetivo “aberto” e o significado do verbo intransitivo “abrir” é que o significado de “abrir” conceitualiza o evento cujo estado final é o de estar aberto. A função apropriada para representar o significado da frase “A estrada alcançou Goiânia”, isto é, o significado de que a estrada finalmente (ou rapidamente) se estendeu até a cidade de Goiânia, é a função INCH, que codifica o evento de X se estender até Y, como mostra (53).²⁹

(53) [Event INCH ([State EXT ([Thing ESTRADA], [Path TO ([Thing GOIÂNIA])])])]

Há, por fim, funções que atuam na dimensão causativa, que incluem CAUSE (CAUSAR), LET (PERMITIR), HELP (AUXILIAR), nas dimensões *Event* e *State*. A função CAUSE (X,Y) tem dois argumentos obrigatórios, o Agente/Causativo e o Efeito, e um argumento opcional, o Paciente, como em “O vento me fez cair”, em (54).

(54)

CAUSAR (VENTO, ME, [Evento EU CAIR])

Causativo Paciente Efeito

‘O vento, agindo sobre mim, causou o evento da minha queda’.

²⁹ Para montar a estrutura conceitual da frase “A estrada alcançou Goiânia”, poderia se pensar em recorrer à função GO. Essa função, porém, representaria o significado de “a estrada viajou/mudou para a cidade de Goiânia”, interpretação não desejável. A função GO indica mudança ao longo do tempo e não um processo de extensão não temporal.

Essas funções é que são usadas para construir a representação do significado verbal. Por exemplo, ao esquema frasal (55a), que contem o verbo “ir”, corresponde a estrutura conceitual (55b).

(55)

a.. Estrutura Sintática

[John]_i foi_j [para_k [dentro de_l [a sala]_m]

b. Estrutura Conceitual

[Event GO ([Thing JOHN]_i, [Path TO ([Place IN ([Thing ROOM]_m)]_l)]_k]_j

Nessa estrutura, a forma verbal "foi" expressa a função GO (X,Y).³⁰ O evento GO, indexado com o índice *j*, possui dois argumentos: o argumento externo, JOHN, um SN na sintaxe, indexado com o índice *i*, e o argumento complexo preenchido pela função TO (Path) (*para dentro de a sala*), um SP na sintaxe, indexado pelo índice *k*. A função TO, por sua vez, tem como argumento a função IN (Place) (*dentro da sala*), um outro SP na sintaxe, indexado pelo índice *l*; a função IN, por fim, tem como argumento ROOM (Thing), também um SN na sintaxe, indexado pelo índice *m*.

A estrutura conceitual da frase correspondente do inglês, *John went into the room*, em que *into* significa “para (*to*) dentro (*in*)”, é a mesma já apresentada em (55b) e repetida em (56b). Em (56c) e (56d) apresentam-se, separadamente, as entradas de *go* e *into*, cujas estruturas conceituais se combinam para formar (56b).

³⁰ Abstraindo-se do tempo, do modo, da voz, do aspecto, entre outras categorias gramaticais de que o verbo é suporte.

(56).

a. Estrutura Sintática

$$[John]_i \text{ went }_j [into]_k [the \text{ room}]_m]$$

b. Estrutura Conceitual

$$[Event \text{ GO } ([Thing \text{ JOHN}]_i, [Path \text{ TO } ([PlaceIN ([Thing \text{ ROOM}]_m)]))]_k)]_j]$$
c. Entrada lexical de *into*

$$\left(\begin{array}{l} into_i \\ P_i \\ \text{---} SN_m \\ [Path \text{ TO } ([PlaceIN ([Thing]_m)]))]_i \end{array} \right)$$
d. Entrada lexical de *go*

$$\left(\begin{array}{l} go_j \\ V_j \\ \text{---} SP_k \\ [Event \text{ GO } ([Thing]_i, [Path]_k)]_j \end{array} \right)$$

Na estrutura conceitual em (56b), o verbo lexicaliza a função *Event* e a preposição *into*, as categorias ontológicas *Path* e *Place*. Quando na estrutura sintática não houver um SP, o argumento *Path* não é especificado, mas continua parte da estrutura conceitual.

Há uma alternativa gráfica de representação da estrutura conceitual de (56b), dada na fig. 26. Nessa representação, as funções são ligadas por linhas duplas, seus argumentos por linhas simples e a contribuição da estrutura conceitual de cada constituinte está incluída nas elipses pontilhadas. Observe que *John* e *room*, núcleos dos respectivos SNs, são tratados como funções que não requerem argumentos, são semelhantes às “constantes” de Levin e Rappaport-Hovav (1995) mencionadas na subseção 3.1.

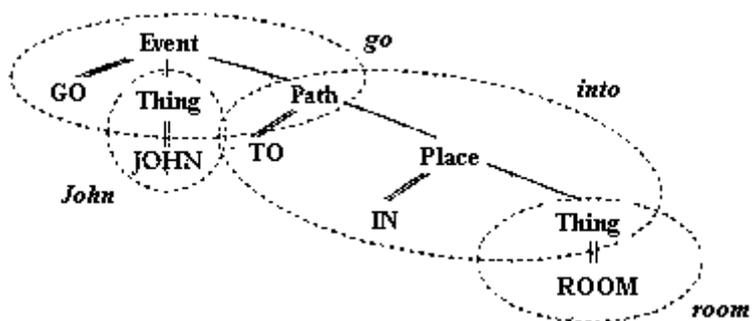


Figura 26. Representação da correspondência entre as estruturas conceitual e sintática da frase *John went into the room*.

Além disso, a estrutura conceitual dada em (56b) pode também ser expressa na frase do inglês *John entered the room* (“John entrou na sala”). Nessa forma de expressão, o verbo *enter* incorpora as funções TO e IN, que estão expressas separadamente por *into*, configurando a estrutura conceitual em (57b).

(57)

a. Estrutura Sintática

[John]_i entered _j [the room]_k

b. [Event GO ([Object JOHN]_i, [Path TO ([Place IN ([Object ROOM]_k))])]_j

c. Entrada lexical de *enter*

$$\left(\begin{array}{l} \text{enter } j \\ V_j \\ \text{---} \langle \text{SN}_k \rangle^{31} \\ [\text{Event GO } ([\text{Thing}]_i, [\text{Path TO } ([\text{Place IN } ([\text{Thing}]_k)]))]_j \end{array} \right)$$

Conforme mostra (57c), *enter* subcategoriza um complemento opcional SN, não SP, como *go*.

³¹ Os sinais <> indicam opcionalidade de argumentos.

No português, na frase, “João entrou na sala, o verbo “entrar” incorpora as funções GO e TO, (58d), deixando, para a preposição “em”, a tarefa de lexicalizar a função IN, conforme mostra (58c).

(58)

a. Estrutura Sintática

[João]_i entrou_j [em_k[a sala]_m]

b. Estrutura Conceitual

[Event GO ([Thing JOHN]_i, [Path TO ([Place IN ([Thing ROOM]_m)]_k)]_j]

c. Entrada lexical de “em”

$$\left(\begin{array}{l} \text{em}_i \\ P_i \\ \text{--- SN}_j \\ [\text{Place IN } ([\text{Thing }]_j)]_i \end{array} \right)$$

d. Entrada lexical de “entrar”

$$\left(\begin{array}{l} \text{entrar}_j \\ V_j \\ \text{---} \langle \text{SP}_k \rangle \\ [\text{Event GO } ([\text{Thing }]_i, [\text{Path TO } ([\text{Place }])]_k)]_j \end{array} \right)$$

Observe-se que o significado nuclear das frases “John foi para dentro da sala”, *John went into the room*”, “John entrou na sala” e “*John entered the room*” (55b) (Fig.26) pode ser descrito pela mesma estrutura conceitual. As indexações nos exemplos descritos demonstram a constatação de que a todo constituinte sintático de uma frase (S, SN, SAdj, SP, etc)

corresponde uma parcela da estrutura conceitual, mas nem todo constituinte conceitual realiza-se sintaticamente (Jackendoff, 1990, p.23).

Há ainda outros exemplos interessantes de verbos que “incorporam” o significado de preposições em seu significado lexical. Em “Paulo embolsou o dinheiro”, o verbo representa a classe dos verbos ‘denominais’ que sintaticamente incorporam a pseudo-preposição *en-* e semanticamente “incorporam” o significado dessa pseudo-preposição a seu significado lexical. Como mostra a fig.27, o verbo realiza as funções CAUSE, INCH, BE e IN.

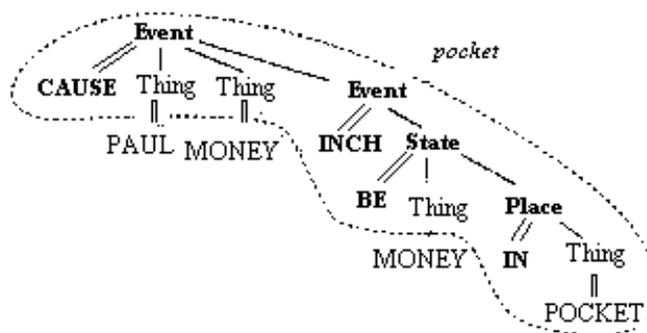


Figura 27. Representação da estrutura conceitual da frase *Paulo embolsou o dinheiro*.

A proposta de Jackendoff (1990, 2002), delineada nesta seção, é referência para a busca de funções semânticas primitivas que agem na construção da decomposição dos verbos, constructos essenciais para o estudo de como a semântica afeta as alternâncias sintáticas dos argumentos dos verbos. Sua proposta é também pertinente para pesquisas em PLN, como pode ser verificado em Dorr (1992), trabalho no qual a autora propõe a junção entre a semântica lexical e a frasal, utilizando-se dos componentes do significado propostos por Jackendoff para gerar frases que encontrem correspondentes em outras frases de outras línguas, criando um mecanismo de tradução automática.

Desse modo, como será feito na subseção 4.3, propõe-se que a representação léxico-semântica dos verbos constituintes de synsets das redes wordnets seja descrita por meio das funções e categorias ontológicas definidas por Jackendoff (1990, 2002).

3.6 A interdependência sintaxe-semântica: uma advertência

Jackendoff (1990, 2002) reflete o grau em que o comportamento sintático de um item lexical pode ser previsto a partir de sua semântica. Para ele, grande parte do comportamento sintático de um item reflete sua estrutura conceitual, embora afirme que nem todas as configurações sintáticas estão estritamente relacionadas a aspectos semânticos.

Para respaldar essa hipótese, exploram-se alguns fatos lingüísticos que focalizam a interface entre sintaxe e semântica, como, por exemplo, o número de argumentos sintáticos do verbo, que, como já foi dito, são as manifestações de seus argumentos semânticos.

Supõe-se que o número de argumentos sintáticos do verbo deve ser igual ou menor que o número dos seus argumentos semânticos (JACKENDOFF,2002); porém, essa afirmação não se sustenta, pois, em casos como os apresentados em (59), verifica-se que há mais argumentos sintáticos que semânticos.

(59)³²

- a. **Eles₁ se₂** casaram.
- b. **Ela₁** sonhou **um sonho horrível₂**.
- c. Choveu **uma chuva fina₁**.

³² Os algarismos subscritos sinalizam os argumentos.

³³ Excluindo-se o sentido em que “casaram-se” é recíproco

Em (59a), o pronome reflexivo “se” pode ser retirado, sem alteração substantiva do significado da frase³³; em (59b) e (59c), há objetos cognatos, que ocorrem somente quando os objetos possuem a mesma raiz do verbo: **sonh_** e **chov_**. Exemplos desse tipo sugerem que, em alguns casos, o número de argumentos sintáticos pode ser maior que o número de argumentos semânticos. Portanto, não se verifica a suposição do parágrafo anterior.

Um outro aspecto dos argumentos sintáticos que não pode ser totalmente previsto pela semântica do verbo é a determinação da categoria sintática em que um argumento semântico se realiza. Muitas vezes, a escolha do argumento sintático é indicada pelo argumento semântico. Por exemplo, quando o argumento semântico for um objeto concreto, incluindo pessoas, ele será representado na sintaxe por um SN. Entretanto, as proposições podem ser expressas por um SN ou por uma Oração, conforme ilustrado em (60a), e propriedades podem ser expressas por um SAdj ou por um SN, como mostra (60b).

(60)

- a. “João mencionou $\left. \begin{array}{l} \text{que Pedro havia traído Inês.} \text{ (Oração)} \\ \text{um boato desagradável.} \text{ (SN)} \end{array} \right\}$ ”
- b. “João ficou $\left. \begin{array}{l} \text{furioso.} \text{ (SAdj)} \\ \text{um leão.} \text{ (SN)} \end{array} \right\}$ ”

O valor semântico dos argumentos em (60a) e também em 60b, em uma determinada interpretação e em um certo nível de abstração, pode ser considerado o mesmo: “João mencionou X” (60a), em que X é uma proposição e “João ficou Y” (60b), em que Y é uma propriedade. Assim, observa-se que um mesmo argumento semântico é expresso por

categorias sintáticas diferentes. Entretanto, somente nesse nível de abstração é que a sintaxe não reflete a semântica.

Outra questão que se coloca para a investigação do mapeamento sintaxe-semântica diz respeito às preposições regidas por verbos e substantivos, conforme ilustram os exemplos (61a) e (61b).

(61)

- a. “Contamos **com** você” vs “Confiamos **em** você.”
- b. “Assistimos **ao** filme” vs “Vimos **o** filme.”

Os membros desses pares se diferem pouco quanto ao significado, mas apresentam preposições bem diferentes; a escolha das preposições pode ser, em alguns casos, semanticamente motivada, mas em verbos estreitamente relacionados como em (61a), a semântica parece não interferir na determinação da preposição, do contrário, as preposições seriam as mesmas. Ainda em (61b) “assistir” rege a preposição “a” e “ver” não rege nenhuma, apenas pede um artigo.

A posição dos argumentos sintáticos também nem sempre é motivada pela estrutura conceitual dos verbos. Quando os argumentos pertencem a categorias sintáticas diferentes, eles seguem a regra, estritamente sintática, definida em (62), ordem que é sempre seguida, exceto em situações em que a prosódia impõe sua própria ordem.

(62) Ordem dos argumentos sintáticos³⁴

SN > SAdj > SP > ORAÇÃO

Os argumentos que pertencem à mesma categoria sintática, por outro lado, (i) podem ser ordenados livremente, quando são SPs: “Conversamos longamente com Marta sobre

³⁴ O símbolo > indica precedência.

Pedro”/ “Conversamos longamente sobre Pedro com Marta” e (ii) devem obedecer a uma ordenação quando são SNs: “Considero Ana um gênio”/ *”Considero um gênio Ana”.

A ordem dos argumentos SNs múltiplos é determinada pelo papel temático do argumento semântico correspondente que, em geral, segue o disposto em (63) (GRIMSHAW, 1990). Trata-se de uma regra de associação que não é nem puramente semântica nem puramente sintática, mas está na interface desses dois domínios.

(63) *Hierarquia de associação para os SNs argumentos.*

Agente > Recipiente (objeto indireto de “dar”, por exemplo) > Tema (entidade que sofre mudança) > Local > SN predicativo

De acordo com (63), é preciso apenas listar os argumentos semânticos e verificar se são obrigatórios ou opcionais na expressão dos argumentos sintáticos. Como exemplos, podem ser retomados os verbos “comer” e devorar”, mencionados na subseção 2.1. Observe-se, em (64), que o índice *ob* indica a obrigatoriedade do argumento e a não ocorrência desse índice indica opcionalidade.

- (64)
- a. [DEVORAR (X_{ob}, Y_{ob})]
 - b. [COMER (X_{ob}, Y)]

Como conclusão, constata-se que o número, o tipo e a ordem dos argumentos sintáticos de uma frase decorrem, mas em parte, da estrutura semântica. Há aspectos do número e da categoria dos argumentos sintáticos que são consequência de características idiossincráticas dos itens lexicais. Além disso, há também aspectos da estrutura sintática que são consequências de princípios que agem na interface sintaxe-semântica. Essas

considerações relativizam a hipótese de que a semântica determina a sintaxe e, seguindo Jackendoff (1990,2002), sugerem que não há supremacia nem da sintaxe e nem da semântica na explicação do que ocorre na estrutura de argumentos do um verbo. A estrutura de argumentos só pode ser compreendida ao se investigarem as regras de formação sintática, as regras de formação conceitual e os princípios de interação entre essas duas dimensões (sintática e semântica).

Seção 4 A sintaxe e semântica dos verbos do tipo “roubar” da WordNet.Br

4.1 Análise preliminar *Synset 3742*

Conforme destacado na Introdução, um dos objetivos deste trabalho é o refinamento e aperfeiçoamento dos synsets de verbos da rede WordNet.Br com base na hipótese de que é possível prever-se o comportamento sintático dos verbos a partir de suas estruturas conceituais. Assim, um conjunto de verbos do português do Brasil, o **Synset 3742**, e nocionalmente correspondente a uma subclasse específica de verbos originalmente isolada por Levin (1993), foi selecionado para aplicação da estratégia de análise exposta nas seção 2, em especial na subseção 2.1, para que se verifique sua validade como método de descrição das propriedades léxico-semânticas e léxico-sintáticas dessa classe de verbos do português, bem como sua aplicabilidade na construção e no refinamento de bases de verbos de redes wordnets.

A subclasse denominada *Steal Verbs* de verbos do inglês, retirada da classe original de Levin (1993) e por ela denominada “Verbos de Destituição de Posse-Verbos do Tipo ‘Roubar’” (*Verbs of Possessional Deprivation-Steal Verbs*), aproximadamente corresponde ao **Synset 3742** do português.³⁵

Steal Verbs: {*abduct* (“seqüestrar”, “raptar”), *cadge* (“filar”), *capture* (“capturar”), *confiscate* (“confiscar”), *cop* (“furtar”, “surrupiar”), *embezzle* (“desfalcar”), *extort* (“estorquir”), *extract* (“arrancar”), *filch* (“surrupiar”), *grab* (“surrupiar”), *impound* (“confiscar”), *kidnap* (“seqüestrar”), *liberate* (“pilhar”), *lift* (“furtar”), *pilfer* (“surrupiar”),

³⁵ Essa subclasse resultou da exclusão dos verbos do inglês que deveriam apropriadamente ser classificados como verbos de ‘restituição de posse’ (e não de “Destituição de Posse”), como *emancipate* (“resgatar”), *recover* (“reaver”), *redeem* (“resgatar”), *reclaim* (“reclamar”), *regain* (“recuperar”), *repossess* (“reempossar”), *rescue* (“resgatar”), *retrieve* (“reaver”) e *weasel* (“ludibriar”).

pinch (“furtar”), *pirate* (“piratear”), *plagiarize* (“plagiar”), *purloin* (“afanar”), *rustle* (“roubar animais de fazendas”), *seize* (“empossar”), *smuggle* (“contrabandear”), *snatch* (“pegar”), *sneak* (“roubar”), *sponge* (“explorar”), *steal* (“roubar”), *swipe* (“furtar”), *take* (“tomar”, “pegar”), *thieve* (“roubar”), *winkle* (“arrancar”), *withdraw* (“retirar”), *wrest* (“tirar a força”)}

Synset 3742 {abafar, afanar, agadanhar, arrancar, arrapinar, arrebatado, bater, bifar, bispar, comer, defraudar, despojar, desvalijar, empalmar, escamotar, escamotear, escorchar, esponjar, furtar, gatear, gatunar, gatunhar, ladroar, larapiar, limpar, palmar, palmear, pelar1, rapar, rapinar, rapinhar, roubar, safar, saltear, saquear, sobnegar, sonegar, subtrair, surripiar, surrupiar, tirar, tomar, unhar}

Os verbos compilados por Levin descrevem a remoção de algo (Localizado/Tema) de algum lugar ou da posse de alguém (Local) ou para o benefício de alguém (Beneficiário). No inglês, esses verbos podem ocorrer com a preposição *from*, que sinaliza o Local, mas não participam da **alternância locativa**, conforme mostra (65); ocorrem com a preposição *for*, que sinaliza o Beneficiário, mas não participam da **alternância benefactiva** (66)³⁶.

(65). *The thief stole*[Localizado *the painting*] [*Local from the museum*].

(“O ladrão roubou [Localizado a pintura] [Local do (que estava no) museu].”)

**The thief stole*[*Local the museum*] [*Localizado of the painting*].

(* “O ladrão roubou [Local o museu] [Localizado da pintura].”)³⁷

(66). *The thief stole*[Tema *the painting*] [*Beneficiário for Mr. Smith*].

(“O ladrão roubou [Tema a pintura] [Beneficiário para Mr. Smith].”)

**The thief stole* [*Beneficiário Mr. Smith*] [*Tema the painting*].

(* “O ladrão roubou [Beneficiário Mr. Smith] [*Tema a pintura*].”)

³⁶ Essa alternância se caracteriza por apresentar uma variante com sintagma preposicional SN1 V SN2 [_{SP} for SN3] e outra variante apresentando o objeto duplo SN1 V SN3 SN2, como exemplo *Bill carved a toy for Tom/Bill carved Tom a toy*.

³⁷ Observe-se que, no português, a construção com a preposição “a” é possível: “Jamais roubarão [_{Local} ao tempo] [_{Localizado} a faculdade de dilatar-se nos momentos de felicidade]” (HOUAISS, 2001).

De acordo com a metodologia de análise proposta na seção 2, analisam-se as alternâncias das quais os verbos do Synset 3742 participam, de modo que seja feita a reestruturação desse synset.³⁸ Observe-se que os verbos desse synset apresentam um comportamento variado quanto à expressão sintática dos argumentos.

A filtragem preliminar do Synset 3742, recorrendo-se às informações lexicográficas registradas no dicionário Michaelis. (WEISZFLOG, 1998), permite que se excluam desse synset formas que codificam o registro informal, o sentido figurado, regionalismos, gírias e variantes.³⁹ Como resultado, o synset reduz-se para 12 verbos, conforme mostra o Synset 3742 _Resultante (**Synset 3742R**).

Synset 3742R: {comer; despojar; escamotear; escorchar; furtar; rapinar; ratonear roubar; safar; saquear; subtrair; tirar; tomar}

O passo seguinte consiste na análise do synset por meio da seleção de tipos de frases-exemplo adaptadas do dicionário Michaelis e de cópulas,⁴⁰ para cada um desses verbos, com o objetivo de (i) isolar as diferentes estruturas de argumentos em função das alternâncias de que os verbos participam e, como resultado, (ii) reagrupar os verbos em novos synsets.

Como resultado dessa tarefa inicial, discriminam-se quatro synsets: **Synset_3742a**, **Synset_3742b**, **Synset_3742c**, **Synset_3742d**,⁴¹ que participam, respectivamente, das seguintes alternâncias.

³⁸ Observe-se que esse synset contém 43 verbos, o que dificulta a apreensão do significado nele implicitamente codificado.

³⁹ É importante ressaltar que se faz esse recorte apenas por simplificação metodológica. Não se desconhece o fato de que essas formas têm sua importância e, como as outras, também participam de alternâncias.

⁴⁰ Por se tratar apenas de um exercício exploratório não-exaustivo, adaptaram-se abonações do dicionário, observando-se, entretanto, o uso com o recurso da consulta ao Córpus do Nilc e a textos de português do Brasil disponibilizados na Web. O primeiro está disponível para consulta no site <http://acdc.linguatca.pt/acesso/>.

⁴¹ As frases-exemplo estão dispostas seguindo a seguinte ordem: do verbo intuitivamente considerado mais prototípico para o menos prototípico.

Synset 3742_a: {furtar, rapinar, roubar}

Variante intransitiva:

[A1=Agente-Suj]

Subcategorização: _ V_j⁴²

Frases-exemplo:

Aquelas crianças (A1) só **roubavam**.

Maria não (A1) **furtou** mais

Bush (A1) **rapina** nos EUA, na Europa, na África, na Ásia e nas Américas Central e do Sul.

Synset 3742b: {despojar, furtar, saquear, roubar}

Variante transitiva com objeto Local:

[A1=Agente-Suj, A3=Local-Obj]

Subcategorização: _ V SN_n

Frases-exemplo:

Um grupo de quatro indivíduos (A1) agrediu e **roubou um rapaz**_n (A3).

Uma quadrilha (A1) prendeu dois padres e **roubou a igreja**_n (A3).

Polícia prendeu ontem Silvia Helena da Silva (A1), que **furtou 50 residências**_n (A3).

Travestis (A1) são presos após **saquear turista**_n (A3).

Os vikings (A1) **saquearam as igrejas**_n (A3), apoderando-se das riquezas.

Os gatunos (A1) **despojaram a igreja**_n (A3).

Synset 3742c: {comer, escamotear, furtar, rapinar, roubar, saquear, subtrair, tirar, tomar}

Variante transitiva com objeto Localizado:

[A1=Agente-Suj, A2=Localizado-Obj]

Subcategorização: _ V SN_m

Frases-exemplo:

O piloto (A1) **roubou a caixa preta do avião**_m (A2).

Os ladrões (A1) **furtaram os computadores**_m (A2).

Eles (A1) **tiraram meu dinheiro**_m (A2), deixando minha carteira vazia!

⁴² Nesta subseção propõe-se uma descrição preliminar da representação da dimensão sintática da entrada lexical dos verbos, representação que será mais detalhada na subseção seguinte.

Eles (A1) **tomaram** meu dinheiro_m (A2).

Michael Jackson (A1), por exemplo, **rapinou** as poções_m (A2).

Tropas da UNITA (A1) **saquearam** os bens da população_m (A2).

O PRI (A1) **subtraiu** US\$ 120 milhões_m (A2).

Algumas empreiteiras (A1) **comeram** muito dinheiro público_m (A2).

Os ladrões (A1) **escamotearam** uma carteira_m (A2).

Synset_3742d: {roubar, subtrair, safar}

Variante transitiva com os objeto Local e Localizado:

[A1=Agente-Suj, A2=Localizado-Obj, A3=Local-Obl]

Subcategorização: _ V SN_m SP_k

Frases-exemplo:

O governo (A1) **roubou** a herança_m (A2) ao herdeiro_k (A3).

O moleque (A1) **subtraiu** uma maçã_m (A2) ao quitandeiro_k (A3).

Os gatunos (A1) **safaram** a carteira_m (A2) ao transeunte_k (A3)

Com esse exercício, espera-se ter demonstrado a relevância e a aplicabilidade desse modelo de análise lexical na construção e no refinamento dos synsets da base de verbos da rede WordNet.Br.

4.2 Proposta de descrição da dimensão léxico-sintática

Recorde-se que, conforme exposto na subseção 3.1, a descrição lingüística do verbo deve contemplar, pelo menos, duas dimensões: a léxico-semântica e a léxico-sintática. Assim, nesta subseção, sugere-se uma representação léxico-sintática para os verbos dos synsets reagrupados da WordNet.Br, com no que se discutiu nas subseções 3.2, 3.3 e 3.4.

Para cada um dos contextos sintáticos em que o verbo “roubar” ocorre, isto é, para cada uma das alternâncias de que ele participa, propõe-se uma frase-tipo que reflita as

configurações sintáticas e semânticas isoladas na subseção anterior.⁴³ Em (67a), (67b), (67c), (67d), apresentam-se as frases-tipo, bem como as respectivas alternâncias de que participam os verbos e os seus esquemas de subcategorização.

(67)

a. **Synset 3742_a:** {furtar, rapinar, roubar}

Variante intransitiva:

[A1=Agente-Suj]

Subcategorização: _ V_j

Frase-exemplo: [Políticos]_i [roubam]_j

b. **Synset 3742_b:** {despojar, furtar, saquear, roubar}

Variante transitiva com objeto Local:

[A1=Agente-Suj, A3=Local-Obj]

Subcategorização: _ V_j SN_n

Frase-exemplo: [Políticos]_i [roubam]_j [a nação]_n

c. **Synset 3742_c:** {comer, escamotear, furtar, rapinar, roubar, saquear, subtrair, tirar, tomar}

Variante transitiva com objeto Localizado:

[A1=Agente-Suj, A2=Localizado-Obj]

Subcategorização: _ V_j SN_m

Frase-exemplo: [Políticos]_i [roubam]_j [a esperança]_m

d. **Synset 3742_d:** {roubar, subtrair, safaram}

Variante transitiva com os objeto Local e Localizado:

[A1=Agente-Suj, A2=Localizado-Obj, A3=Local-Obl]

Subcategorização: _ V_j SN_m SP_k

Frase-exemplo: [Políticos]_i [roubam]_j [a esperança]_m [à nação]_k

⁴³ Essa estratégia visa a facilitar a compreensão e a construção das estruturas sintáticas e conceituais para os verbos dos synsets.

Os verbos do **Synset_3742a**, em (67a), apresentam um argumento externo A1, semanticamente Agente e não apresentam nenhum objeto. Desse modo, não apresentam especificador. A fig. 28 ilustra a representação para “roubar”⁴⁴ na frase-exemplo “Políticos roubam”. Note-se que se trata de uma estrutura monádica, já que o argumento externo, sujeito sintático, na faz parte dos argumentos internos projetados pelo item lexical.

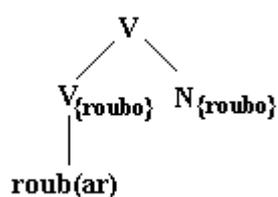


Figura 28. Estrutura de argumentos do verbo “roubar” no Synset_3742a.

A fig.28 demonstra que **roub_**, que contém uma matriz fonológica como parte de sua entrada lexical, é inserido diretamente na posição verbal, por meio da Inserção Vocabular, desobrigando a necessidade de haver Fusão. Além disso, os índices que estão entre as chaves na estrutura (como ocorre na fig.28) denotam uma seleção semântica entre o verbo “roubar” e o objeto, que no caso está implícito.

O synset **Synset_3742b**, em (67b), agrupa verbos que apresentam um argumento externo e um argumento interno Local, que é o especificador. A fig.28 ilustra a configuração do verbo “roubar” na frase-exemplo “Políticos roubam a nação.”

⁴⁴ Escolheu-se o verbo “roubar” para a aplicação das teoria pois ele é intuitivamente considerado o verbo mais prototípico. As estruturas sintáticas e conceituais propostas para cada um dos significados desse verbo poderão ser estendidas para os demais verbos dos synsets.

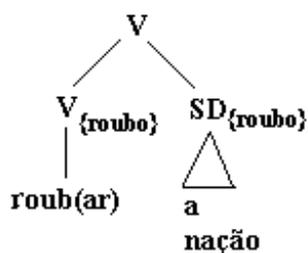


Fig.29. Estrutura de argumentos do verbo “roubar” no Synset_3742b.

Do mesmo modo que na estrutura representada na fig.29, o item lexical, sem categoria definida a priori, **roub_** é inserido diretamente no núcleo verbal, tornando-se verbo. Por meio da representação dos índices, há a classificação semântica entre o verbo “roubar” e seu objeto sintático, o especificador “a nação”, que nesse caso é explícito.

Os verbos do **Synset_3742c**, em (67c), também apresentam um argumento interno, mas que é classificado como Localizado, o que não modifica a estrutura de argumentos do verbo em relação à fig.29. A fig.30 ilustra a configuração do verbo “roubar” na frase-exemplo “Políticos roubam a esperança”.

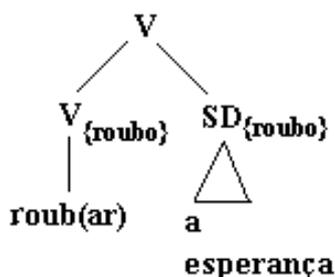


Fig.30. Estrutura de argumentos do verbo “roubar” no Synset_3742c.

Por fim, o **Synset_3742d**, em (67d), é constituído por verbos que apresentam dois argumentos internos, um Local e outro Localizado. A estrutura de argumentos da fig.31 ilustra o verbo “roubar” na frase-exemplo “Políticos roubam a esperança à nação”. Essa representação tem dois índices que mostram a relação semântica entre o verbo “roubar” (cuja

raiz **roub_** foi inserida em V) e o substantivo “roubo”, indexado com {}, sinalizando os objetos explícitos da ação de roubar: “a esperança” e “a nação”.

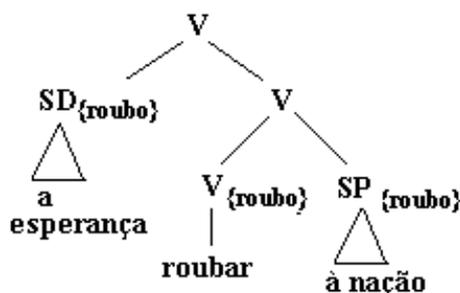


Figura 31. Estrutura de argumentos do verbo “roubar” no Synset_3742d.

Assim, finaliza-se a representação da dimensão léxico-sintática dos verbos que compõem os synsets da WordNet.Br.

Evidencia-se, portanto, a importância da teoria de Hale e Keyser (2002) na descrição dessa dimensão lexical dos verbos.

4.3 Proposta de descrição da dimensão léxico-semântica

Conforme demonstrado na subseção 4.1, um synset inicialmente contendo 43 unidades foi reestruturado em termos de quatro novos synsets. Essa reanálise do synset original foi feita seguindo-se a orientação discutida nesta dissertação de se agrupar, em classes semânticas, os verbos que compartilham uma mesma configuração sintática, no que se refere às alternâncias de expressão dos seus argumentos. Assim, como cada um dos synsets deve apontar para um conceito distinto e todos os verbos do synset devem conformar-se a esse conceito, dado que eles estão agrupados no mesmo synset por manterem entre si a relação de sinonímia, cada um dos synsets deve, então, estar associado a uma estrutura conceitual específica, que os defina e que seja comum a todos os verbos que o constituem. Assim, nesta última subseção do

exercício exploratório, propõe-se a descrição conceitual projetada pelo verbo “roubar”, ilustrando, assim, a representação léxico-semântica dos mesmos synsets acima referidos neste parágrafo.⁴⁵

O verbo “roubar” pode realizar-se com apenas um argumento externo, caracterizado como um SN Agente (“os políticos”), conforme (67a); ou pode apresentar, além desse argumento externo, um argumento interno, um SN Local (“a nação”), conforme (67b), ou um SN Localizado (“a esperança”), conforme (67c), ou ainda dois argumentos internos, um SN Localizado (“esperança”) e um SP Local (“à nação”), conforme (67d). Sugere-se, então, que cada uma das variantes listadas em para “roubar” [(67a), (67b), (67c) e (67d)] e, conseqüentemente, cada um dos demais verbos dos synsets, deve estar associada a uma estrutura conceitual, já que cada alternância remete a um significado do verbo “roubar”. Desse modo, em (68) representam-se as estruturas conceituais para cada um dos quatro synsets, que são estruturalmente idênticas, mas distinguem-se uma das outras no que se refere às indexações, o que sugere que a realização sintática reflete o componente de significado que está sendo indexado.

(68)

a. [Event CAUSE ([Thing X]_i, [Event INCH ([State NOT BE ([Thing], [Place IN ([Thing])]))])]_j

Esquema sintático: [x_{Agente}]_i [roubam]_j

b. [Event CAUSE ([Thing X]_i, [Event INCH ([State NOT BE ([Thing], [Place IN ([Thing y]] n)])])]_j

Esquema sintático: [x_{Agente}]_i [roubam]_j [y_{Local}]_n

⁴⁵ Este exercício de aplicação do modelo de representação da estrutura conceitual na descrição do significado dos verbos dos synsets selecionados apresentará um esboço de como podem ser representadas as estruturas conceituais básicas.

c. [Event CAUSE ([Thing X]_i, [Event INCH ([State NOT BE ([Thing z]_m, [Place IN ([Thing])]))])])]_j

Esquema sintático: [x_{Agente}]_i [roubam]_j [z_{Localizado}]_m

d. [Event CAUSE ([Thing X]_i, [Event INCH ([State NOT BE ([Thing z]_m, [Place]_k)]))])]_j⁴⁶

Esquema sintático: [x_{Agente}]_i [roubam]_j [z_{Localizado}]_m [y_{Local}]_k

Assim, para os significados de “roubar”, aqui discutidos, propõe-se a entrada lexical ilustrada em (70).

(70) Entrada lexical de “roubar”

	<p>roubar_j</p> <p>V_j</p> <p>— <SN_n> <SN_m> <SP_k></p> <p>a. [Event CAUSE ([Thing]_i, [Event INCH ([State NOT BE ([Thing], [Place IN ([Thing])])])])])]_j</p> <p>b. [Event CAUSE ([Thing]_i, [Event INCH ([State NOT BE ([Thing], [Place IN ([Thing])] n)]))])]_j</p> <p>c. [Event CAUSE ([Thing]_i, [Event INCH ([State NOT BE ([Thing]m, [Place IN ([Thing])])])])])]_j</p> <p>d. [Event CAUSE ([Thing]_i, [Event INCH ([State NOT BE ([Thing]m, [Place]k)]))])]_j</p>	
--	---	--

Conforme a discussão promovida nesta subseção, verifica-se que os modos alternativos de expressão dos argumentos do verbo “roubar” exigem uma representação conceitual distinta, para demonstrar os diferentes significados desse verbo. Assim, propõe-se uma única entrada para “roubar”, representada em (70); essa entrada estabelece as possíveis configurações sintáticas dos diferentes significados elementares do verbo, por meio da indexação dos argumentos opcionais.

⁴⁶ Entrada lexical da preposição “a”

<p>a_j</p> <p>P_j</p> <p>— SN_k</p> <p>[Place IN ([Thing] n)]_j</p>

Seção 5 Considerações finais

Apresentou-se, nesta dissertação, um estudo das teorias que analisam a associação entre semântica e sintaxe, a partir da investigação da hipótese de que a estrutura de argumentos projetada por um verbos reflete aspectos de sua estrutura conceitual.

Primeiramente, apresentou-se a teoria de Levin e Rappaport-Hovav (1991), que, a partir da comparação entre os verbos que participam do mesmo grupo de alternâncias sintáticas, identificam a classe semântica à qual esses verbos pertencem. Nesse sentido, reagrupam-se os verbos do tipo “*remove*”, que pareciam pertencer à mesma classe semântica, em três novas classes semânticas: a classe dos verbos do tipo “*remove*”, a classe dos verbos do tipo “*clear*” e a classe dos verbos do tipo “*wipe*”. Discutiu-se também a *Hipótese da Inacusatividade*, que preconiza a existência de duas classes de verbos intransitivos: a classe dos verbos inergativos e a classe dos verbos inacusativos, cada qual associada a uma representação sintática distinta. O fenômeno da Inacusatividade foi considerado semanticamente motivado e sintaticamente codificado, contribuindo, portanto, para a investigação da interface semântico-sintática.

De Levin (1993), extraiu-se uma classe de verbos, classificada de acordo com a hipótese de que há aspectos do significado que são sintaticamente relevantes, que serviu como parâmetro para a escolha da classe de verbos do português analisada nesta dissertação: a classe dos verbos do tipo “roubar”, semanticamente coerente com a classe dos verbos do tipo “*steal*”, discriminada por Levin. A classe dos verbos do tipo “roubar”, nocionalmente equivalente à classe dos verbos do tipo “*steal*”, foi retirada da base de verbos da rede WordNet.Br, que agrupa, por meio da relação léxico-semântica de sinonímia, itens lexicais do

português em synsets, que são deverão ser relacionados entre si por relações de hiponímia, hiperonímia, causa, acarretamento, entre outras relações léxico-conceituais.

Seguindo a proposta das autoras citadas, aplicou-se a análise dos verbos que apresentam os argumento Local e Localizado na análise do Synset 3742, da WordNet.Br, e, desse modo, desmembrou-se esse synset em quatro synsets distintos.

A teoria de Levin e Rappaport-Hovav (1995) contribuiu, ainda, com a proposta de que existem duas dimensões de representação lexical dos verbos: uma dimensão léxico-sintática e outra léxico-semântica. Essa proposta encaminhou o estudo da teoria de Hale e Keyser (2002) e culminou com a proposta de um esquema de representação sintática da estrutura de argumentos projetada pelo verbo. A teoria de Hale e Keyser foi utilizada na proposição da representação da estrutura léxico-sintática projetada pelos verbos de cada synset desmembrado do Synset 3742.

Jackendoff (1990, 2002), por sua vez, contribui para a investigação das funções semânticas primitivas que são utilizadas na descrição da dimensão léxico-semântica dos verbos. Em conformidade com esse autor, enfatizou-se, nesta dissertação, que aspectos do significado do verbo estão refletidos na sua sintaxe. A partir dessa discussão, propôs-se a representação da dimensão léxico-semântica para os verbos constituintes dos synsets Synset3742-a, Synset 3742-b, Synset 3742-c e Synset 3742-d da WordNet.Br.

Do ponto de vista lingüístico, esta dissertação contribui com um estudo exploratório das propriedades léxico-semânticas e léxico-sintáticas de um conjunto de verbos, a partir do estudo das teorias acima mencionadas. Contribui também para uma primeira descrição das dimensões léxico-semântica e léxico-sintática da classe dos verbos de “roubar” do português.

Do ponto de vista do domínio lingüístico-computacional, o estudo da interface entre semântica e sintaxe proporciona um método de refinamento de bases de verbos computacionais, em particular, da base de verbos da rede WordNet.Br.

Ressalta-se ainda que essa pesquisa de Mestrado se desenvolveu em consonância com um projeto ainda maior que é a construção da base da rede WordNet.Br, coordenada pelo Prof. Dr. Bento Carlos Dias da Silva (FCL/Unesp). É uma pesquisa que não se desenvolveu separadamente de outros estudos que fazem parte desse projeto maior, portanto, seus resultados, teóricos e práticos, podem ser aplicados e aprofundados no desafio de refinar e construir a WordNet.Br.

Referências

ÁVILA, M. C. Escalpar, escarpelar e escarpelizar devem ser termos sinônimos na WordNet.Br.? In: **XV Congresso de Iniciação Científica da Unesp**. Marília-SP, 2003. 1CD-ROM

BAKER, M. C. **Incorporation**. A theory of Gramatical Function Changing. Chicago: The University of Chicago Press, 1988.

BROWN, K.; MILLER, J. (Ed.) **Concise encyclopedia of syntactic theories**. Oxford: Elsevier Science, 1996.

BORBA, F.S. **Uma gramática de valências para o português**. São Paulo: Ática, 1996.

_____. **Dicionário de Usos do Português do Brasil**. São Paulo: Ática, 2002.

CRUSE, D. A. **Lexical semantics**. New York: Cambridge University Press, 1986.

CHOMSKY, N. **Lectures on Government and Binding**. Dordrecht: Foris, 1991.

_____. **Knowledge of Language: it's nature, origin and use**. Westport: Praeger Publishers, 1986.

DIAS-DA-SILVA, B. C. **A face tecnológica dos estudos da linguagem: o processamento automático das línguas naturais**. Araraquara, 1996, 272 f. Tese (Doutorado em Letras) - Faculdade de Ciências e Letras, UNESP, Araraquara, 1996.

_____. Reusability of dictionaries in the compilation of NLP lexicons. In: N.J.MAMEDE; J.BAPTISTA; I.TRANCOSO; M.G.V.NUNES (eds.) **Computational processing of the portuguese language**. Berlin: Springer-Verlag, p.78-85, 2003a.

_____. Human language technology research and the development of the Brazilian Portuguese Wordnet. In: Hajičová, E., Kotěšovcová, A., Mírovský, J. (Ed.).

INTERNATIONAL CONGRESS OF LINGUISTS, 17, 2003, Prague. **Proceedings...** UK: Matfyzpress, 2003b. 1 CD

DIAS-DA-SILVA, B.C.; OLIVEIRA, M.F., MORAES, H.R. Groundwork for the development of the Brazilian Portuguese Wordnet. In: E.M. RANCHHOD; N.J. MAMEDE (eds.) **Advances in natural language processing**. Berlin: Springer-Verlag, 2002. p.189-196.

DIAS-DA-SILVA, B.C.; OLIVEIRA, M.F.; HASEGAWA, R.; MORAES, H.R.; AMORIM, D.; PASCHOALINO, C.; NASCIMENTO, A.C. A construção de um thesaurus eletrônico para o português do Brasil. In: PROPOR - ENCONTRO PARA O PROCESSAMENTO COMPUTACIONAL DA LÍNGUA PORTUGUESA ESCRITA E FALADA, 5, 2000, Atibaia. **Proceedings...** Atibaia, 2000.

DORR, B. J. **Machine Translation**: a view from the lexicon. Cambridge, Mass.: The MIT Press, 1992.

FELLBAUM, C. (ed.) **WordNet**: an electronic lexical database. Cambridge, Mass.:The MIT Press, 1998.

FERNÁNDEZ, Ana et al. Establishing Semantic Oppositions for the Typification of Predicates: On the Concept of Diathesis Alternations as Semantic Oppositions. **Language Design**, Granada, vol.II, 1999. Disponível em <<http://ashda.ugr.es/laboratorio/language.htm>> Acesso em 23 nov de 2004

FERREIRA, A.B. **Dicionário Aurélio eletrônico século XXI** (Versão 3.0). São Paulo: Lexikon Informática Ltda.,1999.1 CD-ROM

FILLMORE, C. The case for case. In: _____BACH, E.; HARMS, R. **Universals in Linguistic Theory**. New York: Holt, Rinehart and Winston Inc., 1968.

GOLBERG, A.E. Construction grammar. In: BROWN, K; MILLER; J. **Concise encyclopedia of syntactic theories**. Cambridge: Cambridge University Press, 1996. p.68-71

GRIMSHAW, J. **Argument structure**. Cambridge, Massachusetts: The MIT Press, 1994.

GRUBER, J. **Studies in lexical relations**.Cambridge, Massachusetts: The MIT Press, 1965.

HAIMAN, J. Dictionaries and encyclopedias. **Lingua**, p. 329-357, 1980.

HALE, K.; KEYSER, S.J. **Prolegomenon to a Theory of Argument Structure**. Cambridge: The MIT Press, 2002.

HOUAISS, A. **Dicionário Eletrônico do Português (Versão 1.0)**. Rio de Janeiro, 2001. 1 CD-ROM

JACKENDOFF, R. **Semantic Structures**. Cambridge: The MIT Press. 1990

_____. **Foundations of Language**. Oxford: Oxford University Press 2002.

JURAFSKY, D.; MARTIN, J. H. **Speech and Language Processing: an introduction to natural language processing, computational linguistics and speech recognition**. United states of America: Prentice Hall, 2000.

LEVIN, B. **English Verb Classes and Alternations : a preliminary investigation**. Chicago: The University of Chicago Press, 1993.

LEVIN, B.; HOVAV, M. R. **Unaccusativity: at the syntax-lexical semantics interface**. Cambridge: The MIT Press, 1995.

LEVIN, B.; RAPAPPORT HOVAV, M. Wiping the slate clean: a lexical semantic exploration. **Cognition**, vol. 41, p. 123-151, 1991.

MILLER, G. A. et al. **Five Papers on Wordnet**. 1993. Disponível em <<http://www.cogsci.princeton.edu/~wn>>. Acesso em 12 set-2003.

MILLER, G.; FELLBAUM, C. Semantic networks of English. **Cognition**. vol. 41, p. 197-229, 1991.

OLIVEIRA, A. E. B., Ávila, M.C. Subsídios para a construção da rede Wordnet para o português do Brasil. In: 51º Seminário do GEL (Grupo de Estudos Linguísticos do Estado de São Paulo), 2003, Taubaté. **Caderno de Resumos do GEL**, 2003.

OLIVEIRA, A. E. B., Ávila, M.C. A face linguística da rede Wordnet para o português do Brasil: uma análise léxico-semântica e lógico-conceitual de verbos e adjetivos. In: IV Encontro de Iniciação Científica - VII Mostra de Pós-Graduação da Universidade Presbiteriana Mackenzie, 2003, São Paulo. **Anais - IV Encontro de Iniciação Científica**, 2003. p.306

PUSTEJOVSKY, J. **The Generative Lexicon**. Massachusetts: MIT Press, 1996.

RADFORD, A. **Syntactic theory and the Structure of English: a minimalist approach**. Cambridge: Cambridge University Press, 1997.

ROSEN, C.G. The Interface between Semantic Roles and Initial Gramatical Relations. In: PERLMUTTER & ROSEN (eds) **Studies in Relational Grammar**. Chicago: University of Chicago Press, 1984

TALMY, L. Lexicalization patterns: semantic structures in lexical forms. IN: SHOPEN, T. (ed). **Language Typology and Syntactic Description**. Cambridge. Cambridge University Press, 1995. v3.

THE RANDOM HOUSE Webster's- Unabridged Eletronic Dictionary. Novell Inc., 1996. 1 CD-ROM

WEBSTER's Dicionário Eletrônico (inglês-português/português-inglês). Lexikon Informática Ltda., 1982. 1 CD-ROM

TSUNODA, T. Transitivity. In _____. BROWN, K.; MILLER, J. (Ed.) **Concise encyclopedia of grammatical categories**. Oxford: Elsevier Science, 1999.

VAN VALIN, R. D. Jr. Parameters of Split Intransitivity. **Language**, v. 66, n. 22, p 221-260, 1990.

VOSSEN, P. Introduction to EuroWordNet. **Computers and the humanities**. vol. 32, p. 73-89, 1998

WEISZFLOG, W (ed.) **Michaelis português-moderno dicionário da língua portuguesa** (Versão 1.0). São Paulo: DTS Software Brasil Ltda. 1998.