

# Descrição do verbo *pegar* para identificação de verbo suporte

Lorena Silva Teixeira da Cruz

<sup>1</sup>Universidade Federal do Espírito Santo  
Vitória - Espírito Santo 29.075-910, Brasil

<sup>2</sup>Fundação de Amparo à Pesquisa do Espírito Santo - FAPES  
Vitória - Espírito Santo 29.075-910, Brasil

[lorenasteixeira@hotmail.com](mailto:lorenasteixeira@hotmail.com)

## RESUMO

Este trabalho apresenta uma descrição do verbo *pegar* a partir de suas propriedades morfo-sintáticas, semânticas e pragmáticas em contexto de uso observando-se as exigências e restrições dessas propriedades e levando-se em conta o sujeito e a natureza dos complementos que se relacionam com ele. O referencial teórico utilizado será a teoria do Léxico-gramática (Gross, 1975) que tem como objetivo reconhecer padrões de regularidades nas línguas por meio da aplicabilidade de critérios sintático-semânticos formais, considerando o julgamento de aceitabilidade por parte de falantes nativos em relação à gramática da língua e ao sentido reconhecido pelo uso.

## 0 INTRODUÇÃO

Esta pesquisa pretende identificar construções com o verbo *pegar* funcionando como verbo suporte, construção esta definida como “um tipo específico de construção complexa, constituído fundamentalmente em torno de um verbo e de um nome com função predicativa” (Athayde, 2001).

Além da identificação de verbo suporte, o presente trabalho pretende demonstrar a importância da descrição de itens lexicais para o processamento automático de linguagem natural na forma de dicionário eletrônico, demonstrando a relevância da elaboração de softwares que lidam com processamento de linguagem natural e requerem a transformação da língua para ser lida por uma máquina.

O referencial teórico utilizado será a teoria do Léxico-gramática (Gross, 1975) que tem como objetivo reconhecer padrões de regularidades nas línguas por meio da aplicabilidade de critérios sintático-semânticos formais, considerando o julgamento de aceitabilidade por parte de falantes nativos em relação à gramática da língua e ao sentido reconhecido pelo uso.

Para a identificação de verbo suporte será utilizado como base para os exemplos construções com o verbo *pegar*. Os exemplos analisados foram construídos por falantes nativos da Língua Portuguesa, prática esta respaldada na Teoria do Léxico-gramática.

O objetivo principal deste estudo é demonstrar a importância da pesquisa linguística em diversas áreas, e as possibilidades de estudo que a linguística pode proporcionar aos estudiosos da área.

## 1 A TEORIA DO LÉXICO-GRAMÁTICA

A teoria do Léxico-gramática foi desenvolvida pelo linguista francês Maurice Gross (1975, 1981, 1994), juntamente com a equipe do Laboratoire d'Automatique Documentaire ET Linguistique (LADL) (SANTOS 2011, p. 27) com o intuito de instituir uma metodologia capaz de descrever de forma morfossintático-semântica os componentes do léxico de uma determinada língua.

De acordo com LAPORTE (2008 p. 32), para dar uma fundamentação teórica à sua Teoria, Maurice Gross valeu-

se dos estudos do linguista Zellig Harris (1964, 1976). Gross estabeleceu, juntamente com algumas exigências formais, uma vasta lista de transformações e uma lista de verbos para a aplicação das transformações.

De acordo com Smarsaro (2004, p. 28), o Léxico-Gramática

[...] tem como meta investigar os procedimentos lexicais e gramaticais que levam ao reconhecimento de padrões de palavras nos quais está baseado o processo de entendimento para representação das propriedades de sequências linguísticas computacionalmente. o(s) detentor(es) destas marcas. Isso deve ser feito na primeira vez que a(s) marca(s) registrada(s) aparecer(em).

A base teórica que fundamenta o Léxico-gramática é o Distribucionalismo, oriundo da gramática Transformacional do linguista Zellig Harris (1964, 1976) que de acordo com (GROSS, 1991, p.2 apud SANTOS), a construção da teoria teve como ponto de partida a Teoria da Gramática Transformacional de Harris.

Mantendo os estudos da teoria, os linguistas que fazem a descrição de uma língua produzem vastas descrições de verbos e nomes dessa língua e, diante da enormidade de sua pesquisa, é interessante e necessário utilizar as vantagens da informática no programa científico e técnico da teoria do Léxico-gramática. Essas vantagens consistem no armazenamento dos dados linguísticos pesquisados por meio do computador, que se tornou um aliado importante para o linguista.

Dessa forma, o computador tornou-se uma ferramenta de suma importância para os pesquisadores dessa área, pois permite o armazenamento de forma segura e eficiente, e o pesquisador pode realizar modificações a qualquer momento. Antes da utilização das máquinas para o armazenamento de informações, as pesquisas não eram armazenadas com tanta segurança e não permitiam modificações rápidas e precisas.

Portanto, alguns linguistas direcionaram o estudo do léxico na linguagem para a área computacional, com intuito de desenvolver técnicas modernas, tentando “associar conhecimentos linguísticos a conhecimentos computacionais numa atividade interdisciplinar”

(BIDERMAN, 2001, p. 79). Com isso, pode-se utilizar o computador como ferramenta para manipulação da linguagem humana. Dessa interação nasce a linguística computacional ou processamento de linguagem natural.

## 2 A DESCRIÇÃO DE ITENS LEXICAIS E O PROCESSAMENTO AUTOMÁTICO DE LINGUAGEM NATURAL

O avanço no campo científico da humanidade tornou-se algo inquestionável na sociedade atual. Cada vez mais o homem tem produzido máquinas poderosas para desempenhar atividades diversas que, anteriormente, poderiam ser feitas apenas pelo ser humano. Tal avanço ocorre também nos estudos linguísticos. A utilização do computador com o auxílio da inteligência artificial (IA), que, segundo Winston (1987, apud Smarsaro 2000, p. 78), é um estudo de conceitos que permitem aos computadores serem inteligentes, tem gerado grandes avanços para a linguística computacional.

De acordo com Smarsaro (2000, p. 79), a Linguística Computacional surge, portanto, da pesquisa em inteligência artificial *stricto sensu* – processamento de linguagem natural, *lato sensu* – pesquisas e usos com o auxílio do computador. Ela tem o objetivo de codificar as pesquisas feitas sobre o léxico de uma língua para eliminar possíveis ambiguidades na tradução feita pelo computador.

Hoje em dia, praticamente todos os textos são produzidos utilizando computadores, o que comprova que os arquivos podem ser armazenados em formato computacional. Os programas de computador podem escanear os textos desses arquivos à busca de informações específicas. Mas a informação dada em configuração linguística, ou seja, por meio de palavras, sempre apresenta as mesmas dificuldades de interpretação das palavras: ambiguidade e composicionalidade (GROSS, 1986).

Quando falamos em Linguística computacional pensamos em dicionários eletrônicos utilizados nos tratamentos do léxico de uma língua. Esses dicionários são utilizados para a tradução linguística de diversas línguas, inclusive o português. Porém, como citado acima, essa tradução feita pelo dicionário eletrônico gera ambiguidades e, com isso, problemas de interpretação.

O computador é capaz de realizar as mais distintas funções quanto à tradução de línguas, porém a tarefa de pensar e codificar as palavras dessa língua é feita pelo homem.

Um programa de tradução reconhece palavras, como combiná-las em frases, o que as palavras significam, mas a máquina não é capaz de eliminar ambiguidades e fornecer ao usuário uma tradução fiel ao esperado. Isso porque tais ambiguidades são eliminadas no contexto de uso de determinada palavra ou sentença, o que não é considerado pela máquina.

Com o intuito de reduzir problemas de ambiguidades em traduções automáticas, a Linguística Computacional vale-se da formalização. O linguista formaliza os dados da língua, pois o computador só pode lidar com conteúdos formalizados, com informações precisas. Logo, o processamento da linguagem requer o domínio e a compreensão, primeiramente, da língua escrita.

Para que a máquina efetue a leitura dos dados, antes de tudo está a importância do trabalho realizado pelo linguista, que, ao estudar a língua, cria os mecanismos

necessários ao computador, ou seja, a codificação, para que este exerça sua função.

Observa-se claramente, portanto, que somente a leitura feita pelo computador não é suficiente para determinar sentidos diversos para situações de falas distintas. Para que o computador seja eficiente a formalização torna-se essencial, pois

não se pode esquecer que a linguagem em computação é altamente formalizada, temos que dar fórmulas para explicar os fatos da língua se a queremos inserir num ambiente computacional. Para isso, antes, é preciso reconhecer as generalizações, regularidades e irregularidades de um reconhecimento linguístico. Só podemos incluir num sistema computacional informações objetivas, ou seja, regras formalizadas. Em outras palavras, o computador só pode lidar com conteúdos formalizados, com informações precisas, através de regras bem explícitas, pois qualquer nível de ambiguidade dificulta o funcionamento de qualquer programa que utilize uma linguagem natural. Assim sendo, a linguagem deve ser o mais determinística possível, caracterizando todos os símbolos linguísticos através de seus traços sintáticos, morfológicos, semânticos e pragmáticos (SMARSARO 2000, p. 81).

Logo, a formalização é uma prática imprescindível para a Linguística Computacional, pois é através dela que os linguistas podem desenvolver programas de tradução e leitura cada vez mais eficientes que eliminem problemas como a ambiguidade que comprometem a leitura adequada de uma sentença da língua.

## 3 IDENTIFICAÇÃO DE VERBO SUPORTE

Existem diferentes testes utilizados para comprovar que um verbo pode ser classificado como verbo suporte. Alguns destes testes serão mostrados a seguir utilizando o verbo *pegar* como referência.

Para verificar se um verbo pode ser classificado como suporte em determinada sentença, a partir de uma frase simples elaboramos uma oração relativa e em seguida subtraímos da relativa o verbo em estudo. Se o resultado não apresentar prejuízo para o sentido da frase, afirmamos que o verbo é um verbo-suporte.

Exemplo:

- (1) João *pegou uma gripe* muito forte.
- (1.1) A gripe que João *pegou* era muito forte.
- (1.2) A gripe de João era muito forte.

Na sequência *pegar gripe* destacada em (1), temos um claro exemplo do verbo *pegar* funcionando como verbo-suporte, pois a retirada do mesmo não causou prejuízo semântico à frase matriz. O que também ocorre no exemplo a seguir:

- (2) Henrique *pegará pena* de dois anos de prisão.
- (2.1) A pena de prisão que Henrique *pegará* será de dois anos.
- (2.2) A pena de prisão de Henrique será de dois anos.

Portanto, em *pegar gripe* e *pegar pena*, o verbo *pegar* é classificado como verbo-suporte.

Quando aplicamos o mesmo teste utilizado em (1) e em (2) e essa subtração do verbo em estudo causa algum prejuízo semântico para a frase matriz, dizemos que a sequência não caracteriza um caso de verbo suporte. Nesses casos a sequência será classificada como expressão fixa ou uma sequência com verbo pleno. Vejamos alguns exemplos:

- (3) João pegou uma onda violenta.  
 (3.1) A onda que João pegou era violenta.  
 (3.2) ?A onda de João era violenta.

No exemplo (3), o verbo *pegar* não caracteriza um caso de verbo suporte, pois a oração formada com a retirada do verbo não mantém o sentido desejado na oração matriz. A interrogação utilizada indica uma oração duvidosa, pois pode gerar o entendimento de que João é dono da onda, e não de que ela apenas a surfou. Por isso o verbo *pegar* em *pegar onda* não pode ser classificado como verbo suporte.

Outro exemplo em que não há verbo suporte:

- (4) A polícia pegou um ladrão perigoso.  
 (4.1) O ladrão que a polícia pegou era perigoso.  
 (4.2) \*O ladrão da polícia era perigoso.

A sequência *pegar ladrão* demonstrada em (4) não caracteriza um caso com verbo suporte, pois ao retirar o verbo em estudo da oração relativa formada, o sentido da oração matriz não se mantém, ou seja, há prejuízo semântico, pois perde-se totalmente a informação de que a polícia tenha prendido um ladrão perigoso. Portanto, no exemplo acima não há verbo suporte, mas sim um verbo pleno, pois o mesmo está em seu sentido prototípico, que significa prender, segurar, agarrar ou tomar para si.

O asterisco utilizado na oração formada após a retirada do verbo em questão, demonstrada em (4.2), indica que é uma oração inaceitável para este caso, ou seja, não mantém a informação dada na oração matriz, que é o que se deseja encontrar na análise.

Outra forma de identificar um verbo suporte é a utilização de um critério sintático. Segundo Laporte, uma construção com verbo suporte sempre tem uma variante na qual o verbo suporte é ausente, sem modificação imprevisível do sentido.

Exemplo:

- (5) Meu gato *pegou o vício* de fazer sujeira no banheiro.  
 (5.1) Depois de (*E + pegar*) mania de fazer sujeira no banheiro, meu gato melhorou.  
 (5.2) A mania do gato de fazer sujeira no banheiro o fez melhorar.  
 (6) João *pegou amor* pela mulher.  
 (6.1) Depois do (*E + pegar*) amor pela mulher, João se acalmou.  
 (6.2) O amor de João pela mulher o acalmou.

Nos exemplos (5) e (6), tínhamos frases com o verbo *pegar* funcionando como verbo suporte. Para comprovar esta classificação, tentamos criar uma variante do verbo, da qual ele estaria ausente, para ver se haveria modificação no sentido.

Em ambos os casos, o verbo em questão pôde ser retirado sem causar prejuízo semântico às orações utilizadas como exemplo.

Portanto, é importante conhecer as variantes de construções com verbo suporte nas quais o verbo suporte é ausente para poder reconstituir em novos exemplos e testar.

#### 4 PROBLEMAS NA IDENTIFICAÇÃO DE VERBO SUPORTE: AMBIGUIDADE

A aplicação do teste da relativa para a identificação de verbo suporte é um dos mais eficientes para tal análise. Porém, ele pode apresentar problemas para o pesquisador que o utiliza.

Ao aplicarmos o teste citado em algumas construções com o verbo *pegar*, alguns problemas para a classificação de verbo-suporte podem gerar dúvidas, como a ambiguidade.

Exemplo:

- (7) Maria *pegou uma carona* perigosa.  
 (7.1) A carona que Maria *pegou* era perigosa.  
 (7.2) A carona de Maria era perigosa.

Com a retirada do verbo *pegar* gera-se uma frase ambígua, pois Maria estava dando ou pegando carona? Apenas com a aplicação do teste da relativa não é possível afirmar se o verbo *pegar* em *pegar carona* caracteriza um caso de verbo suporte.

Uma forma de desfazer a ambiguidade é acrescentar à frase novos argumentos:

- (7.3) A carona de Maria era perigosa, porque ela não conhecia o motorista.  
 (7.4) A carona de Maria era perigosa porque seu carro era velho.

Com os novos argumentos representados em (7.3) e em (7.4), não há dúvidas de que na primeira frase Maria está pegando carona, e o verbo *pegar* neste caso é classificado como verbo suporte. Já na segunda frase era ela quem estava dando carona a alguém, o que não mantém a informação dada na frase matriz, não caracterizando um caso de verbo suporte.

Outro exemplo que apresenta ambiguidade após a aplicação do teste de verificação de verbo-suporte é a sequência *pegar gosto*:

- O malte *pegou o gosto* da fumaça de um fogo próximo.  
 Luiz *pegou o gosto* pelas novas tecnologias.

Aplicando o teste para verificar se o verbo é um verbo-suporte temos:

- (8) O malte *pegou o gosto* da fumaça de um fogo próximo.  
 (8.1) O gosto de fumaça que o malte *pegou* era de um fogo próximo.  
 (8.2) O gosto de fumaça do malte era do fogo próximo.  
 (9) Luiz *pegou o gosto* pelas novas tecnologias.  
 (9.1) O gosto que Luiz *pegou* pelas novas tecnologias é surpreendente.  
 (9.2) O gosto de Luiz pelas novas tecnologias é surpreendente.

Através do acréscimo de argumentos diferentes, a ambiguidade presente nos exemplos de *pegar o gosto* foi desfeita e os dois sentidos possíveis formaram frases com verbo-suporte.

Ao contrário do que ocorreu com os exemplos de *pegar carona*, em que, ao desfazer a ambiguidade, um dos sentidos formou uma frase com verbo-suporte e o outro não.

Os casos de ambiguidade representados acima podem gerar problemas também para os programas de tradução de linguagem natural, pois o computador não é capaz de desfazer essas ambiguidades sozinho. O falante da língua é capaz de fazer isso porque ele consegue distinguir e selecionar contextos diferentes para a utilização de cada frase da língua.

Com isso, mais uma vez fica clara a importância da descrição feita pelo linguista para a Linguística computacional ou Processamento Automático de Linguagem Natural.

## 5 CONCLUSÃO

Este artigo apresentou um breve estudo sobre a Teoria do Léxico-gramática e sobre a identificação de verbo suporte em construções utilizadas por falantes da Língua Portuguesa.

Além disso, foi apresentada a importância da descrição de itens lexicais para o Processamento Automático de linguagem natural e para a criação de programas de tradução automática da língua.

Para exemplificar as construções com verbo suporte e os casos de ambiguidade foram utilizados exemplos construídos por falantes nativos da Língua Portuguesa. Tal prática está amparada na Teoria do Léxico-gramática que defende a noção de que o falante nativo tem autonomia linguística para criar frases e julgá-las como aceitáveis ou não.

Com a análise dos exemplos pôde-se representar a identificação de verbo suporte em exemplos construídos por falantes nativos da Língua Portuguesa. Além disso, demonstrou-se casos em que o teste de identificação de verbo suporte gera ambiguidade e como desfazer tal problema.

Portanto, a aplicação de testes formais para a identificação do verbo suporte é de fundamental importância, mesmo podendo gerar dúvidas. Dúvidas estas que foram desfeitas com o acréscimo de novos argumentos permitindo a classificação do verbo sem maiores problemas.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [1] ATHAYDE, Maria Francisca, "Construções com verbo-suporte (Funktionsverbgefüge) do Português e do Alemão", in: *Construções com Verbo-Suporte (Funktionsverbgefüge) do Português e do Alemão*, cadernos do cieq, nº 1, 2001, p. 5-68.
- [2] BIDERMAN, M. T. C. Teoria Linguística: teoria lexical de linguística computacional. São Paulo: Martins Fontes, 2001.
- [3] GROSS, M. Méthodes empiriques em syntaxe. Paris : Hermann, 1975.
- [4] GROSS, Maurice. Lexico-Grammar. The Representation of Compound Words. In: COLING – 1986 Proceedings, Bonn, 1986.
- [5] LAPORTE, E. Exemplos atestados e exemplos construídos na prática do Léxico-Gramática. Revista (com) textos linguísticos. Vitória – ES, n. 2 p. 26-51, 2008.
- [6] NEVES, M. H. M. Estudo das Construções com Verbo-suporte em Português. In: INGEDORE V. K. Gramática do português falado. Campinas, SP. Editora da Unicamp, 2002.
- [7] SANTOS, Vangevaldo Cardoso dos. Expressões fixas com substantivos de partes do corpo do português brasileiro para processamento automático de linguagem natural. (Dissertação de mestrado). Universidade Federal do Espírito Santo, 2011.
- [8] SMARSARO, Aucione das Dores. O Processamento Automático da Linguagem Natural: Uma Introdução à Linguística Computacional. In: Revista Recortes Linguísticos. Vitória: Saberes, n.1 p. 77-90, 2000
- [9] SMARSARO, Aucione D., **Descrição e Formalização de palavras compostas do português do Brasil para a elaboração de um dicionário eletrônico**. 2004, 154 f. Tese (Doutorado em Letras) Programa de Pós-Graduação em Estudos da Linguagem, PUC-Rio, Rio de Janeiro, 2004.