

SEMÂNTICA E PRAGMÁTICA

Leonardo Teixeira de Freitas Ribeiro Vilhagra



SOLUÇÕES
EDUCACIONAIS
INTEGRADAS



Semântica computacional e semântica cultural

Objetivos de aprendizagem

Ao final deste texto, você deve apresentar os seguintes aprendizados:

- Identificar a semântica computacional e a sua interface com as ciências linguísticas.
- Resumir as divisões da semântica computacional.
- Explicar os níveis de sentido no âmbito da semântica cultural.

Introdução

Neste capítulo, você vai estudar duas vertentes da semântica, a computacional e a cultural. Primeiramente, você vai conhecer os pressupostos básicos da semântica computacional e ver como se configurou a relação entre a linguística e a informática. Ademais, vai conferir quais foram os impactos dessa relação na pesquisa sobre a linguagem.

Depois, você vai se debruçar sobre os conceitos básicos da semântica cultural, uma vertente relativamente nova da área da semântica. Diferentemente das outras vertentes, ela privilegia os aspectos culturais e identitários no estudo do sentido nas línguas naturais.

1 Semântica computacional

Para começar a compreender a semântica computacional, você precisa se familiarizar com a linguística computacional. Afinal, a semântica computacional é uma das ramificações da linguística computacional. A seguir, você vai conhecer os pressupostos básicos da linguística computacional e verificar a sua relação com a semântica computacional.

Semântica e linguística computacional

A linguística computacional é, como descrevem Vieira e Lima (2001, p. 1), “[...] a área de conhecimento que explora as relações entre linguística e informática, tornando possível a construção de sistemas com capacidade de reconhecer e produzir informação apresentada em linguagem natural”. Othero (2006), por sua vez, destaca que o conhecimento produzido pela linguística computacional é utilizado para tentar processar (isto é, compreender e produzir) as línguas naturais e dominar o conhecimento linguístico envolvido em uma linguagem natural. A linguística computacional teve início em 1950, com Warren Weaver, matemático estadunidense.

Weaver, aproximando a matemática da comunicação, iniciou o campo de conhecimento da tradução automática (*machine translation*) a partir da projeção de um computador voltado para a tradução, denominado *memorandum*. Assim, a linguística computacional, ao aliar essas duas áreas do conhecimento, promove um tratamento computacional da linguagem, ou seja, o linguista computacional aplica os princípios computacionais, principalmente o de processamento de dados, à linguística.



Saiba mais

Para saber mais sobre a tradução automática, leia o artigo de Ronaldo T. Martins e Maria das Graças V. Nunes intitulado “Noções Gerais de Tradução Automática”. O artigo está disponível *on-line* e você pode encontrá-lo por meio de uma busca na internet.

O processamento de dados consiste basicamente no conjunto de atividades executadas de modo planejado para coletar e ordenar dados, a fim de executar algo com determinado fim (BONIATI; PREUSS; FRANCISCATTO, 2014). A produção de texto no computador é um exemplo de processamento de dados. O programador de um *software* de edição de texto fornece uma série de comandos, os quais são inseridos no computador. Quando o usuário aciona o teclado e o *mouse*, vários dados são gerados; tais dados são coletados e transformados em texto digital.

Agora, considere essa mesma lógica no âmbito da linguística computacional. Ao articular os conhecimentos da computação à linguística, o linguista computacional aplica o processamento de dados aos diversos fenômenos linguísticos, envolvendo desde o armazenamento até a interpretação. Um exemplo é o ReGra, um revisor gramatical para o português desenvolvido

pelo Núcleo Interinstitucional de Linguística Computacional de São Carlos. Esse programa tem a função de verificar desvios ortográficos, utilização inadequada da crase, regência e concordância verbal e nominal, tudo de acordo com a norma padrão da língua portuguesa (RINO *et al.*, 2002).

Portanto, a linguística computacional busca explorar os conhecimentos da informática e aplicá-los à linguística, de modo que os fenômenos linguísticos sejam tratados pela computação — tanto no que diz respeito ao armazenamento e à organização quanto no que se refere à compreensão e à interpretação.

Princípios da semântica computacional

Como você viu, a linguística computacional é o ramo interdisciplinar que abarca tanto os conhecimentos da informática quanto os da linguística. A partir disso, ela busca sistematizar os fenômenos linguísticos, ora na perspectiva do armazenamento, ora na perspectiva da interpretação.

Em síntese, a semântica computacional é o braço da linguística computacional cujo princípio é unir os conhecimentos da informática aos da semântica. Dessa forma, segundo Pagani (2013, p. 1), a semântica computacional possui uma:

[...] tendência maior para a explicação dos fenômenos (tanto linguísticos, em relação à capacidade semântica dos falantes de alguma língua, quanto computacionais, em relação à implementação desses fenômenos em computadores) ou para a sua metacompreensão (novamente, tanto linguisticamente na implementação computacional de modelos linguísticos quanto computacionalmente discutindo-se a computabilidade das teorias ou dos fenômenos linguísticos).

Ou seja, a semântica computacional almeja estudar os fenômenos semânticos com a contribuição da informática, seja para compreender tais fenômenos por meio da informática, seja para utilizar esta última com a finalidade de colher, tratar e organizar os dados que envolvam fenômenos linguísticos.

Um exemplo disso pode ser observado nos programas do tipo *chatbots*. Os *chatbots* são programas de computador projetados para simular uma conversa (*chat* significa “conversa”, enquanto *bot* é a abreviação de “robô”). Veja o que Othero (2006, p. 348) comenta a respeito dessa ferramenta:

[...] acreditamos que a próxima geração de *chatbots* deva exigir dos programadores um profundo conhecimento [...] da semântica de uma língua (que tornará possível fazer com que o programa seja capaz de interpretar o significado do *input* linguístico dado a ele pelo usuário humano).

Note que a semântica computacional vincula os conhecimentos da própria semântica aos da informática. A ideia é que a partir disso seja possível estudar e pesquisar os processos de significação de maneira interdisciplinar. Ademais, embora a semântica computacional tenha como característica a interdisciplinaridade com a informática, isso não quer dizer que ela esteja longe da linguística. Na verdade, o que existe é um aprofundamento do estudo dos fatos da linguagem, mas com contribuições da informática, de modo que se possa entender os fenômenos linguísticos pela óptica dos computadores.

2 Áreas da semântica computacional

A partir de agora, você vai conhecer a divisão existente na semântica computacional. Pode-se dizer que essa divisão está mais ligada à maneira pela qual os conhecimentos da semântica e os da informática são aplicados e ao fim que eles possuem. Assim, tais conhecimentos são divididos em dois grandes grupos: a linguística de *corpus* e o Processamento de Linguagem Natural (PLN).

Linguística de *corpus*

Essa área da semântica computacional está associada ao “[...] trabalho a partir de *corpora* eletrônicos que contenham amostras de linguagem natural” (OTHERO, 2006, p. 342). Ou seja, essa categoria está preocupada em coletar, armazenar e analisar dados e informações de ordem semântica. Em suma, a linguística de *corpus* almeja reunir um conjunto de dados referentes a fenômenos semânticos e tratá-los para determinado fim. Um exemplo disso são os modelos semânticos descritos por Hanis e Noller (2012).

Os modelos semânticos são programas de computador capazes de classificar e relacionar as informações de um banco de dados qualquer, de modo que se possa categorizar e cruzar essas informações. Ou seja, a partir de definições e conceitos previamente programados, é possível agrupar dados similares e fazer com que eles interajam com outros.

Para compreender melhor, considere o caso dos aplicativos de tráfego (HANIS; NOLLER, 2012). Em um aplicativo de tráfego inteligente, dados são fornecidos em tempo real a partir de sensores de semáforo, sensores de velocidade, relatórios meteorológicos e de acidentes, eventos do calendário (feriados, por exemplo), etc. Porém, é necessário categorizar tais dados para que eles sejam funcionais. Além disso, existem situações nas quais pode haver ambiguidade, a exemplo da palavra “veículo”. Essa palavra sozinha pode não

representar a realidade, ou seja, o veículo pode ser um carro, uma caminhonete, um caminhão, uma motocicleta, etc.

É nesse momento que o modelo semântico pode ser inserido. Nesse contexto:

[...] um modelo semântico poderia ajudar a entender relações como as dos sensores de semáforo com as intersecções que eles monitoram, qualquer sensor de semáforo com outros sensores na mesma estrada ou a relação das estradas sobre as quais temos dados de sensor específicos com outras estradas que as cruzam e, coletivamente, como alimentadores para as principais rodovias. O modelo também pode gerar informações parecidas sobre linhas de ônibus ou de metrô. Pode descrever os tipos de serviço disponíveis com os locais atendidos. As relações entre as estações e endereços e linhas de serviço e rotas de carro forneceriam a base para entender as implicações sobre o tráfego rodoviário de perturbações específicas no serviço de transporte coletivo (HANIS; NOLLER, 2012, documento *on-line*).

Portanto, a linguística de *corpus* se volta diretamente para os *corpora* de natureza semântica, com o intuito de tratar, organizar e analisar esse conjunto de dados por meio da articulação entre os saberes da semântica e da informática.

PLN

Outra divisão da semântica computacional diz respeito ao PLN. Como afirma Othero (2006, p. 342), essa divisão:

[...] preocupa-se diretamente com o estudo da linguagem voltado para a construção de *softwares*, aplicativos e sistemas computacionais específicos, como tradutores automáticos, *chatbots*, *parsers*, reconhecedores automáticos de voz, geradores automáticos de resumos, etc.

Assim, esse tipo de semântica computacional está diretamente associado à interpretação das línguas, sobretudo no nível do sentido. Isso pode ser atestado por Vieira (2002, p. 20), que diz que, “[...] para o processamento da língua natural, vários subsistemas são necessários para dar conta dos diferentes aspectos da língua: sons, palavras, sentenças e discurso nos níveis estruturais, de significado e de uso”.

Um caso que representa isso, inclusive já abordado aqui, é o dos *chatbots*. Assim, o PLN se mostra um campo bastante fértil para a semântica computacional, principalmente quando atrelado à Inteligência Artificial (IA). Nesse sentido, McDonald e Yazdani (1990, p. 176) pontuam que “[...] a pesquisa em PLN pode proporcionar *insights* bastante úteis sobre processos e representações da linguagem na mente humana, apontando, assim, para a verdadeira IA”.

3 Semântica cultural

Agora, você vai conhecer mais uma vertente da semântica, a semântica cultural. Para Ferrarezi Junior e Basso (2013, p. 1), essa vertente “[...] estuda a relação entre os sentidos atribuídos às palavras ou demais expressões de uma língua e a cultura em que essa mesma língua está inserida”.

De certa forma, a semântica cultural traz uma abordagem nova, pois atrela o aspecto do sentido à cultura da língua. Souza e Vianna (2015, p. 111) destacam que a semântica cultural atesta:

[...] que as estruturas das línguas naturais não são meros agrupamentos de regras gramaticais, mas reflexos da cultura, das organizações sociais e do meio. Nesse sentido, a Semântica Cultural vê a língua como sendo mais do que uma herança passada de geração a geração, tendo em vista que ela interfere, diretamente, na forma como enxergamos e vivenciamos o mundo, uma vez que é através dela, da língua, que expressamos esse mundo.

Portanto, na perspectiva da semântica cultural, uma língua não é um produto artificial e isolado socialmente. Na verdade, ela é viva e contextualizada, de modo que influencia e é influenciada pela cultura. Aqui, a cultura é entendida a partir do conceito proposto por Eagleton (2005). Para esse autor, a cultura é um conjunto de valores, costumes, crenças e práticas que configuram a forma de viver de um grupo específico. Tal conjunto é responsável por gerar conhecimento implícito, permitindo que os integrantes do grupo ressignifiquem e se apropriem de modos de agir em contextos específicos.

Considerando esses pressupostos, você consegue depreender como a semântica cultural concebe o processo de significação das línguas humanas? Para essa corrente, as línguas possuem duas dimensões: a visão de mundo dos falantes e a dimensão linguística.

Dimensão da visão de mundo do falante

Para Novais e Ferrarezi Junior (2015, p. 3), a dimensão da visão de mundo do falante “[...] é construída por meio e entremeada com a construção cultural de cada comunidade (onde se inserem os significados)”. Ou seja, ela é o conjunto de crenças, costumes e práticas singulares de uma região, o qual particulariza uma forma de se enxergar a realidade, de acordo com a perspectiva do falante. É nesse nível que os significados estão inseridos.

Dimensão linguística

Já a dimensão linguística, segundo Novais e Ferrarezi Junior (2015), é a manifestação da visão de mundo do falante. É na dimensão linguística que o sentido está inserido. Para a semântica cultural, existe uma distinção entre o significado (nível da visão de mundo do falante) e o sentido (nível linguístico).

Entretanto, para essa vertente da semântica, é mais pertinente estudar o nível linguístico:

Se o significado não é objeto do estudo da Semântica, mas suas manifestações é que o são — quais sejam elas, as múltiplas formas pelas quais a língua o manifesta em seu sistema e nas funções para as quais esse sistema foi criado [...] podemos confirmar a adoção do termo “sentido”, que me parece ser, na literatura corrente, aquele que mais se aproxima da ideia que pretendo expressar quando falo de manifestações linguísticas do significado (FERRAREZI JUNIOR, 2010, p. 59).

Com o intuito de tornar esses dois conceitos mais palpáveis, Almeida e Fossile (2016) formulam o exemplo apresentado no Quadro 1, a seguir. Esse exemplo parte do princípio de que cada região do Brasil encara a figura da mulher de maneira diferente. Afinal, cada lugar, devido às suas singularidades, possui uma visão de mundo específica. Ademais, existe também a prática dessa visão de mundo, isto é, a manifestação verbal dela por meio de uma língua. Considerando isso tudo, o Quadro 1 mostra o caso da palavra “rapariga”, indicando como ela é utilizada em cada região do Brasil.

Quadro 1. Dimensão da visão de mundo do falante e dimensão linguística utilizando como exemplo a palavra “rapariga”

Região do Brasil	Sentido
Amazônia	Menina virgem, donzela
Região Sul	Moça jovem de boa família
Região Nordeste	Concubina geralmente sustentada por um homem casado
Região Norte e Região Centro-Oeste	Prostituta

Fonte: Adaptado de Almeida e Fossile (2016).

Como você pode inferir, ainda que esteja em jogo a mesma palavra, o sentido vinculado a ela (dimensão linguística) está em função de um significado (dimensão da visão de mundo) contextualizado em determinada cultura. Inclusive, alguns sentidos podem ser mais estabilizados, uma vez que podem ser mais comuns a determinadas culturas do que outros sentidos, os quais são mais singulares. Nesse viés, destacam-se as expressões idiomáticas, as quais, segundo Rocha (2012), são unidades lexicais que contêm uma linguagem metafórica e, muitas vezes, são fruto de uma comunidade específica. Aliás, elas são frequentemente estudadas por essa vertente semântica.



Fique atento

Ao longo deste capítulo, você estudou os princípios básicos da semântica computacional e cultural. Como você pôde notar, a semântica computacional reúne conhecimentos tanto da própria semântica quanto da informática, sendo dividida em linguística de *corpus* e PLN. Já a semântica cultural insere o fator cultural na investigação do significado, permitindo que os traços identitários sejam levados em consideração no momento da pesquisa e da análise semântica.



Referências

ALMEIDA, N. R.; FOSSILE, D. K. Semântica cultural: um estudo acerca da atribuição de sentidos às palavras e expressões da língua. *Raído*, v. 10, n. 24, 2016.

BONIATI, B. B.; PREUSS, E.; FRANCISCATTO, R. *Introdução à informática*. Frederico Westphalen: Universidade Federal de Santa Maria, 2014. Disponível em: http://estudio01.proj.ufsm.br/cadernos/cafw/tecnico_agroindustria/introducao_informatica.pdf. Acesso em: 5 maio 2020.

EAGLETON, T. *A idéia de cultura*. São Paulo: Unesp, 2005.

FERRAREZI JUNIOR, C. *Introdução à semântica de contextos e cenários: de la langue à la vie*. Campinas: Mercado de Letras, 2010.

FERRAREZI JUNIOR, C.; BASSO, R. *Semântica, semânticas: uma introdução*. São Paulo: Contexto, 2013.

HANIS, T.; NOLLER, D. *A função dos modelos semânticos nas operações industriais mais inteligentes*. North Castle: IBM, 2012. Disponível em: <https://www.ibm.com/developerworks/br/industry/library/ind-semanticmodels/index.html>. Acesso em: 8 abr. 2020.

MCDONALD, C.; YAZDANI, M. *Prolog programming: a tutorial introduction*. Oxford: Blackwell Scientific, 1990.

NOVAIS, S. L. Ó.; FERRAREZI JUNIOR, C. Estudo do falar de Paraguaçu, MG e região: expressões idiomáticas. *Rev (Entre Parênteses)*, v. 1, n. 4, 2015.

OTHERO, G. de Á. Linguística computacional: uma breve introdução. *Letras de Hoje*, v. 41, n. 2, p. 341-351, 2006.

PAGANI, L. A. Semântica computacional. In: FERRAREZI JUNIOR, C.; BASSO, R. *Semântica, semânticas: uma introdução*. São Paulo: Contexto, 2013. p. 57-69.

RINO, L. H. M. *et al.* Aspectos da construção de um revisor gramatical automático para o português. *Est Linguíst*, v. 31, p. 1-6, 2002.

ROCHA, C. C. A formação do português brasileiro pela observação de expressões idiomáticas. *Rev ContraPonto*, v. 2, n. 2, p. 14-27, 2012.

SOUZA, D. S.; VIANNA, A. L. T. Semântica, semânticas: uma introdução ao estudo do significado para estudantes brasileiros. *Entrelinhas*, v. 9, n. 1, 2015.

VIEIRA, R. Linguística computacional: fazendo uso do conhecimento da língua. *Entrelinhas*, ano 2, n. 4, 2002.

VIEIRA, R.; LIMA, V. L. S. Linguística computacional: princípios e aplicações. In: NEDEL, L. (org.). *IX Escola Regional de Informática - 2001 (ERI2001)*. Porto Alegre: SBC- Regional Sul, 2001. v. 1, p. 27-58.

Leitura recomendada

MARTINS, R. T.; NUNES, M. G. Noções gerais de tradução automática. *NILC - ICMC-USP*, n. 68, p. 1-26, 2005. Disponível em: http://www.nilc.icmc.usp.br/nilc/download/Notas-DidaticasILCMC_68.pdf. Acesso em: 5 maio 2020.



Fique atento

Os links para sites da web fornecidos neste capítulo foram todos testados, e seu funcionamento foi comprovado no momento da publicação do material. No entanto, a rede é extremamente dinâmica; suas páginas estão constantemente mudando de local e conteúdo. Assim, os editores declaram não ter qualquer responsabilidade sobre qualidade, precisão ou integralidade das informações referidas em tais links.

Encerra aqui o trecho do livro disponibilizado para esta Unidade de Aprendizagem. Na Biblioteca Virtual da Instituição, você encontra a obra na íntegra.

Conteúdo:



SOLUÇÕES
EDUCACIONAIS
INTEGRADAS