

Saussure na era da IA: modelagem semântica de word embeddings à luz da teoria do valor

Leonardo Giamarusti¹

Resumo

Este estudo explora a relação entre a Teoria do Valor (TdV) de Ferdinand de Saussure e a modelagem semântica dos Word Embeddings (WE), especialmente através da Semântica Distribucional (SD). Neste trabalho, propomos analisar alguns frameworks de embeddings de palavras gerados pelo modelo Word2Vec, buscando possíveis articulações entre a TdV e a captura de significados em modelos de linguagem baseados na Semântica Distribucional. Partindo da hipótese de que a noção de valor em Saussure reside nas relações de semelhança e dessemelhança entre um signo e sua totalidade sistêmica, objetivamos demonstrar que este mesmo princípio em torno do valor linguístico parece também estar presente na Semântica Distribucional, preconizada pelos linguistas Z. Harris, nos EUA, e J. R. Firth, na Inglaterra, em meados da década de 50. Nesse sentido, apresentamos alguns indícios de que, tal qual Saussure, a SD também parte das relações do sistema para propor uma teoria da significação na língua. Essa teoria, por sua vez, tem sido amplamente aplicada no Processamento de Linguagem Natural para a modelagem semântica de WE, o que possibilita discussões acerca da possível pertinência do saussurianismo para o desenvolvimento de novas técnicas de PLN no século XXI.

Palavras-chave: Saussure; Linguística Computacional; Processamento de Linguagem Natural; Teoria do Valor; Word Embeddings

Data de submissão: novembro. 2024 – Data de aceite: novembro. 2024

<http://dx.doi.org/10.5335/rdes.v20i3.16461>

¹ Mestrando em Estudos Linguísticos pelo Programa de Pós-Graduação em Estudos Linguísticos (PPGEL/UFU). É membro do Grupo de Pesquisa Ferdinand de Saussure (CNPq) desde 2018, onde realiza pesquisas sobre a fortuna saussuriana e seus impactos no Processamento de Linguagem Natural. <http://lattes.cnpq.br/5597723489199237> E-mail: lenardogiamarusti@gmail.com

Introdução

A Teoria do Valor (TdV), proposta por Ferdinand de Saussure principalmente no Terceiro Curso de Linguística Geral (TCLG) entre 1910 e 1911, ocupa um papel central na compreensão da língua como um sistema de signos. Nesse sistema, entende-se que a significação surge por relações diferenciais entre o signo e sua totalidade sistêmica, gerando valores linguísticos que são expressos por meio das relações de “semelhança” e “dessemelhança” (cf Saussure, 2012, p.160) que o signo contrai com os outros.

Nesta pesquisa, apresentamos alguns indícios teóricos de que esse princípio saussuriano que vê o valor linguístico como produto das relações do sistema parece ter sido reinterpretado por movimentos estruturalistas do século XX, desdobrando-se numa teoria da significação conhecida como “Semântica Distribucional” (SD) (Firth, 1957; Harris, 1954). A SD, devido à sua aproximação com a matemática e a lógica, hoje, é considerada um dos principais aportes teóricos, no Processamento de Linguagem Natural (PL), para a modelagem semântica de modelos vetoriais de linguagem, notadamente os Word Embeddings (WE).

Portanto, propomos um estudo exploratório, em tom ensaístico, em que analisamos algumas possíveis articulações entre a modelagem semântica de WE e a possível relação com a teoria dos valores linguísticos do mestre genebrino. Nossa hipótese é de que a noção de *valor* pode ser percebida na Semântica Distributiva como mecanismo pelo qual se chega à significação; mecanismo este emulado por alguns sistemas de IA para a captura de sentidos e posterior execução de tarefas de PLN.

Assumimos, bem como fazem Seno et al (2023) e Freitas (2023), que, atualmente, uma IA não é capaz, ainda, de compreender significados isolados, isto é, compreender itens lexicais por eles mesmos. Contudo, as Inteligências Artificiais baseadas na SD parecem conseguir reconhecer o espaço que cada palavra, ou signo, ocupa na língua e, a partir disso, inferir possíveis relações entre palavras semelhantes – que chamaremos de *valores positivos* – e dessemelhantes, que trataremos como *valores negativos*. Ao analisar esses valores, os algoritmos conferem uma significação às relações entre as palavras, de modo que a IA se torna capaz de emular a língua humana em níveis semântico-lexicais .

Para esclarecermos o que entendemos por valor linguístico, nos pautamos no capítulo IV da segunda parte do Curso de Linguística Geral (CLG) e em trabalhos de pesquisadores saussurianos interessados por esse tema (Normand, 2009; Silveira, 2009; Coelho, 2013; 2015; Coelho e Lima, 2014). Após esse esclarecimento, analisaremos alguns *frameworks* de WE desenvolvidos por meio do modelo Word2Vec, buscando evidenciar a importância da noção de valor – e o que dela deriva, a saber, as relações na língua – para a modelagem semântica de IAs baseadas na SD. A partir dessa pesquisa, esperamos

possibilitar discussões acerca da pertinência de pressupostos da linguística formalista do século XX como base teórica para a modelagem semântica de algumas IAs no século XXI.

1 Relações na língua e o que delas derivam: os valores

Como dissemos, a TdV de Saussure foi essencialmente apresentada no último dos três cursos que o mestre lecionou no início do século XX na Universidade de Genebra. Coelho e Lima (2014) possibilitam inferirmos que essa teoria pode ser uma forte candidata a representar o amadurecimento das ideias linguísticas de Saussure, uma vez que a teoria do valor só pôde ser desenvolvida por ele

a partir da delimitação de todos os outros aspectos e princípios linguísticos por ele [Saussure] expostos nos cursos, tais como a arbitrariedade do signo, a linearidade do significante, a definição do significado e do significante como constituintes do signo linguístico e a distinção entre “língua”, “linguagem” e “fala” (Coelho e Lima, 2014, p.347).

Para nós, a chave para entender os valores linguísticos propostos pela linguística saussuriana reside, primeiro, na noção de que a língua é um sistema de signos: “[...] poder-se-ia dizer que não é a linguagem que é natural ao homem, mas a faculdade de constituir uma língua, vale dizer: *um sistema de signos distintos correspondentes a ideias distintas*” (Saussure, 2012, p.42, grifo nosso). Dito isso, afirmar que a língua é um sistema é, ao mesmo tempo, tomá-la como um conjunto de elementos que estão intimamente relacionados. Esse entendimento, por sua vez, revela-nos que, necessariamente, a noção de sistema está atrelada à totalidade de signos linguísticos de uma língua; mais precisamente, a todas as relações possíveis que um determinado signo pode admitir.

Nesse sentido, para além de compreendermos que a língua é um sistema, na Teoria do Valor, também é essencial compreendermos que a língua é “um sistema de relações” (Saussure, 2012 [1916], p.100). Sendo assim, a ideia básica construída por Saussure sobre a TdV, a qual podemos acompanhar no CLG nos capítulos dedicados ao valor, é de que a noção de *relação* é um princípio que atravessa a constituição dos valores linguísticos, como já observado anteriormente por Marques (2016).

No CLG, são recorrentes as vezes em que a noção de valor é articulada com a de relação. No entanto, é no capítulo IV da segunda parte que esses pressupostos parecem dialogar de forma mais evidente, levando à formulação de conceitos como o valor em seus aspectos conceitual e material. É também neste capítulo que acompanhamos, no CLG, distinções entre as noções de valor e de significação.

Embora não seja o objetivo deste artigo tratar das possíveis diferenças entre valor e significação, vale pontuarmos que ambos os termos admitem implicações teóricas na teoria saussuriana, não sendo, portanto, equivalentes. Em outras palavras, o valor está

imbricado na constituição da significação, como bem relembra o mestre genebrino: “o *valor* constitui, sem dúvida, um elemento da *significação*, e é difícilimo saber como esta se distingue dele, apesar de estar sob sua dependência” (Saussure, 2012 [1916], p.161, grifo nosso). Diferentemente dos significados, que são produtos da relação arbitrária entre os constituintes dos signos, “o valor é dado por outros dados; pela situação recíproca das partes na língua; e assim por diante”² (Saussure apud Riedlinger, 1997, p. 29, traduzido por Coelho, 2013, p.8)

Dessa maneira, o valor de um signo passa a ser determinado por sua “parte na língua”, a saber, a *posição* que ele ocupa em relação aos outros signos do sistema linguístico. Essa posição e essa relação recíproca das partes geram o *valor*, que, portanto, depende da rede de *diferenças e semelhanças* na língua para a sua delimitação. Por meio da anotação de aula de Riedlinger supracitada, feita durante o TCLG, vemos que Saussure enfatiza que, mesmo sendo parte da significação, o valor possui uma natureza relacional única, distinta do processo pelo qual os significados são construídos.

Nesse sentido, tal natureza relacional única dos valores pode ser melhor entendida quando Saussure explica a constituição do valor via dois tipos de relações, a saber, a semelhança e a dessemelhança:

[...] verifiquemos inicialmente que, mesmo fora da língua, todos os valores parecem estar regidos por esse princípio paradoxal. Eles são sempre constituídos:

1º - por uma coisa *dessemelhante*, suscetível de ser *trocada* por outra cujo valor resta determinar:

2º - por coisas *semelhantes* que podem *comparar* com aquela cujo valor está em causa. (Saussure, 2012, p.162)

Para além desse trecho, Saussure apresenta uma metáfora, no CLG, que parece explicar que tanto as relações de semelhança quanto as de diferença entre os signos são essenciais para definir um valor. Para justificar essa proposição, o mestre genebrino utiliza o exemplo de uma moeda de 5 francos (cf. Saussure, 2012 [1916], p.162). Para estabelecer o valor deste signo “5 francos”, é necessário, primeiramente, trocá-lo por algo distinto, como “pão”, e depois, colocá-lo em relação a outro elemento de valor semelhante dentro do sistema, como uma moeda de 1 franco.

Saussure também observa que os valores em sistemas linguísticos não são fixos, já que é preciso uma rede dinâmica de relações tanto com signos semelhantes quanto com signos distintivos para defini-los. Ele parece reforçar essa ideia afirmando que “o conteúdo [o valor] só é realmente determinado pela interação com o que existe fora [da palavra]” (Saussure, 2012 [1916], p.162), indicando que o valor é relacional e que também ultrapassa as relações internas de cada signo, entre o significante e o significado;

² No original: Il y a un phénomène déjà par le fait que cette différence est une des choses qui contribuent à la signification.

estabelecendo relações externas - valores positivos e negativos - que dependem da posição de cada elemento no sistema.

As referidas ideias em torno do valor como fruto de relações parece sugerir que, na língua, os valores demarcam “espaços” para cada signo. Ou seja, cada signo ocupa uma parte na língua que lhe é dada pelas diferenças entre o signo e a sua totalidade sistêmica. Esse mesmo entendimento, assim, pode ser notado por meio de outra metáfora contida no CLG, a saber, o Jogo de Xadrez (cf. Saussure, 2012 [1916]).

A metáfora do jogo de xadrez, em linhas gerais, pode servir de plano de fundo para diferentes pressupostos saussurianos, como relembra Normand (cf. 2009). Porém, neste artigo, trabalhamos com a interpretação de que essa metáfora ilustra a ideia de que o valor dos signos em um sistema depende de suas posições e das relações que eles mantêm com os outros elementos. No xadrez, cada peça tem um papel específico, e seu valor é definido não apenas por suas características intrínsecas, ou seja, o fato do cavalo andar em “L”, mas principalmente pela sua posição relativa e pelas interações com as demais peças no tabuleiro. Em outras palavras, importa que o movimento do cavalo é um; e o da torre é diferente do movimento do cavalo; assim como o movimento do bispo é diferente do cavalo e da torre; e assim por diante. Portanto, uma peça no xadrez só adquire sua verdadeira funcionalidade e importância no contexto das outras peças.

Além disso, mais relevante do que entender como as peças se movimentam no jogo, importa saber a posição delas no final de cada jogada. Ou seja, os movimentos de cada peça, que podemos associar às relações na língua, geram, como produto da jogada, uma posição única no tabuleiro. Analogamente, na língua, essas relações geram, também, uma posição única no sistema para cada signo, isto é, os seus valores.

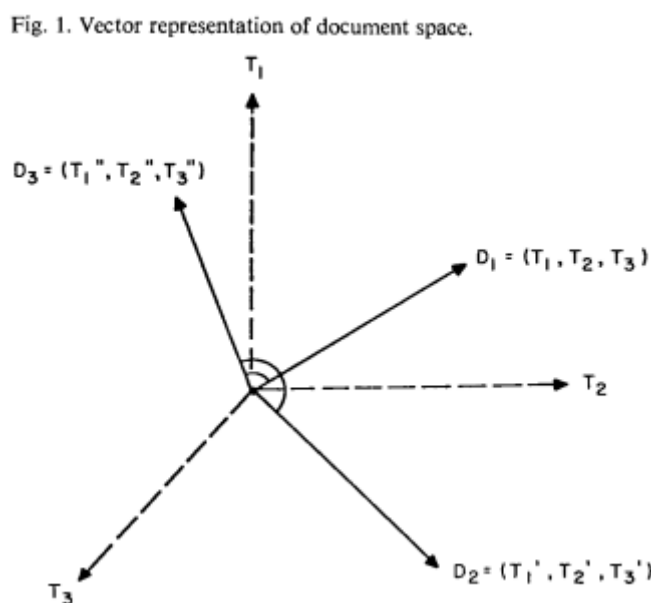
Sendo assim, assim como no xadrez, em que a posição de cada peça é essencial para definir seu valor no contexto de cada partida, no sistema linguístico, cada signo adquire uma posição na língua, um valor, que lhe confere características únicas frente à totalidade sistêmica. O valor, então, torna-se relacional e dependente das posições e da interação do signo com os demais, o que faz com que o sistema da língua engendre, por sua própria natureza, um “sistema de relações” (Saussure, 2012 [1916], p.100).

Em síntese, a teoria segundo a qual o valor dos signos é determinado por suas relações de semelhança e dessemelhança, gera, portanto, uma posição para cada signo dentro do sistema. Esse entendimento, assim, parece ter sido de grande importância para a formulação da Hipótese Distributiva (HD), uma teoria da significação preconizada por Zellig Harris, nos EUA, e por John R. Firth, na Inglaterra, na metade do século XX, em que o aspecto semântico da língua está relacionado à distribuição dos significados dentro de um *corpus* a partir das relações entre os termos do sistema. Como efeito da HD, a maneira de entender as relações dos signos como essenciais para a construção da significação

alcançou o Processamento de Linguagem Natural a partir da década de 70, com a publicação do artigo “A Vector Space Model for Automatic Indexing”, de Gerard Salton, A. Wong e C. S. Yang (1975).

O artigo, em suma, explora um modelo vetorial para indexação automática de documentos, voltado para recuperação de informações. A abordagem busca representar documentos como vetores³ em um espaço tridimensional, em que cada dimensão corresponde a um termo indexado. Como consequência, a similaridade entre documentos é calculada através do produto interno entre os vetores, ou pelo ângulo entre eles, o que indica a proximidade semântica, como vemos na figura a seguir:

Figura 1. Representação vetorial de um conjunto de documentos.

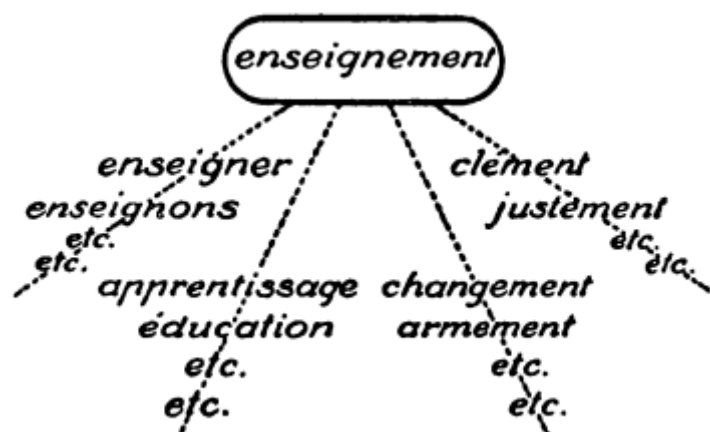


Fonte: Salton; Wong; Yang, 1975, p.614.

Nesse sentido, a Figura 1 mostra como signos podem ser convertidos em vetores num espaço vetorial, ou seja, em um gráfico onde cada palavra admite relações de semelhança e dessemelhança com as outras, marcadas pelo cálculo do ângulo de abertura dos vetores. Na linguística saussuriana, encontramos uma ilustração do CLG que parece muito próxima dessa representação vetorial da língua proposta por Salton, Wong e Yang (1975), a saber:

³ Em Semântica Vetorial, um vetor é uma maneira de representar o significado de palavras em forma numérica, o que ajuda o computador a "entender" e manipular o significado de palavras. Por exemplo, se imaginarmos que cada palavra é um vetor no espaço tridimensional (como um mapa em 3D), o tamanho desses vetores, e o ângulo entre eles, indicará a semelhança entre os significados. Assim, palavras que têm significados parecidos, como "gato" e "cachorro", ficam próximas nesse "mapa", enquanto palavras com significados diferentes, como "gato" e "computador", ficam distantes.

Figura 2. Rede de relações da palavra “enseignement”.



Fonte: Saussure, 1995, p.175.

Com base nesse esquema contido no CLG, em que Saussure discute a rede de analogias de “enseignement”, demonstrando linhas que podem ser interpretadas também como vetores, Gastaldi (2020) indica articulações diretas entre o saussuriano e a modelagem semântica de WE, chegando a afirmar que “É impossível não reconhecer no diagrama de Saussure o mecanismo presente nos espaços de embeddings, conforme caracterizado anteriormente [...]”⁴ (Gastaldi, 2020, p.199)

Voltando-nos à pesquisa de Salton, Wong e Yang (1975), foram considerados três termos para análise dos autores, representados por: T1T_1T1, T2T_2T2 e T3T_3T3. Estes termos, então, são como “temas” ou “palavras-chave” importantes para identificar o conteúdo dos documentos. Cada termo, por conseguinte, está representado por uma linha que aponta em uma direção diferente. Já as notações D1D_1D1, D2D_2D2 e D3D_3D3 estão representadas como documentos nos quais se buscam as similaridades temáticas. Com isso, inferimos que:

- (i) O Documento D1D_1D1 está mais próximo de T1T_1T1, o que significa que ele provavelmente usa muito esse termo ou é mais relacionado a esse tema;
- (ii) O Documento D2D_2D2 está mais próximo de T2T_2T2, indicando que esse termo é mais importante para o conteúdo do D2D_2D2.
- (iii) O Documento D3D_3D3 está entre T1T_1T1, T2T_2T2 e T3T_3T3, indicando que ele contém uma mistura desses temas.

Nesse sentido, o modelo vetorial proposta por Salton, Wong e Yang (1975), anos mais tarde, foi retomado por autores como Mikolov et al. (2013), com o desenvolvimento do Word2Vec, um dos primeiros modelos de linguagem amplamente usados para gerar representações de palavras em alta dimensionalidade - vetores com várias direções - que

⁴ No original: “It is impossible not to recognize in Saussure’s diagram the mechanism at play in embedding spaces as we characterized them earlier [...]”

preservam informações semânticas e contextuais mais complexas e profundas.

Portanto, o aporte teórico que sustenta o modelo de linguagem apresentado por Mikolov et al (ibidem) é o mesmo de Salton, Wong e Yang (1975), a saber, a Hipótese Distributiva. Porém, a inovação de Mikolov e seus colegas reside na inovação da representação vetorial em um espaço multidimensional, enquanto que, em Salton, Wong e Yang, tem-se uma representação tridimensional. De igual forma, ambos os modelos possuem um cerne teórico comum, isto é, a ideia de que a posição relativa das palavras dentro do sistema é fundamental para a delimitação dos sentidos e para, posteriormente, uma máquina poder recuperar significados e informações. Nas palavras de Firth: “Reconhecerás uma palavra pela companhia que ela mantém” (Firth, 1955, p.108).

Ademais, vale dizer que a SD, hoje, mantém-se como um dos principais aportes teóricos para o desenvolvimento de modelos de linguagem vetorial, como os Word Embeddings, como bem aponta Seno et. al (2023):

A semântica distribucional tem sido atualmente a principal abordagem de representação do significado lexical adotada nas mais diversas tarefas do processamento de linguagem natural. Nessa abordagem, os itens lexicais (palavras) são representados por meio de vetores de *valores* reais, conhecidos por vetores semânticos, que codificam o significado das palavras a partir de sua distribuição em textos (Seno et. al, 2020, p.1, grifo nosso).

Nessa perspectiva, nossa hipótese investigativa é de que a conexão entre a SD e o desenvolvimento de *word embeddings* sugere articulações, em níveis teóricos, entre a modelagem semântica de WE e o saussurianismo, mediadas por reinterpretações da noção de valor e de relação que a SD promove. Em outras palavras, defendemos, neste trabalho, que o que a SD entende por mecanismo de significação - aporte teórico utilizado na modelagem semântica de WE - parece, na verdade, a determinação de valores linguísticos tal como já fora previsto por Saussure no início do século 20.

No próximo tópico, então, vamos nos dedicar a esclarecer alguns aspectos importantes da SD e suas possíveis relações com a noção de relação e de sistema presentes no saussurianismo, preparado-nos para a seção subsequente sobre a análise de *frameworks* de WE à luz da TdV.

2 Semântica distribucional: uma releitura estruturalista das noções saussurianas de *valor* e *relação*

A Semântica Distribucional (SD) é uma teoria da significação desenvolvida em meio à efervescência dos estruturalismos estadunidense e inglês, a qual, hodiernamente, é amplamente utilizada no Processamento de Linguagem Natural. Na SD, entende-se que a importância do contexto e da distribuição das palavras em textos é a chave para o sentido

da língua. Essa teoria foi proposta inicialmente por linguistas como Zellig Harris (1909-1992)⁵ e J.R. Firth (1890-1960), que buscavam uma maneira de entender a semântica das palavras com base na co-ocorrência em diferentes contextos.

Harris, em sua obra seminal intitulada *Distributional structures* (1954), infere que o significado de uma palavra não deve ser entendido de forma isolada, mas em relação à sua distribuição em um corpus textual. O autor utiliza o exemplo de “cão” e “gato”, que, em níveis de semântica distributiva, possuem traços de marcações comuns: ambos animais domésticos, de quatro patas, que comem ração etc. Essas marcações comuns indicam que, num sintagma como “O cachorro fugiu”, o sujeito poderia ser substituído por “O gato fugiu” sem muitos prejuízos semânticos. Por essas semelhanças, por exemplo, “cão” e “gato” possuem valores aproximados no sistema, o que, em PLN, significa que ambas as palavras estariam próximas num mesmo espaço vetorial.

No mesmo entendimento, Saussure (2012 [1916]) já havia previsto que o sentido na língua não está contido na palavra, mas no concurso que existe fora dela, a saber, as relações que ocorrem no exterior dos signos, formando uma rede de valores positivos, semelhantes, e negativos, dessemelhantes.

Porém, é, no mínimo, intrigante, que, embora Harris tenha tido contato com o Curso de Linguística Geral, são poucas as vezes em que o linguista norteamericano faz referência direta ao mestre genebrino, deixando em aberto os motivos para esse suposto apagamento de Saussure na escola norteamericana estruturalista; principalmente porque muitos pressupostos saussurianos são notadamente perceptíveis numa leitura rápida da linguística harriana, como a noção de sistema e a ênfase às relações que dela derivam. No *Distributional Structures*, contudo, o nome de Saussure aparece uma única vez, numa pequena nota de rodapé em que Harris explica seu interesse por uma ciência autônoma da linguagem, tal como já havia sido proposto por Saussure: “O estudo independente da estrutura descritiva foi esclarecido em grande parte pelo “Cours de linguistique générale” de Ferdinand de Saussure.”⁶ (Harris, 1954, p.149).

Firth, igualmente importante para a SD, por sua vez, faz recorrentes menções a Saussure (Firth, 1934; 1949; 1950). Nesse sentido, a maior aproximação teórica entre o mestre genebrino e o mestre londrino pode ser evidenciada no trabalho de Firth intitulado *Linguistics and the functional point of view*, em que o autor parte da máxima saussuriana de que *o ponto de vista cria o objeto* (cf. Saussure, 2012 [1916]) para fundamentar a sua visão estruturalista e funcional para a língua:

⁵ A título de observação, Zellig Harris foi o orientador da tese de doutorado de Noam Chomsky, como aponta Brunilla e La Violette (2023, p.3): “Few linguists contributed more to linguistic theory than Zellig Harris (1909–1992), and not just by serving as Noam Chomsky’s doctoral advisor. In fact, the two came to share little in common (Goldsmith, 2005; Nevin, 2010).”

⁶ No original: “The independent study of descriptive structure was clarified largely by Ferdinand de Saussure’s Cours de linguistique générale”.

Qualquer sistema de conhecimento é uma forma de pensamento sobre os fatos. As estrelas não fazem a astronomia. As estrelas podem ser o objeto, os fatos, mas é o astrônomo quem faz a astronomia. Alguns de nossos astrônomos estão, inclusive, criando seus próprios universos, compondo, por assim dizer, sua própria música das esferas. E isso não é tão absurdo quanto parece à primeira vista. Sem conhecimento prévio, o que um homem vê que realmente importa? Como observa Saussure: “Bem longe de o objeto preceder o ponto de vista, poder-se-ia dizer que é o ponto de vista que cria o objeto”⁷ (Firth, 1934, p.18).

Entretanto, vale dizer que as várias menções a Saussure feitas por Firth, nem sempre, estão no nível da concordância. No mesmo trabalho *Linguistics and the functional point of view*, por exemplo, o autor chega a afirmar que a sua hipótese funcionalista para a língua, isto é, a ideia de que o sentido está relacionado ao uso real em diferentes contextos, necessita de uma radical superação da linguística saussuriana. Em outras palavras, ainda que ambos partam das relações sistêmicas para formular suas teorias sobre a natureza da língua, para Saussure (2012 [1916]), o indivíduo não é capaz de alterar as leis de funcionamento do sistema; enquanto que, para Firth, as leis linguísticas não são exteriores à massa de falante, mas partes dela, materializadas no uso humano da língua, que, na linguística firthiana, é contínua, dinâmica, intencional e criativa.

O ponto de vista funcional exige uma unidade de investigação totalmente diferente e muito mais abrangente, rompendo todas as fronteiras de Saussure. Esse ponto de vista não consegue observar a linguagem em funcionamento se estiver restrito apenas às relações opostas ou aos termos de um código estático de signos. Sob a perspectiva funcional, vê-se o ser humano em sua totalidade em ação com seus semelhantes, e a linguagem como diferentes modos de ação em contextos de situação. A linguagem não é uma espécie de vontade geral linguística intelectualizada, mas sim parte do próprio ser humano e de sua atividade contínua, dinâmica, intencional e criativa (Firth, 1934, p.20)⁸

A Semântica Distributiva, portanto, se distingue de outras abordagens semânticas por seu foco nas relações contextuais e na co-ocorrência de palavras. Em matéria de PLN, a pesquisa em SV é concomitante ao desenvolvimento das teorias de Firth e de Harris, em meados da década de 50, investigações essas que foram “impulsionadas pela convicção de que o significado de uma palavra pode ser definido a partir da sua distribuição nos contextos linguísticos em que ela ocorre” (Seno et. al, 2023, p.2).

⁷ No original: “Any system of knowledge is a form of thought about facts. The stars do not make astronomy. The stars may be the subject matter, the facts, but it is the astronomer who makes astronomy. Some of our astronomers are making their own universes as well, composing as it were their own music of the spheres. And this is not so absurd as it seems at first sight. Without previous knowledge what man sees anything that matters ? As de Saussure remarks "Bien loin que l'objet précède le point de vue, on dirait que c'est le point de vue qui crée l'objet”.

⁸ No original: “The functional point of view requires an altogether different and much more comprehensive unit of investigation, violating all de Saussure's frontiers. It cannot see language functioning if it confines itself to the opposite relations or the terms of a static code of signs. From the functional point of view you see your whole man in action with his fellows, and his language as various modes of action in contexts of situation. Language is not a sort of intellectualised linguistic general will — but part of the man himself and his continuing activity, dynamic, purposive, creative.”

Precisamente, esses “contextos linguísticos” a que se referem Seno et. al (ibidem), posteriormente, serviram de base para a criação de um modelo de linguagem baseado numa noção espacial de língua; ou, como tratamos neste trabalho, numa visão de língua que reside na noção de um sistema de valores em que cada signo ocupa um espaço no sistema. Essa noção espacial da língua, preconizada, como vimos, por Salton, Wong e Yang (1975), tornou-se fundamental para a proposição de Mikolov et al (2013) para o PLN, a saber, a criação de um modelo de linguagem capaz de capturar significados por meio das relações estabelecidas entre as palavras dentro de um espaço multidimensional, modelo este conhecido como Word2Vec⁹.

Nessa perspectiva, os WE, em linhas gerais, são representações vetoriais de palavras que buscam capturar similaridades e disparidades semânticas com base em contextos de uso a partir de um espaço multidimensional, isto é, mais de três dimensões (tridimensional). Os Word Embeddings, amplamente aplicados para tarefas como análise de sentimento, tradução e classificação de corpora, codificam cada palavra como um vetor em que a proximidade entre o ângulo de abertura dos vetores reflete a semelhança de significados (cf. Mikolov et al, 2013).

Em modelos vetoriais de linguagem, a modelagem semântica de palavras pode ser realizada através de duas abordagens distintas: por vetores densos e por vetores esparsos (Seno et. al, 2023). Essas representações são fundamentais para o processamento de linguagem natural (PLN) e têm implicações significativas nas performances de diversos algoritmos.

Os vetores densos são caracterizados por uma representação onde a maioria de suas entradas é “não nula”¹⁰. Essa forma de representação é tipicamente utilizada em modelos como Word2Vec, GloVe e FastText, que produzem vetores de tamanho variados (por exemplo, 100 ou 300 dimensões¹¹) para cada palavra.

Essa abordagem de vetores densos permite que as relações semânticas e contextuais entre as palavras sejam capturadas de maneira eficaz, visto que cada dimensão do vetor denso encapsula informações sobre características específicas da palavra em seu contexto.

Os vetores esparsos, por sua vez, apresentam uma estrutura onde a maioria das

⁹ Word2Vec é um modelo de aprendizado de máquina desenvolvido para gerar *word embeddings*, ou representações vetoriais de palavras, com o objetivo de capturar suas relações semânticas. Criado por Tomas Mikolov e sua equipe no Google em 2013, Word2Vec utiliza redes neurais para aprender associações entre palavras com base em grandes corpora de texto. Esse modelo transforma cada palavra em um vetor em um espaço multidimensional, de modo que a proximidade entre vetores representa a similaridade semântica das palavras.

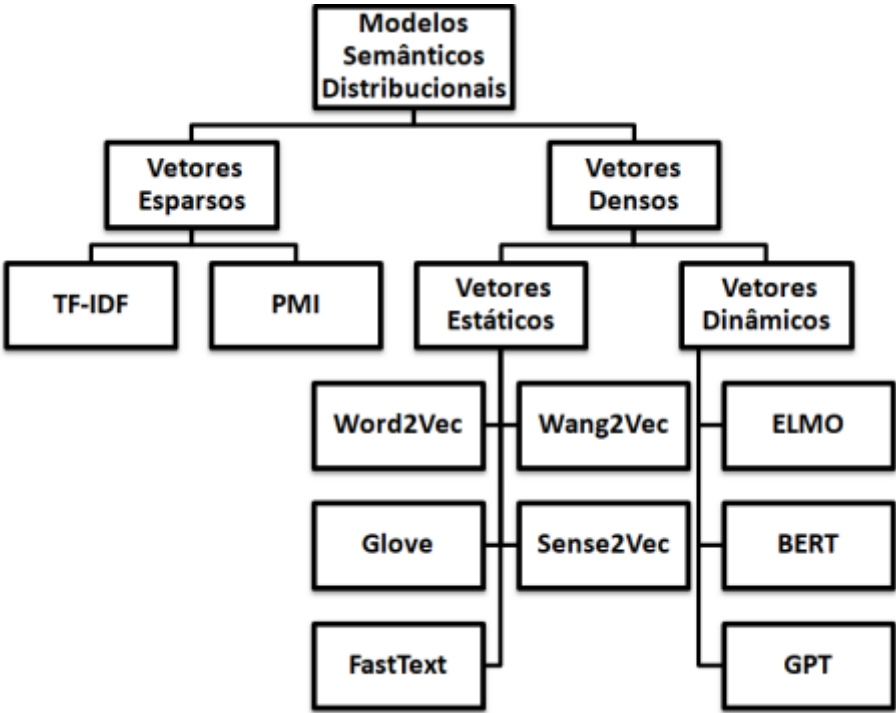
¹⁰ Considerando o vetor $v=[1,2,0,4,5]$, tem-se que 4 de suas 5 entradas são não-nulas (1, 2, 4 e 5), portanto, é um vetor denso.

¹¹ Com mais dimensões, Mikolov et al (2013) demonstraram que um vetor pode representar relações semânticas mais complexas. Por exemplo, em modelos de embeddings de palavras, como o Word2Vec, as dimensões podem capturar diferentes aspectos do significado, como gênero, número, ou contextos de uso. Assim, vetores com mais dimensões podem melhor refletir similaridades semânticas e relações entre palavras.

entradas é zero¹². Essa representação é comumente associada a métodos como "bag of words" (saco de palavras) e "TF-IDF" (Term Frequency-Inverse Document Frequency), que capturam menos informações da palavra do que os vetores densos. Nesse contexto, cada dimensão do vetor esparsos corresponde a uma palavra do vocabulário, e os valores refletem a frequência dessa palavra em um documento ou seu peso em relação a outros termos.

Na Figura a seguir, Seno et. al (2023) apresentam um diagrama que demonstra os modelos de linguagem vetoriais de vetores densos e esparsos atualmente disponíveis:

Figura 3. Esquema de modelos semânticos distribucionais.



Fonte: Extraído de Seno et. al (2023, p.3).

Essa técnica de modelagem semântica de modelos de linguagem vetoriais, assim, permite que algoritmos entendam diferentes relações semânticas a partir da noção de que palavras usadas em contextos similares tendem a ter vetores próximos no espaço vetorial,. No contexto de nossa pesquisa, entendemos que as palavras usadas em contextos similares tendem a ter esse comportamento porque, primeiramente, possuem valores linguísticos (Saussure, 2012 [1916]) parecidos. Na próxima seção, então, analisaremos algumas representações vetoriais geradas pelo Word2Vec buscando evidenciar a importância da determinação dos valores semelhantes e dessemelhantes para a modelagem semântica de WE.

¹² Um vetor como $w=[0,0,0,3,0]$ é considerado esparsos, pois a maioria de suas entradas (4 de 5) são nulas (zero).

3 Modelagem semântica de WE: um possível retorno à TdV mediado pela semântica distributiva

A modelagem semântica de WE refere-se ao processo de representar o significado de palavras, frases e textos em forma de vetores ou estruturas que podem ser manipuladas por algoritmos, permitindo uma emulação mais profunda e contextual da língua humana. Essa abordagem é amplamente utilizada em PLN e por sistemas de recomendação, busca semântica, entre outras áreas.

Um dos modelos de embedding de palavras mais conhecidos, como já dissemos, é o Word2Vec, que opera com vetores densos. Sobre ele, Church (2016) pontua que:

Word2vec não é o primeiro, o último ou o melhor [modelo de linguagem] para discutir espaços vetoriais, embeddings, analogias, métricas de similaridade etc. Mas o word2vec é simples e acessível. Qualquer um pode baixar o código e usá-lo em seu próximo artigo. E muitos o fazem (para o melhor e para o pior) (Church, 2016, 156)¹³.

No mesmo entendimento, Aguiar e Prati argumentam sobre a “simplicidade” do Word2Vec, assumindo que:

O modelo [Word2Vec] consiste em criar representações vetoriais de palavras utilizando técnicas de redes neurais, e fornece duas arquiteturas eficientes de implementação, sendo elas uma baseada em Bag-of-Words, cujo objetivo é prever a probabilidade de um termo ocorrer, baseado em uma janela de palavras próximas e o Skip-Gram, em que procura-se maximizar a predição de quais termos ocorrem próximas a um determinado termo (Aguiar; Prati, 2015, p.3)

Na figura a seguir, apresentamos a representação em semântica vetorial de algumas palavras em Inglês geradas pelo modelo Word2Vec.

¹³ No original: “Word2vec is not the first, last or best to discuss vector spaces, embeddings, analogies, similarity metrics, etc. But word2vec is simple and accessible. Anyone can download the code and use it in their next paper. Any many do (for better and for worse).”

Figura 4. Representação vetorial pelo modelo Word2Vec

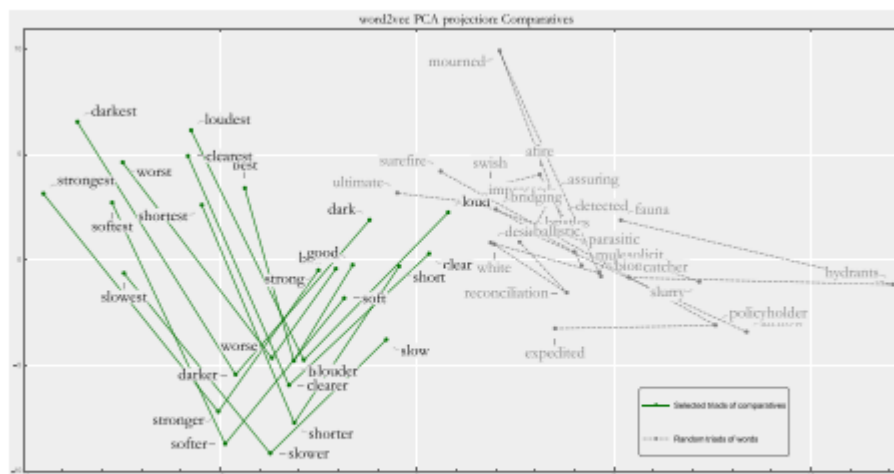


Fig. 3 Pattern in the embedding space (word2vec) corresponding to the comparative category (base, comparative and superlative forms)

Fonte: extraído de Gastaldi, 2020, p.160

Nesse sentido, a Figura 4, extraída do trabalho *Why Can Computers Understand Natural Language? The Structuralist Image of Language Behind Word Embeddings*, do linguista computacional Gastaldi (2020, p.160), mostra uma análise comparativa, utilizando o Word2Vec, entre diferentes formas de um mesmo adjetivo, a saber, *Strong*, *stronger* e *strongest*. As linhas verdes, assim, quase paralelas, buscam representar as possíveis relações semânticas entre as referidas palavras, e as relações que delas derivam. Os pontos cinzas, por sua vez, representam palavras aleatórias, que estão dispersas no espaço sem padrões evidentes de agrupamento, servindo como um fundo para evidenciar as palavras da categoria comparativa em destaque.

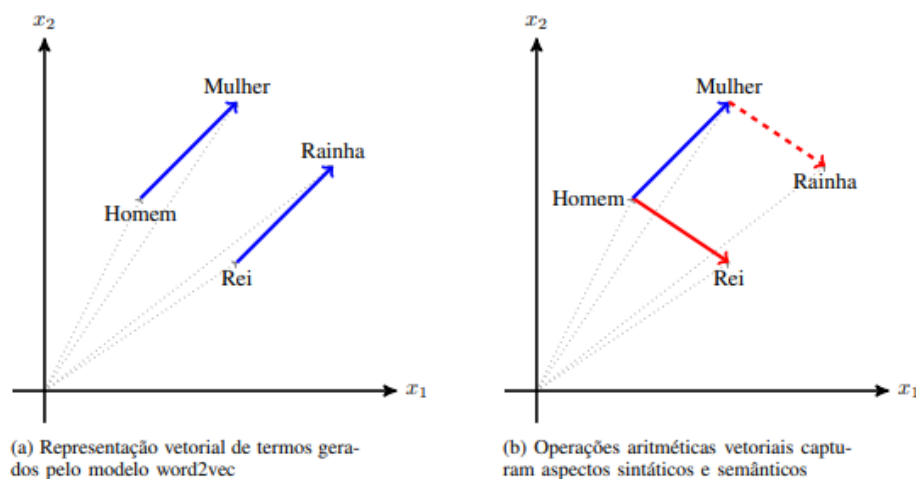
Ao relacionar esse framework gerado pelo Word2Vec com a TdV de Saussure, podemos observar que o modelo captura as relações diferenciais entre as palavras, isto é, os seus valores positivos (semelhantes) e negativos (dessemelhantes). Observa-se, assim, duas vias de relações para delimitar qual o espaço da palavra - do vetor - neste mapa vetorial. Primeiramente, temos as relações de semelhança, representadas pelas palavras que possuem marcações semânticas comuns, as quais assumem um espaço mais próximo à medida que há maior similaridade semântica. E, não obstante, as relações de dessemelhança, representadas pelas palavras em cor cinza, que possuem marcações semânticas distintas.

O framework acima, portanto, pode ilustrar a maneira como o word2vec utiliza o contexto das palavras para apreender essas relações diferenciais, conferindo ao signo, isto é, ao vetor, um espaço próprio - um valor - que é necessariamente constituído de semelhanças e diferenças.

Apresentamos, agora, um outro framework para análise, também gerado pelo

Word2Vec, extraído do trabalho *Incorporação de representação vetorial distribuída de palavras e parágrafos na classificação de SMS SPAM*, de Aguiar e Prati (2015):

Figura 5. Exemplo de representação vetorial criada pelo word2vec. Estão representados os vetores dos termos “homem”, “mulher”, “rei” e “rainha”. Observa-se que, em (a), a orientação e norma do homem→mulher é a mesma de rei→rainha, e em (b), a operação vetorial rei-homem+mulher = rainha



Fonte: extraído de Aguiar e Prati, 2015, p.4

Como se nota, a figura ilustra a representação vetorial de palavras usando o modelo word2vec. Nela, observamos como o modelo organiza semanticamente os signos de valor positivo, semelhantes, como "homem," "mulher," "rei," e "rainha". A primeira parte da imagem (a) mostra a similaridade de orientação e magnitude entre os vetores "homem → mulher" e "rei → rainha", enquanto a segunda (b) demonstra que operações aritméticas, como a soma vetorial "rei - homem + mulher," resultam em um vetor próximo ao vetor de "rainha". Do ponto de vista da semântica distributiva, a Figura 5 retoma a alguns princípios da hipótese distributiva, isto é, a similaridade vetorial entre "homem" e "mulher" e entre "rei" e "rainha" reflete a semelhança de contexto, indicando que esses termos frequentemente compartilham papéis gramaticais ou semânticos comuns.

Do ponto de vista do saussurianismo, cada termo na Figura 5 ganha seu significado não de forma isolada, mas em relação aos outros. Relações como "homem → mulher" e "rei → rainha" mostram que o significado de uma palavra é construído pelo contraste com outras palavras. Nesse mesmo sentido, a operação vetorial que aproxima "rei - homem + mulher" de "rainha" exemplifica como o valor das palavras pode ser representado matematicamente a partir, também, de valores negativos, isto é, diferenciais. Isso mostra que palavras usadas em contextos semelhantes têm não só valores linguísticos aproximados, como também, por lógica, ocupam espaços semelhantes no sistema.

Nessa perspectiva, podemos afirmar que a Teoria do Valor de Saussure parece oferecer uma base sólida para compreender a modelagem semântica de Word Embeddings

(WE). Em outras palavras, a noção saussuriana de que o valor de um signo linguístico é determinado pelas suas relações diferenciais e similares dentro de um sistema encontra eco na hipótese distributiva de Harris e Firth, em que o contexto e a co-ocorrência das palavras são fundamentais para a construção de significado. A semântica distributiva, ao valorizar as relações entre as palavras nos diferentes contextos, evidencia a interdependência entre os elementos do sistema linguístico, conforme sugerido por Saussure.

Dessa forma, o desenvolvimento de WE, como o Word2Vec, beneficia-se diretamente desses princípios, permitindo que os modelos de IA capturem relações semânticas complexas por meio de operações vetoriais que traduzem valores linguísticos de semelhança e dessemelhança em vetores próximos ou distantes num espaço multidimensional. Por isso, argumentamos que a modelagem semântica em IA pode ser vista, também, como uma espécie de "retorno" e uma aplicação prática da Teoria do Valor, a partir da ênfase ao sistema de relações promovido por modelos de linguagem baseados na Semântica Vetorial de Harris e Firth, como os Word Embeddings..

Considerações finais

Este estudo procurou explorar a influência da Teoria do Valor (TdV) de Ferdinand de Saussure na modelagem semântica de Word Embeddings (WE), em especial através da semântica distributiva. Argumentamos que a TdV, com seu foco nas relações diferenciais e de semelhança entre os signos, antecipa princípios fundamentais utilizados na Semântica Distribucional (SD), uma abordagem que guia o desenvolvimento de WE no Processamento de Linguagem Natural (PLN).

A partir da hipótese distributiva de Harris (1954) e Firth (1934; 1939), a SD considera o contexto e a coocorrência de palavras como chaves para a construção do significado, algo que também aparece na linguística saussuriana na medida em que o valor de um signo é definido pelas relações que mantém com outros signos dentro de um sistema. Em modelos como o Word2Vec, essa visão é implementada matematicamente por meio de vetores que representam palavras em um espaço multidimensional, onde a proximidade entre vetores indica similaridade semântica. Assim, esses modelos capturam o significado das palavras ao localizar cada termo em relação a outros, refletindo as interdependências diferenciais e de semelhança propostas por Saussure.

Dessa forma, propomos que a modelagem semântica de WE, especialmente em abordagens baseadas na SD, pode ser vista como uma extensão prática da TdV. Essa ligação sugere que a TdV ainda possui relevância e aplicabilidade na compreensão e no desenvolvimento de tecnologias linguísticas, servindo como um fundamento teórico para

sistemas de IA capazes de emular aspectos complexos da língua humana. Em última instância, o estudo aponta para a pertinência de alguns conceitos saussurianos no século XXI, especialmente na criação de modelos de linguagem que exigem uma modelagem semântica relacional para a construção de significados por Inteligências Artificiais.

Saussure in the age of AI: semantic modeling of word embeddings in light of value theory

Abstract

This study explores the relationship between Ferdinand de Saussure's Theory of Value (TdV) and the semantic modeling of Word Embeddings (WE), particularly through Distributional Semantics (SD). In this work, we propose an analysis of some word embedding frameworks generated by the Word2Vec model, seeking possible connections between TdV and meaning capture in language models based on Distributional Semantics. Starting from the hypothesis that Saussure's notion of value lies in the relations of similarity and dissimilarity between a sign and its systemic entirety, we aim to demonstrate that this same principle regarding linguistic value also appears in Distributional Semantics, advocated by linguists Z. Harris in the USA and J. R. Firth in England in the 1950s. In this sense, we present evidence that, like Saussure, SD also relies on systemic relations to propose a theory of meaning in language. This theory, in turn, has been widely applied in Natural Language Processing for WE semantic modeling, enabling discussions about the potential relevance of Saussurean ideas for developing new NLP techniques in the 21st century.

Keywords: Saussure; Computational Linguistics; Natural Language Processing; Theory of Value; Word Embeddings

Referências

AGUIAR, R; PRATI, R. Incorporação de representação vetorial distribuída de palavras e parágrafos na classificação de SMS SPAM. In: **XII Encontro Nacional de Inteligência Artificial e Computacional**, 2015. Disponível em: https://www.researchgate.net/publication/281348546_Incorporacao_de_representacao_vetorial_distribuida_de_palavras_e_paragrafos_na_classificacao_de_SMS_SPAM. Acesso em: 4 nov. 2024.

CHURCH, K. W. Word2Vec. **Natural Language Engineering**, v. 23, n. 1, p. 155–162, 2017. Disponível em: <https://doi.org/10.1017/S1351324916000334>. Acesso em: 4 nov. 2024.

COELHO, M. P. “Significação” em Saussure: os três cursos de linguística geral. **Anais do SILEL**, Uberlândia, v. 3, n. 1, p. 943-956, 2013b. Disponível em: https://www.ileel.ufu.br/anaisdosilel/wp-content/uploads/2014/04/silel2013_943.pdf. Acesso em: 8 out. 2024.

COELHO, M. P. Sistema e relação na Teoria do Valor de Ferdinand de Saussure. **Estudos Linguísticos**, São Paulo, v. 42, n. 1, p. 378-391, 2013a. Disponível em: <https://revistas.gel.org.br/estudos-linguisticos/article/view/878/1179>. Acesso em: 8 out. 2024.

COELHO, M. P. **A noção de sistema na fundação da linguística moderna**. 2015. 129 f. Dissertação (Mestrado em Estudos Linguísticos) - Universidade Federal de Uberlândia, Uberlândia, 2015. DOI: <https://doi.org/10.14393/ufu.di.2015.307>.

COELHO, M. P.; SILVA E LIMA, T. R. Língua, linguagem e fala na “Teoria do Valor” de Ferdinand de Saussure. **Estudos Linguísticos** (São Paulo. 1978), [S. l.], v. 43, n. 01, p. 347–357, 2015. Disponível em: <https://revistas.gel.org.br/estudos-linguisticos/article/view/443>. Acesso em: 9 out. 2024.

DUCROT, O. **Estruturalismo e Linguística**. São Paulo: Cultrix, 1968.

FIRTH, J. R. Linguistics and the functional point of view. **English Studies**, v. 16, n. 1-6, p. 18–24, 1934. Disponível em: <https://doi.org/10.1080/00138383408596618>. Acesso em: 4 nov. 2024.

FIRTH, J. R. Personality and Language in Society. **The Sociological Review**, v. a42, n. 1, p. 37-52, 1950. Disponível em: <https://doi.org/10.1111/j.1467-954X.1950.tb02460.x>. Acesso em: 4 nov. 2024.

FIRTH, J. R. The semantics of linguistic science. **Lingua**, v. 1, p. 393-404, 1949. Disponível em: [https://doi.org/10.1016/0024-3841\(49\)90085-6](https://doi.org/10.1016/0024-3841(49)90085-6). Acesso em: 4 nov. 2024.

FLORES, V. do N. **A Linguística Geral de Ferdinand de Saussure**. São Paulo, SP: Editora Contexto, 2023.

GASTALDI, J. L. Why Can Computers Understand Natural Language? **Philosophy & Technology**, v. 34, p. 149–214, 2021. Disponível em: <https://doi.org/10.1007/s13347-020-00393-9>. Acesso em: 4 nov. 2024.

HARRIS, Z. S. Distributional Structure. **WORD**, v. 10, n. 2-3, p. 146-162, 1954. Disponível em: <https://doi.org/10.1080/00437956.1954.11659520>. Acesso em: 4 nov. 2024.

MARQUES, A. C. M. **A noção de relação na teoria linguística de Ferdinand de Saussure**. 2016. 117 f. Dissertação (Mestrado em Estudos Linguísticos) - Universidade Federal de Uberlândia, Uberlândia, 2016. DOI: <http://doi.org/10.14393/ufu.di.2016.180>.

MIKOLOV, T.; CHEN, K.; CORRADO, G.; DEAN, J. **Efficient estimation of word representations in vector space**. arXiv preprint arXiv:1301.3781, 2013. Disponível em: <https://arxiv.org/abs/1301.3781>. Acesso em: 4 nov. 2024.

NORMAND, C. **Saussure**. Trad. Ana de Alencar e Marcelo Diniz. São Paulo: Estação Liberdade, 2009.

SALTON, G.; WONG, A.; YANG, C. S. **A vector space model for automatic indexing**. *Communications of the ACM*, v. 18, n. 11, p. 613-620, 1975.

SAUSSURE, F. **Curso de linguística geral**. Trad. de A. Chelini; J. P. Paes e I. Blikstein. 27ª ed. São Paulo: Cultrix, 2012. Cours de linguistique générale. Charles Bally e Albert Sechehaye (orgs.), com a colaboração de Albert Riedlinger, [1916].

SAUSSURE, F. **Cours de linguistique générale**. Grande Bibliothèque Payot. Paris: Editions Payot & Rivages, 1995.

SAUSSURE, F. **Troisième Cours de Linguistique Générale (1910-1911): d'après les cahiers d'Emile Constantin / Saussure's third course of lectures on general linguistics (1910-1911): from the notebooks of Emile Constantin**. French text edited by Eisuke Komatsu e English text edited by Roy Harris. Pergamon Press, 1993.

SENO, E. R. M.; CLARO, D.; MOTA, L.; RODRIGUES, J. Semântica Distribucional. In: Caseli, H.M.; Nunes, M.G.V. (org.) **Processamento de Linguagem Natural: Conceitos, Técnicas e Aplicações em Português**. 2 ed. BPLN, 2024. Disponível em: <https://brasileiraspln.com/livro-pln/2a-edicao>.

SILVEIRA, E. **As marcas do movimento de Saussure na fundação da Linguística**. Campinas, SP: Mercado de Letras, 2007.

SILVEIRA, E. A teoria do valor no Curso de Linguística Geral. *In*: Revista **Letras & Letras**, Uberlândia, Edufu, v. 25, n. 1, p. 39-54, 2009.