

**LEXICOGRAFIA PEDAGÓGICA EM TEORIA SENTIDO-TEXTO**  
**PEDAGOGICAL LEXICOGRAPHY IN MEANING-TEXT THEORY**  
**LEXICOGRAFÍA PEDAGOGICA EN TEORÍA SENTIDO-TEXTO**

Caroline de Castro Pires<sup>1</sup>

**RESUMO:** O presente artigo propõe uma discussão sobre o uso de abordagens pedagógicas ao que se refere às ferramentas lexicográficas de descrição de léxico de uma língua, quando considerado o processo de aprendizagem de léxico. No entanto, diferente do que é comumente trabalhado em linha didática, isto é, propor produtos lexicográficos que visam a aprendizagem de segunda língua, especificamente, com vistas à proficiência linguística nos moldes do ensino de língua estrangeira. Neste artigo, optou-se por apresentar a perspectiva pedagógica da Teoria Sentido-Texto (TST), teoria lexicográfica de base semântica que lança mão da Lexicografia Explicativa e Combinatória para gerar produtos lexicográficos que oferecem uma descrição linguística complexa e bastante completa das unidades lexicais de uma língua, seja pensando a descrição do léxico de língua comum, seja a descrição do léxico das diversas áreas de especialidade, tanto em língua materna, quanto em segunda língua. Tal teoria apresenta uma linha mais didática ao “simplificar” a descrição lexicográfica com foco na “popularização” científica. Assim, traremos os exemplos do Léxico Ativo do Francês (LAF) e dos dicionários desenvolvidos a partir do banco de dados DiCo, que é uma versão eletrônica do Dicionário Explicativo e Combinatório (DEC), principal produto da TST, ambas ferramentas lexicográficas desenvolvidas através desse viés pedagógico.

**Palavras-chave:** Lexicografia Pedagógica. Teoria Sentido-Texto. Lexicografia Explicativa e Combinatória. Produtos Lexicográficos.

**ABSTRACT:** This article proposes a discussion on the use of pedagogical approaches to the lexicographic tools of lexicon description of a language, when considering the lexicon learning process. However, differently from what is commonly studied on in the didactic line, that is, proposing lexicographic products that aim at second language learning, specifically, with a view to linguistic proficiency along the lines of foreign language teaching. In this article, we opted to present the pedagogical perspective of the Meaning-Text Theory (MTT), a semantic based lexicographic theory that makes use of Explanatory and Combinatorial Lexicography to generate lexicographic products that offer a complex and complete linguistic description of the lexical units of a language, whether considering the description of the lexicon of common language, or the description of the lexicon of the various areas of specialty, both in the mother tongue and in the second language. Such a theory presents a more didactic line by “simplifying” the lexicographic description with a focus on scientific “popularization”. Thus, we will bring the examples of the French Active Lexicon (LAF) and the dictionaries developed from the DiCo database, which is an electronic version of the Explanatory and Combinatorial Dictionary (ECD), MTT's main product, both lexicographic tools developed through this pedagogical bias.

**Keywords:** Pedagogical Lexicography. Meanig-Text Theory. Explanatory and Combinatorial Lexicography. Lexicographic Tools.

**RESUMEN:** Este artículo propone una discusión sobre el uso de enfoques pedagógicos de las herramientas lexicográficas de descripción léxica de una lengua, al considerar el proceso de aprendizaje del léxico. Sin embargo, a diferencia de lo que comúnmente se trabaja en la línea didáctica, es decir,

<sup>1</sup> Mestre em Estudos da Linguagem pela UFRGS, Mestre em Linguística pela PUCRS. Doutoranda em Estudos da Linguagem pela UFRGS. Professora no IFRS. caroline.castro@alvorada.ifrs.edu.br. Orcid: <https://orcid.org/0000-0002-0570-9604>.

proponer productos lexicográficos que tengan como objetivo el aprendizaje de una segunda lengua, específicamente, con miras al dominio lingüístico en la línea de la enseñanza de lenguas extranjeras. En este artículo optamos por presentar la perspectiva pedagógica de la Teoría del Texto-Sentido (TST), una teoría lexicográfica de base semántica que hace uso de la Lexicografía Explicativa y Combinatoria para generar productos lexicográficos que ofrecen una descripción lingüística compleja y bastante completa de las unidades léxicas de una lengua, ya sea considerando la descripción del léxico del lenguaje común, o la descripción del léxico de las distintas áreas de especialidad, tanto en la lengua materna como en la segunda lengua. Tal teoría presenta una línea más didáctica al “simplificar” la descripción lexicográfica con un enfoque en la “popularización” científica. Así, traeremos los ejemplos del French Active Lexicon (LAF) y los diccionarios desarrollados a partir de la base de datos DiCo, que es una versión electrónica del Explanatory and Combinatorial Dictionary (DEC), producto principal de TST, ambas herramientas lexicográficas desarrolladas a través de este sesgo pedagógico.

**Palabras Clave:** Lexicografía Pedagógica. Teoría Sentido-Texto. Lexicografía Explicativa y Combinatoria. Productos Lexicográficos.

## CONSIDERAÇÕES INICIAIS

Há uma variedade de pesquisas na literatura lexicográfica voltados para a aprendizagem de léxico de uma língua quando se fala em Lexicografia Pedagógica. Muitos desses estudos têm por intuito proporcionar ao aprendiz ferramentas linguísticas para a descrição de léxico com o objetivo de suprir as necessidades do aprendiz, em seu processo de aprendizagem, ao que competem as tarefas de produção e de compreensão na língua-alvo. Uma ampla parcela desses estudos e ferramentas lexicográficas desenvolvidas é voltada para a aquisição, em grande parte, de segunda língua. Por outro lado, a literatura também apresenta pesquisas que refletem a aplicabilidade do valor pedagógico de uma teoria lexicográfica que pensam a descrição linguística do léxico de língua materna, visando aprendizes que necessitam de um grau de compreensão, por exemplo, de léxico específico de uma determinada área de especialidade.

A busca por uma teoria lexicográfica que permite tal reflexão, ou seja, o uso de ferramentas lexicográficas para a descrição linguística do léxico de língua materna, seja para a descrição de léxico em língua comum ou léxico para fins específicos, levou à escolha da Teoria Sentido-Texto (TST) como perspectiva teórica proposta neste artigo. TST é uma teoria lexicográfica de base semântica, iniciada em 1965, que tem por princípio a descrição das línguas naturais, que se dá por meio de um aparato formal, ou Modelo Sentido-Texto, com vistas à descrição da materialidade linguística, tendo a tradução automática como finalidade primeira. Ademais, a valorosa aplicação da TST no Processamento da Linguagem Natural (PLN) (MILIĆEVIĆ, 2006, p.187-9) proporcionou o desenvolvimento de produtos lexicográficos que visam à exaustiva descrição linguística, os chamados Dicionários Explicativos e Combinatórios (DECs).

Devido à complexidade descritiva da teoria e com o aumento do número de ferramentas e produtos lexicográficos desenvolvidos com base na TST com a finalidade da “popularização”<sup>2</sup> da descrição lexicográfica, houve a necessidade da TST apresentar uma perspectiva de cunho pedagógico para esses produtos e ferramentas lexicográficas, ofertando ao usuário desses produtos uma descrição menos complexa, se comparada à exaustiva análise presente nos DECs. Esse viés didático está presente, por exemplo, nos estudos e projetos de Mel’čuk e Polguère (cf. Polguère; Mel’čuk, 2007, Polguère, 2007), cujo trabalho tem foco na análise do léxico da língua comum francesa, e nas pesquisas e plataformas virtuais de L’Homme, focadas nas áreas

<sup>2</sup> O vocábulo “popularização”, aqui, está sendo usado com o sentido ‘permitir que a descrição linguística proposta pela teoria (TST) atinja tanto usuários leigos quanto especialistas’.

especializadas. Isto posto, passemos a organização deste artigo, a primeira seção é dedicada à caracterização da TST e da Lexicografia Explicativa e Combinatória, teoria lexicográfica por traz dos produtos que iremos apresentar na seção seguinte. A segunda seção faz uma apresentação das ferramentas linguísticas desenvolvidas a partir da lexicografia pedagógica em TST. Por fim, trago minhas considerações finais.

Feitas essas considerações iniciais e apresentada a organização deste artigo, passemos a tratar de forma mais específica a teoria lexicográfica escolhida como fundamentação da discussão proposta por este artigo.

## TEORIA SENTIDO-TEXTO E A LEXICOGRAFIA EXPLICATIVA E COMBINATÓRIA

A Teoria Sentido-Texto (TST) é uma teoria lexicográfica de base semântica, iniciada na década de sessenta, na Rússia, tendo por objetivo a descrição da materialidade linguística das línguas naturais e que tem na tradução automática sua finalidade primeira, já que seu desenvolvimento propiciou sua ampla aplicação no Processamento da Linguagem Natural (PLN) (MILIĆEVIĆ, 2006, p.187-9). Isso contribuiu para o desenvolvimento de produtos lexicográficos, os ditos DEC's, que apresentam uma exaustiva descrição das lexias das línguas naturais. Segundo Mel'čuk e Milićević (2014), nessa teoria, o objeto de estudo da Linguística é a língua natural e, no processo de comunicação linguística, há uma correspondência biunívoca na relação sentido-texto. Partindo do ponto de vista do locutor, o sentido ou conteúdo é toda e qualquer troca de informação entre locutor e destinatário, e o texto é toda e qualquer concretização perceptível do conteúdo, ou seja, sua materialização. Isto quer dizer que o locutor, no processo de síntese linguística (do sentido ao texto) expressa um determinado sentido X que pode ser realizado de n maneiras, isto é, por meio de inúmeras paráfrases possíveis. Ao destinatário cabe compreender e estabelecer uma correspondência entre a paráfrase (realização) e o sentido que ela pode apresentar. Para a teoria, toda língua é um sistema de códigos de sentidos conhecidos, sendo a noção de língua extraída da comunicação linguística, pois, "língua é um conjunto estruturado de conhecimentos e habilidades empregados em situação de comunicação humana, quando duas pessoas falam" (MEL'ČUK; MILIĆEVIĆ, 2014, p. 30).

A TST amplia a noção saussuriana de signo linguístico ao propor que o signo nada mais é do que uma associação entre a ideia e a forma que é capaz de transmitir essa ideia. Assim, a ideia é expressa pelo significado do signo, e a forma que exprime essa ideia, o significante do signo linguístico. Por se tratar de uma situação de comunicação (atos de fala linguísticos) há um referente que, por sua vez, é o objeto no mundo a que o signo linguístico refere, dada uma situação de comunicação real, tendo na denotação de um signo o conjunto de referentes possíveis. Ademais, todo o signo linguístico possui um caráter composicional que se dá pela união de três fatores: o significante, o significado e a sintaxe, havendo, assim, uma tríplice relação do signo linguístico (MEL'ČUK; MILIĆEVIĆ, 2014, p. 30-35).

A TST propõe uma arquitetura ou estruturação própria que representa o processo de síntese-análise linguística, em que a síntese parte do sentido ou Representação Semântica (RSem) e vai até o texto ou Representação Fonológica (RPhon), sendo o caminho inverso o gerador do processo de análise. Para a teoria, a síntese linguística é a maneira mais simples de representação de seu funcionamento e da sua estrutura ao expressar como um dado sentido (*input*) pode ser realizado por meio de um conjunto de paráfrases possíveis (*output*).

Outra característica importante da TST é seu caráter funcional. Para a teoria, um modelo funcional é aquele considerado relevante se, ao se descrever um objeto, essa descrição seja a mais próxima da realidade de uma dada língua (MEL'ČUK, 1997, p.3). Por esta razão, o Modelo Sentido-Texto (MST) é considerado um aparato formal que tem por objetivo aproximar ou

simular a correspondência entre sentido e texto, isto é, o percurso de síntese linguística da teoria. Dessa maneira, na descrição de língua proposta pela TST, o MST simula a atividade linguística do falante nativo de uma determinada língua (locutor), sendo essa descrição a correspondência entre o sentido e o texto; se caracterizando, assim, como uma máquina lógica virtual, formada por componentes, sendo eles: (i) o componente declarativo (constituído pela gramática e o dicionário) e (ii) o componente processual que tem por finalidade a descrição linguística, isto é, a Lexicologia Explicativa e Combinatória.

As unidades lexicais (ou lexias) descritas pela TST são divididas entre lexemas (ou palavras) e fraseas (agrupamentos). Segundo Mel'čuk (2006), um frasea é uma expressão linguística formada por pelo menos dois lexemas sintaticamente relacionados (MEL'ČUK, 2015), podendo ser considerados composicionais ou não. Essas estruturas são diretamente relacionadas a dois conceitos importantes para a teoria: (i) composicionalidade do signo linguístico complexo e (ii) pivô ou eixo do sentido<sup>3</sup>. Isto posto, conclui-se que a Lexicografia Explicativa e Combinatória nada mais é do que o produto da modelização linguística, ou seja, são os materiais produzidos a partir da descrição de uma língua. O produto mais conhecido é o Dicionário Explicativo e Combinatório (DEC). A descrição das unidades lexicais de uma língua, no DEC, é feita a partir de duas perspectivas: (i) a paradigmática, que estabelece uma relação de seleção entre unidades lexicais; e (ii) a sintagmática, que estabelece uma relação de combinatória restrita entre as unidades lexicais.

No DEC, então, segundo Mel'čuk *et al.* (1992), as unidades lexicais que aparecem como entradas no dicionário, chamadas de artigo, verbete, ou expressão vedete, são as lexias que podem ser, como já mencionado, de dois tipos: (i) lexema, se a lexia corresponder a uma palavra, e (ii) frasea, se a lexia corresponder a um agrupamento do tipo locução. Isso quer dizer que, de acordo com a tipologia de fraseas<sup>4</sup>, colocações, nominemas e clichês não representam unidades lexicais presentes nas entradas de artigos, mas aparecem na descrição do verbete, que contém a caracterização completa da lexia; isto é, possui todas as informações necessárias para a descrição do artigo (entrada lexical) no dicionário.

As informações constantes na descrição de uma a lexia, no DEC, são organizadas por meio de cinco zonas, sendo elas: (i) a Zona de Introdução que apresenta a informação morfológica e a informação sintática da expressão vedete (entrada lexical); (ii) a Zona Semântica que apresenta a forma proposicional da expressão vedete, sua definição e as suas conotações, ou seja, mostra como o conceito emerge a partir de todos os sentidos possíveis (paráfrases) que o lexema possui em uma determinada língua; (iii) a Zona de Combinatória

<sup>3</sup> Sobre isso, Mel'čuk (2015) assume que um signo linguístico complexo AB é composicional se  $AB = A + B$ , ou seja, pela definição, um signo linguístico é composicional ou não. Quanto ao pivô ou eixo semântico, o significado 'σ' é dividido em duas partes 'σ 1' e 'σ 2' ('σ' = 'σ 1' + 'σ 2'). A parte 'σ 1' do significado 'σ' é chamado de pivô semântico de 'σ' se a outra parte 'σ 2' é um predicado de qual 'σ 1' é o argumento: 'σ' = 'σ 2' ('σ 1'). O pivô semântico do significado 'σ' é logicamente diferente do componente comunicativamente dominante de 'σ', que é a paráfrase mínima de 'σ'. Assim, no frasea tomar uma ducha cujo significado é 'lavar-se sob um chuveiro', o pivô semântico é "chuveiro", enquanto o componente predominantemente comunicativo é "lavar-se".

<sup>4</sup> Segundo Mel'čuk (2015), pela tipologia dos fraseas, os fraseas podem ser lexicais e semântico-lexicais. Os lexicais referem-se às estruturas que apresentam um sentido que é construído livremente pelo falante; mas seus lexicais constituintes são selecionados de forma restrita (cf. *bater um bolão*) e são não-composicionais (opacos); logo, por serem lexicais, configuram como unidades lexicais plenas, e a esse tipo fazem parte as locuções ou idiomatismos (cf. *chutar o pau da barraca*). Já os fraseas do tipo semântico-lexical que tanto o sentido quanto os lexicais são selecionados de maneira restrita (cf. *por outro lado*), são composicionais (transparentes), e a esse tipo pertencem as colocações (cf. *educação básica*), nominemas (cf. *Cristo Redentor*) e clichês (cf. *quer casar comigo?*). Por essa razão, estes últimos não são considerados unidades lexicais plenas, no sentido de que, as colocações, por exemplo, são vistas, pela TST, como algo que faz parte da língua, uma forma de materialidade linguística; ou os clichês que são formas sociais com esvaziamento semântico. Assim, fraseas lexicais e semântico-lexicais se diferenciam pela natureza das restrições e pela composicionalidade. Isso acarreta que apenas locuções (idiomatismos) são entradas lexicais no DEC.

Sintática que apresenta o esquema de regime (ER), as restrições do ER, os exemplos ilustrativos do ER e das restrições do ER e as modificações sintáticas da entrada lexical, isto é, esta zona mostra quantos actantes estão em relação com a lexia dependendo da paráfrase observada; (iv) a Zona de Combinatória Lexical que apresenta as FLs (paradigmáticas, sintagmáticas e não padrão) e os exemplos ilustrativos de uso — os sentidos e as combinatórias — do artigo do dicionário; e (v) a Zona Fraseológica que apresenta as locuções ou expressões congeladas (expressões idiomáticas, ditos populares, provérbios, etc.) das quais a entrada lexical faz parte. Para ilustrar o funcionamento das zonas do DEC, apresentamos a descrição completa de um lexema. Apesar de todas essas zonas fazerem parte do que é preciso para se realizar uma descrição completa de um lexema ou de um frasema do tipo locução, nem todas as lexias apresentam as informações necessárias para o preenchimento das informações solicitadas por todas as zonas do DEC. Como exemplo de entrada lexical, nos moldes do DEC, escolhemos como expressão vedete o lexema ENSINO, como mostra a figura a seguir:

Expressão Vedete: ENSINO <sup>5</sup>						
<b>1. Zona de Introdução</b>	ENSINO, nome, masculino, passível de plural					
<b>2. Zona Semântica</b>	ENSINO por X de Y em Z					
<b>3. Zona de Combinatória Sintática</b>	<u>Esquema de Regime (ER)<sup>6</sup></u>					
	<table border="1" style="width: 100%; text-align: center;"> <thead> <tr> <th>1 = X</th> <th>2 = Y</th> <th>3 = Z</th> </tr> </thead> <tbody> <tr> <td>1. por N</td> <td>1. de N 2. A poss</td> <td>1. em N 2. de N</td> </tr> </tbody> </table>	1 = X	2 = Y	3 = Z	1. por N	1. de N 2. A poss
1 = X	2 = Y	3 = Z				
1. por N	1. de N 2. A poss	1. em N 2. de N				
<b>4. Zona de Combinatória Lexical<sup>7</sup></b>	<u>Restrições sobre o ER</u>					
	<p>C<sub>3.1</sub> e C<sub>3.2</sub> : N = algo transferível e objeto de conhecimento  C<sub>1</sub> sem C<sub>2</sub> : impossível  C<sub>2.1</sub> : N = indivíduo ou grupo</p> <p style="text-align: center;"><u>Exemplos ilustrativos do ER e restrições do ER</u></p> <p>C<sub>2.1</sub> : o e. de jovens e adultos  C<sub>2.2</sub> : seu e.  C<sub>3.2</sub> : o e. de língua inglesa  C<sub>1</sub> + C<sub>2</sub> : o e. dos alunos por tutores  C<sub>2</sub> + C<sub>3</sub> : o e. de crianças em língua portuguesa  C<sub>1</sub> + C<sub>3</sub> : impossível [o e. pelo professor em espanhol]</p>					
	Syn <sub>0</sub> = alfabetização, aprendizagem, escolarização Syn <sub>c</sub> = letramento Gener = instrução V <sub>0</sub> = ensinar S <sub>1</sub> = educador, professor, tutor, docente S <sub>2</sub> = aprendiz, aluno, estudante, discente S <sub>instr</sub> = programa [de N], projeto [de N], política [de N]					

<sup>5</sup> O exemplo é sugestão nossa. O objetivo é propor uma entrada lexical, em língua portuguesa, na mesma estrutura de artigo encontrada no DEC.

<sup>6</sup> No Esquema de Regime (ER), X, Y e Z representam os argumentos que compõem a relação apresentada na Zona Semântica; N representa o nome e A poss o adjetivo possessivo.

<sup>7</sup> Na Zona de Combinatória Lexical, as relações lexicais expressas são atestadas em todas as línguas naturais modeladas dentro da TST, sendo representadas na teoria pelas Funções Lexicais (FLs). Estas funções, no sentido matemático, admitem um argumento que gera um conjunto de valores igual [f(x)=y]. Assim, uma FL é uma ferramenta formal que visa à descrição (sistemática e compacta) de todos os tipos de relações lexicais existentes em unidades lexicais de qualquer língua; sendo, uma FL aplicável a várias lexias, chamada de padrão ou *standard*; e, uma FL aplicável a apenas uma lexia (ou a poucas lexias fechadas semanticamente) chamada de degenerada ou *não-standard*.

<p><b>5. Zona de Fraseológica</b></p>	<p>A<sub>0</sub> = ensinado  A<sub>1</sub> : (X por N) bom, experiente, sábio, capaz  A<sub>2</sub> : (Y de N) rebeldes, curiosos, esforçados  A<sub>3</sub> : (Z em N) língua materna, língua estrangeira  Não Padrão : ensino de geografia / ensino geográfico  e. a distância; e. médio; e. fundamental; e. técnico</p>
---------------------------------------	--

**Figura 1:** Quadro com a expressão vedete ENSINO

Fonte: Adaptado de MEL'ČUK *et.al.* (1992)

Como observado no quadro acima, a descrição de uma expressão vedete, no DEC, ao mesmo tempo em que apresenta uma exaustiva e completa análise da lexia, para aprendizes de uma língua, por exemplo, tal descrição é considerada muito complexa. Por isso o esforço dos pesquisadores da linha pedagógica da TST em tornar essa descrição acessível ao grande público, reconsiderando a modelagem descritiva de maneira que o produto resultante seja uma ferramenta adequada para o público-alvo daquele produto lexicográfico.

## LEXICOGRAFIA PEDAGÓGICA EM TEORIA SENTIDO-TEXTO (TST)

A perspectiva que adotaremos da lexicografia pedagógica, neste artigo, não é a que visa o ensino do léxico com vistas à proficiência em segunda língua, mas a que surge como uma vertente da própria TST, uma vez que a teoria não é aplicada apenas no âmbito da tradução, mas também como ferramenta de descrição linguística visando, como público, usuários/aprendizes de léxico de língua materna, principalmente ao que compete às linguagens de especialidade em língua materna.

A partir do exposto, faremos uma discussão direcionada a uma prática lexicográfica guiada por uma teoria lexicográfica (TST) que apresenta uma vertente pedagógica com vistas ao ensino de léxico de língua. Nossa intenção, nessa seção, é trazer algumas propostas que desenvolvem projetos com uma perspectiva mais pedagógica da TST. Assim, a título de ilustração, iremos expor o projeto *Léxico Ativo do Francês* (LAF), desenvolvido por vários pesquisadores, dentre eles Mel'čuk e Polguère, e os projetos construídos a partir do banco de dados DiCo (versão eletrônica do DEC) e da DiCoUèbe (base de dados apenas de derivações semânticas e colocações do francês, aos moldes do LAF), como o Dicionário Fundamental de Informática e Internet (DiCoInfo), o Dicionário Fundamental de Meio Ambiente (DiCoEnviro), o Dicionário de Terminologia Relacionado à Lei (JuriDiCo) e o Grande Acesso Público ao Dicionário de Combinatórias do Francês (DiCoPop), todos desenvolvidos e coordenados por L'Homme (disponíveis em <http://olst.ling.umontreal.ca>), sendo este último, o DiCoPop, totalmente pensado a partir da proposta de lexicografia pedagógica via LEC. Tanto o LAF quanto os projetos do banco de dados DiCo são produtos da vertente pedagógica da TST, uma vez que, são aplicadas adaptações na nomenclatura de maneira que o usuário não familiarizado com a teoria compreenda a formalização proposta pela TST. O LAF, por exemplo, tem como usuários estudantes de linguística, linguistas, aprendizes de francês e professores do idioma, isto é, o público-alvo varia entre especialistas e aprendizes, mas todos com algum grau de domínio da língua francesa.

### LÉXICO ATIVO DO FRANCÊS – LAF

O LAF é um produto lexicográfico de análise de léxico de língua francesa, porém, de acordo com Polguère & Mel'čuk (2007, p.14-7), não é um dicionário comum, uma vez que

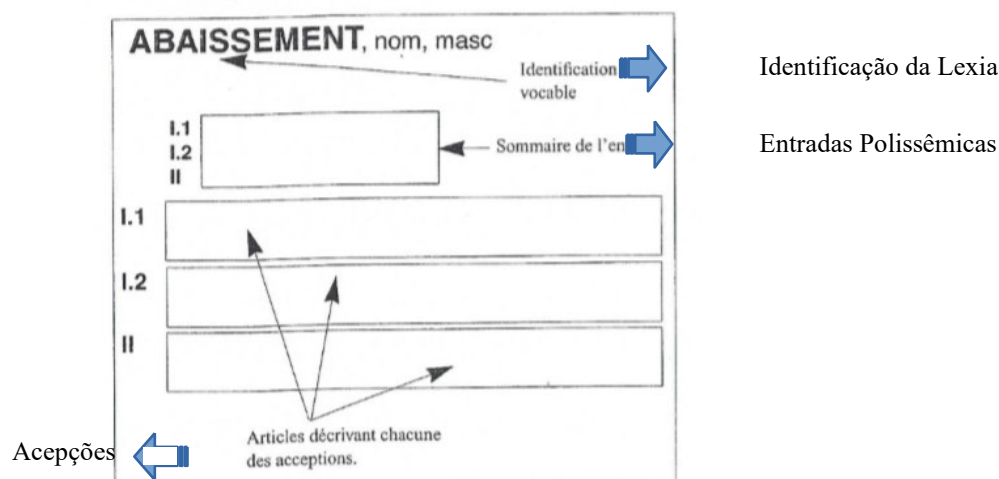
possui quatro características que o distingue, são elas: (i) é um dicionário especializado, pois se foca na descrição linguística de vinte mil derivações semânticas e colocações francesas, tendo como abordagem teórica a Lexicografia Explicativa e Combinatória; (ii) o LAF é uma ferramenta de codificação linguística do pensamento do locutor, isto é, é um dicionário de produção linguística, já que visa fornecer ao usuário todos os meios lexicais necessários para expressar, da maneira mais idiomática possível, uma determinada ideia; (iii) o LAF é um dicionário formalizado, ou seja, a necessidade de apresentar no LAF uma grande quantidade de informações de forma explícita, torna inevitável o uso de mais formalismos do que é corrente nos dicionários atuais, ademais esses formalismos garantem rigor e sistemática na descrição, tornando-a mais compacta; e (iv) o LAF é um dicionário experimental, uma vez que propõe uma nova forma de descrição lexical e serve de modelo de referência para o desenvolvimento de uma pedagogia para o ensino de conhecimentos lexicais (POLGUÈRE; MEL'ČUK, 2007, p.14-17). Cabe ressaltar que a originalidade do LAF está na “popularização” da abordagem formal explicativa e combinatória a fim de tornar a descrição linguística acessível ao usuário, seja ele especializado ou não.

Em seu artigo sobre o LAF, Polguère (2007, p.1) explica que seu objetivo, além de oferecer uma apresentação geral do projeto, é dar especial atenção às tensões geradas pela dualidade da descrição linguística, pois, para o autor, descrever rigorosamente as relações lexicais e, ao mesmo tempo, tornar essas informações acessíveis ao leigo, exige algumas reflexões acerca do que é basilar em uma descrição linguística. A partir desse argumento, o autor nomeia dois dispositivos descritivos que para ele são essenciais para a análise linguística utilizada no LAF, ou seja, o autor dá destaque à hierarquia das etiquetas semânticas e às fórmulas metalinguísticas que codificam as relações das funções lexicais presentes. Esses dispositivos, para o autor, são basilares para que haja uma descrição de língua que seja considerada adequada, porém, não completa, como veremos mais adiante.

Como já destacado, tem sido prática comum entre linguistas formais, desde o auge da Inteligência Artificial, ter no Processamento da Linguagem Natural (PLN) a aplicação final de modelos linguísticos formais. Porém, Polguère (2007, p.1-2) destaca que tal crença é um mito e defende que apenas a aplicação de modelos formais com vistas ao ensino/aprendizagem de línguas sejam um ambiente adequado para testar a validade de tais modelos, especialmente, no que diz respeito ao léxico. Dessa maneira, por ter um objetivo que apresenta uma dualidade, isto é, ser pedagógico e ser descritivo, linguisticamente falando, o LAF não é um dicionário real, o projeto é, antes de tudo, um manual de lexicologia, baseado em uma prática e abordagem orientadas a dados, tendo na primeira parte do volume uma introdução aos conceitos gerais de Lexicologia Explicativa Combinatória (LEC) e, na segunda parte do volume, o conteúdo e estrutura do dicionário postos em prática.

O LAF possui uma macroestrutura convencional, semelhante à estrutura de um dicionário padrão com entradas lexicais ordenadas alfabeticamente, mas apresenta uma microestrutura (organização interna) desenvolvida de acordo com os princípios da Lexicografia Explicativa e Combinatória. Na microestrutura de um artigo do LAF são observados os seguintes aspectos (POLGUÈRE; MEL'ČUK, 2007, p.31-37): (i) as características gramaticais da expressão vedete, isto é, a parte do discurso, classe gramatical e outras características relevantes que o artigo possa apresentar; (ii) as etiquetas semânticas, ou seja, marcações que identifiquem o tipo de significado geral veiculado à expressão vedete; (iii) as fórmulas actanciais, isto é, expressões linguísticas constituídas da expressão vedete e de seus actantes; e (iv) as derivações semânticas e colocações, ou seja, descrição das relações derivacionais e colocacionais presentes na expressão vedete; (v) os exemplos para abonação; (vi) a lista de locuções que incluem a expressão vedete. Além do exposto, microestruturalmente, ainda há tanto a possibilidade de organizar as lexias tanto por meio de hierarquia das etiquetas semânticas (HES) quanto por campos semânticos (CS) ou redes semânticas (RS). A seguir, está

a figura 5 que apresenta a estrutura de um artigo do LAF a partir da descrição proposta por Polguère e Mel'čuk, 2007.



**Figura 2:** Estrutura de um Artigo do LAF  
**Fonte:** POLGUÈRE; MEL'ČUK, 2007, p. 30.

Estruturalmente, um artigo do LAF apresenta a lexia descrita em negrito em maiúsculas, no caso a palavra francesa ABAISSEMENT (‘abaixamento’, ‘queda’, etc.) e sua identificação gramatical básica, no caso, substantivo masculino. Na segunda parte, o artigo dá destaque às entradas polissêmicas, que são as multiplicidades de sentido de um lexema ou frase. Assim, na lexia ABAISSEMENT são apresentados os seguintes valores polissêmicos: (i) diminuição, queda (cf. ‘*on constate un abaissement de la température*’); (ii) comovente (cf. ‘*les ouvriers ont effectué l’abaissement des trottoirs*’); e (iii) forma (cf. ‘*le début de la zone urbaine coincide avec un abaissement du relief*’). Na última seção da estrutura do artigo encontram-se as acepções que descrevem os valores semânticos e servem de abonação para cada um dos significados.

Sobre o processo de uso de recursos que permitam a simplificação da descrição linguística, ou seja, uso de elementos que permitam apresentar uma descrição de língua que seja adequada, mas não completa, Polguère (2007, p.5-6) destaca que o principal problema durante a construção do LAF foi encontrar o equilíbrio adequado entre o que Igor Mel'čuk chama de “verdade (científica)” - lógica científica crua – e a inteligibilidade para o leigo. Assim, para simplificar a descrição linguística presente no LAF, o autor partiu da identificação do subconjunto de noções essenciais da Lexicografia Explicativa e Combinatória (LEC), isto é, partiu do entendimento de que existem elementos descritivos via LEC que não poderiam ser dispensados em um dicionário de derivação semântica e colocações, sendo essas noções: (i) o conteúdo semântico principal da estrutura actiológica da palavra-chave, (ii) o padrão de regências da palavra-chave, (iii) o valor sintático-semântico das ligações paradigmáticas e sintagmáticas e (iv) o padrão de regência das colocações da palavra-chave (POLGUÈRE, 2007, p.5-6). Segundo o autor, essas noções são o mínimo para que se possa fornecer uma descrição lexicográfica que seja adequada e suficiente (embora incompleta) no contexto do ensino e aprendizagem de vocabulário. Por fim, é importante dizer que além da estrutura exposta nos parágrafos anteriores, o site do LAF<sup>8</sup> inclui recursos de acesso aos dados linguísticos da plataforma LAF e sugestões de atividades pedagógicas de aprendizagem do vocabulário.

<sup>8</sup> Disponível em <http://olst.ling.umontreal.ca/laf/about/lang-pref/fr/>.



## DICIONÁRIOS INTERATIVOS DESENVOLVIDOS A PARTIR DA BASE DE DADOS DICO

Marie-Claude L’Homme é uma linguista que desenvolve pesquisas que visam a construção de plataformas com base na Lexicografia Explicativa e Combinatória (LEC) para a apreensão de léxico de âmbitos especializados a partir do banco de dados DiCo, versão eletrônica aos moldes do DEC; e da base DiCoUèbe, versão eletrônica com o mesmo propósito do LAF, isto é, a descrição de derivações semânticas e colocações da língua francesa. De acordo com Jousse *et al.* (2011, p.135-6), nos últimos anos, os dicionários (impressos ou não) e os bancos de dados lexicais tendem a apresentar uma variedade de critérios de descrição linguística. No entanto, é verificada uma falta de modelagem adequada para a descrição das relações semânticas entre unidades lexicais (lexemas e frases) de uma língua. De acordo com os autores, isso acarreta, por parte dos lexicógrafos, em classificar as relações semânticas *ad hoc* e, muitas vezes, a recorrer a métodos introspectivos (JOUSSE *et al.* 2011, p. 135-6).

A partir do exposto, é possível dizer que uma descrição linguística só será considerada mais próxima da verdade, do ponto de vista de todo o léxico de uma língua, se houver uma sólida formalização das relações lexicais, uma vez que tal formalização permite uma generalização da classificação para partes mais amplas do léxico (com base no desenvolvimento de um modelo que parta de uma amostra representativa). Essa amostra pode ser o vocabulário associado a um âmbito especializado, o léxico de um idioma específico ou um grupo de idiomas. A formalização torna-se necessária, então, para processar e manipular esses dados (JOUSSE *et al.* 2011, p. 136).

Segundo Jousse *et al.* (2011, p.136), o sistema de Funções Lexicais (FL) utilizadas em análises linguísticas via LEC está perfeitamente adaptado para esse tipo de formalização, destacada no parágrafo anterior. Assim, com a formalização do léxico via TST, vários recursos lexicais baseados nessa estrutura foram desenvolvidos nas últimas décadas; sendo que muitos assumem a forma de bancos de dados lexicais, outros de ferramentas para o aprendizado de idiomas, mas em todos há a aplicação de FLs para representar propriedades sintático-semânticas das relações lexicais existentes entre as unidades que compõem uma lexia.

Essas foram as razões que levaram à construção dos produtos lexicais DiCo, como os dicionários DiCoInfo, DiCoEnviro, JuriDiCo e DiCoUèbe, construídos a partir da Lexicografia Explicativa e Combinatória, com terminologia específica, respectivamente, das áreas Computação, Meio Ambiente, Direito e *Internet*. Todas essas plataformas já são construídas com elementos que visam a facilitar a compreensão do usuário, mas ainda exigem um entendimento prévio do léxico pesquisado por parte do usuário/aprendiz. Dessa maneira, com vistas à “popularização”, L’Homme e sua equipe vem desenvolvendo plataformas ainda mais “popularizadas”, pensadas para aprendizes, como a versão mais popular da DiCoUèbe, o DiCoPop, que visa dar acesso ao grande público ao dicionário de combinatórias do francês, e a ferramenta NeoVisual, desenvolvida como suporte à plataforma DiCoEnviro.

A plataforma DiCoPop<sup>9</sup> disponibiliza as informações por meio da “simplificação” da modelização descritiva, tendo suas lexias organizadas da seguinte maneira: (i) nomenclatura, com 393 vocábulos, de onde observam-se 799 acepções; (ii) hierarquia das etiquetas semânticas, com 790 etiquetas, das quais 706 são nominais, 40 verbais e 38 adjetivais; (iii) funções lexicais, de onde observam-se 3368 funções com 20479 relações lexicais. Já a plataforma DiCoEnviro apresenta recursos de “popularização” que facilitam o entendimento da estrutura organizacional do âmbito do Meio Ambiente. Essa plataforma apresenta uma peculiaridade que é um sistema organizacional que é considerado tanto um dicionário padrão,

<sup>9</sup> Disponível em <http://olst.ling.umontreal.ca/dicopop/>

quanto um *frame* por campos semânticos e uma ferramenta que permite a visualização das redes semânticas que constituem a área do Meio Ambiente, chamada de NeoVisual. O NeoVisual<sup>10</sup> é uma ferramenta de “popularização” do conhecimento ao “simplificar” a compreensão do mesmo para o leitor leigo, por meio de um apelo visual, de modo que o usuário/aprendiz identifique a estrutura.

Esta seção teve o intuito de apresentar as plataformas criadas para determinadas áreas do conhecimento. As quais foram desenvolvidas em moldes que permitam uma “simplificação” da modelização descritiva, com a intenção de “popularizar” o conhecimento, tornando essas plataformas, bem como o LAF, apresentado na seção anterior, produtos da lexicografia pedagógica. Utilizar elementos que visam uma “popularização” da descrição linguística em uma ferramenta pedagógica baseada na LEC é beneficiar um maior público, de forma que esses usuários/aprendizes tenham acesso à perspectiva da TST, tão complexa, permitindo que o aprendiz entenda que a teoria parte do dado semântico (sentido) para chegar aos contextos possíveis de uso (texto), isto é, às realizações daquela lexia na língua.

## CONSIDERAÇÕES FINAIS

Como visto, a vertente pedagógica da Teoria Sentido-Texto tem por intuito “simplificar” a descrição lexicográfica presente na Lexicografia Explicativa e Combinatória (LEC), cujo principal produto são os Dicionários Explicativos e Combinatórios (DECs), que são produtos lexicográficos que apresentam uma descrição exaustiva e complexa das lexias de uma língua, o que torna muitas vezes difícil sua compreensão descritiva, dando vasão de compreensão apenas ao linguista iniciado em TST. Por conta disso, a vertente pedagógica tem por finalidade “popularizar” a descrição lexicográfica a fim de aproximar os usuários/aprendizes de língua à descrição linguística proposta pela LEC. Cabe ter claro que proporcionar ao usuário uma modelização da língua que parta do dado semântico (sentido), aquilo que o locutor quer dizer, às possíveis formas de realizá-lo (texto), as possíveis formas do locutor dizer, permite ao usuário/aprendiz um entendimento do processo de síntese linguística do ponto de vista do locutor em um ato comunicativo, uma vez que o objeto da linguística é a língua, o que acarreta em uma reflexão sobre a finalidade desse ato comunicativo e consequente consciência linguística.

Ademais, como destaca Polguère (2007), o intuito nessa “popularização” não é oferecer ao usuário/aprendiz uma descrição completa e complexa da língua, que seja uma “verdade” científica, mas munir o aprendiz de elementos que sejam indispensáveis para a LEC e para a compreensão daquilo que é essencial à teoria. Com isso, ao aprendiz, é ofertada uma descrição linguística suficiente para a compreensão de língua.

## REFERÊNCIAS

JOUSSE, Anne-Laure; L’HOMME, Marie-Claude; LEROYER, Patrick; ROBICHAUD, Benoît. **Presenting collocates in a dictionary of computing and the Internet according to user needs**. Barcelona, 8-9, September, 2011.

<sup>10</sup> Disponível em <http://olst.ling.umontreal.ca/dicoenviro/neovisual/>

L'HOMME M.-C., POLGUÈRE A. Mettre en bons termes les dictionnaires spécialisés et les dictionnaires de langue générale. In: F. Maniez, P. Dury (dir.), **Lexicologie et terminologie: histoire de mots. Hommage à Henri Béjoint**. Lyon: Travaux du CRTT, p. 191-206, 2008.

MEL'ČUK, Igor. **Semantics: from meaning to text**. Vol. 3. Philadelphia: John Benjamins, 2015.

MEL'ČUK, Igor; MILIČEVIĆ, Jasmina. **Introduction à la linguistique**. Vol. 1. Paris: Hermann Éditeurs, 2014.

MEL'ČUK, Igor; POLGUÈRE, Alain. **Lexique actif du français: L'apprentissage du vocabulaire fondé sur 20000 dérivations sémantiques et collocations du français**. Paris: Duculot, 2007.

MEL'ČUK, Igor; SICA, G, ed. Explanatory combinatorial dictionary. In: **Open Problems in Linguistics and Lexicography**. Monza: Polimetrica, p. 222-355, 2006.

MEL'ČUK, Igor. **Ver une linguistique Sens-Texte. Leçon Inaugurale**. Paris: Collège de France, 1997.

MEL'ČUK, I. A.; ARBATCHEWSKY-JUMARIE, N.; IORDANSKAJA, L.; MANTHA, S. **Dictionnaire explicatif et combinatoire du français contemporain: Recherches lexicosémantiques III**, Montréal, Les Presses de l'Université de Montréal, 1992.

MILIČEVIĆ, Jasmina. A short-guide to the Meaning-Text Linguistic Theory. In: **Journal of Koralex**, vol. 8, p. 187-233, 2006.

POLGUÈRE, Alain. Lessons from the Lexique actif du français. In: **MTT 2007**, Klagenfurt, May, p. 21-24, 2007.