

## RESSALVA

Atendendo solicitação da autora, o texto completo desta dissertação será disponibilizado somente a partir de 19/10/2022.



UNIVERSIDADE ESTADUAL PAULISTA  
"JÚLIO DE MESQUITA FILHO"  
Campus de São José do Rio Preto

Giseli Aparecida Cecílio

**IDENTIFICAÇÃO E ANÁLISE DE COLOCAÇÕES ACADÊMICAS NAS  
ÁREAS DE LINGUÍSTICA, LETRAS E ENGENHARIAS**

São José do Rio Preto

2021

Giseli Aparecida Cecílio

**IDENTIFICAÇÃO E ANÁLISE DE COLOCAÇÕES ACADÊMICAS NAS  
ÁREAS DE LINGUÍSTICA, LETRAS E ENGENHARIAS**

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestra em Estudos Linguísticos, junto ao Programa de Pós-graduação em Estudos Linguísticos, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, campus de São José do Rio Preto.

Orientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Adriane Orenha-Ottaiano.

Coorientadora: Prof.<sup>a</sup> Dr.<sup>a</sup> Tanara Zingano Kuhn.

São José do Rio Preto

2021

C388i

Cecílio, Giseli Aparecida

Identificação e análise de colocações acadêmicas nas áreas de linguística, letras e engenharias / Giseli Aparecida Cecílio. -- São José do Rio Preto, 2021  
148 f. : tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp), Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto

Orientadora: Adriane Orenha-Ottaiano

Coorientadora: Tanara Zingano Kuhn

1. Linguística Aplicada. 2. Descritor. 3. Descritor. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Giseli Aparecida Cecílio

## **IDENTIFICAÇÃO E ANÁLISE DE COLOCAÇÕES ACADÊMICAS NAS ÁREAS DE LINGUÍSTICA, LETRAS E ENGENHARIAS**

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestra em Estudos Linguísticos, junto ao Programa de Pós-graduação em Estudos Linguísticos, do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, campus de São José do Rio Preto.

### **COMISSÃO EXAMINADORA**

Prof.<sup>a</sup> Dr.<sup>a</sup> Adriane Orenha Ottaiano  
UNESP – Câmpus de São José do Rio Preto  
Orientador

Prof.<sup>a</sup> Dr.<sup>a</sup> Tanara Zingano Kuhn  
Universidade de Coimbra – Portugal

Prof.<sup>a</sup> Dr.<sup>a</sup> Cleci Regina Bevilacqua  
UFRGS – Rio Grande do Sul

Prof.<sup>a</sup> Dr.<sup>a</sup> Margarita Maria Correia Ferreira  
Universidade de Lisboa – Portugal

São José do Rio Preto

19 de abril de 2021

Dedico este trabalho aos grandes amores da minha vida, meu pai (*in memoriam*), à minha mãe, meus irmãos e ao meu lindo filho, Rafael.

## **AGRADECIMENTOS**

Agradeço a Deus por ter sonhado meu sonho, e me dado a coragem e a persistência para nunca desistir.

À minha orientadora, Prof.<sup>a</sup> Dr.<sup>a</sup> Adriane Orenha-Ottaiano, uma das pessoas que mais admiro na vida, exemplo de dedicação e esforço constantes. Serei eternamente grata por ter acreditado em mim. Pelos momentos de conversa, pela amizade e por todo conhecimento que adquiri ao seu lado.

À minha coorientadora, Prof.<sup>a</sup> Dr.<sup>a</sup> Tanara Zingano Kuhn. Tê-la conhecido foi um dos melhores presentes que este mestrado me trouxe, sua gentileza, seu conhecimento e seu desprendimento em passá-lo adiante me encantam. Obrigada por todo o tempo dedicado à “nossa pesquisa”, pois se cheguei até aqui é porque você e a Prof.<sup>a</sup> Adriane estiveram sempre presentes na construção desta dissertação.

Agradeço à Prof.<sup>a</sup> Dr.<sup>a</sup> Stella Esther Ortweiller Tagnin, pelas inúmeras contribuições dadas a esta pesquisa no debate do Selin da Unesp/Ibilce, em 2019.

Meu muito obrigada à Prof.<sup>a</sup> Dr.<sup>a</sup> Cristiane Krause Kilian, pela participação na qualificação, cujas contribuições trouxeram excelentes melhorias ao trabalho.

À Prof.<sup>a</sup> Dr.<sup>a</sup> Cleci Bevilacqua, que gentilmente aceitou o convite para qualificação e que com seus elogios, comentários e sugestões trouxeram um outro olhar à pesquisa.

Sou profundamente grata aos membros da banca de defesa e à Prof.<sup>a</sup> Dr.<sup>a</sup> Cleci Bevilacqua, que novamente aceitou o convite.

À Prof.<sup>a</sup> Dr.<sup>a</sup> Margarita Maria Correia Ferreira, por aceitar o convite e contribuir para a finalização deste trabalho.

Agradeço a todos os amigos que fiz neste período, principalmente aos da linha de pesquisa que faço parte, que muito me ajudaram desde minha entrada no programa de pós-graduação. Descrever os nomes poderia ser injusto, já que todos foram essenciais para meu crescimento intelectual.

E, por último e não menos importante, agradeço à minha família, que sempre me incentivou e acreditou em mim. Ao meu pai, que embora não esteja fisicamente presente, esteve e estará sempre torcendo por mim; à minha mãe, que diante de suas dores e sofrimentos soube sorrir com minhas conquistas; a meu irmão, que sempre que possível conseguiu, em meio aos meus desesperos e dúvidas, fazer-me sorrir; à

minha irmã, que nunca duvidou que eu conseguiria e, ao meu filho, que teve a paciência e amor por uma mãe que precisou fazer a difícil escolha pelos estudos.

“O conhecimento científico gerado em pesquisa se transpõe para o currículo e, mais adiante, chega à indústria por meio da linguagem. Nessa figura, a linguagem é representada como o sistema que, na educação e na ciência, constrói discursivamente os fenômenos e o conhecimento, em níveis crescentes de teorização, abstração e technicalidade”. (MOTTA-ROTH, 2006, p. 845).

## RESUMO

Tendo em vista a crescente demanda de publicações de textos acadêmico-científicos e objetivando divulgar seus trabalhos em revistas, congressos, simpósios e seminários, muitos alunos de graduação e pós-graduação de diferentes áreas do conhecimento devem produzir, em português, resumos, artigos, pôsteres, além de suas dissertações ou teses. Todavia, escrever textos acadêmicos de qualidade é desafiador. Não raramente, os estudantes sentem-se limitados na produção textual, mesmo tendo a língua portuguesa como língua materna. Essa situação é agravada no caso dos estrangeiros que se propõem a estudar e escrever em português. Assim, obras lexicográficas e materiais didáticos baseados em *corpus* podem auxiliar no processo de aprendizagem, proporcionando aos discentes brasileiros ou estrangeiros ferramentas que visem o aprimoramento de sua escrita. Portanto, este estudo objetiva o levantamento das colocações acadêmicas de estruturas verbais e adjetivas que encontramos em comum, quanto à forma, nas áreas de Engenharias e Letras/Linguística, mas que não possuem o mesmo sentido, especializando-se em função da área de conhecimento. Além disso, visamos uma análise dessas colocações, encontradas nas duas áreas de investigação: no *Subcorpus* de Engenharias (CoPEP – KUHN; FERREIRA, 2020), com 441.591 vocábulos, e no *Corpus* da grande área de Letras e Linguística da língua portuguesa do Brasil, contendo 476.191 palavras, compilado para esta investigação. Para isso, temos como fundamentação teórica a Linguística de *Corpus* (BIBER; CONRAD; REPPEN, 1998; BEVILACQUA, 1996; KUHN; FERREIRA, 2020), a Fraseologia (CORPAS PASTOR, 1996; NESSELHAULF, 2005; TAGNIN, 2013; ORENHA-OTTAIANO, 2004, 2015, 2017, 2020) e a Lexicografia (BIDERMAN, 1998; ATKINS, 2008; RUNDELL, 2008). No que diz respeito aos aspectos metodológicos, contamos com o auxílio da ferramenta *Sketch Engine* (KILGARRIFF *et al.*, 2004). Por meio do referido *software*, extraímos as colocações acadêmicas comuns às duas áreas e, na sequência, as descrevemos e analisamos sob o ponto de vista sintático-morfológico e léxico-semântico, com a finalidade de inseri-las na Plataforma *On-line Dicionários de Colocações Acadêmicas (DICA)* monolíngues em *inglês e português do Brasil e Europeu*, cujo projeto de pesquisa está investigação está inserida. Dessa maneira, a existência de um dicionário de colocações que se especialize na área acadêmica, com exemplos específicos em cada campo científico, poderá ser um instrumento valioso para graduandos e pós-graduandos brasileiros ou estrangeiros que almejem desenvolver sua escrita acadêmica na língua portuguesa do Brasil em uma área específica.

**Palavras-chave:** Fraseologia. Colocações. Dicionário. Linguagem acadêmica. Língua portuguesa.

## ABSTRACT

Considering the growing demand for academic-scientific text publications and aiming to disseminate their work in journals, conferences, symposia, and seminars, many undergraduate and graduate students from different areas of knowledge must produce abstracts, articles, posters in Portuguese, in addition to their dissertations or theses. However, writing quality academic texts is challenging. Not rarely, students feel limited in textual production, even in Portuguese as their mother tongue. This situation is aggravated in the case of foreigners who intend to study and write in Portuguese. Thus, corpus-based lexicographic works and didactic materials could assist in the learning process, providing Brazilian or international students with tools to improve their writing. Therefore, this study aims to survey the collocations we find in common, concerning the form, in the areas of Engineering and English/Linguistics, but which do not have the same meaning, specializing according to the area of knowledge. In addition, we aim to analyze these collocations found in the two research areas: one in the Engineering area corpus (CoPEP – KUHN; FERREIRA, 2020), with 441,591 words and another corpus in the area of English and Linguistics in the Brazilian Portuguese, containing 476,191 words, exclusively compiled for this investigation. Therefore, we have as theoretical framework: Corpus Linguistics (BIBER; CONRAD; REPPEN, 1998; BEVILACQUA, 1996; KUHN, 2018), Phraseology (CORPAS PASTOR, 1996; NESSELHAULF, 2005; TAGNIN, 2013; ORENHA-OTTAIANO, 2004, 2015, 2017, 2020) and Lexicography (BIDERMAN, 1998; ATKINS, 2008; RUNDELL, 2008). Regarding the methodological aspects, we use the Sketch Engine tool (KILGARRIFF *et al.*, 2004). Using this software, we extract the academic collocations common to both areas, and we describe and analyze them under a syntactic-morphological and lexical-semantic perspective, with the purpose of inserting them in *Online Platform* of Monolingual Academic Collocations Dictionaries of Brazilian and European Portuguese and English, whose research project this investigation is part of. Thus, the existence of a collocations dictionary that specializes within the academic area, and with specific examples in each scientific field, can be a valuable tool for Brazilian and foreign undergraduate and graduate students who wish to develop their academic writing in Brazilian Portuguese in a specific area.

**Keywords:** Phraseology. Collocations. Dictionary. Academic language. Portuguese language.

## LISTA DE FIGURAS

|                                                                                                                                     |     |
|-------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figura 1 – Plataforma <i>On-line</i> de Dicionários de Colocações – tela de cadastro e <i>login</i>                                 | 41  |
| Figura 2 – Definição de colocações (Orenha-Ottaiano, no prelo).....                                                                 | 41  |
| Figura 3 – Os tipos de colocações existentes .....                                                                                  | 49  |
| Figura 4 – Resultado do encontro das duas áreas: as colocações acadêmicas em comum .....                                            | 55  |
| Figura 5 – <i>Word Sketch</i> do <i>Subcorpus</i> de Engenharias.....                                                               | 56  |
| Figura 6 – <i>Word Sketch</i> do <i>Subcorpus</i> de Engenharias.....                                                               | 57  |
| Figura 7 – <i>Layout</i> da página do <i>Subcorpus</i> Engenharias-Br (SCEs) – CoPEP Engenharias .....                              | 68  |
| Figura 8 – <i>Layout</i> da página inicial da ferramenta <i>Sketch Engine</i> .....                                                 | 71  |
| Figura 9 – <i>Layout</i> da página: <i>Create corpus</i> .....                                                                      | 72  |
| Figura 10 – <i>Corpus</i> Letras/Linguística-Br (CLL) .....                                                                         | 72  |
| Figura 11 – <i>Corpus</i> Letras/Linguística-Br (CLL) .....                                                                         | 73  |
| Figura 12 – <i>Word Sketch</i> do <i>Subcorpus</i> de Engenharias.....                                                              | 75  |
| Figura 13 – <i>Word List</i> do <i>Subcorpus</i> de Engenharias-Br (SCEs).....                                                      | 77  |
| Figura 14 – <i>Word List</i> do <i>Subcorpus</i> de Engenharias-Br (SCEs).....                                                      | 78  |
| Figura 15 – <i>Word Sketch</i> do <i>Subcorpus</i> de Engenharias-Br (SCEs) em destaque CA “sistema computacional” .....            | 85  |
| Figura 16 – <i>Word Sketch</i> do <i>Subcorpus</i> de Letras e Linguística (CLL); em destaque a CA “sistema computacional” .....    | 86  |
| Figura 17 – <i>Concordance</i> do <i>Subcorpus</i> de Engenharias-Br (SCEs) em destaque CA “sistema computacional” .....            | 87  |
| Figura 18 – <i>Concordance</i> do <i>Subcorpus</i> de Letras/Linguística (CLL) em destaque CA “sistema computacional” .....         | 87  |
| Figura 19 – <i>Concordance</i> do <i>Subcorpus</i> de Letras/Linguística (CLL) em destaque a CV “modelar sistema” .....             | 87  |
| Figura 20 – <i>Concordance</i> do <i>Subcorpus</i> de Letras/Linguística (CLL) em destaque CV “ter sistema” .....                   | 88  |
| Figura 21 – Parte da lista de candidatas a colocações acadêmicas comuns (CLL e SCEs).....                                           | 90  |
| Figura 22 – Tela inicial da plataforma – PLATCOL: colocações em ordem alfabética .....                                              | 92  |
| Figura 23 – Tela inicial da plataforma – PLATCOL: taxonomia das colocações .....                                                    | 93  |
| Figura 24 – Tela das plataformas existentes no PLATCOL .....                                                                        | 94  |
| Figura 25 – Recorte parcial do <i>Word Sketch</i> do substantivo “forma” no <i>corpus</i> de Engenharias .....                      | 100 |
| Figura 26 – <i>Word Sketches</i> de “tempo” nas relações gramaticais substantivo + adjetivo e verbo + substantivo no CL e SCEs..... | 102 |
| Figura 27 – Linhas de concordância da colocação “ter tempo” no <i>Corpus</i> CLL .....                                              | 103 |
| Figura 28 – Contexto expandido de uma linha de concordância com “ter tempo” no <i>Corpus</i> CLL .....                              | 103 |

|                                                                                                                                         |     |
|-----------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figura 29 – Linhas de concordância da colocação “ter tempo” no <i>Subcorpus</i> SCEs .....                                              | 104 |
| Figura 30 – <i>Word Sketches</i> de “caso” nas relações gramaticais substantivo + adjetivo e verbo + substantivo no CL e SCEs.....      | 105 |
| Figura 31 – Linhas de concordância da colocação “caso particular” no <i>Subcorpus</i> SCEs .....                                        | 106 |
| Figura 32 – Linhas de Concordância da Colocação “Caso Particular” no <i>Corpus</i> CLL .....                                            | 107 |
| Figura 33 – Linhas de concordância da colocação “haver caso” no <i>Corpus</i> CLL....                                                   | 108 |
| Figura 34 – Linhas de concordância da colocação “haver caso” do <i>Subcorpus</i> SCEs .....                                             | 108 |
| Figura 35 – <i>Word Sketches</i> de “sistema” nas relações gramaticais substantivo + adjetivo e verbo + substantivo no CL e SCEs .....  | 109 |
| Figura 36 – Linhas de concordância da colocação “sistema computacional” no <i>Subcorpus</i> SCEs .....                                  | 110 |
| Figura 37 – Linhas de concordância da colocação “sistema computacional” no <i>Corpus</i> CLL.....                                       | 110 |
| Figura 38 – Linhas de concordância da colocação “sistema operacional” no <i>Subcorpus</i> SCEs .....                                    | 111 |
| Figura 39 – Linhas de concordância da colocação “sistema operacional” no <i>Corpus</i> CLL.....                                         | 111 |
| Figura 40 – Definição de sistema operacional no dicionário Houaiss .....                                                                | 112 |
| Figura 41 – <i>Word Sketches</i> de “processo” nas relações gramaticais substantivo + adjetivo e verbo + substantivo no CL e SCEs ..... | 113 |
| Figura 42 – Linhas de concordância da colocação “envolver processo” no <i>Corpus</i> CLL .....                                          | 114 |
| Figura 43 – Linhas de concordância da colocação “envolver processo” no <i>Corpus</i> CLL .....                                          | 115 |
| Figura 44 – Linhas de concordância da colocação “envolver processo” no <i>Corpus</i> CLL .....                                          | 115 |
| Figura 45 – Linhas de concordância da colocação “ <i>envolver processo</i> ” no <i>Corpus</i> CLL .....                                 | 116 |
| Figura 46 – Linhas de concordância da colocação “envolver processo” no <i>Subcorpus</i> SCEs .....                                      | 116 |
| Figura 47 – Linhas de concordância da colocação “envolver processo” no <i>Corpus</i> CLL .....                                          | 117 |
| Figura 48 – Linhas de concordância da colocação “envolver processo” no <i>Corpus</i> CLL .....                                          | 117 |
| Figura 49 – Linhas de concordância da colocação “envolver processo” no <i>Corpus</i> CLL .....                                          | 117 |
| Figura 50 – Linhas de concordância da colocação “envolver processo” no <i>Corpus</i> CLL .....                                          | 118 |

|                                                                                                                                          |     |
|------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figura 51 – <i>Word Sketches</i> de “resultado” nas relações gramaticais substantivo + adjetivo e verbo + substantivo no CL e SCEs ..... | 119 |
| Figura 52 – Linhas de concordância da colocação “resultado obtido” no <i>Subcorpus</i> SCEs .....                                        | 120 |
| Figura 53 – Linhas de concordância da colocação “resultado obtido” do <i>Subcorpus</i> SCEs .....                                        | 121 |
| Figura 54 – Linhas de concordância da colocação “resultado obtido” do <i>Corpus</i> CLL .....                                            | 121 |

## LISTA DE TABELAS

|                                                                                                         |     |
|---------------------------------------------------------------------------------------------------------|-----|
| Tabela 1 – Grandes áreas.....                                                                           | 66  |
| Tabela 2 – Dados sobre as revistas selecionadas para a compilação do <i>corpus</i> de artigos .....     | 69  |
| Tabela 3 – Dados obtidos a partir da ferramenta <i>Word List</i> do programa <i>Sketch Engine</i> ..... | 78  |
| Tabela 4 – Dados obtidos a partir da ferramenta <i>Word List</i> do programa <i>Sketch Engine</i> ..... | 81  |
| Tabela 5 – Número de colocações encontradas a partir dos 100 primeiros substantivos .....               | 89  |
| Tabela 6 – Dados obtidos a partir da ferramenta <i>Word List</i> do programa <i>Sketch Engine</i> ..... | 95  |
| Tabela 7 – Dados obtidos a partir da ferramenta <i>Word List</i> do programa <i>Sketch Engine</i> ..... | 96  |
| Tabela 8 – Dados obtidos a partir da ferramenta <i>Word List</i> do programa <i>Sketch Engine</i> ..... | 97  |
| Tabela 9 – Dados da tabela elaborada no <i>Excel</i> .....                                              | 120 |
| Tabela 10 – Dados da tabela elaborada no <i>Excel</i> .....                                             | 122 |

## LISTA DE QUADROS

|                                                            |    |
|------------------------------------------------------------|----|
| Quadro 1 – Tipos e gêneros textuais .....                  | 29 |
| Quadro 2 – Colocações verbais .....                        | 52 |
| Quadro 3 – Colocações nominais .....                       | 53 |
| Quadro 4 – Colocações adjetivas.....                       | 53 |
| Quadro 5 – Colocações adverbiais .....                     | 53 |
| Quadro 6 – Periódicos de Letras, Linguística e Artes ..... | 67 |

## LISTA DE GRÁFICOS

|                                                                                        |    |
|----------------------------------------------------------------------------------------|----|
| Gráfico 1 – Total de <i>tokens</i> para cada revista selecionada após limpeza.....     | 70 |
| Gráfico 2 – <i>Word List</i> com os substantivos comuns às duas áreas (SCEs e CLL) ... | 98 |

## **LISTA DE ORGANOGRAMAS**

|                                                                                            |     |
|--------------------------------------------------------------------------------------------|-----|
| Organograma 1 – Etapas da aprendizagem: da alfabetização às colocações especializadas..... | 47  |
| Organograma 2 – Processo para identificação das colocações adjetivas e verbais .           | 91  |
| Organograma 3 – Dados sobre a inserção dos verbetes-piloto .....                           | 123 |

## LISTA DE ABREVIATURAS

|      |                                      |
|------|--------------------------------------|
| CA   | Colocação Adjetiva                   |
| CLL  | <i>Corpus</i> de Letras/Linguística  |
| CV   | Colocação Verbal                     |
| DICA | Dicionários de Colocações Acadêmicas |
| SCEs | <i>Subcorpus</i> de Engenharias      |

## SUMÁRIO

|       |                                                                                    |    |
|-------|------------------------------------------------------------------------------------|----|
| 1     | INTRODUÇÃO .....                                                                   | 19 |
| 2     | LINGUAGEM ACADÊMICA .....                                                          | 26 |
| 2.1   | Os gêneros textuais .....                                                          | 26 |
| 2.1.1 | <i>A importância do gênero acadêmico na construção do conhecimento</i> .....       | 30 |
| 2.2   | A realidade na universidade .....                                                  | 31 |
| 2.3   | Linguagem acadêmica no contexto desta pesquisa .....                               | 36 |
| 2.4   | Linguagem acadêmica e as práticas pedagógicas que a envolvem .....                 | 38 |
| 2.5   | Novas possibilidades pedagógicas por intermédio das colocações acadêmicas .....    | 39 |
| 3     | FUNDAMENTAÇÃO TEÓRICA .....                                                        | 43 |
| 3.1   | Fraseologia .....                                                                  | 43 |
| 3.1.1 | <i>Colocações da língua geral e colocações especializadas</i> .....                | 46 |
| 3.1.2 | <i>Taxonomia das colocações</i> .....                                              | 52 |
| 3.1.3 | <i>Colocações acadêmicas</i> .....                                                 | 54 |
| 3.2   | Linguística de <i>Corpus</i> .....                                                 | 58 |
| 3.2.1 | <i>A aplicabilidade da Linguística de Corpus no estudo das colocações</i> .....    | 60 |
| 3.3   | Lexicografia: as <i>colocações</i> e os <i>dicionários</i> .....                   | 61 |
| 4     | METODOLOGIA DA PESQUISA .....                                                      | 65 |
| 4.1   | A escolha das fontes para a compilação do <i>corpus</i> e do <i>subcorpus</i> .... | 65 |
| 4.2   | Compilação do <i>corpus</i> da grande área de Engenharias a partir do CoPEP .....  | 67 |
| 4.3   | Compilação do <i>corpus</i> de artigos da grande área Letras/Linguística ...       | 68 |
| 4.4   | O Programa <i>Sketch Engine</i> .....                                              | 71 |
| 4.4.1 | <i>Word List</i> .....                                                             | 74 |
| 4.4.2 | <i>Word Sketch</i> .....                                                           | 74 |
| 4.4.3 | <i>Concordance</i> .....                                                           | 76 |
| 4.5   | Procedimentos adotados para seleção dos dados .....                                | 76 |
| 4.5.1 | <i>Identificação dos substantivos em comum no CLL e no SCEs</i> .....              | 78 |
| 4.6   | Procedimentos adotados para identificação das colocações acadêmicas comuns .....   | 88 |
| 4.7   | Microestrutura do dicionário .....                                                 | 92 |
| 5     | ANÁLISE DOS DADOS .....                                                            | 95 |

|              |                                                                                            |            |
|--------------|--------------------------------------------------------------------------------------------|------------|
| <b>5.1</b>   | <b>Análise quantitativa e descritiva dos substantivos presentes no SCEs e no CLL .....</b> | <b>95</b>  |
| <b>5.2</b>   | <b>Delimitação dos tipos de colocações analisadas na pesquisa .....</b>                    | <b>99</b>  |
| <b>5.3</b>   | <b>Análise das colocações no CLL e SCEs.....</b>                                           | <b>101</b> |
| <b>5.3.1</b> | <b><i>Tempo</i>.....</b>                                                                   | <b>101</b> |
| 5.3.1.1      | <i>Colocações adjetivas.....</i>                                                           | 102        |
| 5.3.1.2      | <i>Colocações verbais .....</i>                                                            | 103        |
| <b>5.3.2</b> | <b><i>Caso</i>.....</b>                                                                    | <b>105</b> |
| 5.3.2.1      | <i>Colocações adjetivas.....</i>                                                           | 106        |
| <b>5.3.3</b> | <b><i>Sistema</i>.....</b>                                                                 | <b>109</b> |
| 5.3.3.1      | <i>Colocações adjetivas.....</i>                                                           | 110        |
| 5.3.3.2      | <i>Colocações verbais .....</i>                                                            | 112        |
| <b>5.3.4</b> | <b><i>Processo</i> .....</b>                                                               | <b>113</b> |
| 5.3.4.1      | <i>Colocações adjetivas.....</i>                                                           | 114        |
| 5.3.4.2      | <i>Colocações verbais .....</i>                                                            | 114        |
| <b>5.3.5</b> | <b><i>Resultado</i> .....</b>                                                              | <b>118</b> |
| 5.3.5.1      | <i>Colocações adjetivas.....</i>                                                           | 119        |
| <b>5.4</b>   | <b>Amostragem da estrutura dos verbetes-pilotos na Plataforma.....</b>                     | <b>123</b> |
| <b>6</b>     | <b>CONSIDERAÇÕES FINAIS .....</b>                                                          | <b>125</b> |
|              | <b>REFERÊNCIAS.....</b>                                                                    | <b>128</b> |
|              | <b>APÊNDICES.....</b>                                                                      | <b>134</b> |
|              | APÊNDICE A – Candidatas a colocações .....                                                 | 134        |
|              | APÊNDICE B – Figuras das colocações acadêmicas .....                                       | 136        |

## 1 INTRODUÇÃO

É notável o crescente número de alunos que ingressam em universidades públicas e particulares todos os anos no Brasil. De acordo com o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), em 2017 o número de estudantes ingressantes no ensino superior teve um crescimento de 8,1% em relação a 2016. Se comparados aos anos de 2016 e 2017, na rede pública esse aumento foi de 11,3% e, na particular, foi de 7,3% no mesmo período. Esse crescimento alcança cifras gigantescas se somarmos os últimos 10 anos da pesquisa, que compreende os anos de 2007-2017, nos quais as instituições públicas tiveram um aumento de 41,7% e a rede privada mostrou um acréscimo de 53,1%. Embora o site oficial do INEP não nos forneça dados conclusivos sobre 2018, podemos verificar o crescente interesse dos indivíduos por uma graduação, nas redes pública e privada e nas modalidades presencial e a distância (EAD).

De acordo com Almeida *et al.* (2012), esse aumento de ingressantes no ensino superior se deve a vários fatores, mas principalmente ao fenômeno de “massificação” do ensino superior. Alguns programas voltados para fundos de investimentos na educação, nacionais e transnacionais, ligados a empresas do setor educacional privado e público fortaleceram esse fenômeno. Entre eles estão o Sistema de Seleção Unificado (SiSU), desenvolvido pelo Ministério da Educação (MEC), cujo objetivo é selecionar candidatos para as vagas disponibilizadas pelas universidades públicas por meio do Exame Nacional do Ensino Médio (ENEM).

Outros programas que têm colaborado para esse processo de massificação são: o Fundo de Financiamento Estudantil (FIES), criado também pelo MEC, com o objetivo de financiar a educação acadêmica de graduandos que já estejam estudando; e o Programa Universidade para Todos (PROUNI), desenvolvido em 2005 pelo governo do presidente Luiz Inácio Lula da Silva, tendo como finalidade a concessão de bolsas de estudos de 25% a 50% para estudantes de baixa renda para cursos de graduação em instituições particulares, que seriam beneficiadas com isenções fiscais, cujo processo de concessão se dá a partir da nota obtida no ENEM.

O setor privado de instituições universitárias foi, e ainda é, um dos pilares dessa massificação, com preços mais acessíveis e a disponibilidade da modalidade de EAD,

favorecendo esse crescimento. Segundo dados do jornal Folha de São Paulo<sup>1</sup>, de 2018, o ensino a distância expandiu e já representa 21,2% do total de alunos matriculados no ensino superior. Os estudantes saem da realidade do ensino médio entusiasmados com essa nova etapa, contudo, o processo de transitoriedade da escola básica para a universidade é, por vezes, impactante e pode gerar dificuldades.

Podemos perceber essa dificuldade na linguagem dos textos científicos que, de acordo com Fischer (2007, p. 51), “está presente ou compõe os gêneros discursivos do meio acadêmico” e, tão logo iniciam suas atividades acadêmicas, os alunos de graduação, por exemplo, precisam produzir materiais como: resumos, resenhas, entre outros tipos de trabalhos que demonstrem seu conhecimento acadêmico. Enquanto isso, discentes de pós-graduação precisam desenvolver produções para apresentações em congressos, artigos, bem como suas dissertações ou teses, com o objetivo de divulgar suas pesquisas em revistas, simpósios e seminários, haja vista o crescente número de publicações de textos acadêmico-científicos pelas mais diversas áreas de conhecimento.

Todavia, Biber (2006) destaca em suas pesquisas que o processo para adquirir uma escrita acadêmica adequada é desafiador. Segundo o pesquisador, além de os estudantes precisarem se ajustar à “realidade universitária”, muitos obstáculos devem ser superados para se adaptarem aos padrões acadêmicos esperados. Com efeito, de acordo com o autor, os universitários precisam “aprender” a usar a linguagem de outras maneiras (BIBER, 2006, p. 1) e, quando não preparados para isso, sentem-se limitados na escrita acadêmica, mesmo se tratando de sua língua materna.

Tendo em vista a situação descrita acima, propomos um estudo a partir das colocações acadêmicas verbais e adjetivais encontradas nos artigos científicos em português brasileiro, as quais, segundo Orenha-Ottaiano *et al.* (2021), são combinações convencionais e recorrentes dentro dos textos acadêmicos e, que não possuam um valor terminológico dentro dos textos das áreas selecionadas, que não sejam específicas de uma determinada área e que transitam entre as mais diferentes áreas de estudo.

Dessa maneira, selecionamos duas grandes áreas, seguindo a classificação da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), que a

---

<sup>1</sup> Disponível em: <https://www1.folha.uol.com.br/educacao/2018/09/ensino-superior-volta-a-crescer-no-pais-mas-so-na-modalidade-a-distancia.shtml>. Acesso em: 19 fev. 2019.

define como conjunto de “diversas áreas do conhecimento, em virtude da afinidade de seus objetos, métodos cognitivos e recursos instrumentais refletindo contextos sociopolíticos específicos”<sup>2</sup>. Assim, no que se refere às grandes áreas, escolhemos para nossa pesquisa: “Letras, Linguística e Artes”. Porém, optamos por trabalhar somente com “Letras e Linguística” e a de “Engenharias”.

Para conduzir esta pesquisa, compilamos um *corpus* de artigos do *Scientific Electronic Library On-line (SciELO)*, da grande área de Letras/Linguística (doravante CLL), contendo 476.191 palavras e, a partir do *Corpus* de Português Escrito em Periódicos (CoPEP), desenvolvido por Kuhn e Ferreira (2020), compilamos um *subcorpus* da grande área de Engenharias (doravante SCEs), com 441.591 palavras no português do Brasil, uma vez que o CoPEP abrange o português de Portugal e o português do Brasil.

A decisão de realizar uma pesquisa voltada para os gêneros acadêmicos, especificamente, os artigos científicos, visto que os alunos se deparam com vários gêneros textuais durante a graduação ou pós-graduação, tais como resenhas, fichamentos, resumos, monografias, projetos de pesquisa, artigos científicos, dissertações e teses, se deu por vários motivos, entre eles: 1) as dificuldades que permeiam estudantes, no que diz respeito à escrita acadêmica; 2) a escassez de materiais comprovados por alguns pesquisadores (ORENHA-OTTAIANO, 2020; KUHN, 2017; e, principalmente, 3) a visível necessidade de alternativas pedagógicas para amenizar essas dificuldades na produção de textos científicos.

Pensando nessas “alternativas pedagógicas” e observando a relevância das colocações no contexto acadêmico, propomos, ao analisá-las, verbetes-pilotos para os dicionários em desenvolvimento, o *DICA, Dicionários de Colocações Acadêmicas* monolíngues, cujo *link* de acesso constará na *Multilingual Collocations Dictionaries Online Platform* (doravante *PLATCOL*). Essas e outras colocações acadêmicas serão inseridas, de acordo com as pesquisadoras responsáveis pelo *DICA*, Prof.<sup>a</sup> Dr.<sup>a</sup> Adriane Orenha-Ottaiano e Prof.<sup>a</sup> Dr.<sup>a</sup> Tanara Kuhn, a fim de auxiliar estudantes brasileiros ou estrangeiros a desenvolverem seus textos. Todavia, salientamos que iremos inserir exemplos destinados aos estudantes brasileiros e estrangeiros, a este último será possível encontrar mais informações já que o exemplo pode não ser

---

<sup>2</sup> Disponível em: <https://www.capes.gov.br/avaliacao/instrumentos-de-apoio/tabela-de-areas-do-conhecimento-avaliacao>. Acesso em: 15 fev. 2019.

suficiente para a compreensão de uso de determinada colocação, visto que são públicos diferentes.

Destacamos que, por questões de limitação de espaço, conforme mencionamos, focamos o estudo em duas grandes áreas, Letras/Linguística e Engenharias, selecionando para este trabalho o gênero artigo.

A plataforma para esse dicionário está sendo desenvolvida pela Prof.<sup>a</sup> Dr.<sup>a</sup> Adriane Orenha Ottaiano e pela Prof.<sup>a</sup> Dr.<sup>a</sup> Tanara Zingano Kuhn, no âmbito do grupo de pesquisa do Conselho Nacional de Desenvolvimento Científico e Desenvolvimento Tecnológico (CNPq) *Fraseologia e Colocações Baseado em Corpora* (FRASCORP) e do projeto de pesquisa Desenvolvimento de uma Plataforma *On-Line* de Dicionários Monolíngues de Colocações Acadêmicas em inglês e português do Brasil e Europeu, no qual este trabalho também se insere.

Nossa escolha por artigos científicos se baseou no fato de que eles são vistos como um modelo de escrita acadêmica (HYLAND, 2008, p. 47), por conseguinte, são ideais ao propósito da pesquisa, visto que nosso interesse é propiciar meios para que o estudante consiga melhorar sua escrita e, diante da qualidade que possuem os artigos científicos, acreditamos ser uma das principais fontes para a identificação dos candidatos a verbetes-pilotos.

Assim, ao optarmos por duas áreas distintas, embora saibamos que entre Letras e Linguística, bem como dentro de cada área específica de Engenharias há diferenças na forma de construção dos gêneros acadêmicos, todavia, por estarem dentro de uma grande área, tomamos as análises a partir das áreas de acordo com a hierarquização segundo a Capes,<sup>3</sup> com a finalidade de contemplar como essas colocações se comportam dentro de cada área específica. Assim, nossa abordagem buscará, em ambas áreas, as colocações que se especializam dentro do contexto

---

<sup>3</sup> 1º nível - Grande Área: aglomeração de diversas áreas do conhecimento, em virtude da afinidade de seus objetos, métodos cognitivos e recursos instrumentais refletindo contextos sociopolíticos específicos; 2º nível - Área do Conhecimento (Área Básica): conjunto de conhecimentos inter-relacionados, coletivamente construído, reunido segundo a natureza do objeto de investigação com finalidades de ensino, pesquisa e aplicações práticas; 3º nível - Subárea: segmentação da área do conhecimento (ou área básica) estabelecida em função do objeto de estudo e de procedimentos metodológicos reconhecidos e amplamente utilizados; 4º nível - Especialidade: caracterização temática da atividade de pesquisa e ensino. Uma mesma especialidade pode ser enquadrada em diferentes grandes áreas, áreas básicas e subáreas. Disponível em: <https://www.gov.br/capes/pt-br/acesso-a-informacao/acoes-e-programas/avaliacao/instrumentos/documentos-de-apoio-1/tabela-de-areas-de-conhecimento-avaliacao>. Acesso em: 28 dez. 2020.

acadêmico, a fim de ser um instrumento para contribuir na qualificação da escrita dos estudantes.

Para atingir os propósitos desta pesquisa, criamos uma *Word List* com os substantivos comuns entre elas e, a partir disso, verificarmos como esses substantivos se comportam dentro de cada área. Para tanto, extraímos os dados dos textos compilados por meio da ferramenta *Word Sketch*, do programa *Sketch Engine* (KILGARRIFF *et al.*, 2004), buscando identificar as ocorrências de colocações com os substantivos em comum nas duas áreas e como essas colocações acadêmicas se apresentam no *corpus* e no *subcorpus* das áreas selecionadas para o estudo.

Desse modo, após a identificação dos 100 substantivos de cada área, selecionamos os 27 mais frequentes que ocorriam nas duas áreas e, em uma terceira etapa, com o auxílio da *Word Sketch*, extraímos as colocações que eram comuns no CLL e no SCEs, para, assim, darmos início à nossa análise. Embora tenhamos observado equivalência na forma, não é possível afirmar que os sentidos seriam os mesmos nas duas áreas. Partimos da hipótese de que era possível encontrar colocações com um sentido transversal às duas áreas, enquanto outras poderiam ter significados diferentes, mostrando sua especificidade dentro de cada área.

E, por fim, como nossa investigação vai além do fato de extrair as colocações acadêmicas, adicionando à extração, as análises e a proposta de verbetes-pilotos, justifica-se o número menor de dados para testar a metodologia que adotamos neste trabalho.

Dessa maneira, esta pesquisa sustenta-se pelo fato de que, embora existam trabalhos e cursos que envolvam o ensino de linguagem acadêmica com fins específicos, ainda são escassos os estudos que contemplem o aspecto léxico-fraseológico e os materiais de apoio para o desenvolvimento das habilidades colocacionais (ORENHA-OTTAIANO, 2004, 2009, 2012, 2015, 2016, 2020). Assim, propomos desenvolver uma ferramenta que possa ser utilizada em conjunto com a plataforma já existente, a Plataforma *On-line* de Colocações (PLATCOL), a qual traz uma aba específica com o *link* de acesso para o Dicionário de Colocações Acadêmicas (DICA). A PLATCOL já possui os Dicionários de Colocações nas direções inglês-português e português-inglês e está sendo ampliada nos últimos meses para as línguas francesa, espanhola e chinesa, em cooperação com pesquisadores, conforme consta no site da plataforma, das Universidades de Montreal, de Granada,

A Coruña, de Alcalá e de Santiago de Compostela. Além disso, possibilitará um *link* de acesso para a plataforma do *DICA*, em fase de elaboração, o qual permitirá que alunos brasileiros ou estrangeiros desenvolvam ou melhorem sua escrita acadêmica no português brasileiro.

Esta investigação, portanto, tem como objetivo geral identificar as colocações acadêmicas no CLL e no SCEs e, como questões de pesquisa, temos:

- 1) Quais são as colocações acadêmicas encontradas no *corpus* e no *subcorpus* de artigos das duas áreas?
- 2) Quais são as diferenças/semelhanças sintático-morfológicas e léxico-semânticas e como se comportam as colocações encontradas em cada uma das áreas selecionadas para o estudo?

E, por fim, seguimos os seguintes objetivos específicos:

- i. Compilar um *corpus* de artigos científicos escrito em periódicos da subárea Linguística/Letras;
- ii. Identificar as colocações acadêmicas mais frequentes e comuns no *corpus* compilado e no *Subcorpus* Engenharia do CoPEP;
- iii. Extrair e analisar, do ponto de vista sintático-morfológico e léxico-semântico, as colocações presentes nas áreas estudadas;
- iv. Comparar as colocações encontradas nas duas áreas de estudo;
- v. Propor verbetes-pilotos para alimentar o *Dicionário de Colocações Acadêmicas* em construção (*DICA*), voltados para graduandos, pós-graduandos brasileiros e estrangeiros que almejem desenvolver sua escrita acadêmica na língua portuguesa do Brasil.

Para tanto, este trabalho está organizado da seguinte forma: no capítulo 2, apresentaremos alguns apontamentos sobre o ensino superior no Brasil, incluindo como os alunos encaram essa nova fase da vida e os desafios no letramento acadêmico. No capítulo 3, abordaremos os aspectos teóricos desta pesquisa: a Linguística de *Corpus* e sua interdisciplinaridade com a Fraseologia e a Lexicografia, com a finalidade de sustentar a investigação sobre as colocações empregadas no contexto acadêmico. No capítulo 4, evidenciaremos a metodologia do estudo, detalhando as etapas de compilação do *Corpus* Letras/Linguística e do *Subcorpus* de Engenharias de artigos e o levantamento das colocações. No capítulo 5, demonstraremos as análises das colocações acadêmicas encontradas nas áreas de

Engenharias e Letras/Linguística, bem como um modelo de entrada para o Dicionário de Colocações. Por fim, encerramos com as considerações finais, as referências e os apêndices.

## 6 CONSIDERAÇÕES FINAIS

A presente investigação teve como objetivo o estudo das colocações acadêmicas presentes no *Corpus* de Letras/Linguística e no *Subcorpus* de Engenharias, tendo como hipótese inicial de que era possível encontrar colocações com um sentido transversal às duas áreas, enquanto outras poderiam ter significados diferentes, mostrando sua especificidade dentro de cada área.

Para comprovar e consolidar essa hipótese a proposta deste estudo buscou, extrair e analisar, do ponto de vista léxico-semântico, as colocações acadêmicas do português do Brasil, presentes no *Corpus* compilado da área de Letras/Linguística e um *Subcorpus* da área de Engenharias a partir do CoPEP, a partir de artigos do SciELO.

Para desenvolver a pesquisa, buscamos, por meio da testagem dos procedimentos metodológicos, encontrar os substantivos comuns nas duas áreas e, em seguida, verificar a existência de colocações com os mesmos substantivos. Encontradas e extraídas essas colocações, examinamos como cada uma se comportava dentro de cada área acadêmica, com o auxílio do programa *Sketch Engine* (KILGARRIF *et al.*, 2004).

A pesquisa mostrou que a metodologia adotada para as análises já apresentou dados relevantes para que possamos dar continuidade na busca por padrões colocacionais nas áreas selecionadas deste estudo, bem como para os demais tipos de colocações existentes. Segundo os resultados das análises, encontramos colocações acadêmicas que, apesar de possuírem a mesma forma, apresentavam sentido diferente quando utilizadas em outra área. A hipótese que levantamos no início deste trabalho, de que era possível encontrar colocações com um sentido transversal às duas áreas, enquanto outras poderiam ter significados diferentes, mostrando sua especificidade dentro de cada área, foi confirmada por meio da testagem através dos substantivos encontrados em comum nas duas áreas.

Foram testadas oito colocações acadêmicas adjetivas e verbais a partir de cinco substantivos: *tempo*, *caso*, *sistema*, *processo* e *resultado*. Com a colocação *ter tempo*, foi possível observar que algumas palavras intermediárias, ou seja, o contexto alargado demonstrou o quanto são relevantes os exemplos para conhecer melhor as colocações.

Diante desses contextos, percebemos quando o verbo era mais idiomático, enquanto na segunda ocorrência tratava-se de um sintagma preposicional. Sendo funcional o contexto, as linhas de concordâncias, para tais definições, comprovam a importância de exemplos contextualizados das colocações para favorecer o desenvolvimento acadêmico dos alunos, principalmente dos estrangeiros que precisam de mais informações, já que possuem necessidades diferentes, cujos exemplos podem não ser suficientes. Por isso a importância das análises, pois podem ajudar na criação de notas explicativas, favorecendo a compreensão e oferecendo suporte para a escolha de qual colocação acadêmica usar.

Dos cinco substantivos selecionados para testar nossa metodologia, aquele que mais se destacou foi a ocorrência de *resultado*. Elegemos a CA *resultados obtidos* e a CV *obter resultados*, enfatizando a alta frequência de colocações com esse substantivo. Observamos que, embora tenha ocorrido muito mais a CA *resultado obtido* do que a CV *obter resultados*, foi possível verificar que não podemos unir as duas colocações no dicionário, porque da perspectiva da abordagem *corpus-driven*, indica que ela é preferida, e se é, há uma razão para tal escolha. Assim, mais uma vez se comprova a importância de uma análise seletiva para o desenvolvimento do DICA.

Dessa forma, não há como utilizar uma colocação sem estudar seu contexto, analisando suas possibilidades de uso. Não é possível afirmar que são a mesma coisa e são usadas da mesma forma. Contudo, podemos perceber que tanto o verbo *obter* e sua forma participial, bem como o substantivo *resultado* é frequente em ambas as áreas de estudo, entretanto, prevalecendo a maior frequência no SCEs.

No final, houve um alto número de colocações acadêmicas em comum que tiveram o mesmo sentido nas duas áreas. Por outro lado, tivemos ocorrências com o verbo *ter*, os quais, embora possuam a mesma forma, verificamos a diferença de sentido nas áreas analisadas, como relatamos acima.

Acreditamos que, ao tomar conhecimento dessas possibilidades de significações, o estudante seja auxiliado quando fizer uso da Plataforma. Uma vez que a construção do saber acadêmico é feita de estudos e pesquisas contínuas, ao fazer uso do dicionário o aluno não somente tem contato com uma colocação, mas aprende como e quando utilizá-la, já que, como mencionamos no capítulo da fundamentação teórica, a didática da Plataforma mostrou-se simples e convidativa.

A análise das colocações acadêmicas mostrou-se importante para o desenvolvimento de um dicionário de colocações acadêmicas que já se encontra em desenvolvimento pela Prof.<sup>a</sup> Dr.<sup>a</sup> Adriane Orenha Ottaiano e pela Prof.<sup>a</sup> Dr.<sup>a</sup> Tanara Zingano Kuhn, por meio do projeto *Desenvolvimento de uma Plataforma On-Line de Dicionários de Colocações Acadêmicas (DICA) monolíngues em inglês e português do Brasil e Europeu*.

Analizando os aspectos discutidos e apresentados neste trabalho, esperamos que esta pesquisa possa trazer significativas contribuições no aprendizado das colocações acadêmicas à luz da Linguística de *Corpus*, da Fraseologia e da Lexicografia, auxiliando os estudantes na qualidade de sua escrita acadêmica.

Assim, pretendemos continuar com a extração das colocações acadêmicas em pesquisas futuras, a fim de ampliar o número de dados do estudo e enriquecer o dicionário, utilizando a metodologia testada nas colocações das áreas de Letras/Linguística e Engenharias em outras áreas, assim como para os demais tipos de colocações existentes.

## REFERÊNCIAS

ALMEIDA, L. *et al.* Democratização do acesso e do sucesso no ensino superior: uma reflexão a partir das realidades de Portugal e do Brasil. *Avaliação*, Campinas, v. 17, n. 3, p. 899-920, 2012. Disponível em: doi: 10.1590/S1414- 40772012000300014. Acesso em: 23 fev. 2021.

ALVES, M. F.; MOURA, L. O. B. M. A escrita de artigo acadêmico na universidade: autoria x plágio. *Ilha Desterro*, Florianópolis, v. 69, n. 3, p. 77-93, dez. 2016. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S2175-80262016000300077&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2175-80262016000300077&lng=en&nrm=iso). Acesso em: 23 set. 2020.

ARROJO, R. *Oficina de tradução: a teoria e a prática*. 2. ed. São Paulo: Ática, 1992.

ATKINS, B.T.S. Theoretical lexicography and its relation to dictionary-making. In: T. FONTENELLE (ed.). *Practical lexicography*. Oxford: Oxford University Press, 2008. p. 31-50.

BALLY, C. *Traité de stylistique française*. C. Winter Heidelberg, 1957.

BARBOSA, M. A. *Lexicologia, lexicografia, terminologia, terminografia, identidade científica, objeto, métodos, campos de atuação*. In Anais do II Simpósio Latino-Americano de Terminologia. I Encontro Brasileiro de Terminologia Técnico-Científica. Brasília, 1990. p.152-158.

BEVILACQUA, C. R. *A fraseologia jurídico-ambiental*. 1996. 120 f. Dissertação (Mestrado em Estudos de Linguagem) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 1996.

BÉJOINT, H. *Tradition and innovation in modern English dictionaries*. Oxford: Oxford University Press, 1994.

BERBER SARDINHA, T., Mayer Acunzo, C., & São Bento Ferreira, T. de L. (2016). *Metáforas da economia no dicionário de colocações do português brasileiro: Uma análise multidimensional baseada em corpus*. *Filologia E Linguística Portuguesa*, 18(1), 175-198. <https://doi.org/10.11606/issn.2176-9419.v18i1p175-198>.

BIBER, D. *A corpus-based study of spoken and written registers*. University language. Amsterdam/Philadelphia: John Benjamin's Publishing Company, 2006.

BIBER, D.; CONRAD, S.; REPPEN, R. *Corpus Linguistics: Investigating Language Structure and Use*. Nova Iorque: Cambridge University Press, 1998.

BIDERMANN, T. C. A ciência da lexicografia. *Alfa: Revista de Linguística*, São José do Rio Preto, v. 28, supl., p. 1-26, 1984a.

BIDERMANN, M. T. C. O dicionário padrão da língua. *Alfa: Revista de Linguística*, São José do Rio Preto, v. 28, n. 1, p. 27-43, 1984b.

BIDERMAN, M. T. C. As ciências do léxico. In: OLIVEIRA, A. M. P. P.; ISQUERDO, A. N. (org.). *As ciências do léxico: lexicologia, lexicografia e terminologia*. Campo Grande: Ed. UFMS, 1998. 1 v. p. 11-20.

BRÉAL, M. *Ensaio de semântica*. São Paulo: Fontes/Educ, 1992.

BRONCKART, J. *Atividade de linguagem, textos e discursos: por um interacionismo sócio discursivo*. São Paulo: Educ, 1999.

CABRÉ, M. T. Una nueva teoría de la terminología: de la denominación a la comunicación. In: CABRÉ, M. T. *La terminología: representación y comunicación*. Barcelona: Universitat Pompeu Fabra, 1999. p. 109-127.

CANALE, M. From Communicative Competence to Communicative Language Pedagogy. In: RICHARD, J. C.; SCHMIDT, R. W. (ed.). *Language and Communication*. London: Longman, 1983. p. 2-14.

CORPAS PASTOR, G. *Manual de fraseología española*. Madrid: Editorial Gredos, 1996.

DAYRELL, C. Corpora no ensino do inglês acadêmico: padrões léxico-gramaticais em *abstracts* de pós-graduandos brasileiros. In: VIANA, V.; TAGNIN, S. E. O. (org.). *Corpora no ensino de línguas estrangeiras*. São Paulo: Hub Editorial, 2010, p. 137-171.

ENSINO superior volta a crescer no país, mas só na modalidade a distância. *Folha de São Paulo*. 20 set. 2018. Disponível em: <https://www1.folha.uol.com.br/educacao/2018/09/ensino-superior-volta-a-crescer-no-pais-mas-so-na-modalidade-a-distancia.shtml>. Acesso em: 19 fev. 2019.

FILLMORE, C. Innocence: a second idealization for linguistics. In: BERKELEY LINGUISTICS SOCIETY, 50., 1979, Berkeley. *Proceedings* [...]. Berkeley, 1979. p. 63-76.

FIORIO, N. M. *Quem conta um conto...A metáfora rural de provérbios em língua portuguesa*. Goiânia: Editora UCG, 1995.

FIRTH, J. R. Modes of meaning. In: *Papers in Linguistics – 1934-1951*. Oxford: Oxford University Press, 1957. p. 114-118.

FISCHER, A. *A construção de letramentos na esfera acadêmica*. 2007. 341 f. Tese (Doutorado em Linguística) – Universidade Federal de Santa Catarina, Florianópolis, 2007.

HALLIDAY, M. Lexicology. In: HALLIDAY, M. et al. *Lexicology and corpus linguistics*. London: Continuum, 1997, p. 1-22.

HALLIDAY, M. A.; MARTIN, J. R. (ed.). *Writing science. Literacy and discursive power*. London; Washington, DC: Falmer Press, 1996. Disponível em: doi:

10.1017/CBO9781107415324.004. Acesso em: 23 fev. 2021.

HANKS, P. Corpus evidence and electronic lexicography. In: GRANGER, S.; PAQUOT, M. (ed.). *Electronic Lexicography*. Oxford: Oxford University Press, 2012a. p. 57-82. Disponível em: [www.patrickhanks.com/uploads/5/1/4/9/5149363/hanks\\_2012f.pdf](http://www.patrickhanks.com/uploads/5/1/4/9/5149363/hanks_2012f.pdf). Acesso em: 19 fev. 2019.

HANKS, P. The corpus revolution in lexicography. *International Journal of Lexicography*, v. 25, n. 4, p. 398-436, 2012b.

HAUSMANN, F. J. Un dictionnaire de collocations est-il possible? *Travaux de linguistique et de littérature*, Strasbourg, v. 17, n. 1, p. 187-195, 1979.

HAUSMANN, F. J. Kollokationen in deutschen Wörterbuch. Ein Beitrag zur Theorie des lexikographischen Beispiels'. In: BERGENHOLTZ, H.; MUGDAN, J. (org.). *Lexikographie und Grammatik*. Tübingen: Niemeyer, 1985. p. 22-26.

HEID, U.; MARTIN, W.; POSCH, I. An overview of approaches towards the description of collocations. In: Feasibility of standards for collocational description of lexical items. EUROTRA 7, Report, Stuttgart: Amsterdam, 1991.

HOUAISS. Dicionário. 2020. Disponível em: [houaiss.uol.com.br/corporativo/apps/uol\\_www/v5-4/html/index.php#0](http://houaiss.uol.com.br/corporativo/apps/uol_www/v5-4/html/index.php#0). Acesso em: 2 maio 2020.

HYLAND, K. *Academic discourse*. London: Continuum, 2008.

INEP. *Censo da Educação Superior*. Brasília: INEP, 2017. Disponível em: <http://inep.gov.br/censo-da-educacao-superior>. Acesso em: 15 fev. 2019.

KENNEDY, G. *An introduction to corpus linguistics*. London: Longman, 1998.

KILGARRIFF, A.; BAISA, V.; RYCHLÝ, P.; JAKUBÍČEK, M. Longest-commonest match. In: KOSEM, I.; JAKUBÍČEK, M.; KALLAS, J.; KREK, S. (ed.). *ELECTRONIC LEXICOGRAPHY IN THE 21ST CENTURY: linking lexical data in the digital age*. 2015. *Proceedings* [...]. Ljubljana/Brighton: Trojina, Institute for Applied Slovene Studies/Lexical Computing Ltd. 2015, p. 397-404. Disponível em: [https://www.sketchengine.co.uk/wp-content/uploads/Longestcommonest\\_eLex2015.pdf](https://www.sketchengine.co.uk/wp-content/uploads/Longestcommonest_eLex2015.pdf). Acesso em: 15 fev. 2019.

KILGARRIFF, A. et al. The sketch engine. In: WILLIAMS, G.; VESSIER, S. (ed.). *EURALEX INTERNATIONAL CONGRESS, 11., Lorient. Proceedings* [...]. Lorient: Université de Bretagne-Sud, 2004. p. 105-116.

KILGARRIFF, A. et al. The Sketch Engine: ten years on. *Lexicography: Journal of ASIALEX*, v. 1, p. 7-36, 2014.

KOSEM, I. *Designing a model for a corpus-driven dictionary of academic English*. 2010. 498 f. Tese (Doutorado em Linguística de Corpus) – Aston University, Birmingham, 2010.

KUHN, T. Z. *A Design Proposal of an On-line Corpus-Driven Dictionary of Portuguese for University Students*, 2017. 421 f. Tese (Doutorado em Estudos Linguísticos) – Universidade de Lisboa, Faculdade de Letras, Lisboa, 2017.

KUHN, T. Z.; FERREIRA, J. P. O corpus de português escrito em periódicos - CoPEP. *Delta*, São Paulo, v. 36, n. 2, 2020. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-44502020000200408&lng=pt&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502020000200408&lng=pt&nrm=iso). Acesso em: 23 set. 2020.

LEA, M. R.; STREET, B. V. Student writing in higher education: an academic literacies approach. *Studies in Higher Education*, Abingdon, Oxon, UK, v. 23, n. 2, p. 157-172, 1998.

MACMILLAN. *English Dictionary for Advanced Learners*. 2010. Disponível em: <http://www.macmillandictionary.com/>. Acesso em: 23 set. 2020.

MARCUSCHI, L.A. Gêneros textuais: definição e funcionalidade. In: *Gêneros textuais e ensino*. Rio de Janeiro: Lucena, 2003, p. 20-36

MCENERY, T.; WILSON, A. *Corpus Linguistics: An introduction*. Edinburgh: Edinburgh University Press, 2006.

MCENERY, T.; HARDIE, A. *Corpus linguistics: methods, theory and practice*. New York: Cambridge University Press, 2012.

MOTTA-ROTH, D. Gêneros textuais na análise crítica de Fairclough”. In: MEURER, José Luis; BONINI, Adair; MOTTA-ROTH, Désirée (org.). *Gêneros: teorias, métodos e debates*. São Paulo: Parábola Editorial, 2005.

MOTTA-ROTH, D. Escrevendo no contexto: contribuições da LSF para o ensino de redação acadêmica. In: INTERNATIONAL SYSTEMIC FUNCTIONAL CONGRESS, 33., São Paulo. *Anais [...]*. São Paulo, 2006. p. 828-851. Disponível em: [http://w3.ufsm.br/desireemroth/images/stories/ANAIS\\_DO\\_III\\_SIGET.pdf](http://w3.ufsm.br/desireemroth/images/stories/ANAIS_DO_III_SIGET.pdf). Acesso em: 20 mar. 2020.

NESSELHAULF, N. *Collocations in a learner corpus*. Amsterdam: John Benjamins Publishing Company, 2005. 335 p.

NUNES, J. H. Lexicologia e lexicografia. In: GUIMARÃES, E.; ZOPPI-FONTANA, M. (orgs.). *Introdução às Ciências da Linguagem*. Campinas: Pontes, 2006. 1 v. p. 149-172.

OLIVEIRA, E. F. *Letramentos acadêmicos: o gerenciamento de vozes em resenhas e artigos científicos produzidos por alunos universitários*. 2015. 446 f. Tese (Doutorado em Linguística Aplicada) – Universidade Estadual de Campinas, Campinas, 2015.

ORENHA-OTTAIANO, A. *A compilação de um glossário bilíngue de colocações, na área de jornalismo de negócios, baseado em corpus comparável*. 2004. 246 f. Dissertação (Mestrado em Estudos Linguísticos e Literários) – Universidade de São Paulo, São Paulo, 2004.

ORENHA-OTTAIANO, A. *Unidades fraseológicas especializadas: colocações e colocações estendidas em contratos sociais e estatutos sociais traduzidos no modo juramentado e não-juramentado*. 2009. 282 f. Tese (Doutorado em Estudos Linguísticos) – Universidade Estadual Paulista, São José do Rio Preto, 2009.

ORENHA-OTTAIANO, A. Semelhanças e diferenças entre colocações e colocações especializadas. In: ORTIZ-ALVAREZ, M. L. (org.). *Tendências atuais na pesquisa descritiva e aplicada em fraseologia e paremiologia*. 1. ed. Campinas: Editora Pontes, 2012, 2 v. p. 147-163.

ORENHA-OTTAIANO, A. Collocations workbook: um material de apoio pedagógico on-line e baseado em corpus para o ensino de colocações em inglês. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 23, n. 3, p. 833-881, 2015a.

ORENHA-OTTAIANO, A. Ensino de inglês como LE e contribuições pedagógicas de um glossário bilíngue de colocações. *Signótica*, Goiânia, v. 27, n. 2, p. 485-510, 2015b.

ORENHA-OTTAIANO, A. The compilation of an online corpus-based bilingual collocations dictionary. In: CORPAS PASTOR, G. (org.). *Computerised and corpus - based approaches to phraseology: monolingual and multilingual perspectives*. Genebra: Editions Tradulex, 2016. 1 v. p. 486-493.

ORENHA-OTTAIANO, A. The compilation of an online corpus-based bilingual collocations dictionary: motivations, obstacles and achievements. In: E-LEX Conference, 2017, Leiden. *Proceedings* [...]. Leiden, 2017. p. 458-473.

ORENHA-OTTAIANO, A. *Tela de Abertura da Plataforma On-line de Dicionários Multilíngues de colocações*. 2020. Disponível em: <http://www.institucional.grupogbd.com/dicionario/sobre>. Acesso em: 18 jul. 2020.

ORENHA-OTTAIANO, A. The creation of an On-line English Collocations Platform to help develop collocational competence. *Phrasis. Rivista Di Studi Fraseologici e Paremiologici*, Roma, 2020 (no prelo).

ORENHA-OTTAIANO, A. et al. Proposta de desenvolvimento de uma Plataforma On-line de Dicionários de Colocações Acadêmicas. In: CONGRESSO DE PORTUGUÊS COMO LÍNGUA ESTRANGEIRA, 1., 2021, Nova York. *Anais* [...]. Nova York: Columbia University, 2021.

OXFORD. *Oxford collocations dictionary for students of English*. Oxford: Oxford University Press, 2012. 992 p.

ROJO, R. Letramento e capacidades de leitura para a cidadania. In: CONGRESSO SEE: CENP, 2004, São Paulo. *Anais* [...]. São Paulo: SEE: CENP, 2004. p. 1-8.

RUNDELL, Michael. *Recent Trends in English Pedagogical Lexicography*. *International Journal of Lexicography*, v. 11, n. 4, p. 315-342, 1998.

RUNDELL, M. (ed.). *Macmillan Collocations Dictionary*. Oxford: Macmillan Education, 2008.

SAUSSURE, F. *Curso de Linguística Geral*. São Paulo: Editora Cultrix, 2006.

SILVA FILHO, R. L. L. *et al.* A evasão do ensino superior brasileiro. *Cadernos de Pesquisa*, São Paulo, v. 37, n. 132, p. 641-659, 2007.

SINCLAIR, J. Corpus and text – basic and principles. In: WYNNE, M. (ed.). *Developing linguistic corpora: a guide to good practice*. Oxford: Oxbow Books: 2005. p. 1-16.

SOARES, M. Novas práticas de leitura e escrita: letramento na cibercultura. *Educação e Sociedade*, Campinas, n. 81, dez. 2002.

STREET, B. Os novos estudos sobre letramento: histórico e perspectivas. In: MARILDES, M.; CARVALHO, G. T. *Cultura escrita e letramento*. Belo Horizonte: Editora da UFMG, 2010a.

STREET, B., Dimensões “escondidas” na escrita de artigos acadêmicos. *Revista Perspectiva*, Florianópolis, v. 28, n. 2, 541-567, jul./dez. 2010b.

SWALES, J. M. *Genre analysis: English in Academic and Research Settings*. Cambridge, UK. New York: Cambridge University Press, 1990.

TAGNIN, S. E. O. Collecting data for a bilingual dictionary of verbal collocations: From scraps of paper to corpora research. In: LEWANDOWSKA-TOMASZCZYK, B.; MELIA, P. J. (ed.). *PALC '99: Practical Applications in Language Corpora*. International Conference at the University of Lodz, 15-18 abr. 1999. Frankfurt am Main: Peter Lang GmbH, 2000, p. 399-407.

TAGNIN, S. E. O. O humor como quebra da convencionalidade. *Revista Brasileira de Linguística Aplicada*. v. 5, n. 1, Belo Horizonte, p. 247-257, 2005a.

TAGNIN, S. E. O. *O jeito que a gente diz, expressões idiomáticas e convencionais*. 1. ed. São Paulo: Disal Editora, 2005b.

TAGNIN, S. E. O. *O jeito que a gente diz: combinações consagradas em inglês e português*. Barueri: Disal Editora, 2013.

TODD, R. W. An opaque engineering word list: which words should a teacher focus on? *English for Specific Purposes*, v. 45, p. 31-39, 2017. Disponível em: doi:10.1016/j.esp.2016.08.003. Acesso em: 18 jul. 2020.

VYGOTSKY, L. S. Psicologia concreta do homem. Manuscrito inédito de Vygotsky. *Psikhologiya*, Moscou, n. 1, p. 51-64, 1986.