

RESSALVA

Atendendo solicitação do(a)
autor(a), o texto completo
desta dissertação será
disponibilizado somente
a partir de 30/05/2020.



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

DESENVOLVIMENTO DA COMPETÊNCIA COLOCACIONAL BASEADO
EM CORPORA DE APRENDIZES DO ENSINO FUNDAMENTAL I

SÃO JOSÉ DO RIO PRETO
2018

ELAINE CRISTINA FERREIRA DE OLIVEIRA

DESENVOLVIMENTO DA COMPETÊNCIA COLOCACIONAL BASEADO
EM CORPORA DE APRENDIZES DO ENSINO FUNDAMENTAL I

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre 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”, Câmpus de São José do Rio Preto. (Área de concentração: Linguística Aplicada)

Orientadora: Prof^a. Dr^a. Adriane Orenha-Ottaiano

SÃO JOSÉ DO RIO PRETO
2018

Oliveira, Elaine Cristina Ferreira de .

Desenvolvimento da competência colocacional baseado em corpora de aprendizes do ensino fundamental I / Elaine Cristina Ferreira de Oliveira -- São José do Rio Preto, 2018

200 f.: il.

Orientador: Adriane Orenha-Ottaiano

Dissertação (mestrado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências, Letras e Ciências Exatas

1. Linguística de corpus. 2. Língua portuguesa - Brasil. 3. Ensino fundamental. I. Universidade Estadual Paulista "Júlio de Mesquita Filho". Instituto de Biociências, Letras e Ciências Exatas. II. Título.

CDU - 41

ELAINE CRISTINA FERREIRA DE OLIVEIRA

DESENVOLVIMENTO DA COMPETÊNCIA COLOCACIONAL BASEADO
EM CORPORA DE APRENDIZES DO ENSINO FUNDAMENTAL I

Dissertação apresentada como parte dos requisitos para obtenção do título de Mestre 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”, Câmpus de São José do Rio Preto. (Área de concentração: Linguística Aplicada)

Comissão Examinadora

Prof^a. Dr^a. Adriane Orenha-Ottaiano
UNESP – São José do Rio Preto
Orientadora

Prof^a. Dr^a. Rosemeire Selma Monteiro-Plantin
UFC – Universidade Federal do Ceará

Prof^a. Dr^a. Mariangela de Araujo
USP – Universidade de São Paulo

São José do Rio Preto
30 de maio de 2018

Aos queridos conselheiros, confidentes e pessoas em quem me espelho na honestidade e caráter, mamãe Vera, papai Antonio e também aos caros irmãos Leandro e Joyce.

Ao meu amigo, noivo, amante e o homem pelo qual me apaixono todos os dias: Ravel Gimenes. Obrigada pelas orações constantes, companheirismo ímpar e principalmente por acreditar em mim quando eu própria não achava ser capaz. Essa dissertação só foi possível graças a vocês, meus amores!!!

AGRADECIMENTOS

A Deus, por me conceder persistência, direcionamento e capacidade intelectual. A Ele, sou grata por todos os momentos “epifânicos” em que encontrei forças para desenvolver este árduo e prazeroso trabalho dissertativo.

A querida e prestativa orientadora, professora Dra. Adriane Orenha-Ottaiano. Obrigada por acreditar neste ambicioso projeto; pelas incontáveis trocas de e-mails, áudios e mensagens via *WhatsApp* ‘fora do horário comercial’. Seus ensinamentos foram inestimáveis para que eu atravessasse o rio da Linguística de Corpus e seus afluentes com segurança e, sobretudo, lhe agradeço por me ensinar o poder da humildade científica e intelectual.

As professoras que gentilmente participaram da banca de qualificação, Prof^ª. Dr^ª. Marilei Amadeu Sabino e Prof^ª. Dr^ª. Tanara Zingano Kuhn que muito contribuíram para meu crescimento acadêmico ao sugerirem leituras e oferecerem importantes conselhos nesta escrita dissertativa. Respectivamente, incluo neste parágrafo as docentes Prof^ª. Dr^ª. Rosemeire Selma Monteiro-Plantin e Prof^ª. Dr^ª. Mariangela de Araujo por terem aceito o convite para integrar a da banca de defesa desta dissertação.

Estendo meus agradecimentos à Secretária Municipal de Votuporanga, Profa. Dra. Encarnação Manzano que deferiu os pedidos de licença para fins de estudo, permitindo a minha dedicação às obrigações do Mestrado, que exigiram minha ausência da sala de aula (durante a participação de disciplinas e Congressos), enquanto meus alunos eram assistidos pelas colegas de trabalho.

As diretoras Sandra Poltronieri de Jesus e Diná Mara Filasi Barbosa; e às coordenadoras Roberta Bento, Ana Rodrigues e Miriam Campi, assim como a meus colegas professores e equipe escolar pelo apoio incondicional. Todos esses profissionais foram prestativos para que o processo de estudos fosse realizado em harmonia com o meu trabalho docente, na querida unidade escolar em que leciono, Centro de Educação Municipal Professora Clary Brandão Bertoncini, na cidade de Votuporanga, estado de São Paulo.

Neste sentido, sou grata ainda aos pais e responsáveis pelos alunos que me permitiram coletar as redações que protagonizam esta investigação empírica. Graças a vocês, caros parceiros, foi possível realizar as análises cujos frutos espero que possam ajudar na

compreensão dos fenômenos relacionados ao ensino e aprendizagem da língua materna em nosso país.

Também agradeço aos meus colegas do Grupo de Pesquisa “Pedagogia do Léxico e da Tradução e Linguística de Corpus”. Juntos desenvolvemos pesquisas, estudamos sobre linguística, tomamos muito café e cappuccino na cantina da UNESP enquanto aprendíamos e ensinávamos uns aos outros, caros Jean Michel Pimentel Rocha e Ariane Donizete Delgado Ribeiro Caldas.

Dedico estas linhas a Guilherme de Souza e Vinicius Cineli Alves, pela amizade sincera e torcida fraterna mútua. Os agradáveis momentos de *happy hour*, desabafos na labuta desta escrita e incríveis conselhos de vocês dois estarão guardados para sempre em meu coração!

Por último, mas não menos importante, a caríssima cunhada Danúbia Cristina Lemos da Costa Oliveira e ao meu irmão Leandro Ferreira, que tão fraternalmente me acolheram em seu lar nos meses derradeiros do mestrado. Na casa de vocês encontrei a calma, a paz e o conforto necessários para efetuar as análises finais deste texto, em meio a conversas tão agradáveis, carinhosas e afetuosas, na cozinha, sala e copa do apartamento mais encantador que há em São José do Rio Preto!

A todos, o meu desejo que o Criador lhes abençoe em dobro pelo que fizeram por mim e finalmente, muito obrigada por tudo!!!

Eis o meu segredo. É muito simples: só se vê bem com o coração. O essencial é invisível aos olhos.
(SAINT-EXUPÉRY, 1994, p. 72)

RESUMO

Esta pesquisa tem como embasamento teórico-metodológico a Linguística de Corpus (SINCLAIR, 1991; MEYER, 2004; MCENERY; XIAO; TONO, 2006) e investigações sobre corpora de aprendizes (TRIBBLE, 1990; GRANGER, 1998; GRANGER; HUNG; PETCH-TYSON, 2002; ORENHA-OTTAIANO, 2012, 2016). Relaciona-se ao Grupo de Pesquisa denominado “Fraseologia e colocações a partir de corpora” (FRASCORP), devidamente catalogado no Diretório de Grupos de Pesquisa do CNPq, além de integrar uma das propostas do projeto guarda-chuva “A compilação de materiais didáticos e glossários especializados baseados em corpora e sua contribuição para uma Pedagogia do Léxico e da Tradução”, desenvolvido na UNESP de São José do Rio Preto, pela Profa. Dra. Adriane Orenha-Ottaiano. O trabalho traz as investigações sobre dois Corpora de Redações de Aprendizes do Ensino Fundamental (CRAEF 1 e CRAEF 2). O CRAEF 1 é composto por 123 redações escritas por 30 estudantes do terceiro ano do Ciclo Básico (terceiro ano do Ensino Fundamental), de uma escola pública do centro da cidade de Votuporanga, no noroeste paulista. O CRAEF 2 é composto por 101 redações escritas por 29 estudantes do terceiro ano do Ciclo Básico de uma escola pública da periferia da mesma cidade. Para realizar a análise fraseológica, utilizamos o software WordSmith Tools (SCOTT, 2012), versão 6.0, a fim de verificar os itens de maior representatividade (Keywords) nos referidos corpora de aprendizes. Analisamos o léxico-fraseológico, mais especificamente as colocações, de cada sala mencionada, bem como as particularidades da escrita dos aprendizes de ambas as escolas em foco. Portanto, o presente estudo visa a apresentar uma das aplicações da Linguística de Corpus voltada ao ensino de colocações no Português do Brasil, nas séries iniciais, e, a partir dos resultados propiciados pelas análises com o uso de corpora, desenvolver materiais autênticos de ensino da variante brasileira do português.

Palavras-chave: Colocações. Corpora de Aprendizes. Competência Colocacional. Português do Brasil.

ABSTRACT

The present research is based on the theoretical-methodological framework of Corpus Linguistics (SINCLAIR, 1991; MEYER, 2004; MCENERY; XIAO; TONO, 2006) and on research in learner corpora (TRIBBLE, 1990; GRANGER, 1998; GRANGER; HUNG; PETCH-TYSON, 2002; ORENHA-OTTAIANO, 2012, 2016). It is linked to the Research Group "Corpus-based Phraseology and Collocations" (FRASCORP), duly registered in the Directory of Research Groups of CNPq, besides constituting one of the propositions of the umbrella project "The compilation of didactic materials and specialized glossaries based on corpora and its contribution to a Teaching of the Lexicon and of Translation", developed at UNESP/São José do Rio Preto by Prof. Dr. Adriane Orenha-Ottaiano. This work involves research on two Primary Education Learner Essay Corpora (CRAEF 1 and CRAEF 2). The CRAEF 1 is composed of 123 essays written by 30 students in the third year of the Basic (Elementary) Education of a school placed in downtown Votuporanga, a town in the northwest region of the state of São Paulo. The CRAEF 2 is composed of 101 essays written by 29 students in the third year of the Basic Education of a school placed in the suburb of the same town. To carry out the phraseological analysis, we used the software WordSmith Tools (SCOTT, 2012), version 6.0, in order to verify the most representative items (Keywords) in the aforementioned learner corpora. We analyzed the phraseological units, more specifically the collocations, found on the texts written by each group mentioned, as well as the particularities in the writing of the learners of both schools. Therefore, the present study aims at presenting one of the applications of Corpus Linguistics concerning the teaching of collocations in Brazilian Portuguese, in the initial stages, and, stemming from the results achieved by the analysis using corpora, develop authentic teaching materials for Brazilian Portuguese.

Keywords: Collocations. Learner Corpora. Collocational Competence. Brazilian Portuguese.

LISTA DE FIGURAS

Figura 01 – O uso de corpora no ensino e aprendizagem de segunda língua.....	28
Figura 02 – Níveis de Convencionalidade.....	47
Figura 03 – Alguns tipos de Unidades Fraseológicas	54
Figura 04 – Convencionalidade e Fraseologia	55
Figura 05 – Taxonomia de Corpora.....	74
Figura 06 – IDEB da escola do Centro.....	77
Figura 07 – IDEB da escola da Periferia	79
Figura 08 – <i>Interface do Software WordSmith Tools</i> e as ferramentas utilizadas para a realização das análise	83
Figura 09 – Legenda das siglas	86
Figura 10 – Estatísticas do CRAEF 1	88
Figura 11 – Estatísticas do CRAEF 2	90
Figura 12 – Amostra da Escola 1 (Aluno 18)	98
Figura 13 – Amostra da Escola 1 (Aluno 30)	99
Figura 14 – Amostra da Escola 2 (Aluno 03)	100
Figura 15 – Amostra da Escola 2 (Aluno 07)	100
Figura 16 – Amostra da Escola 1 (Aluno 10)	104
Figura 17 – Amostra da Escola 1 (Aluno 24)	104
Figura 18 – Amostra da Escola 2 (Aluno 17)	105
Figura 19 – Amostra da Escola 2 (Aluno 24)	105
Figura 20 – Amostra da Escola 1 (Aluno 01)	109
Figura 21 – Amostra da Escola 1 (Aluno 26)	110
Figura 22 – Amostra da Escola 2 (Aluno 04)	110
Figura 23 – Amostra da Escola 2 (Aluno 08)	111
Figura 24 – Imagem utilizada como apoio ao Tema 04	114
Figura 25 – Amostra da Escola 1 (Aluno 06).....	116
Figura 26 – Amostra da Escola 1 (Aluno 22).....	117
Figura 27 – Amostra da Escola 2 (Aluno 05).....	118
Figura 28 – Amostra da Escola 2 (Aluno 19)	119
Figura 29 – Ilustração do Tema 05	121
Figura 30 – Amostra da Escola 1 (Aluno 03)	123
Figura 31 – Amostra da Escola 2 (Aluno 04)	123
Figura 32 – Tela das concordâncias com a palavra “BANHO”	127
Figura 33 – Tela das concordâncias com a palavra “PEDRA”	128
Figura 34 – Tela das concordâncias com a palavra “COMIDA”	129
Figura 35 – Tela das concordâncias com a palavra “DENTES”	129
Figura 36 – Tela das concordâncias com a palavra “CHUVA”	131
Figura 37 – Tela das concordâncias com a palavra “FORTE”	131
Figura 38 – Tela das concordâncias com a palavra “PESCOÇO”	132
Figura 39 – Tela das concordâncias com a palavra “MATRIZ”	133
Figura 40 – Tela das concordâncias com a palavra “DEUS”	134

Figura 41 – Tela das concordâncias com a palavra “GREVE”	134
Figura 42 – Tela das concordâncias com a palavra “JOGAR”	135
Figura 43 – Tela das concordâncias com a palavra “PIPA”	136
Figura 44 – Tela das concordâncias com a palavra “ANDAR”	137
Figura 45 – Tela da concordância com a palavra “CORAÇÃO”	138
Figura 46 – Tela das concordâncias com a palavra “PÉ”	139
Figura 47 – Tela das concordâncias com a palavra “RÁPIDO”	140
Figura 48 – Tela das concordâncias com a palavra “BEM”	141
Figura 49 – Tela das concordâncias com a palavra “PÉ”	142
Figura 50 – Tela das concordâncias com a palavra “MILHO”	142
Figura 51 – Tela das concordâncias com a palavra “MUITO”	143
Figura 52 – Tela das concordâncias com a palavra “CAVALO”	145
Figura 53 – Tela das concordâncias com a palavra “FOGO”	146
Figura 54 – Tela das concordâncias com a palavra “SAIR”	148
Figura 55 – Tela das concordâncias com a palavra “FOGO”	149
Figura 56 – Tela das concordâncias com a palavra “NATAL”	151
Figura 57 – Tela das concordâncias com a palavra “BOA”	152
Figura 58 – Tela das concordâncias com a palavra “SORO”	153
Figura 59 – Tela das concordâncias com a palavra “CÂNCER”	154

LISTA DE QUADROS

Quadro 01 – Trabalhos de destaque do grupo de pesquisa “Pedagogia do Léxico e da Tradução e Linguística de Corpus”	34
Quadro 02 – Tipologia das colocações	62
Quadro 03 – Colocações do CRAEF 1 (Tema 01)	130
Quadro 04 – Colocações do CRAEF 2 (Tema 01)	132
Quadro 05 – Colocações do CRAEF 1 (Tema 02)	136
Quadro 06 – Colocações do CRAEF 2 (Tema 02)	138
Quadro 07 – Colocações do CRAEF 1 (Tema 03)	141
Quadro 08 – Colocações do CRAEF 2 (Tema 03)	144
Quadro 09 – Colocações do CRAEF 1 (Tema 04)	147
Quadro 10 – Colocações do CRAEF 2 (Tema 04)	150
Quadro 11 – Colocações do CRAEF 1 (Tema 05)	152
Quadro 12 – Colocações do CRAEF 2 (Tema 05)	154
Quadro 13 – Campo semântico das colocações do Tema 01	160
Quadro 14 – Campo semântico das colocações do Tema 02	162
Quadro 15 – Campo semântico das colocações do Tema 03	164
Quadro 16 – Campo semântico das colocações do Tema 04	165
Quadro 17 – Campo semântico das colocações do Tema 05	167

LISTA DE TABELAS

Tabela 01 – Redações do CRAEF 1	86
Tabela 02 – Rol de Redações coletadas (CRAEF 1)	88
Tabela 03 – Redações do CRAEF 2	89
Tabela 04 – Rol de Redações coletadas (CRAEF 2)	90
Tabela 05 – Particularidades das redações	91
Tabela 06 – Quantidade de textos não coletados	92
Tabela 07 – Somatória do CRAEF 1 e 2	93
Tabela 08 – Comparação do CRAEF 1 e 2 (1º Tema)	96
Tabela 09 – <i>KeyWords</i> geradas pelo software na análise do 1º Tema – CRAEF 1 e 2 ..	97
Tabela 10 – <i>KeyWords</i> das Amostras do 1º Tema de redação	98
Tabela 11 – Comparação do CRAEF 1 e 2 (Tema 02)	102
Tabela 12 – <i>KeyWords</i> geradas pelo software na análise do 2º Tema – CRAEF 1 e 2 ..	103
Tabela 13 – <i>KeyWords</i> das Amostras do 2º Tema de redação	103
Tabela 14 – <i>Tokens, types e type-token ratio</i> das amostras do Tema 02	106
Tabela 15 – Comparação do CRAEF 1 e 2 (Tema 03)	107
Tabela 16 – <i>KeyWords</i> geradas pelo software na análise do 3º Tema – CRAEF 1 e 2 ..	108
Tabela 17 – <i>KeyWords</i> das Amostras do 3º Tema de redação	108
Tabela 18 – <i>Tokens, types e type-token ratio</i> das amostras do Tema 03	112
Tabela 19 – Comparação do CRAEF 1 e 2 (Tema 04)	114
Tabela 20 – <i>KeyWords</i> geradas pelo software na análise do 4º Tema – CRAEF 1 e 2 ..	115
Tabela 21 – <i>KeyWords</i> das Amostras do 4º Tema de redação	116
Tabela 22 – <i>Tokens, types e type-token ratio</i> das amostras do Tema 04	119
Tabela 23 – Comparação do CRAEF 1 e 2 (Tema 05)	121
Tabela 24 – <i>KeyWords</i> geradas pelo software na análise do 5º Tema – CRAEF 1 e 2 ..	122
Tabela 25 – <i>KeyWords</i> das Amostras do 5º Tema de redação	122
Tabela 26 – <i>Tokens, types e type-token ratio</i> das amostras do Tema 05	124

LISTA DE SIGLAS E ABREVIATURAS

AMD	Aprendizagem Movida por Dados
BNC	British National Corpus
CEALE	Centro de Alfabetização, Leitura e Escrita
CEFAM	Centro Específico de Formação e Aperfeiçoamento do Magistério
CHILDES	Child Language Data Exchange System
CRAEF	Corpora de Redações de Aprendizizes do Ensino Fundamental
CRAEF 1	Corpora de Redações de Aprendizizes do Ensino Fundamental – 1
CRAEF 2	Corpora de Redações de Aprendizizes do Ensino Fundamental – 2
DDL	Data Driven Learning
EAP	English for Academic Purposes
FLIV	Festival Literário de Votuporanga
IDEB	Índice de Desenvolvimento da Educação Básica
INAF	Indicador de Alfabetismo Funcional
L1	Primeira língua
L2	Segunda língua
LE	Língua Estrangeira
LC	Linguística de Corpus
MEC	Ministério da Educação
OHE	Observar, “Hipotetizar”, Experimentar
PCNs	Parâmetros Curriculares Nacionais
PNL	Plano Nacional de Educação
SME	Secretaria Municipal da Educação de Votuporanga
TTR	Type-token ratio
UFs	Unidades Fraseológicas

SUMÁRIO

INTRODUÇÃO	17
1. FUNDAMENTAÇÃO TEÓRICA	22
1.1 O uso da língua autêntica à luz da Linguística de Corpus.....	23
1.1.1 Linguística de Corpus direcionada ao ensino	27
1.1.2 Contextualizando o uso de Corpora de Aprendizes em pesquisas que contemplam o universo infantil	35
1.2 O estudo do léxico no ensino de Língua Portuguesa do Brasil nas séries iniciais	39
1.3 Convencionalidade e Fraseologia	45
1.4 Fraseologismos: as Colocações	58
1.5 Aprendizagem Movida por dados para a elaboração de atividades baseadas em corpora	68
2. METODOLOGIA.....	73
2.1 Caracterização dos Corpora de Redações de Aprendizes do Ensino Fundamental (CRAEF 1 e 2).....	73
2.2 Caracterização das escolas A e B (Centro e Periferia)	76
2.3 Procedimentos para coleta dos dados e temáticas para a construção dos corpora de aprendizes	79
2.4 Tratamento de informação adotado nos corpora de aprendizes	82
3. ANÁLISE DOS RESULTADOS.....	85
3.1 Análise de dados do CRAEF	86
3.2 O léxico infantil presente no CRAEF 1 e 2	94
3.3 Análise fraseológica do CRAEF 1 e 2	125

4. PROPOSTAS DE ATIVIDADES PARA O DESENVOLVIMENTO DA COMPETÊNCIA COLOCACIONAL DOS ALUNOS DO ENSINO FUNDAMENTAL I	157
4.1 Procedimentos metodológicos para a seleção das colocações presentes nas atividades	159
4.2 Atividades baseadas em corpora para o Desenvolvimento da Competência Colocacional dos alunos do Ensino Fundamental I	168
5. CONSIDERAÇÕES FINAIS	181
REFERÊNCIAS BIBLIOGRÁFICAS	187
ANEXO	198
Autorização dos pais	

INTRODUÇÃO

As mudanças ocorridas recentemente com o acréscimo de um ano escolar a mais no início do ensino fundamental, efetuadas a partir de 2006 no Brasil, pedem um olhar aguçado sobre os métodos utilizados e reflexões sobre o ensino de língua materna nas primeiras séries do ensino fundamental I, período este que compreende do 1º ao 5º ano.

Os alunos outrora alfabetizados tinham em torno de sete anos de idade completos, ao contrário dos estudantes que agora frequentam o primeiro ano atual, que tem apenas cinco ou seis anos. A lei nº 11.114, de 16 de maio de 2005 altera os arts. 6º, 30 (VETADO), 32 e 87 da Lei no 9.394, de 20 de dezembro de 1996, com o objetivo de tornar obrigatório o início do ensino fundamental aos seis anos de idade.

Conforme descreve o Plano Nacional de Educação, PNE, a determinação legal (Lei nº 10.172/2001, meta 2 do Ensino Fundamental) de implantar progressivamente o Ensino Fundamental de nove anos, pela inclusão das crianças de seis anos de idade, tem duas intenções: “oferecer maiores oportunidades de aprendizagem no período da escolarização obrigatória e assegurar que, ingressando mais cedo no sistema de ensino, as crianças prossigam nos estudos, alcançando maior nível de escolaridade”.

Por isso, a passagem da educação infantil para o ensino fundamental deve ser feita de maneira natural, a fim de evitar traumas em sua infância e na sua singularidade, caso contrário o processo de escolarização será prejudicado.

Em qualquer um destes segmentos de ensino, os docentes precisam desenvolver em suas salas de aula um ambiente favorável à aprendizagem, seja por meio da oralidade, da escrita ou ainda de situações que façam o discente se apropriar da linguagem, compreendendo-a como meio de representar seu cotidiano.

Sabemos que a convencionalidade linguística é adquirida desde a infância, a partir da interação da criança com os indivíduos a sua volta. Assim, no processo de compreender o funcionamento da língua materna, espera-se que os alunos entendam como se articulam os blocos de palavras. Estes são pronunciados em determinadas circunstâncias, sob certos contextos específicos de comunicação.

Quando uma criança diz “Gosto de soltar pipa”, está, sem perceber, expressando uma colocação verbal, ou seja, fazendo relações entre o ato de brincar e a expressão “soltar pipa”, que não pode se realizar com outro sintagma qualquer, como “soltar carrinho”, “soltar

videogame”, dentre outras construções. Saber utilizar colocações verbais, com propriedade, por exemplo, é um dos indicativos de que uma pessoa compreende o funcionamento do seu idioma.

Neste sentido, expressões idiomáticas, colocações, fórmulas de rotina, frases feitas são alguns dos fraseologismos existentes em uma língua, ou seja, construções linguísticas convencionais caracterizadas pelo uso por uma dada comunidade linguística. Dessa forma, ao entender como se articulam os fraseologismos, o falante denota conhecer e manipular o código de seu idioma com propriedade.

Da mesma maneira que a linguagem do estudante brasileiro se modificou com as novas tecnologias, o docente do novo milênio também deve se adequar à era do conhecimento digital. Uma alternativa no trato linguístico das produções textuais dos discentes é o uso da Linguística de Corpus (doravante LC) no estudo do léxico.

A LC se destaca na Linguística Aplicada por analisar a linguagem a partir da compilação de dados reais de comunicação. Aliadas à informática e amparadas por *softwares* de análise lexical, são realizadas pesquisas quantitativas e qualitativas relevantes para a comunidade acadêmica, que também aplica suas descobertas ao público em geral.

Em conjunto com as ferramentas computacionais disponíveis atualmente, é possível observar, por exemplo, padrões lexicais, escolha de determinadas construções sintáticas, identificar estilos, entre outros aspectos relevantes, propiciando novas análises sobre a escrita dos discentes.

A análise de composições textuais infantis por meio de *softwares* informatizados pode, por sua vez, nos dizer muito sobre o conhecimento lexical dos alunos. A LC propicia as condições teórico-metodológicas ideais para investigarmos diversas ocorrências linguísticas, além de possibilitar a análise de um grande número de dados com maior propriedade científica, dados estes que podem dar origem a um corpus eletrônico.

Com base no exposto, o objetivo geral e enfoque desta pesquisa é analisar o léxico geral e fraseológico de duzentas e vinte e quatro redações de estudantes do terceiro ano do Ensino Fundamental, de duas salas de aula de escolas da cidade de Votuporanga (noroeste paulista). Uma delas se localiza no Centro da cidade (Escola A), enquanto a outra pertence à área periférica (Escola B) do município.

Esse material compõe os Corpora de Redações de Aprendizizes do Ensino Fundamental (CRAEF), compilados para fins de estudo léxico-fraseológico e coletados em cinco momentos diferentes do ano de 2015. Cada redação solicitou propostas diferentes e, com a variação da temática, observamos itens lexicais diversos, denotando mudanças na expressão escrita dos

discentes, conforme alterávamos o assunto a ser redigido.

Além disso, esta investigação tem como propósito traçar o perfil das redações de uma escola do centro e uma da periferia, observando as informações no que concerne à variedade lexical e ao uso das colocações presentes nos textos. Nesse sentido, tem como objetivos específicos:

- a) Compilar dois corpora de aprendizes, um proveniente de alunos que estudam em uma unidade escolar próxima ao centro da cidade, denominado CRAEF 1 e outro coletado em uma escola localizada na área periférica do mesmo município, chamado de CRAEF 2.
- b) Analisar o léxico presente nas cinco propostas de textos sugeridas do CRAEF 1 e 2.
- c) Extrair e analisar as colocações existentes nos corpora mencionados.
- d) Elaborar atividades que propiciem o desenvolvimento da competência colocacional destes estudantes do Ensino Fundamental 1.

Nesta pesquisa, entendemos por competência colocacional, a aptidão da criança em combinar palavras do seu idioma materno na enunciação de discursos. Este entendimento perpassa a compreensão de como a comunidade linguística realiza as combinações lexicais, já convencionalizadas e rotineiramente empregadas pela comunidade falante português do Brasil.

Considerando que a autora desta dissertação também é a professora responsável pelas duas salas de aula mencionadas, propomos duas perguntas de pesquisa iniciais, a fim de investigar como as crianças provenientes de locais sociais distintos redigem e de que maneira escrevem os fraseologismos que expressam na oralidade, a saber:

- a) Há variação lexical nas redações escritas pelos alunos da escola do centro e da periferia? Se sim, de que forma ela ocorre?
- b) Embora sejam sugeridos os mesmos temas para a produção escrita, como cada grupo (centro e periferia) explora estes temas, no que diz respeito aos fraseologismos empregados, especificamente às colocações?

Contudo, após as realizações das análises, observamos que o número de fraseologismos apresentados foi abaixo do esperado, pois diversos estudantes não conseguiram transmitir na escrita as colocações que utilizam na fala. O número de 189 colocações redigidas pelos alunos das duas escolas ao longo do ano de 2015 nos levou a refletir sobre os motivos pelos quais o aluno, seja estudando no centro da cidade ou na

periferia do município, possui dificuldades em redigir tais construções linguísticas.

Como professora de uma escola pública, compartilhamos essas inquietações de pesquisa com a professora orientadora, Doutora Adriane Orenha-Ottaiano. Juntas, decidimos incluir mais um capítulo à dissertação: propostas de atividades que possam melhorar a fluência das crianças, no que tange aos fraseologismos, mais especificamente às colocações. O intuito é possibilitar que os estudantes se tornem mais conscientes das possibilidades de combinações lexicais na língua portuguesa do Brasil e, assim, adquirir a uma fluência linguística mais ampla na escrita de redações. Outro propósito é auxiliar as crianças a entender de que forma as palavras se articulam entre si, e têm seu uso consagrado de acordo com a convenção da comunidade de falantes do seu idioma.

Dessa maneira, os resultados desta investigação buscam oferecer contribuições para a disseminação da linha de pesquisa “Pedagogia do Léxico e da Tradução Baseada em Corpora”. O conhecimento produzido nesta dissertação poderá auxiliar a preencher as lacunas existentes neste rol de pesquisas acadêmicas, concomitantemente em que traz à discussão aspectos relacionados à realidade educacional brasileira, vivenciada por inúmeros docentes do Ensino Fundamental I. A realização desta proposta servirá, por fim, de incentivo para possíveis proposições de estudos e futuros trabalhos nesta área em expansão no Brasil, aproximando teoria e prática docente.

Infelizmente, são poucas as faculdades e universidades brasileiras que possuem disciplinas voltadas ao estudo de corpora na grade curricular de seus programas. Por esse motivo, acreditamos que a presente pesquisa seja pertinente, devido a escassez de estudos sobre a variação lexical baseada em corpus de aprendizes formado por redações do Ensino Fundamental I no Brasil. A maioria dos trabalhos acadêmicos desta faixa etária ora pautam-se na descrição de desvios ortográficos e gramaticais dos estudantes (LUFT, 1994); ora focalizam gêneros textuais específicos (DOLZ; NOVERRAZ; SCHNEUWLY, 2004) ou incluem, ainda, descrições dos aspectos fonológicos e fonéticos (KATO, 2013).

Outras lacunas existentes nessa área se referem à existência de poucos estudos sobre a variação lexical e a produção de fraseologismos em textos escritos por crianças do Ensino Fundamental I no panorama nacional; ao baixo número de redações coletadas de estudantes do Noroeste Paulista, principalmente na faixa etária de 7 a 9 anos, enquanto estão consolidando a aprendizagem da língua materna.

Portanto, para atingir os propósitos descritos acima, o percurso teórico desta pesquisa faz revisões de pesquisas que utilizam corpora em seu escopo de investigações e de contribuições daquelas que se pautam na análise de Corpora de Aprendizes. Ademais, versa

sobre o ensino do léxico em língua materna, bem como os subsídios oferecidos pela Fraseologia no ensino dos fraseologismos e especificamente das colocações, assim como trata das vantagens da “Aprendizagem movida por dados” na elaboração de atividades baseada em corpus.

Metodologicamente, versará sobre a descrição dos dois grupos analisados, assim como a composição dos corpora de estudo aqui investigados. Na realização das análises, utiliza o *software WordSmith Tools* (SCOTT, 2012), versão 6.0, a fim de manusear os dados e verificar os itens descritos com maior precisão.

No que se refere às atividades didáticas descritas neste trabalho, estas irão compor a página do *Online English Collocations Workbook*¹ em português. Em projetos posteriores, uma das propostas é avaliar de que forma estes exercícios podem ser aplicados ao público-alvo destinado e quais os resultados observados provenientes da interação dos alunos com este material de ensino.

Em seguida, as considerações finais tratam do que foi exposto ao longo da dissertação e descrevem possíveis caminhos a serem ainda trilhados na pesquisa, já que é impossível discorrer sobre todas as nuances de um tema em um único trabalho investigativo. Finalmente, as referências bibliográficas encerram a escrita, elencando os autores citados no presente texto.

¹ Trata-se de um material de apoio *on-line*, que visa a auxiliar os aprendizes a compreender as colocações e seu uso em atividades práticas e lúdicas, como atestam os jogos da memória (*Memory Game*) e as atividades de preencher lacunas (*Gap Fill*), sob a responsabilidade da Profa. Dra. Adriane Orenha Ottaiano. Disponível em: <<http://www.gbd.institucional.ibilce.unesp.br/workbook/index>>. Acesso em 30 nov. 2017

5. CONSIDERAÇÕES FINAIS

A partir da nossa experiência enquanto aluna do magistério normal do Ensino Médio (CEFAM), podemos perceber que a formação para o exercício da profissão educativa deve ser ampla, completa, bem como abarcar diversas vertentes da alfabetização para que esta seja uma prática eficaz na aprendizagem do aluno.

Como docente lecionando na educação básica, também observamos essa dificuldade em conversas informais com colegas de profissão, pelo desconhecimento e despreparo que algumas têm para preparar aulas de língua portuguesa ou, ainda, discutir alternativas para auxiliar crianças com dificuldades de redigir produções textuais.

Afinal, se as dificuldades de expressão oral e escrita não forem sanadas nos primeiros anos do ensino fundamental I, a probabilidade dessas dúvidas se perpetuarem durante sua jornada educacional (e fora dela, na idade adulta) é muito alta.

As consequências desta triste realidade já são bem conhecidas por nós brasileiros, haja vista que estampam várias revistas e órgãos governamentais especializados em educação: o analfabetismo funcional, uma triste realidade existente nos bancos escolares públicos de norte a sul do país.

Os dados do Indicador de Alfabetismo Funcional, INAF, (2011) mostram que:

“(...) o Brasil já avançou principalmente nos níveis iniciais da alfabetização, mas não conseguiu progressos visíveis no alcance de níveis mais altos, que são hoje condição para inserção plena na cultura letrada e direito de todos os cidadãos, a quem se assegura o acesso pelo menos ao ensino fundamental completo.”

Para atingir esse objetivo, é preciso refletir sobre o que torna um nativo de língua portuguesa apto a atingir os níveis mais altos da alfabetização, bem como quais as habilidades linguísticas necessárias a serem trabalhadas pelos docentes das séries iniciais de escolarização.

Esta necessidade ainda é falha no ensino do idioma materno, conforme indica Tagnin (2013), ao apontar as lacunas existentes para o estudo do léxico em nosso idioma: “Cumprer ressaltar, no entanto, que na maior parte das vezes a bibliografia indicada será para a língua inglesa, já que muito pouco se tem feito na área da convencionalidade no âmbito do português.” (p. 15).

Dessa forma, nesta dissertação, buscamos relatar uma faceta das possibilidades e

relevância da utilização de corpora de aprendizes para investigações sobre o léxico infantil. Tendo em vista que traz à comunidade de investigadores o conhecimento sobre pesquisas em LC, sobretudo a respeito de investigações sobre corpora de aprendizes, este trabalho teve o propósito de demonstrar mais uma perspectiva científica desta vertente linguística, em termos de investigação acadêmica e pedagógica que possam preencher as lacunas observadas no ensino do português do Brasil.

É sabido que o processo de aprendizagem não é arbitrário e nem inerente ao ser humano, pois envolve diversos aspectos do indivíduo: emocionais, sociais, culturais, econômicos, dentre outros. Por isso, reiteramos a nossa justificativa em atuar com alunos em processo de aprendizagem da escrita, haja vista que a etapa inicial da escolarização é de suma importância na vida dos pequenos discentes.

Dessa forma, a sala de aula torna-se um ambiente propício para o desenvolvimento do léxico da criança, uma vez que as constantes interações sociais que a escola proporciona ajudam o discente a entender as interpretações acerca de si mesmo, sua família, a sociedade e o mundo como um todo. E os Corpora de Redações de Aprendizes do Ensino Fundamental (formado por CRAEF 1 e CRAEF 2) podem ser um meio de evidenciar e comprovar estes aspectos, a serem compartilhados com outros professores do Ensino Fundamental, bem como pesquisadores da área.

Dessa maneira, o percurso desta pesquisa iniciou-se com a introdução do trabalho, assim como a apresentação dos objetivos e as perguntas de pesquisa, que versavam sobre a possibilidade de haver variação lexical entre as redações dos dois grupos de estudo, centro e periferia, além do questionamento acerca de como cada turma empregou as colocações durante a escrita das cinco propostas temáticas.

Em seguida, realizamos a revisão das principais teorias que respaldaram a dissertação, a partir da área mais geral, a LC voltada ao ensino, perpassando pelas investigações sobre corpora de aprendizes infantil e o estudo do léxico de língua materna. Em seguida, tratamos dos assuntos mais específicos, como a convencionalidade do signo linguístico, culminando na fraseologia e na descrição do fraseologismo focado nesta dissertação, as colocações. Devido à necessidade de incluir mais um objetivo a este trabalho, também acrescentamos outro tópico conceitual, a respeito do uso da AMD, como embasamento teórico para a preparação de atividades baseadas em corpus.

A próxima etapa efetuada neste trabalho foi a descrição metodológica das fases de coleta e sistematização dos dados. Dessa maneira, o segundo capítulo iniciou-se com o relato sobre as especificidades dos corpora que compõem o CRAEF (Corpora de Redações de

Aprendizes do Ensino Fundamental), além das características da escola e do bairro no qual se encontram os estudantes do CRAEF 1 e CRAEF 2. Posteriormente, elencamos os processos de agrupamento das redações, bem como o tratamento realizado para transpor as composições escritas em informações digitalizadas.

Na sequência, tratamos de descrever em detalhes as razões pelas quais não foi possível coletar as trezentas redações que havíamos previsto antes da realização deste trabalho. Infelizmente, fatores que fogem de nossa alçada de pesquisadores, como o excesso de faltas, transferência de alunos, recebimento de estudantes durante o ano letivo e ainda a presença de estudantes não alfabetizados prejudicaram no total de textos coletados e efetivamente analisados neste estudo.

Após a contagem final de todas as proposições, o CRAEF está composto por um total de duzentos e vinte e quatro textos, contabilizando 20.435 palavras. Está é a somatória do CRAEF 1, que possui cento e vinte e três produções digitadas, num montante de 12.436 palavras; acrescidas do CRAEF 2, formado por cento e uma composições, com 7.999 palavras, ao todo.

De posse destes dados, realizamos as análises por meio do programa *WordSmith Tools*, observando as palavras-chave e o léxico disposto nas redações. Lexicalmente, o bairro em que vivem (centro ou periferia) influenciou parcialmente a escrita, posto que somente nos temas 02 e 04 pudemos diferenciar a visão as crianças possuem da cidade de Votuporanga, assim como as relações sociais retratadas no tema Solidariedade.

Os itens lexicais aqui analisados apresentaram a maneira pela qual as crianças compreendem o mundo, seja por meio da descrição de um filme infantil, relato sobre animais que conhecem ou gostariam de conhecer, seja pelas impressões que possuem da cidade onde vivem. Outras propostas tratadas nesta dissertação versaram sobre a imaginação em pensar o que brotaria de uma semente mágica, percepções do que pode acontecer se um personagem folclórico estivesse entre nós e, ainda, como descrever uma criança internada no hospital ou alguém que precise de doações.

Por oportunizar a autonomia linguística, a exposição do ponto de vista sobre diversos assuntos, a leitura reflexiva e escrita autônoma conferem ao estudante subsídios para debater ética, cidadania, dentre outros temas transversais. Dessa forma, é no confronto da sua linguagem com a de outras pessoas que a criança solidifica seu conhecimento sobre a língua, comparando os diferentes modos de expressão e formulando um linguajar mais elaborado. Mesmo que esses processos sejam realizados a nível inconsciente, é inegável o amadurecimento em seu progresso cognitivo, garantindo um letramento mais profícuo aos

estudantes.

Tais aprendizagens, são vistas e compreendidas pelas crianças a partir das lentes da realidade em que vivenciam. Os alunos, por sua vez, emolduram-se a estas experiências, enraizadas na casa em que moram, na família e comunidade às quais pertencem, enfim.

No que diz respeito às colocações, ambas as salas as apresentaram, tanto combinatórias mais simples, oriundas da linguagem infantil, quanto colocações mais elaboradas nos dois corpora: “entrar em greve” e “guarda temporária da criança”, por exemplo, apareceram respectivamente no CRAEF 1 e CRAEF 2. Mostramos, dessa maneira, a influência que os alunos recebem da família, ou seja, do meio externo, na formação dos padrões léxicos gramaticais de sua expressão linguísticas.

A análise total deste item resultou na produção de cento e oitenta e nove colocações no CRAEF, sendo cento e uma encontradas no CRAEF 1 e oitenta e oito combinatórias no CRAEF 2. Ainda na esfera comparativa, por apresentar menos composições, os alunos da periferia escreveram mais colocações do que os alunos do centro somente em uma das cinco propostas de produção textual.

Deixamos claro que os resultados apresentados neste texto não tiveram como objetivo fazer afirmações com teor valorativo em relação às escritas dos dois grupos de alunos. Primeiramente, independente de pertencer a Escola 1 ou 2, os indivíduos em estudo são seres humanos em formação, com personalidades diferentes e histórias de vida distintas uns dos outros. Isso significa que o professor jamais pode “rotular” uma sala ou um grupo escolar fundamentado na condição familiar, econômica ou de qualquer outra espécie discriminatória de seus pupilos.

Ademais, fatores como força de vontade, disciplina, interesse em aprender e determinação coexistem em todos os meios sociais, alterando o que pode ser chamado de padrão ou norma. Por isso, o docente deve atuar de maneira a dinamizar as potencialidades de seus educandos, incentivando-os a se envolverem no processo de ensino e aprendizagem.

Outra situação digna de nota encontrada nesta investigação é a escrita por parte das crianças de colocações não convencionalizadas pelos falantes da língua portuguesa do Brasil, como “múmia sem cabeça”, para se referir ao ser folclórico ou “erguer pipa” para explicar a tradicional brincadeira infantil. Neste sentido, apontamos aqui novamente o desconhecimento dos alunos de ambas as turmas frente a algumas combinatórias de palavras, que inibe uma melhor articulação nas ideias que as crianças tentam transmitir por meio da escrita.

Por isso, justificamos a presença do quarto capítulo no presente trabalho, visto que possibilita reflexões sobre o ensino de colocações no Ensino Fundamental I, além de

apresentar propostas para que este trabalho seja realizado na prática pelos professores deste segmento.

Fazemos tal afirmação, pois acreditamos que diferentes métodos de tarefas aumentem a probabilidade de uma maior aceitação dos alunos com o conteúdo colocacional, posto que atividades mais difíceis são motivadoras para o grupo. Ao identificar e compreender as propostas mais fáceis, é necessário elevar o grau de dificuldade inicial, principalmente para motivar aqueles discentes que têm condições de aprender novas combinatórias de palavras. Conseqüentemente, exigindo um grau maior de interpretação e reflexão das colocações, oferece-se condições linguísticas ao aluno para ampliar a capacidade de se expressar na língua portuguesa do Brasil.

Por todo o exposto, alicerçado em um estudo voltado à análise de produções escritas de aprendizes, é possível desenvolver trabalhos interdisciplinares de profundidade crítica e que proporcione reflexões significativas para os estudantes. Além disso, destacamos a riqueza deste trabalho em ampliar a proposta inicial desta pesquisa, com base no baixo número de colocações observada nas redações. A partir da reflexão acerca desta constatação, efetuamos neste estudo, uma alternativa pedagógica e didática para que os professores possam auxiliar seus discentes a transpassarem as dificuldades que por ventura possuam sobre as hipóteses de escrita e o entendimento dos fraseologismos.

Complementando a ideia iniciada no parágrafo anterior, apontamos ainda a necessidade de refletir sobre a formação docente no ensino da língua materna. Se queremos que o aluno compreenda a importância dos fraseologismos para expressar o seu cotidiano, também é necessário orientar este entendimento aos professores das séries iniciais, para que propiciem condições aos seus educandos de refletirem sobre a maneira de articular a linguagem. Infelizmente, os conhecimentos proporcionados pela LC, corpora de aprendizes e fraseologia ainda não estão disseminados o suficiente nos cursos de licenciatura de Letras e Pedagogia, o que dificulta na aplicação de seus postulados em aulas de língua materna no Brasil.

Pelos exemplos de atuações já realizadas no âmbito de trabalhos acadêmicos com redações de crianças já apontados nesta dissertação, acreditamos que muito já foi feito na disseminação dos estudos e pesquisas com corpora no Brasil. Contudo, também defendemos a ênfase na divulgação dos resultados e investigações de nossos colegas do Estudo do Léxico para um aprofundamento das reflexões e discussões desta importante área em expansão.

Por fim, como encaminhamentos futuros no projeto de doutorado, pretendemos analisar as palavras-chaves mais significativas no Corpus do Português e, em seguida, efetuar

a seleção das mesmas pelo *Sketch Engine*. Em uma etapa posterior, acrescentar outras colocações aos exercícios futuros que serão incluídos na “Plataforma de Ensino de Colocações em Português” (versão do português do Brasil para o *Online English Collocations Workbook*) e efetivamente aplicá-los a estudantes do ensino fundamental I de diferentes escolas, analisando os resultados sobre desempenho dos discentes, frente às propostas realizadas neste texto.

REFERÊNCIAS

BALLY, C. *Traité de stylistique française*. Paris: Klincksieck, 1951. v. 2.

BASÍLIO, M. *Formação e classes de palavras no português do Brasil*. São Paulo: Contexto, 2004.

BERBER SARDINHA, T. Usando WordSmith Tools na investigação da linguagem. *DIRECT Papers 40*. São Paulo: LAEL, PUC/SP, 1999. Disponível em: <http://www2.lael.pucsp.br/direct/DirectPapers40.pdf/>. Acesso em: 4 jan. 2017.

BERNARDINI, S. Corpora in the classroom. In: SINCLAIR, J. M. (Ed.). *How to use corpora in language teaching*. Amsterdam: John Benjamins, 2004. p. 15-36.

BIBER, D.; CONRAD, S.; REPPEN, R. *Corpus linguistics: investigating language structure and use*. Cambridge: Cambridge University Press, 1998.

BOULTON, A. Data-driven learning: The perpetual enigma. In: GOŹDŹ-ROSZKOWSKI, S. (ed.). *Explorations across Languages and Corpora*. Frankfurt: Peter Lang, 2011, p. 563-580.

BOWKER, L.; PEARSON, J. *Working with specialized language: a practical guide to using corpus*. London: Routledge, 2002.

BRASIL. Lei n. 10.172, de 09 de janeiro de 2001. Aprova o Plano Nacional de Educação (PNE). *Diário Oficial da União [da] República Federativa do Brasil*, Brasília, DF, 10 jan. 2001. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/leis_2001/110172.htm Acesso em: 20 dez. 2017.

_____. Lei nº 11.114, de 16 de maio de 2005. Altera os artigos 6, 30, 32 e 87 da Lei nº 9.394, de 20 de Dezembro de 1996, com o objetivo de tornar obrigatório o início do ensino fundamental aos 6 (seis) anos de idade. *Diário Oficial da União*, Brasília, DF, 17 maio 2005. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2005/lei/111114.htm. Acesso em: 20 dez. 2017.

_____. Ministério da Educação. *Parâmetros Curriculares Nacionais*. Brasília: MEC, 2000. Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/livro01.pdf> Acesso em: dez. 2017.

BRITISH NATIONAL CORPUS. *BNC*. Oxford: Oxford University, 2015. Disponível em: <<http://www.natcorp.ox.ac.uk/>>. Acesso em: 30 abr. 2017.

CALDAS, A. D. D. R. *A identificação de colocações especializadas extraídas do corpus CSI e do corpus comparável criminal para a elaboração de atividades didáticas*. Dissertação. (Mestrado em Estudos Linguísticos). São José do Rio Preto: UNESP/IBILCE, 2017.

CHILDES. *Child Language Data Exchange System*. [S. l.], [s. d.]. Disponível em: <<http://childes.talkbank.org/>>. Acesso em: 13 abr. 2017.

COBB, T; BOULTON, A. Classroom applications of corpus analysis. In: BIBER, D.; REPPEN, R. (Ed.). *The Cambridge handbook of corpus linguistics*. Cambridge: Cambridge University Press, 2015. p. 478-497.

CORDER, P. S. *Error Analysis and Interlanguage*. Oxford: Oxford University Press, 1981.

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

COSTA, I. O.; MIRANDA, N. S. A construção superlativa de expressão corporal: uma análise baseada em corpora. In: X Encontro de Linguística de Corpus: Aspectos metodológicos dos estudos de corpora. DUTRA, D. P.; MELLO, H. R. (Org.). *Anais do X Encontro de Linguística de Corpus: aspectos metodológicos dos estudos de corpora*. Belo Horizonte: Faculdade de Letras da UFMG, 2012.

COSTA VAL, M. da G. *Da frase ao discurso*. Belo Horizonte: FALE/UFMG. (mimeo), 1997, p. 1-19.

CRISTÓFARO-SILVA, T. et al. *Corpus do e-Labore (Laboratório Eletrônico de Oralidade e Escrita)*. Belo Horizonte: Faculdade de Letras, Universidade Federal de Minas Gerais, 2009. Disponível em: <www.projetoaspa.org/elabore> Acesso em: 3 jan. 2017.

CRYSTAL, D. *Dicionário de lingüística e fonética*. Rio de Janeiro: Jorge Zahar Editor, 2000, trad. Maria Carmelita Pádua Dias, 1985.

DAVIES, M.; FERREIRA, M. J. (2006). *Corpus do Português: 1 billion words of data*. [S. l.], [s.d.]. Disponível em: <<http://www.corpusdoportugues.org/>>. Acesso em: 01 jan. 2018.

DEMUTH, K. Collecting spontaneous production data. In: MCDANIEL, D.; MCKEE, C.; CAIRNS, H. S. (Ed.). *Methods for assessing children's syntax*. p. 3-22. Cambridge, MA: MIT Press, 1996.

DIESSEL, H. Corpus linguistics and first language acquisition. In: LÜDELING, A.; KYTÖ, M. (Ed.). *Corpus linguistics*. Berlin: Mouton de Gruyter, 2009. p. 1197-1212.

DOLZ, J.; NOVERRAZ, M.; SCHNEUWLY, B. Sequências didáticas para o oral e a escrita: apresentação de um procedimento. In: _____. *Gêneros orais e escritos na escola*. Tradução de Roxane Rojo e Gláís Sales Cordeiro. p. 95-147. Campinas: Mercado de Letras, 2004.

ESPÍNDOLA, S. *Análise de um corpus de produção escrita em português por crianças e adultos indígenas bilíngues/monolíngues de Dourados/MS a partir da linguística de corpus*. 2014. 167 p. Tese (Doutorado em Linguística Aplicada) –Universidade Federal do Rio Grande do Sul, Porto Alegre, 2014.

FERREIRA, A. B. de H. *Novo dicionário da língua portuguesa*. Rio de Janeiro: Nova Fronteira, 2000.

FILLMORE, C. *Inocence: a second idealization for linguistics*. Proceedings of the Fifth Berkeley Linguistics Society, 1979.

FIRTH, J. R. *Papers in linguistics 1934-1951*. Londres: Oxford University Press, 1957.

FLIV. FESTIVAL Literário de Votuporanga (FLIV). Votuporanga, 2016. Disponível em: <<http://www.flivotuporanga.com.br/2016/fliv/>>. Acesso em: 13 nov. 2016.

FLOWERDEW, L. The exploitation of small learner corpora in EAP materials design. In: GHADESSY, M.; HENRY, A.; ROSEBERRY, R. (Ed.). *Small corpus studies and ELT: theory and practice*. Amsterdam: John Benjamins, 2001.

GRANGER, S. The computer learner corpus: a versatile new source of data for SLA research. In: _____. *Learner english on computer*. London: Longman, 1998. p. 3-17.

GRANGER S.; HUNG, J.; PETCH-TYSON, S. (Ed.). *Computer learner corpora: second language acquisition and foreign language teaching*. Amsterdam: John Benjamins, 2002.

GRANGER, S.; TRIBBLE, C. Learner corpus data in the foreign language classroom: form-focused instruction and data-driven learning. In: GRANGER, S. (Org.) *Learner english on computer*. New York: Longman, 1998.

GRIES, S. T. Cognitive linguistics and functional linguistics: common assumptions and methods. In: GRIES, S. T.; STEFANOWITSCH, A. (Ed.). *Corpora in cognitive linguistics: corpus-based approaches to syntax and lexis*. Berlin: Mouton de Gruyter, 2006.

HAUSMANN, F. J. O dicionário de colocacións: criterios de organización. In: *Actas... 1. COLOQUIO GALEGO DE FRASEOLOXÍA*. Santiago de Compostela: Centro Ramón Piñeiro, p. 63-81, 1997.

ILARI, R. O Estruturalismo lingüístico: alguns caminhos. In: MUSSALIN, F., BENTES, A. C. (Org.) *Introdução à linguística*. v. 3: Fundamentos epistemológicos. São Paulo: Cortez, 2004, p. 53-92.

INDICADOR de analfabetismo funcional. *INAF Brasil 2011: principais resultados*. São Paulo: Ação Educativa; Instituto Paulo Montenegro, 2011. Disponível em: <http://www.ipm.org.br/download/inf_resultados_inaf2011_ver_final_diagramado_2.pdf>. Acesso em: 30 mar. 2017.

JOHNS, T. Data-driven learning: the perpetual challenge. In: KETTEMANN, B.; MARKO, G. (Ed.). *Teaching and learning by doing corpus analysis*. Amsterdam: Rodopi, 1991.

_____. From printout to handout: grammar and vocabulary teaching in the context of data-driven learning. In: ODLIN, T. (Ed.). *Perspectives on pedagogical grammar*. Cambridge: Cambridge University Press, 1994. p. 293-313.

JORDÃO, S. et al. Violência no Imaginário da Criança. In: ZYNGIER, S.; VIANA, V; FAUSTO, F. (org.). *Venturas e Desventuras*. Rio de Janeiro: Faculdade de Letras da UFRJ, 2005, p. 165-172.

KATO, M. A. *No mundo da escrita: uma perspectiva psicolingüística*. 7. ed. São Paulo: Ática, 2003.

KENNEDY, G. Between and through: The company they keep and the functions they serve. In: AIJMER, K; ALTENBERG, B. (Eds.). *English Corpus Linguistic*. London: Longman, 1991.

KILGARRIFF, A. et al. The Sketch Engine. In WILLIAMS, G.; VESSIER, S. (eds.). *Proceedings of the 11th Euralex International Congress*. Lorient: Universite de Bretagne-Sud, 2004, p. 105-116.

KOESTER, A. Building small specialised corpora. In: KEEFFE, A. O.; MCCARTHY, M. (Ed). *The Routledge handbook of corpus linguistics*. London: Routledge, 2010.

LEECH, G. Preface: learner corpora: what they are and what can be done with them. In: GRANGER, S. *Learner English on computer*. London: Longman, 1998. p. xiv – xx.

LÉON, J. A Lingüística de Corpus: história, problemas, legitimidade. *Filologia e Linguística Portuguesa*. São Paulo, n. 8, 2006, p. 51-81.

LEWIS, M. *Implementing the lexical approach: putting theory into practice*. Hove: Language Teaching Publications, 1997.

_____. *Teaching collocation: further developments in the lexical approach*. Hove: Language Teaching Publications, 2000a.

_____. *The lexical approach: the state of ELT and a way forward*. London: Language Teaching Publications. 2000b.

LUFT, C. P. *Língua e liberdade*. 3. ed. São Paulo: Ática, 1994.

MARTELLI, A. *Lexical collocations in learner English: a corpus-based approach*. Alessandria: Edizioni dell'Orso, 2007.

MCENERY, T.; XIAO, R.; TONO, Y. *Corpus-based language studies: an advanced resource book*. London: Routledge, 2006.

MEYER, C. F. Corpus analysis and linguistic theory. In: _____. *English Corpus Linguistics – An Introduction*. Cambridge, 2004, p. 1-29.

MONTEIRO-PLANTIN, R. S. *Fraseologia: era uma vez um Patinho Feio no ensino de língua materna*. v. I, Fortaleza: Imprensa Universitária, 2014.

_____. Fraseologia: uma mão na roda na construção do sentido. In: *Synergies Tunisie*, Sylvains-les-Moulins, n. °3, p. 161-168, 2011.

_____. ReVEL na Escola: Fraseologia e Paremiologia: para que ensinar, se todo o mundo sabe? *ReVEL*, vol. 15, n. 29, 2017

NOIMANN, A. *Um olhar sobre os fraseologismos (locuções) em um dicionário bilíngue escolar espanhol-português / português-espanhol*. 2007. Dissertação (Mestrado em Estudos da Linguagem) – Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2007.

O'KEEFFE, A.; MCCARTHY, M. J. (Eds). *The Routledge handbook of corpus linguistics*. London: Routledge, 2010.

OLIMPIO DE OLIVEIRA SILVA, M. E. *Fraseografía teórica y práctica: bases para un diccionario de locuciones verbales español-portugués*. Tese de Doutorado, Universidad de Alcalá de Henares, 2004.

OLIVEIRA, E. C. F. de. *Aplicações da linguística de corpus em aulas de língua portuguesa no ensino fundamental*. Monografia. São José do Rio Preto: UNESP/IBILCE, 2014.

OLIVEIRA, E. C. F.; OTTAIANO, A. Harry Potter e a linguística de corpus em aulas de língua inglesa para o ensino fundamental II. *Desenredo: revista do programa de pós-graduação em letras da Universidade de Passo Fundo*, Passo Fundo, v. 13. n. 3. p. 609-627, set./dez. 2017. Disponível em: <<http://seer.upf.br/index.php/rd/article/view/7182>>. Acesso em: 24 dez. 2017.

OLIVEIRA, E. C. F.; OTTAIANO, A. O.; ALVES, V. C. O léxico nas traduções de história em quadrinhos: uma experiência baseada em corpus. *Entrepalavras*, Fortaleza, v. 7, p. 141-159, jan./jun. 2017. Disponível em: <<http://www.entrepalavras.ufc.br/revista/index.php/Revista/article/view/846>> Acesso em: 29 nov. 2017.

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), FFLCH/USP. São Paulo, 2004.

_____. Learning and teaching collocations through an online english collocations workbook. In: *Vocab@Tokyo*, 2016, Tóquio. *VOCAB@ TOKYO CONFERENCE Handbook*. p. 59-60.

_____. Compilação de um corpus de aprendizes de tradução e análise de aspectos colocacionais. In: *ABRALIN EM CENA*, 2012, Cuiabá. Anais, 2012.

_____. *Unidades fraseológicas especializadas: colocações e colocações estendidas em contratos sociais e estatutos sociais traduzidos no modo juramentado e não-juramentado*. Tese de doutorado em Linguística Aplicada. São José do Rio Preto: IBILCE/UNESP, 2009. 274 f.

_____. *Online English Collocations Workbook*. Disponível em:
<<http://www.gbd.institucional.ibilce.unesp.br/workbook/index>>. Acesso em: 30 nov. 2017.

ORENHA-OTTAIANO, A.; ROCHA, J. M. P. Ensino de inglês como LE e contribuições pedagógicas de um glossário bilíngue de colocações. *Signótica*, Goiânia, v. 27, 2015, p. 485-509.

PAWLEY, A.; SYDER, F. H. Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In: RICHARDS, Jack C.; SCHMIDT, Richard W. (Eds.). *Language and communication*. London: Longman, 1983. p.191- 225.

PITKOWSKI, E. F.; GAMARRA, J. V. El uso de los corpus lingüísticos como herramienta pedagógica para la enseñanza y aprendizaje de ELE. *TINKUY n°11*, Section d'études hispaniques, Mayo 2009, Université de Montréal, 2009.

RAGAN, P. H. Classroom use of a systemic functional small learner corpus. In: GHADESSY, A.; HENRY, A.; ROSEBERRY, R. L. (Ed.). *Small corpus studies and ELT: theory and practice*. Amsterdam: John Benjamins, 2001.

RODRÍGUEZ, M. Á. S. *Unidades fraseológicas francesas*. Estudio en un corpus: la Pentalogía de Belleville de Daniel Pennac. Planteamiento didáctico. Tese (Doutorado em Linguística). Universidade de Murcia, 2004.

RÖMER, U. Corpus Research Applications in Second language Teaching. *Annual Review of Applied Linguistics*, v. 31, march, 2011, p. 205-225.

RÖMER, U.; WULFF, S. Applying corpus methods to writing research: explorations of MICUSP. *Journal of Writing Research*, Antwerp, v. 2, n. 2, 2010, p. 99-127.

RUIZ GURILLO, L. Aspectos de fraseología teórica española. In: _____. *Cuadernos de filología*. Valencia: Universidad de Valencia, 1997, p. 1-125.

SAINT-EXUPÉRY, A. de. *O pequeno príncipe*. Rio de Janeiro: Agir, 1994.

SANDMANN, A. J. *Competência lexical*. São Paulo: Editora da UFPR, 1991.

SAUSSURE, F. *Curso de lingüística geral*. São Paulo: Cultrix, 1975.

SCOTT, M. *WordSmith Tools*: versão 6.0. Oxford: Oxford University Press, 2012.

SHEPHERD, T; ZYNGIER, S.; VIANA, V. Feixes lexicais e visões de mundo: um estudo sobre corpus. *Matraga: estudos linguísticos e literários*, Rio de Janeiro, n. 19, jul./dez. 2006, p. 125-140.

SILVA, M. B. Uma palavra só não basta: um estudo teórico sobre as unidades fraseológicas. In: *Revista de Letras*, Fortaleza, n. 28 v. 1/2 – jan./dez. 2006, p. 11-20.

SINCLAIR, J. *Corpus, concordance, collocation*. Oxford: Oxford University Press, 1991.

_____. Corpus evidence in language description. In: WICHMANN, A. et al. (Ed.). *Teaching and language corpora*. London: Longman, 1997, p. 27-39.

SOUZA, G. *Gay language: análise colocacional e a proposta de um glossário bilíngue de colocações baseado em corpus*. Dissertação (Mestrado em Estudos Linguísticos). São José do Rio Preto, Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, 2018.

SRIPICHARN, P. How can we prepare learners for using language corpora? In: O'KEEFFE, A.; MCCARTHY, M. (ed.). *The Routledge Handbook of Corpus Linguistics*. London/New York: Routledge, 2010, p. 371-384.

TAGNIN, S. E. O. A linguística de corpus na e para a tradução. In: VIANA, V.; TAGNIN, S. E. O. *Corpora na tradução*. São Paulo: Hub Editorial, 2015. p. 19-56.

_____. *Convencionalidade e produção de texto: um dicionário de colocações verbais inglês/português; português/inglês*. Tese (Livre-Docência) Universidade de São Paulo, São Paulo, 1998.

_____. *O. Expressões idiomáticas e convencionais*. São Paulo: Ática, 1989.

_____. *O jeito que a gente diz: expressões convencionais e idiomáticas*. São Paulo: Disal Editora, 2005.

_____. *O jeito que a gente diz: combinações consagradas em inglês e português*. Barueri: DISAL, 2013.

_____. Os corpora: instrumentos de auto-ajuda para o tradutor. *Cadernos de Tradução IX*, Tradução e Corpora. Florianópolis: UFSC/NUT, 2002, p. 191-219.

TEIXEIRA, L. F.; CHISHMAN R. L. de O. Aprendizado Movido por Dados para Iniciantes: O Uso de Corpus de Aprendizes para o Estudo da Ordem do Adjetivo em Língua Inglesa. *Revista Intercâmbio*, volume XVII, São Paulo: LAEL/PUC-SP, 2008, p. 246-264.

TEIXEIRA, L. G. *Colocações criativas presentes no corpus literário paralelo Memórias póstumas de Brás Cubas sob a perspectiva de um novo olhar*. Dissertação. (Mestrado em Estudos Linguísticos). São José do Rio Preto: UNESP/IBILCE, 2016.

TOMASELLO, M. *Constructing a language: a usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press, 2003.

TRASK, R. L. *Dicionário de linguagem e lingüística*. Tradução de Rodolfo Ilari. São Paulo: Contexto, 2004.

TREVILLE, M.C.; DUQUETTE, L. *Enseigner le vocabulaire en classe de langue*. Paris: Hachette, 1996.

TRIBBLE, C. *Concordancing and an EAP writing programme*. *CÆLL Journal*, Washington, v. 1, n. 2, p. 10-15, 1990.

VIANA, V. Linguística de corpus: conceitos, técnicas & análises. In: VIANA, V.; TAGNIN, S. E. O. (Org.). *Corpora no ensino de línguas estrangeiras*. São Paulo: Hub Editorial, 2010. p. 25-96.

WILLIS, D. *The lexical syllabus*. London: Collins, 1990.

WRAY, A. *Formulaic language and the lexicon*. Cambridge: Cambridge University Press, 2002.

ZAPPAROLI, Z. M. Um pouco da história da análise informatizada do léxico no Brasil. In: NUNES, J. H.; PETTER, M. (Org.). *História do Saber Lexical e Constituição de um Léxico Brasileiro*. São Paulo / Campinas: Humanitas / Pontes, 2002, p. 223-253.

REFERÊNCIAS DAS IMAGENS UTILIZADAS NAS ATIVIDADES BASEADAS EM CORPORA

BANHEIRO.

Disponível em: <<http://azcolorir.com/desenho/83283>> Acesso em: 9 fev. 2018.

BICICLETA.

Disponível em: <<http://veiculos.colorir.com/outros-veiculos/bicicleta-para-as-criancas.html>>. Acesso em: 7 jan. 2018.

CABELO RUIVO.

Disponível em: <<http://cavaleirodastrevasedaluz.blogspot.com.br/2016/01/snoopy-charlie-brown-peanuts-o-filme.html>>. Acesso em: 29 jan. 2018

CAFÉ PRETO.

Disponível em: <<https://il.wp.com/emagrecerjadicas.com.br/wp-content/uploads/2016/12/caf%C3%A9-preto-engorda-ou-ajuda-a-emagrecer.jpg>>. Acesso em: 29 jan. 2018

CEBOLA ROXA.

Disponível em: <<http://feirinhaexpress.com.br/produto/cebola-roxa-500-g/>>. Acesso em: 29 jan. 2018

CHICLETE TUTI FRUTTI.

Disponível em: <https://mariliaescobar.files.wordpress.com/2011/09/chiclete_de_bola.jpg>. Acesso em: 29 jan. 2018

COZINHA.

Disponível em: <<http://www.desenhosparacolorir.org/desenhos/desenhos.php?id=6129>> Acesso em: 9 fev. 2018.

GIRAFÁ.

Disponível em: <<https://super.abril.com.br/ciencia/existem-quatro-especies-de-girafa-e-nao-uma-so/>>. Acesso em: 29 jan. 2018.

HOMEM DAS CAVERNAS.

Disponível em: <http://vignette3.wikia.nocookie.net/powerlisting/images/3/3d/Grug_Crood.png/revision/latest?cb=20130330234146>. Acesso em: 29 jan. 2018.

LEÃO.

Disponível em: <<https://veja.abril.com.br/economia/aprenda-a-domar-o-leao/>>.

Acesso em: 29 jan. 2018.

PESSOA DOENTE.

Disponível em: <https://www.123rf.com/photo_14226213_sick-person-lying-in-bed.html>.

Acesso em: 29 jan. 2018.

PIPA.

Disponível em: <<https://professoraivaniferreira.blogspot.com.br/2014/03/projeto-ziraldomenino-maluquinho-o.html>>. Acesso em: 7 jan. 2018.

QUARTO.

Disponível em: <<http://www.colorir.blog.br/imagens/desenhos-para-colorir-colorir-quarto-1.gif>> Acesso em: 9 fev. 2018.

SINAL VERDE.

Disponível em:

<http://img.bhs4.com/b6/5/b659c14d47bff3f67a956b5343cd73c33b6c6fd4_large.jpg>.

Acesso em: 29 jan. 2018.

VIDEOGAME.

Disponível em: <<https://www.istockphoto.com/br/vetor/controle-de-video-game-desenho-gm165943066-22220642>>. Acesso em: 7 jan. 2018.

ZEBRA.

Disponível em: <<https://www.pexels.com/photo/zebra-fur-70376/>>. Acesso em: 29 jan. 2018

REFERÊNCIAS DOS SITES CONSULTADOS PARA A ELABORAÇÃO DE ATIVIDADES

GERADOR DE CAÇA PALAVRAS.

Disponível em: <<http://www.atividadeseducativas.com.br/index.php?id=6115>>. Acesso em: 8 jan. 2018.

PARA ELABORAR CRUZADINHAS.

Disponível em: <<https://www.educolorir.com/crosswordgenerator/por/>>. Acesso em: 8 jan. 2018.