



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

1. FUNDAMENTAÇÃO TEÓRICA

Nesta pesquisa, os estudos discutidos visam a corroborar para o entendimento de que trabalhos com foco na exploração de corpora para o ensino e aprendizagem de línguas mostram-se relevantes para os estudos linguísticos, em especial ao estudo do léxico-fraseológico infantil no português do Brasil.

Neste capítulo, tratamos do embasamento teórico da LC, abordando suas ideias centrais e suas peculiaridades, que a distinguem de outras abordagens. Para tanto, está contextualizada pelas pesquisas com o uso de corpora – notadamente no campo do ensino de língua portuguesa do Brasil, sobretudo aquelas cuja coleta de produções textuais norteiam as análises.

Continuando o percurso teórico, fazemos uma revisão da literatura sobre o léxico. Acreditamos que, dessa maneira, o leitor deste texto compreenderá os motivos pelos quais escolhemos investigar a língua portuguesa neste texto, voltada principalmente ao aluno de Ensino Fundamental I, educando das séries iniciais. Em seguida, abordamos a Fraseologia no escopo deste trabalho, em especial as colocações, itens lexicais que foram analisados nos dois corpora de estudo coletados nas salas de aula das escolas descritas.

Dessa forma, adentramos a temática da LC, descrevendo as principais características dos estudos com corpora, conforme os postulados de Costa e Miranda (2011), Léon (2006), Viana (2010) e demais autores. Posteriormente, o tópico sobre a LC voltada ao ensino contextualiza alguns trabalhos do grupo de pesquisa “Pedagogia do Léxico e da Tradução e Linguística de Corpus” relacionados, como atestam os trabalhos de Oliveira, Orenha-Ottaiano e Alves (2017), Teixeira (2016), dentre outros.

Em seguida, tratamos de estudos baseados em corpora de aprendizes, mostrados sob a perspectiva de Tribble (1990) e Granger e Tribble (1998); além da conjuntura de corpora de língua materna e infantil, pelo olhar de teóricos como Demuth (1996), Biber, Conrad e Reppen (1998) e Diessel (2009), por exemplo.

Seguindo a nossa trajetória teórica, relacionamos investigações sobre o léxico pela ótica de Ilari (2004), Costa Val (1997) e outros, bem como considerações sobre a competência lexical e colocacional, por meio dos pesquisadores Basílio (2004), Sandmann (1991), Olímpio de Oliveira Silva (2004). Também elencamos a Fraseologia, a partir das contribuições das autoras Orenha-Ottaiano (2004), Monteiro-Plantin (2014) e Ruiz Gurillo (1997) e, por último,

da temática das colocações em Orenha-Ottaiano (2004, 2009, 2015), Lewis (2000a), Tagnin (1998, 2013) e Martelli (2007); além do aporte da Abordagem Dirigida por Dados na elaboração de tarefas baseadas em corpus, com o subsídio de Kennedy (1991), Teixeira e Chishman (2008), Cobb e Boulton (2015) e Boulton (2011), conforme constam na sequência.

1.1 O uso da língua autêntica à luz da Linguística de Corpus

A LC surgiu da necessidade que os pesquisadores tinham em incluir o uso real da língua nos estudos linguísticos. Por esse motivo, analisa a língua por um viés empírico, assumindo padrões demonstrados por estruturas identificáveis nos dados coletados. Seu objeto de estudo, assim, é o corpus linguístico, ou seja, um conjunto de dados autêntico e capaz de atestar a representatividade da língua.

É de consenso geral que, a partir do advento das tecnologias digitais, diversos campos do conhecimento foram revolucionados, tanto pelas transformações estruturais em seus procedimentos metodológicos, quanto pela disseminação possibilitada pelos recursos da informática. Com a associação da informática e o uso de ferramentas eletrônicas, os elementos coletados possibilitam a pesquisa e o estudo sistemático das informações, possibilitando análises objetivas e mensuráveis de interpretação.

Reconsiderações sobre a linguagem ocorrem em inúmeros contextos do conhecimento humano. Ao tratar o lexema como unidade transitória e relacionadora múltipla de discursos, a ciência linguística amplia seus horizontes para abranger novas perspectivas do material semântico.

Seja apoiando educadores no processo de ensino-aprendizagem ou em parcerias com grandes empresas de tecnologia, a LC pode ser uma ferramenta de apoio a pesquisadores interessados nas áreas de Fraseologia e Lexicografia, bem como empresas que desenvolvem pesquisas no campo da linguagem natural.

Assim como as demais vertentes linguísticas, a LC também sofreu modificações, principalmente após o *boom* da Era Digital. Textos antes compilados manualmente, comprovados pelas infindáveis fichas catalográficas armazenadas em caixas, quartos e escritórios, cederam espaço aos arquivos informatizados, facilitando exponencialmente o arquivamento e manuseio dos dados.

Ao utilizar o computador na catalogação, identificação e análise dessas informações, a

nova vertente possibilitou infinitos campos de estudos e pesquisa de dados reais, oferecendo outros enfoques sobre a linguagem. Na área da Linguística Aplicada, a premissa não poderia ser diferente. Seja em Ensino e Aprendizagem de Línguas, Estudos da Tradução, Lexicologia e Lexicografia, Terminologia e Terminografia ou ainda Pedagogia do Léxico e da Tradução Baseada em Corpora, linha de pesquisa cuja dissertação está inserida, o foco das pesquisas nunca se encontra dissociadas umas das outras.

Zapparoli (2002, p. 223), defende a pesquisa com corpora ainda pela praticidade com a qual se é feita “a busca, organização e análise de dados linguísticos, tornando, por conseguinte, mais rápido e seguro o acesso ao material de análise”. Esta característica é fulcral neste segmento, pois, com a coleta de um grande número de informações de maneira manual, há o risco de perda deste material ou manipulação equivocada do mesmo.

Visto já existir uma bibliografia extensa de LC², esse texto tratará com maior ênfase das particularidades dos estudos com corpora, assim como as diversas aplicabilidades existentes deste arcabouço teórico, principalmente no que tange ao ensino e a análise de dados por meio da exploração de corpora de aprendizes.

Neste sentido, Costa e Miranda (2011) pontuam como:

vantajosa a utilização de corpora na investigação da linguagem, não só por oferecer acesso a informações inacessíveis à introspecção do pesquisador, mas também por permitir descrições mais precisas, e reais, de um determinado objeto, já que as informações emergem naturalmente dos dados. (p. 168)

Este aspecto é de extrema importância para os alunos, já que, pelo conhecimento empírico, é possível conhecer com mais propriedade a linguagem da sociedade, não só as contemporâneas, mas também a de qualquer população, a qual tenha-se acesso a amostras de fala ou de escrita, seja ela de tempos atuais ou então da antiguidade. Para Sinclair (1997, p.31), “a língua não pode ser inventada; somente capturada³”.

Esta frase icônica nos chama a atenção para o caráter analítico e questionador desta vertente teórica, uma vez que ‘capturar a linguagem’ significa olhar as variantes possíveis e

² Por atuar em diversas frentes disciplinares, a LC abrange métodos variados, responsáveis por investigar inúmeros paradigmas dentro de contextos distintos. Suas análises também possibilitam a aplicação em diferentes segmentos, respectivamente na Aprendizagem de Língua Estrangeira, Estudos da Tradução, Corpora de Aprendizes, Fraseologia de língua geral e especializada, dentre outras vertentes. Para os interessados em se aprofundar nestes tópicos, segue um pequeno rol de autores especializados nas temáticas citadas: Sinclair (1997); Granger; Hung; Petch-Tyson (2002); McEnery; Xiao; Tono (2006); Orenha-Ottaiano (2004; 2009); Viana (2010); O’Keeffe; Meyer (2004); Tagnin (2015) entre outros.

³ No original: *Language cannot be invented; it can only be captured*. [Todas as traduções são de nossa responsabilidade.]

identificar quais são as escolhas mais eficazes no processo de comunicação. Também assume que a mudança e a variação linguística se fazem presentes na linguagem.

Para arrematar o exposto, Léon (2006, p. 54-56) elenca os objetivos da LC:

- Colocar os corpora à disposição

Se caracterizam principalmente por serem corpora de referência, ou seja, que trazem em seu bojo a representatividade de um dado idioma. São relevantes por atestarem a língua em uso de uma comunidade de falantes, geralmente compilados por universidades, em parceria com os cursos de tecnologia e sistemas de informação. A título de exemplificação, temos o *Brown Corpus*, da Universidade Brown, nos Estados Unidos, mencionado outrora.

- Confeccionar instrumentos linguísticos

A coleta de dados também auxilia na elaboração de material didático. Como exemplo, podemos citar o livro de gramática: *Longman Grammar of Spoken and Written English*, publicado em 2007, e o *English Collocations in Use: Advanced*, da *Cambridge University Press*, datado do ano de 2008. Nesta categoria, incluem-se dicionários baseados em corpora (*Les dictionnaires bilingues*, da *Aupelf-Uref Editions Duculot*, de 1996) e glossários de língua especializada (como por exemplo, o Glossário CEALE – composto por termos de Alfabetização, Leitura e Escrita para educadores⁴).

- Efetuar descrições linguísticas

Essa jornada envolve o método empírico-indutivo, que trata as informações coletadas do particular para o geral, ou seja, a partir do corpus se chega a conclusões empíricas acerca do tema em estudo.

Utilizando este método, é possível descobrir padrões léxico-gramaticais e demais particularidades. Além disso, ainda podemos observar a mutabilidade da língua (ou das línguas), realizando interpretações diacrônicas e contrastando informações em determinadas décadas ou séculos de linguagem.

- Estudos da variação

Finalizando o item anterior, com a descrição de um fato linguístico, temos a

⁴ Elaborado pelos membros do Centro de Alfabetização, Leitura e Escrita (CEALE), da Faculdade de Educação da UFMG. Disponível em: <<http://www.ceale.fae.ufmg.br/app/webroot/glossarioceale/>> Acesso em: 26 dez. 2017.

confiabilidade de efetuar afirmações no que concerne a características de uma dada comunidade de falantes ou ainda diferenciações linguísticas entre grupos distintos que falam o mesmo idioma.

- Instrumentos de tratamento automático das línguas

Com a coleta de informações, a autora cita a possibilidade de “parametrizar algoritmos probabilísticos, como o reconhecimento da fala, a correção ortográfica, a etiquetagem automática ou instrumentos de auxílio à tradução” (LÉON, 2006).

Para a realização desta pesquisa, os objetivos de efetuar descrições linguísticas e estudar a variação serão tratados em tópicos posteriores. Uma vez que os dois grupos de estudantes elencados nesta pesquisa se relacionam ao item efetuar descrições linguísticas, tem-se a representação do perfil dos alunos no capítulo metodológico.

Já o detalhamento da exposição dos dados coletados é o foco das análises, associado ao objeto dos estudos da variação, notadamente a comparação entre a escrita da escola do centro e da periferia.

De posse do relato dos objetivos de corpora, acrescentamos outra característica deste estudo investigativo: o caráter quantitativo e qualitativo das análises. A partir das contribuições de Viana (2010) temos:

[...] muitos dos estudos de *corpora* têm seu início numa metodologia quantitativa – como a observação de frequências de uso – (...) Nesse panorama, a abordagem qualitativa se revela fundamental para decodificar os padrões evidenciados pela ferramenta computacional, contextualizar os resultados em termos tanto da população investigada como da macroárea em que uma pesquisa se insere, explicar o que os números revelam, e interpretar os achados de uma forma geral. (p. 35)

Devido à peculiaridade das informações advindas dos *softwares* aliados aos dados matemáticos, o trabalho com corpora naturalmente envolve um teor estatístico, já que são resultados dos próprios itens manuseados pelo computador. Contudo, além das informações exatas, o que irá demonstrar a riqueza da pesquisa será a forma como a expressão linguística será verificada qualitativamente.

Ou seja, a identificação dos padrões existentes na linguagem coletada, a reflexão sobre termos de comunicabilidade e intencionalidade do falante, culminando, mais adiante, nas relações semânticas entre tais padrões e a comunidade linguística aos quais os alunos que estudam nas escolas onde foram coletadas as redações estão inseridos.

Por este motivo, Tagnin (2015), Sinclair (1991), McEnery, Xiao e Tono (2006) elencam a importância de compilar corpora de maneira criteriosa, a fim de que haja relevância dos dados para comprovar (ou refutar) as intuições linguísticas do pesquisador.

No próximo tópico, tratamos com mais ênfase sobre o aspecto do ensino, a fim de aprofundar o tema do nosso trabalho, voltado a analisar redações de crianças no Ensino Fundamental I.

1.1.1 Linguística de Corpus direcionada ao ensino

A LC esteve presente nos espaços acadêmicos, descrevendo a linguagem por meio de corpora compilados, sendo posteriormente utilizado em análises linguísticas, tradução e inclusive com finalidades pedagógicas. Corroborando o que já foi dito anteriormente, com base nos preceitos da LC, os dados linguísticos são selecionados, analisados e mantidos em forma eletrônica.

Por volta de 1960, a criação do primeiro corpus linguístico eletrônico, o *Brown University Standard Corpus of Present-day American English*, representou grande inovação na área, com destaque para a grande quantidade de dados para a época: 1 milhão de palavras. A criação do *Corpus Brown* deu-se logo após a publicação de *Syntactic Structures*, de Noam Chomsky, gerando discussões sobre a validade do corpus.

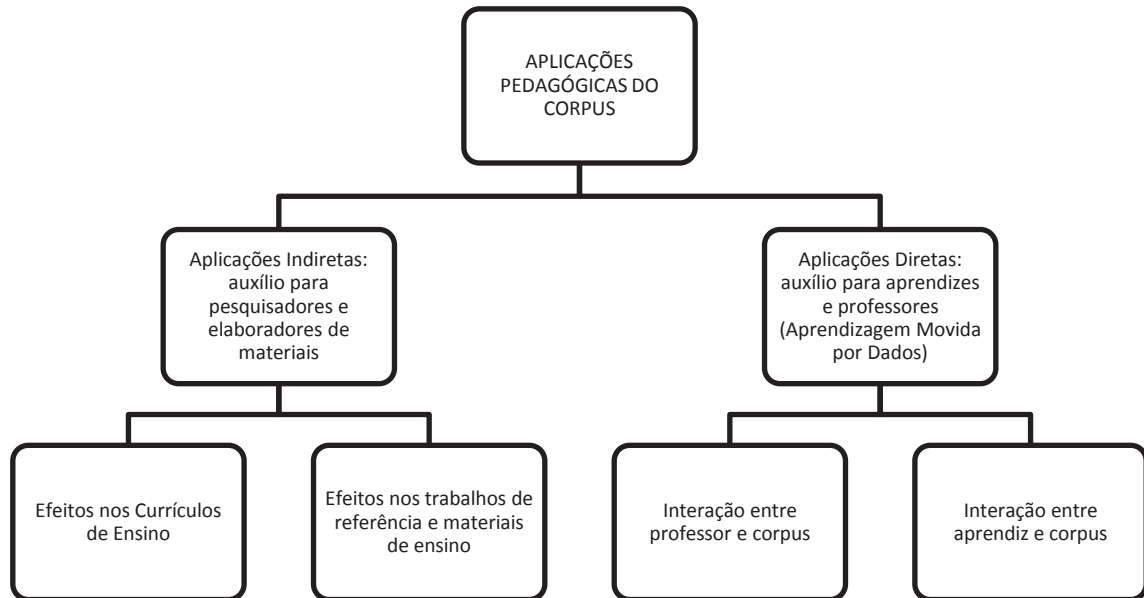
Nesta obra, Chomsky (1950) afirma que os dados linguísticos de que o pesquisador necessita estão em sua mente, sendo acessíveis por meio da introspecção. Tal visão de língua se choca com o que advoga a LC, baseada no empirismo e na coleta de dados produzidos por outras pessoas, não pelo próprio pesquisador.

Nos anos 1980, com a criação dos computadores pessoais, várias mudanças se processaram, havendo um fortalecimento nos trabalhos baseados em corpus, um pouco esquecidos após a mudança de paradigma linguístico, que passou a ser, a partir do final dos anos de 1950, de perfil gerativista.

A partir das inovações tecnológicas que surgiram nas últimas décadas, se faz necessário que o professor use tecnologias que vão além do giz e lousa em sala de aula. Como a utilização da informática corresponde a um hábito para a maioria dos alunos, o papel da escola também passa a ser o de orientadora no contato dos discentes com esse elemento, que está sendo largamente incorporado à vida moderna de todos.

Para compreendermos melhor como a LC está atrelada ao ensino, segue abaixo o esquema de Römer (2011), que ilustra bem a articulação deste processo:

Figura 01 – O uso de corpora no ensino e aprendizagem de segunda língua



Fonte: Römer (2011, p. 207), tradução nossa

Observe que o objetivo principal de aplicações pedagógicas de corpus se divide em dois: o primeiro, chamado de aplicações indiretas, envolve pesquisadores e produtores de material didático. Este, por sua vez, possui ainda, dois desmembramentos – um tratado pela influência exercida nos programas dos cursos e outro relacionado à atuação que estes desempenham sob trabalhos de referência e materiais de ensino propriamente ditos.

Já a segunda conexão do organograma, traz os aspectos diretos relacionados à aplicabilidade de corpora, interligados especialmente aos agentes participantes do processo ensino aprendizagem – alunos e professores. Paralelamente ao anterior, aqui também está seccionado em duas partes, uma versada na interação professor-corpus e o último se interessa pela relação entre aprendiz e corpus.

Haja vista que o interesse deste trabalho está nas “Aplicações Diretas” dos aspectos pedagógico do corpus, é neste cenário que destacamos a utilidade da LC enquanto ferramenta pedagógica.

O professor, tanto de idiomas quanto de língua materna, tem um suporte pedagógico eficiente, que possibilita ao aluno ter acesso ao uso real da língua, em vez de somente recorrer aos exemplos das gramáticas tradicionais.

Ao inserir novas tecnologias em suas aulas, pode despertar a atenção dos alunos, aproximando-se da realidade deles, em um mundo cercado de dispositivos eletrônicos. Além disso, também entra em contato com vários registros de língua, que podem servir de modelo para sua produção.

Durante as aulas em que se dá o uso de corpus, torna-se muito importante a função do professor no processo de ensino-aprendizagem, pois é ele quem irá nortear os alunos em suas pesquisas, fazendo a mediação entre os estudantes e as tecnologias. Há ainda a falta de habilidade de alguns estudantes em lidar com ferramentas tecnológicas num contexto pedagógico, daí a necessidade de uma diretriz que auxiliá-lo-á a transformar essa informação transmitida pelos dados em efetiva aprendizagem.

Dessa forma, os alunos deixam de ser passivos no processo de aprendizagem e passam a ter papel ativo. Tornam-se investigadores, pesquisadores cujo desenvolvimento de certas habilidades, como atenção e discernimento linguístico, lhes possibilitem lidar com os resultados obtidos nas pesquisas realizadas.

Por isso, o objetivo deste tópico é destacar um dos usos de corpora mais disseminados pela comunidade acadêmica: o ensino de idiomas, haja vista que por meio da compilação de escritas e áudios, valoriza-se a comunicação em línguas. Assim, os estudantes entram em contato com porções de linguagem natural, em situações que pedem o conhecimento de blocos de palavras, como as colocações, por exemplo, um dos temas a ser desenvolvido nos próximos tópicos.

A partir do segmento analisado, abordaremos alguns pesquisadores que tratam do ensino do léxico por meio de uso de corpus em sala de aula.

Destes, destacamos na presente pesquisa o *Data Driven Learning* – DDL (JOHNS, 1991); um segmento que trata da elaboração de materiais didáticos como o *Lexical Syllabus* (WILLIS, 1990) e ainda; outro que discute uma abordagem centrada no ensino do léxico, *Lexical Approach* (LEWIS, 2000b).

A primeira delas, DDL, é denominada no Brasil como Aprendizagem Moviada por Dados (AMD) e foi idealizada por Tim Johns nos livros: *Data-driven learning: the perpetual challenge* (1991) e *From printout to handout: grammar and vocabulary teaching in the context of Data-Driven Learning* (1994).

O mote para as atividades de ensino são as linhas de concordância dos corpora coletados e apresentadas pelos *softwares*. O intuito é incentivar os estudantes a buscar padrões no uso da linguagem. Neste processo, há três etapas metodológicas: identificação, classificação e generalização dos dados.

Relacionando a Aprendizagem Movida por Dados à pesquisa desta dissertação, podemos traçar um paralelo entre as linhas de concordância geradas por meio de uma linguagem autêntica (as redações das crianças) e a propostas de atividades citadas na análise deste texto. A partir dos corpora coletado, é possível desenvolver um material que atenda a necessidade dos alunos, a saber, desenvolver a competência colocacional da língua portuguesa do Brasil.

Nos últimos anos, muitos pesquisadores valorizam esta abordagem pelas vantagens que ela apresenta. Bernardini (2004), por exemplo, ilustra esse método citando a metáfora do *aprendiz viajante* para justificar o valor do percurso transposto pelo discente, ou seja, o quanto o aluno aprende enquanto busca compreender as estruturas linguísticas.

Contudo, é necessário fazermos algumas ressalvas a DDL, refletindo sobre a realidade educacional do país: se o foco é trabalhar com o discente identificando padrões léxico-gramaticais de determinados fatos linguísticos, o educador precisa ter à disposição um laboratório de informática com acesso à Internet – local nem sempre disponível na maioria das escolas públicas brasileiras.

Outro fator dificultador referente a aprendizagem da língua materna aponta para a necessidade de o aluno ter o domínio da norma culta, como uma das convenções de escrita, a fim de ser o protagonista neste processo de ensino. Os alunos aprendem português na escola para desenvolver a competência discursiva. Nos dizeres de Monteiro-Plantin (2017) esta se refere à “capacidade de reconhecer e também de selecionar, entre as estruturas linguísticas, paralinguísticas e epilinguísticas disponíveis, as que atendam aos propósitos dos falantes de forma mais satisfatória” (p. 13). Ou seja, o professor deve proporcionar condições pedagógicas para que o estudante aprenda outras modalidades discursivas, a fim de expressar seu pensamento e linguagem nos mais diversos contextos de comunicação.

O segundo segmento, denominado *Lexical Syllabus*, ou Currículo Lexical, possui o pesquisador Dave Willis como elaborador, a partir do livro *The Lexical Syllabus: a new approach to language teaching* (1990). Nele, o autor prega a alta frequência das palavras como fator para inclui-las ou eliminá-las no conteúdo programático do ensino de uma LE. Foca no ensino do vocabulário e do léxico, conforme pautado no corpus do *Cobuild*, centralizando a aprendizagem em tarefas (*task-based*). A partir desse projeto, o autor desenvolveu e publicou no início da década de 1990 uma coleção de livros didáticos conhecidos como *Cobuild English Course*. A preocupação do autor está centrada na fluência, visto que considera necessário que o material formulado descreva criteriosamente o que ocorre no corpus coletado.

Conforme já citado na Introdução deste texto, devido ao fato de os textos dos alunos estudados nesta pesquisa não terem apresentado uma quantidade significativa de fraseologismos, contrariando o que havíamos suposto inicialmente, decidimos incluir nesta dissertação atividades focadas no vocabulário e no léxico.

Uma vez que os postulados do Currículo Lexical possibilitam condições para realizarmos tais tarefas, pretendemos, em pesquisas posteriores, analisar se a aplicação desses exercícios baseados em corpora será efetiva na aprendizagem de fraseologismos pelas crianças em língua materna.

Dentre as vantagens discutidas sobre a proposta de Willis (1990), destaco aquelas mencionadas por Pitkowski e Gamarra (2009):

Este instrumento tecnológico é uma grande ferramenta para a criação de materiais didáticos. Outro atributo dos corpora, que não se pode esquecer, é o acesso fácil e gratuito. Por sua vez, o uso do corpus nos permite criar atividades indutivas. Desta maneira, os estudantes seriam aprendizes ativos na aula e em casa, facilitando, assim, a motivação e a participação⁵. (p. 46)

Contextualizando para a realidade de sala de aula nas séries iniciais, foco desta dissertação, avaliamos positivamente atentar os aspectos de autoestima e participação dos discentes de ensino fundamental I. Além disso, o professor deve considerar o meio social em que o aluno se encontra e não poupar esforços para integrar a família sobre os avanços (ou eventuais dificuldades) que o aluno perpassa no seu desenvolvimento escolar.

Mesmo em um contexto social em que a maioria das crianças não dispõe de um ambiente favorável a leitura/escrita, é preciso que o discente se sinta motivado a aprender e a transformar a sua realidade a partir dos conhecimentos que aprende na escola. Durante o cotidiano escolar, espera-se que o professor incentive o potencial dos alunos, estimule os pais a elogiar os filhos e observar o seu desenvolvimento. Em reuniões de pais, é importante que o professor mostre (por meio de fotos, filmagens, diagnósticos dos alunos) como as crianças estão compreendendo os conteúdos, pedir sugestões dos pais sobre as tarefas, atividades extraclasse, dentre outros.

Uma ressalva ao Currículo Lexical, contudo, diz respeito ao favorecimento do léxico não priorizar os conteúdos gramaticais. Por este aspecto não apresentar-se como prioridade, a aprendizagem destes itens linguísticos pode comprometer a aprendizagem da língua nos

⁵ *Este instrumento tecnológico es una gran herramienta para la creación de materiales didáticos. Otro atributo de los corpus, que no se puede obviar, es su acceso fácil y gratuito. A su vez, el uso de los corpus nos permite la creación de actividades inductivas. De esta manera, los estudiantes serían aprendices activos en el aula y en casa, facilitándose así la motivación y la participación.*

âmbitos gramaticais.

O terceiro tópico, *Lexical Approach*, traduzido aqui por Abordagem Lexical, tem como autor Michael Lewis, e é descrito nas obras *The lexical approach: the state of ELT and a way forward* (2000b) e *Implementing the Lexical Approach: Putting Theory into Practice* (1997).

A Abordagem Lexical, como indica o nome, também traz o léxico como prioridade. Contudo, ao contrário do Currículo Lexical, o *Lexical Approach* agrega em seus conteúdos as formações de palavras compostas, trazendo as colocações, por exemplo, no foco da aprendizagem de LE. Haja vista que um dos pontos centrais desta dissertação versa a respeito da fraseologia, a discussão a respeito destas combinatórias será realizada em tópicos posteriores.

A vantagem principal da abordagem é a defesa de um ensino articulado, fazendo com que o estudante pense em associar as palavras necessárias a elaborar frases e conceitos, concebendo a língua como um todo, e não em termos separados. No ensino de idiomas, essa propriedade é tratada pelo paradigma OHE – Observar, “Hipotetizar”, Experimentar: o aluno observa o idioma alvo, cria hipóteses sobre o uso e por fim, faz experimentações sobre as hipóteses levantadas. Caso haja erros neste processo, é necessário retomar o primeiro passo.

Em relação às críticas, estas citam a informalidade pela qual o corpus é descrito e ao modo como os conceitos são transmitidos. A obra *The lexical approach: the state of ELT and a way forward*, publicada originalmente em 1993 foi bastante criticada por esses fatores, o que levou o autor a reorganizar a sua escrita e publicar em 1997 o livro *Implementing the Lexical Approach: Putting Theory into Practice*, com o intuito de esclarecer alguns aspectos não compreendidos em seu trabalho anterior.

Neste âmbito trazemos as contribuições de Sripicharn (2010), ao sugerir atividades introdutórias de LC antes de iniciar qualquer atividade desta natureza em sala de aula:

Uma vasta gama de técnicas de pesquisa pode ser usada para obter informações sobre o que os alunos já conhecem sobre o tema (diagnóstico inicial). O professor pode começar a aula com uma discussão em classe pedindo aos alunos para relatar o que eles sabem sobre a linguagem corpora e sua experiência de usá-los. Além de uma discussão oral, questionários, listas de verificação ou testes podem ser administrados para dar ao professor algumas ideias sobre o conhecimento prévio dos alunos em relação com o corpus⁶. (p. 371)

⁶ A wide range of survey techniques can be used to elicit information about learners' backgrounds. The teacher may begin a lesson with a class discussion asking the students to report what they know about language corpora and their experience of using them. Apart from an oral discussion, a written survey in the form of

Ao gerar questionamentos sobre assuntos diversos, o estudante terá condições de continuar suas indagações em um plano mais profundo em buscas de significados e compreensão do tema em estudo. O nível de interpretação e entendimento da linguagem será amplificado, desenvolvendo o letramento em um grau maior, além de uma argumentação sólida e a criticidade necessária para compreender textos de qualquer espécie.

Com base nessas propostas, o grupo de pesquisa outrora chamado “Pedagogia do Léxico e da Tradução a partir de corpora” atualmente denominado “Fraseologia e colocações a partir de corpora” (FRASCORP), devidamente catalogado no Diretório de Grupos de Pesquisa do CNPq tem desenvolvido pesquisas tendo como foco a Aprendizagem Movidada a Dados.

Esta dissertação, inclusive, também compõe uma das áreas 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 mesma instituição e coordenado pela professora Adriane Orenha-Ottaiano. Neste contexto, o Currículo Lexical também se faz presente nas investigações da equipe, pautado pela investigação do ensino articulado ao léxico e tradução, além da formação de material didático a partir de corpora eletrônico.

Para ilustrar o exposto, cito alguns estudos desenvolvidos na próxima página, bem como o mote dos estudos relacionados e os pesquisadores responsáveis pelas investigações:

Quadro 01 – Trabalhos de destaque do grupo de pesquisa “Pedagogia do Léxico e da Tradução e Linguística de Corpus”

Produção de material didático	<ul style="list-style-type: none"> • Orenha-Ottaiano e Rocha (2015) compilaram as legendas em português e inglês do seriado <i>House MD</i>, para propor atividades pedagógicas para as aulas de inglês, com foco no ensino de vocabulário. • Oliveira e Orenha-Ottaiano (2017) descreveram a compilação de dois corpora comparável (livro <i>Harry Potter and the Sorcerer’s Stone</i> e legendas do filme de mesmo título) para elaborar atividades de ensino da língua inglesa. • Caldas (2017) compilou legendas do seriado <i>CSI – Crime Scene Investigation</i>, a fim de inserir atividades de <i>Memory game</i> e <i>Gap fill</i> dispostos na <i>website Online English Collocations Workbook</i>.
Produção de material didático disponível via Web	<ul style="list-style-type: none"> • Orenha-Ottaiano idealizou e atualiza a <i>Website</i> intitulada: <i>Online English Collocations Workbook</i>, cujo enfoque é disponibilizar diversas atividades elaboradas por meio da compilação de corpora de aprendizes. A proposta é inovadora por explorar colocações de língua geral e também especializada, além de divulgar os resultados a alcance global, por meio da plataforma.
Tradução	<ul style="list-style-type: none"> • Oliveira, Orenha-Ottaiano e Alves (2017) aplicaram uma atividade de tradução de histórias em quadrinhos, para investigar tanto as escolhas lexicais dos discentes, quanto aos itens lexicais e expressões idiomáticas.
Glossário	<ul style="list-style-type: none"> • Souza (2018) realizou uma análise comparativa das colocações extraídas de cinco temporadas do seriado <i>Queer as Folk</i>, em relação a um corpus comparável, propondo a elaboração de um Glossário de Colocações da Comunidade Homossexual.
Ensino de Língua Materna	<ul style="list-style-type: none"> • Oliveira (2014) compilou um corpus de trinta e uma redações em língua portuguesa, feitas por alunos do quinto ano, para investigar desvios de ortografia e construções verbais encontrados, assim como trazer algumas sugestões pedagógicas de intervenção para sanar tais dificuldades.
Literatura Brasileira baseada em corpus	<ul style="list-style-type: none"> • Teixeira (2016) analisou as traduções das colocações presentes em corpus literário paralelo, formado pela obra em português de Memórias Póstumas de Brás Cubas, escrito por Machado de Assis e suas três versões para língua inglesa.

Além disso, os membros do grupo de pesquisa citado também desenvolvem projetos de pesquisa, tais como “A Compilação de um Dicionário Bilingue de Colocações Baseado em Corpus”, que visa a compilar variadas colocações (verbais, nominais, adjetivas e adverbiais) de um dicionário bilingue em corpus, nos formatos eletrônico e impresso, nas direções Inglês-Português/Português-Inglês, dentre outras concepções correlatas.

Ainda sobre emprego de pequenos corpora na área educacional, faz-se necessário abordar uma das aplicabilidades utilizada com bastante frequência pelos estudiosos da LC, a coleta de textos de alunos, dando origem aos denominados Corpora de Aprendizes, uma das áreas as quais se aplica esta pesquisa. Por isso, a fim de contextualizar este uso e aprofundar a discussão principal neste texto, trataremos de Corpora de Aprendizes no próximo item.

1.1.2 Contextualizando o uso de Corpora de Aprendizes em pesquisas que contemplam o universo infantil

Historicamente, Tribble (1990) foi um dos precursores na utilização de pequenos corpora de aprendizes ao investigar a relação de verbos usados na fala de seus alunos. Em suas análises, recomendou a exploração das linhas de concordância pelos próprios estudantes, como forma de ajudá-los a reformular as sentenças escritas por eles, incentivando-os, assim, a ampliação do vocabulário outrora escrito.

Mais um estudo seminal com incursões para práticas pedagógicas utilizado para o ensino de língua estrangeira é o de Granger e Tribble (1998). Os autores confrontaram o subcorpus francês do *International Corpus of Learner English* (ICLE), composto por redações de aprendizes de inglês e um subcorpus de 1 milhão de palavras do *British National Corpus* (BNC). Por intermédio destes textos é possível perceber como estudantes não nativos se expressam no idioma já citado.

Ainda sobre este tema, é digno de nota relembrar os postulados de Granger (2008), quando afirma que os corpora de aprendizes ajudarão os pesquisadores a ter mais certeza sobre os fatos da língua, e dentre outras vantagens: “eles ajudarão a desenvolver novas ferramentas pedagógicas e práticas de salas de aula que atingirão com maior precisão as necessidades dos aprendizes⁷” (p. 17).

Seja para investigar aprendizagem da linguagem, aspectos morfológicos ou discursos orais e escritos, os corpora de aprendizes de Ensino Fundamental compilados nesta pesquisa podem ser utilizados para diversos fins pedagógicos e didáticos, principalmente na compreensão do funcionamento da língua materna.

Neste sentido, a linguagem das crianças possui características sensíveis ao seu desenvolvimento que merecem ser abordadas em algumas palavras neste texto. A influência dos pais, neste caso, é primordial neste processo de construção da linguagem. A fruição do vernáculo pela criança ocorre principalmente pela repetição, por meio do exemplo. Haja vista que os bebês seguem o padrão que ouvem, a variedade vocabular auxiliará o entendimento da norma popular, que será aprimorada com o passar do tempo na fase escolar, conforme já mencionado anteriormente acerca da competência discursiva.

⁷ they will help to develop new pedagogical tools and classroom practices which target more accurately the needs of the learner.

Em face aos conflitos cognitivos, a criança reconstrói o sistema linguístico de acordo com seu entendimento e, por esse motivo, enxerga a língua por um prisma empírico, assumindo padrões demonstrados por estruturas identificáveis nos dados coletados.

Concomitantemente, há outro elemento relevante em corpus de língua materna: a oportunidade de verificar a escrita dos mesmos alunos em um determinado período de tempo, a fim de verificar se houve alteração significativa na forma de se expressarem ou não.

Leech (2008) aborda o estudo longitudinal de corpora, conforme segue:

Existem muitos aperfeiçoamentos e elaborações de um programa de investigação que pode ser prevista - tais como a coleta de corpora dos mesmos alunos em diferentes fases de aprendizagem (um corpus longitudinal, de fato) ou a coleta de um (corpus de aprendiz longitudinal preferencialmente) de dados de aprendizes individuais em vez de partir de um grupo homogêneo⁸. (p. xv)

A partir da coleta de diferentes textos de um mesmo aluno, podemos traçar um perfil diacrônico da escrita deste discente. De caráter indutivo, o enfoque parte da compilação de material autêntico de linguagem a fim de confirmar ou refutar as hipóteses linguísticas do pesquisador.

Ainda em Diessel (2009), depreende-se que os estudos longitudinais mostram a diferença individual das crianças em comparação umas com as outras. Contudo, devem ser observados outros elementos no corpus, tais como a condição sócio-econômica da criança, o local social, dentre outros fatores.

Conforme explicado anteriormente, pesquisas desta natureza são importantes por contribuir para a investigação do estudo de Corpora de Aprendizes na região em que são coletadas. Dessa forma, propicia diálogo a respeito da expressão infantil em um contexto de variação linguística e evolução da linguagem – observadas nas redações coletadas.

Diessel (2009) elenca ainda duas formas principais de estudar este público alvo: testar o conhecimento linguístico das crianças e examinar a linguagem infantil (p. 1197). Demuth (1996, p. 6) enfoca ainda a análise com dados espontâneos de crianças, no tocante aos aspectos morfológicos e sintáticos. Com a tecnologia atuando a favor do pesquisador, enxerga as vantagens deste uso no entendimento de questões voltadas à aprendizagem da língua materna.

Utilizando corpora infantis, Biber, Conrad e Reppen (1998, p. 173) apontam três

⁸ *There are many refinements and elaborations of such a research programme which can be envisaged - such as the collection of corpora of the same students at different stages of learning (a longitudinal learner corpus, in fact) or the collection of a (preferably longitudinal) corpus of the data from individual learners rather than from a homogeneous group.*

vantagens:

- i. oferecem bases de dados para análises, a partir da coleção dos textos de crianças e adolescentes em diferentes estágios;
- ii. as análises baseadas em corpus permitem que investiguemos os padrões encontrados no conjunto dessas características;
- iii. de posse destas informações, compreender como ocorre o desenvolvimento da linguagem.

Dadas as circunstâncias desta pesquisa, concordamos com os autores acima quanto ao uso da LC em pesquisas com crianças, pois é perfeitamente possível adequar pesquisas de corpora ao público infantil. Contudo, acrescentamos alguns cuidados neste processo, de cunho metodológico, como: realizar reuniões com os pais para explicar os objetivos da investigação e solicitar a autorização dos mesmos para coletar os textos das crianças por escrito; propiciar um ambiente acolhedor aos discentes, para que se sintam a vontade em redigir as propostas de redação, dentre outros procedimentos.

Ao ampliar o nosso olhar investigativo para os alunos de ensino fundamental I, é justificável estendermos a perspectiva do cenário pedagógico outrora apresentado. Já que o estudo dos textos infantis abre espaço para a discussão de métodos de aprendizagem, é possível formular hipóteses para o entendimento da articulação da fala em textos escritos; refletir sobre as estratégias de produção textual aplicados em sala de aula, além de abordar quais métodos são mais adequados para avaliar a expressão linguística dos alunos.

O percurso descrito acima é pertinente para os agentes envolvidos diretamente no processo de ensino/aprendizagem na escola, tais como professores, coordenadores, diretores e dirigentes de ensino, por exemplo.

Fazemos esta afirmação, pois, baseados em uma linguagem empírica produzida pela criança, estes profissionais podem compreender melhor as singularidades observadas nos textos dos educandos. De posse deste material, os *softwares* possibilitarão a identificação de padrões léxico-gramaticais próprios daquele grupo de estudantes.

Voltando à compilação de corpora infantis, é muito conhecido o projeto CHILDES⁹ (*Child Language Data Exchange System*), que teve início em 1984, por Brian MacWhinney e Catherine Snow. Trata-se de um banco de dados com materiais escritos e orais produzidos por crianças, contendo transcrições, áudios e vídeos em 26 idiomas de 130 diferentes corpora.

⁹ Caso haja interesse em saber mais informações sobre os trabalhos mencionados, confira no item *Referências a Web Site* do referido projeto.

Outro exemplo diz respeito ao *The Arizona Corpus of Elementary Student Writing*, descrito por Biber, Conrad e Reppen (1998, p. 175-176), composto por mais de 5.000 textos coletados em 40 classes diferentes, totalizando uma amostra de 15 cidades do estado do Arizona, nos Estados Unidos. Os alunos que redigiram as redações foram divididos em três grupos linguísticos: nativos de inglês, de Navajo e de espanhol. Cada classe escreveu nove composições de temas variados dentro da sala de aula, sem nenhum tipo de interferência de professores.

Dessa forma, no que diz respeito a pesquisas que envolvem a coleta de redações de alunos com a utilização dos preceitos de LC, há inúmeros trabalhos expressivos e reconhecidos nesta área. Predominam, neste rol de pesquisas, o estudo e análise da aquisição de língua estrangeira (GRANGER, 1998; LEECH, 1998) ou ainda investigações destas escritas, a fim de produzir material didático que atenda as necessidades do estudante que está adquirindo um segundo idioma, por exemplo (FLOWERDEW, 2001; ORENHA-OTTAIANO, 2016, 2015).

Contudo, pesquisas voltadas para a compreensão da escrita de crianças sobre a língua materna a partir de corpora ainda são escassas (DEMUTH, 1996; TOMASELLO, 2003). No contexto brasileiro, também há poucas referências (CRISTÓFARO-SILVA et al, 2009; SHEPHERD; ZYNGIER; VIANA, 2006).

Em nosso país, ainda são raros os trabalhos baseados em corpus, ou seja, que compilam a linguagem produzida por crianças, para demonstrar como podemos melhorar os aspectos linguísticos por meio da elaboração de material didático autêntico.

Com o escopo voltado a corpora, mencionamos, na sequência, alguns trabalhos voltados à coleta e análise de redações em língua materna, oriundos de alunos do Ensino Fundamental I e II:

– PROJETO E-LABORE (UFMG) visa coletar e cadastrar um banco de dados das composições de estudantes do 1º ao 6º ano do Ensino Fundamental (6 a 12 anos de idade). A preocupação com a catalogação, digitalização e transposição dos textos físicos informatizados é digno de nota, com o trabalho criterioso da equipe de apoio ao tratamento das informações. O objetivo é esquematizar o vocabulário infantil do vernáculo português brasileiro atual. As discussões, assim, apontam a “interação entre a linguagem adulta e infantil, em um contexto de mudança linguística e evolução de linguagem” (CRISTÓFARO-SILVA, 2009).

- Tese de Espíndola (2014), da Universidade Federal do Rio Grande do Sul (UFRGS) - aborda as dificuldades que crianças e adultos indígenas possuem em redigir textos na língua portuguesa. Dessa forma, relata a constituição de um corpus com 483 textos de crianças e 349 textos de adultos – escritos por indígenas e não indígenas, para análises. Assim, compara a escrita dos dois grupos etários e o caminho linguístico vivenciado nesse período.

- Pesquisas de Shepherd (2006), da Universidade Estadual do Rio de Janeiro (UERJ) e Viana (representando a PUC-Rio) aliam os estudos de corpora à Análise do Discurso, em 2006, verificando os ‘feixes lexicais’ mais utilizados no discurso escrito de alunos da 5ª e 6ª série. Nesta investigação, destacam-se a escrita dos habitantes da área rural (Tocantins/MG) em contraste com os moradores da zona urbana (Favela da Maré/RJ).

- Jordão et al. (2005) utilizaram a ferramenta *Wordlist* do *software WordSmith Tools* para identificar como as crianças do Rio de Janeiro e de Minas Gerais com idades entre 10 a 14 anos se expressavam sobre os temas violência, educação, esporte, religião, dentre outros. O foco da pesquisa foi analisar o contraste das informações dos diferentes grupos frente aos assuntos tratados nas redações.

Visto ser impossível dissociar o ser humano de seu tempo, sua realidade cronológica ou cultural, debates acerca da pluralidade de redações infantis configuram-se como um grande aprendizado político e cultural dentro da área de humanas.

O próximo tópico teórico versa sobre o estudo do léxico em língua materna, além da convencionalidade da linguagem, o que nos direciona aos estudos sobre Fraseologia.

1.2 O estudo do léxico no ensino de Língua Portuguesa do Brasil nas séries iniciais

Em virtude de esta dissertação confrontar tessituras oriundas de diferentes grupos sociais, é necessário abordar algumas particularidades na escrita infantil, como por exemplo, o papel do professor de língua materna em auxiliar o aluno a desenvolver a competência colocacional na aprendizagem da sua língua mãe. Ou seja, quais os tópicos do ensino de língua portuguesa nas séries iniciais que precisam de uma interpretação mais profunda do educador, para que o aprendiz avance significativamente na compreensão das colocações,

ampliando os níveis de entendimento do português do Brasil.

É válido lembrar que o ensino do léxico nem sempre foi o enfoque das aulas de Língua Portuguesa do Brasil nas séries iniciais. A partir do momento em que o foco das pesquisas brasileiras foi estabelecer a linguagem como um fim em si próprio, numa visão imanente da mesma, a linguística estruturalista tornou-se o grande foco dos estudos relacionados à linguagem.

Ilari (2004) chama a atenção para o fato de os ensinamentos de Saussure terem chegado ao Brasil tardiamente, por volta de 1960. A disseminação de seus postulados teve grande influência sobre os gramáticos e teóricos da época, principalmente no que tange ao tratamento do léxico em sala de aula, desde o Ensino Fundamental até o Ensino Superior (p. 53).

Com os novos estudos da linguística, contudo, o objeto observacional mudou da descrição da linguagem para a compreensão da intencionalidade do falante, marcado pelos traços de enunciação. Agora sim, é preciso incluir o falante e os aspectos extralinguísticos nas análises, conforme postula Costa Val (1997):

As tendências mais marcantes nos estudos linguísticos contemporâneos situam-se nessa família teórica, cujo quadro foi se delineando numa trajetória histórica que vai da língua para a fala, da competência para o desempenho, do enfoque exclusivo da forma para o reconhecimento da função, da frase para o texto, do enunciado para a enunciação (p. 5)

A fim de seguir a tendência da qual se vale a autora, é importante oferecer ao estudante, independentemente da sua classe social, a chance de ler e escrever textos diversos ao longo de toda a sua jornada escolar. Neste contexto, o professor tem a missão de auxiliá-los a superar obstáculos que os discentes possuem enquanto adquirem a linguagem e expressão.

Um desses empecilhos diz respeito a dificuldade das crianças em compreender a combinação de palavras em sua língua materna, o que gera dúvidas em interpretar os fraseologismos, e principalmente em redigi-los em suas produções textuais, conforme apresentado em detalhes no capítulo 3 desta dissertação, acerca da análise das redações coletadas.

Dessa forma, nos baseamos nos Parâmetros Curriculares Nacionais (PCNs) do ensino fundamental para alicerçar este tópico, por ser um documento que indica as recomendações para o ensino do português do Brasil. Em especial, acerca da função do professor em ser um agente mediador no processo, “inclusive sobre aspectos não percebidos inicialmente pelo

grupo – intenções, valores, preconceitos que veiculam, explicitação de mecanismos de desqualificação de posições – articulados ao conhecimento dos recursos discursivos e lingüísticos”. (PCN, 1997, p. 48).

A fim de nortear o trabalho docente e atestar o princípio da isonomia perante a Constituição Federal (1998), os PCNs (1997) destacam o papel do docente em orientar os discentes a perceberem os aspectos pragmáticos da língua. Elencamos, assim, alguns objetivos do ensino da língua portuguesa no primeiro ciclo:

- compreender o sentido nas mensagens orais e escritas de que é destinatário direto ou indireto: saber atribuir significado, começando a identificar elementos possivelmente relevantes segundo os propósitos e intenções do autor;
- produzir textos escritos coesos e coerentes, considerando o leitor e o objeto da mensagem, começando a identificar o gênero e o suporte que melhor atendem à intenção comunicativa;
- considerar a necessidade das várias versões que a produção do texto escrito requer, empenhando-se em produzi-las com ajuda do professor. (PCNs, 1997, p. 68)

Ou seja, os esforços do professor em proporcionar ao aluno condições para que este produza textos melhores em sua língua materna deve contemplar também uma mudança positiva da criança frente à linguagem. Para atingir este objetivo, é necessário inseri-lo em um ambiente com intervenções didáticas, em que a opinião do discente seja ouvida e respeitada.

Para que o discente seja capaz de realizar o proposto descrito no excerto acima, é preciso que o mesmo conheça a fundo o seu idioma materno, incluindo as nuances da língua, como, por exemplo, as combinatórias ou colocações.

Haja vista que este conteúdo será tratado com detalhes nos tópicos posteriores e que esta pesquisa também se articula com o entendimento de que é importante tratar do desenvolvimento colocacional em estudantes do ensino fundamental I, outro tema a ser explorado neste espaço diz respeito à competência lexical dos alunos. Segundo Basílio (2004), este conceito é denotado como o “conhecimento internalizado do falante nativo sobre o léxico de sua língua, abrangendo itens lexicais, relações lexicais e processos de formação” (p. 90).

Neste contexto, o professor de ensino fundamental I tem papel fundamental em auxiliar o aluno a transcorrer tal percurso. Isso ocorre, pois, ao oferecer um ensino de qualidade, que crie condições para que as crianças se expressem em diversos contextos de comunicação, as probabilidades de o docente ser bem sucedido em ampliar a competência lexical e, no caso desta pesquisa, a competência colocacional de seus alunos serão muito mais

altas.

Ademais, Sandmann (1991) salienta que “a competência lexical do usuário de uma língua se compõe de dois momentos: o da análise e interpretação das unidades estabelecidas no léxico” (p. 23). Tais tentativas de “reconstrução” da linguagem denotam significativas formas de compreensão do código linguístico da língua portuguesa. Os erros provenientes dessa elaboração permitem ao professor importantes dicas de como proceder na sala de aula, quais as estratégias para redirecionar o caminho seguido pelos alunos.

Aprofundando esta questão, trazemos à baila as proposições de Trévillle e Duquette (1996, p. 98), que apresentam cinco componentes presentes na competência lexical:

- 1) linguístico – engloba palavras e frases, além da ligação da palavra propriamente dita, sua forma, seus significados e seus contextos de ocorrência;
- 2) discursivo – relacionado à coesão, coerência e coocorrência, através das possíveis combinações com outras unidades lexicais que possuem afinidades entre si;
- 3) referencial – diz respeito às experiências subjetivas do falante com o mundo e, conseqüentemente, suas relações;
- 4) sociocultural – composto do sentido das palavras e o registro das mesmas, alinhados ao significado cultural e ao ambiente em que ocorre a interação comunicativa;
- 5) estratégico – que é a capacidade de utilizar as unidades lexicais para inferir situações de acordo com o contexto, também relacionado à paráfrases e produção de sentenças.

Em um sentido amplo, a competência lexical torna-se uma das bases da atividade comunicativa, já que parte da compreensão que as pessoas possuem da língua em seus aspectos mais individuais; expande para a interação existente na comunicação entre indivíduos e perpassa o conhecimento empírico sobre objetos, pessoas e sentimentos.

Outras etapas deste desenvolvimento envolvem ainda aspectos culturais e sociais, dado o fato de que interagimos de maneiras distintas com as pessoas, de acordo com os contatos que possuímos em nosso círculo de convívio ou então em consonância com os papéis que exercemos na comunidade. Por fim, o nível estratégico abarca a intencionalidade do

falante, que pode agir sobre o léxico para deduzir fatos ou ideias, ao enunciar hipóteses linguísticas.

Tais aprendizagens, contudo, 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.

Por este viés, é possível levar o aluno a desenvolver a autonomia linguística, ou ainda expor seu ponto de vista sobre qualquer assunto. Fazemos essa afirmação por acreditar que a leitura reflexiva e escrita consciente são capazes de propiciar ao discente o desenvolvimento de atributos como ética e cidadania em sua escolarização.

Por falar em autonomia, não podemos finalizar este tema sem antes, adentrar a competência fraseológica, visto que se trata de um dos segmentos deste trabalho. Para Olímpio de Oliveira Silva (2004, p. 5): “o conhecimento adequado do fluxo fraseológico é fundamental para que os aprendizes possam ter uma atuação linguística próxima à de um nativo e, portanto, possam alcançar uma ótima competência comunicativa¹⁰”.

Contextualizando para o ensino de língua portuguesa do Brasil, temos que o ensino das unidades fraseológicas (a ser explicado com propriedade nas próximas páginas) é responsável por possibilitar um falante mais desenvolvido em seu próprio vernáculo, uma vez que este domina as diversas particularidades da língua. Por sua vez, o conhecimento fraseológico exige um domínio extralinguístico do indivíduo para que este interprete a informação de maneira correta, sem a deturpação de sentidos.

Consequentemente, a linguagem como meio de expressão da história social do homem, tem na língua a forma mais cristalina para realizar tal tarefa. Através das palavras, pode-se revelar o próprio indivíduo como também revelar um povo; além de registrar os fatos e eventos que se constituem em uma história individual ou em uma história coletiva.

Ainda segundo os PCNs (1997):

O domínio da língua, oral e escrita, é fundamental para a participação social efetiva, pois é por meio dela que o homem se comunica, tem acesso à informação, expressa e defende pontos de vista, partilha ou constrói visões de mundo, produz conhecimento. Por isso, ao ensiná-la, a escola tem a responsabilidade de garantir a todos os seus alunos o acesso aos saberes linguísticos, necessários para o exercício da cidadania, direito inalienável de todos. (p. 16)

¹⁰ [...] *el conocimiento adecuado del caudal fraseológico es fundamental para que los aprendices puedan tener una actuación lingüística cercana a la de un nativo y, por lo tanto, puedan alcanzar una competencia comunicativa óptima.*

O professor de educação básica, portanto, deve relacionar a linguagem a cultura e a criatividade do ser humano, pois compreendendo que o aluno se exprime sobre a maneira como concebe o mundo, este utilizará a língua a seu modo para transmitir sua cultura e suas impressões sobre a realidade que o cerca.

Nesta faixa etária estudada, o aluno já compreende informações usando referências de seu cotidiano, bem como realiza inferências extra linguísticas. Além disso, é nessa idade que as noções de identidade, regras, laços afetivos e cognitivos começam a fazer sentido, seja de forma lógica ou imaginária. Tais interações são feitas pelo aluno, enquanto reflete sobre as relações sociais às quais está inserido em sua realidade.

Assim, em um primeiro momento, as unidades fraseológicas podem ser consideradas um conteúdo complexo para alunos do ensino fundamental I. Contudo, defendemos que esta temática é importante para possibilitar aos pequenos a fluência plena do idioma português brasileiro. Afinal, se a competência fraseológica e, de modo mais específico, a competência colocacional, for trabalhada desde as séries iniciais, promoverá ao estudante a capacidade de reconhecer o significado de expressões próprias da língua; além de oferecer condições linguísticas para que realize a escolha lexical mais adequada nas interações comunicativas, a fim de ser compreendido pelas pessoas com as quais convive.

Concomitantemente a este aspecto, observa-se a capacidade deste em expressar seus pensamentos, ações e sentimentos de forma profícua, tanto na oralidade quanto na escrita. Por esse motivo, o ambiente escolar é de extrema importância para o desenvolvimento da criança. Nesse contexto, a interação social auxilia os discentes a compreenderem suas ações em frente a perguntas e comentários dos colegas; na consciência das intenções semânticas presentes nas trocas discursivas cotidianas e interpretação das mesmas – além de posicionar-se frente a um tema com ética e respeito ao próximo.

Se os professores das séries iniciais tivessem consciência da importância do papel que exercem para que seus alunos entendam o código linguístico da língua portuguesa do Brasil, provavelmente os aspectos pragmáticos das atividades de letramento seriam explorados mais enfaticamente pelos docentes do ensino fundamental. Neste sentido enfatizamos a necessidade de haver mais pesquisas de cunho linguístico para encontrarmos formas de aproximar a fraseologia dos manuais eruditos da academia e os cadernos de classe das crianças que estão na sala de aula do ensino fundamental I.

Um dos caminhos possíveis apontados nesta investigação empírica sugere ao professor das séries iniciais utilizar o conhecimento prévio de seus alunos na produção de material

autêntico. Com base na coleta e análise dos textos coletados, o docente deve refletir sobre quais estratégias de ensino seriam mais produtivas em auxiliar o aluno a compreender como as palavras se articulam para formar significados.

Conforme já citado sobre os objetivos do ensino fundamental nas séries iniciais, para levar a criança a “identificar elementos possivelmente relevantes segundo os propósitos e intenções do autor” (PCN, 1997, p. 68), é necessário que o docente pense em estratégias metodológicas para alcançar esta meta, aqui demonstrada em detalhes nas sugestões de atividades baseadas em corpora, descritas no capítulo 4 deste corrente texto.

Para fecharmos o ciclo teórico desta dissertação, passemos ao próximo tópico, que versará com mais propriedade sobre a convencionalidade do signo linguístico e suas partes constituintes. Consequentemente, este tema nos leva às colocações, um dos vários desmembramentos dos fraseologismos e tema do qual nos interessa enquanto corpus deste trabalho, aliado a Aprendizagem Movidada por Dados, que oferece condições satisfatórias no tocante a metodologia utilizada pelo professor para colocar em prática a teoria observada.

1.3 Convencionalidade e Fraseologia

Já que uma das vertentes que analisamos são as colocações de língua geral presentes nos textos das crianças, este tópico é pertinente por contextualizar o leitor acerca da convencionalidade, como propriedade importante desta fundamentação teórica.

O dicionário Aurélio define convencionalidade por “aquilo que é tacitamente aceito, por uso ou geral consentimento, como norma de proceder, de agir, no convívio social; costume; convenção social” (FERREIRA, 2000). Está relacionada, dessa forma, ao uso que fazemos, ao interagir, de uma série de convenções, expressões e blocos de palavras já estabelecidos e consagrados em nossa comunidade (ORENHA-OTTAIANO, 2004) e pode ser evidenciada tanto por fatores sociais quanto linguísticos.

Mais uma definição de convencionalidade, agora no campo semântico linguístico é trazida por Crystal (1985):

A lingüística usa o termo [convenção] em seu sentido geral – com referência a qualquer prática aceita no uso da língua [...]. Existe também um sentido mais restrito, referindo-se à natureza arbitrária da relação entre expressões lingüísticas e suas significações [...] (p. 68).

Outrossim, se vincula a ideia de arbitrariedade do signo, já que não há motivação linguística que relacione o significante de todas as palavras ao seu respectivo significado. Apesar de se comunicar no mesmo código que o seu interlocutor, o ouvinte só compreenderá a mensagem se também compactuar das particularidades desta língua, enquanto for parte constituinte dos usos sociais dos mesmos traços desta linguagem.

Outra contribuição relevante é tratada por Orenha-Ottaiano (2004). Segundo a autora, a convencionalidade “está relacionada com o uso que fazemos, ao interagir, de uma série de convenções, expressões e blocos de palavras já prestabelecidos e consagrados em nossa comunidade” (p. 9).

Paralelo a essa explanação, Tomasello (2003, p. 8) delibera sobre a convencionalidade contrastando os símbolos usados por animais e seres humanos. Enquanto os primeiros visam atingir o comportamento nos seres que os cercam; os símbolos empregados pelos humanos têm o objetivo de afetar os estados mentais e intencionais de outros indivíduos. É nesta dimensão mental que as palavras demonstram um alto poder comunicativo, predicando todos os elementos existentes no mundo.

Ampliando a discussão para a aprendizagem de língua materna, Tomasello (2003) traz as habilidades que as crianças possuem em compreender esse jogo complexo da linguagem. Afirma que a espécie humana é biologicamente preparada para esta tarefa prodigiosa, pois o desafio enfrentado pelas crianças é aprender não somente as palavras e expressões convencionalizadas, mas também todas as formas de padrões abstratos desta língua gramaticalizada historicamente.

Durante a aprendizagem da primeira língua, as crianças entendem essas construções como peças de um quebra-cabeças, posto que, desde pequenos, conseguem reproduzir estes blocos de palavras com discernimento.

Autores como Pawley e Syder (1983) tratam deste tema a partir dos termos *nativelike selection* (seleção típica de falante nativo) e *nativelike fluency* (fluência típica de falante nativo). Enquanto o primeiro trata a habilidade do falante se comunicar em sua língua materna, o segundo versa sobre a capacidade de o enunciador utilizar falas conectadas e espontâneas.

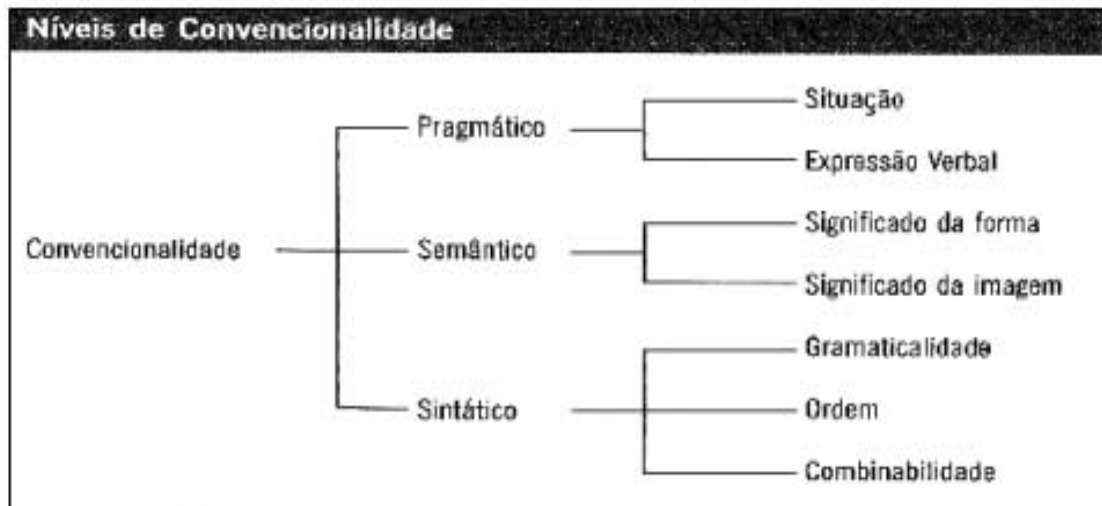
Segundo os referidos autores, o motivo principal de alguém ser fluente em sua própria língua materna é a capacidade de compreender o funcionamento dessas peças linguísticas, ou seja, guardar as sequências ou blocos de palavras na memória e saber quando resgatá-las durante um diálogo ou escrita.

Além destas habilidades serem fulcrais para a aprendizagem de uma língua, denotam que a criança é fluente em sua língua mãe. Tal potencialidade a faz alcançar os níveis de alfabetização plena, caracterizando um falante nativo que expresse, com riqueza de detalhes, seu pensamento e, conseqüentemente, seu idioma.

Por isso, novamente reiteramos a necessidade de abordar a questão da convenção linguística neste texto, como maneira de refletir e reverberar a importância do aprendiz em conhecer o seu vernáculo e escolher com sapiência qual palavra ou conjunto de palavras é mais adequado a cada contexto ou finalidade comunicativa.

Outro trabalho de destaque sobre o assunto, é de Tagnin (2005), cujo destaque é a divisão que a autora faz da convencionalidade em níveis:

Figura 02 – Níveis de Convencionalidade



Fonte: Tagnin (2005, p. 19)

Para ela, a convencionalidade utiliza estruturas recorrentes, subdivididas em fonológica, sintática, semântica e pragmática. A figura acima indica as divisões correspondentes aos três últimos elementos, que serão explanados e exemplificados na sequência.

Tal diferenciação entre as formas de convencionalidade são importantes, pois situam as nuances existentes neste imbricado jogo da linguagem humana. Ao estudarmos com profundidade as partes destacadas pela autora, observa-se que as mesmas não são separadas, mas sim articuladas entre si.

Visto que tratam de aspectos complementares, retomam a coocorrência da utilização do sistema pelos falantes. Conforme dito em páginas anteriores, denota a riqueza do léxico de

uma língua, além de apontar se um nativo é fluente em seu idioma materno pelo conhecimento do uso destes padrões léxico gramaticais ou não.

O primeiro item, a convencionalidade pragmática, versa sobre os valores e práticas inclusas na relação discursiva entre as pessoas. Assim, subdivide-se em situação, ou seja, o contexto social no qual se encontram os indivíduos e a expressão verbal, que se refere ao discurso a ser empregado nestas interações. Exemplificando, temos a situação velório, em que o cumprimento específico esperado (expressão verbal) aos familiares é “Meus pêsames”; enquanto no contexto de festa de casamento, a perspectiva sociointeracionista prega que felicitemos o novo casal com frases do gênero “parabéns” ou ainda “felicidades aos noivos”.

O próximo ponto, convencionalidade semântica, trata da “relação não motivada entre uma expressão e seu significado” (TAGNIN, 2005, p. 19). Tem-se aqui, a explicação para o aspecto da forma, ou seja, literalmente, a ilustração oferecida pela autora *kick the bucket* não se refere a literalmente chutar o balde, mas sim, indica o falecimento de alguém. Visa explicar, dessa maneira, a relação entre as expressões e seus significados adjacentes.

Já a imagem é representativa pela cultura de uma dada comunidade. Pode-se explicar este tópico pela visão ocidental, que possui como prosódia negativa a afirmação de que as pessoas vão para baixo após a morte (relacionando o local inferior como inferno); enquanto o andar de cima descreve o céu, o paraíso. Neste caso, a imagem indica a metáfora presente na linguagem conotativa, característica do significado de ordem semântica.

Por último, a convencionalidade sintática é responsável por interseccionar a combinação dos elementos no eixo sintagmático. Para tanto, Tagnin (2005) efetua três ramificações neste âmbito: a gramaticalidade, a ordem e a combinabilidade, explicados resumidamente a seguir:

- gramaticalidade: apesar de nem sempre representarem o padrão gramatical vigente, certas combinações são consagradas devido ao uso, significando a cristalização do termo pelos nativos do idioma.

- ordem: diz respeito à sistematização dos itens que constituem um conjunto de termos, por ilustração, tem-se: “cama, mesa e banho”. Consequentemente, a convencionalidade sintática não permite a troca na sequência destes sintagmas, como por exemplo: “banho, cama e mesa”.

- combinabilidade: é a propriedade que alguns vocábulos têm de se combinarem uns aos outros, ou seja, o uso de determinadas palavras é tão recorrente pelas pessoas que praticamente não permite outra associação. Para ilustrar essa consequência linguística, Tagnin

(2005) traz a figura do adjetivo “coroca”, cuja recorrência natural para formarmos um bloco de sentido completo é sua união com o substantivo feminino “velha”. Dessa forma, o léxico mental de um falante nativo será sempre “velha coroca”, e não “homem coroca” ou ainda “garota coroca”.

Assim, justifica-se o fato da importância deste tema na revisão teórica desta pesquisa, pois, se o aprendiz não compreender com clareza as nuances de seu código de comunicação, provavelmente terá dificuldades para transmitir seu pensamento em forma de palavras na modalidade oral e escrita.

Após discorrermos sobre a convencionalidade linguística, abordamos, a seguir, a Fraseologia e os tipos de combinações estáveis que ocorrem na língua. Tal diversidade se dá pela variedade das convenções, e que merecem um olhar mais aprofundado, por se tratar do foco das análises desta dissertação.

Haja vista que um mesmo fenômeno pode ser nomeado por vários nomes, dependendo da ótica adotada, é necessário refletir sobre a pluralidade das denominações que envolvem a Fraseologia, afinal, como já dizia Saussure: “Longe de dizer que o objeto precede o ponto de vista, diríamos que é o ponto de vista que cria o objeto” (1975, p. 15). A citação do genebrino indica o motivo pelo qual, ao longo da história, a Fraseologia sofreu inúmeras mudanças epistemológicas em busca de delimitação dos seus estudos (SILVA, 2006; RUIZ GURILLO, 1997).

Dependendo do recorte que os estudiosos da língua realizam sobre as combinatórias de palavras, aliado aos critérios escolhidos para descrevê-las, há uma série de denominações distintas e também suplementares sobre o mesmo assunto.

A título de exemplificação, Orenha-Ottaiano (2004) faz um levantamento das diversas denominações de um fraseologismo. Assim, a pesquisadora prossegue elencando várias designações, conforme demonstram alguns nomes e seus respectivos autores que elencamos aqui. O rol completo encontra-se devidamente registrado na dissertação de mestrado da mesma (Orenha-Ottaiano, 2004, p. 15):

- unidades multipalavras (COWIE, 1998);
- expressões multipalavras (LEWIS, 2000);
- agrupamentos lingüísticos (AZEVEDO, 1986);
- blocos pré-fabricados ou simplesmente *prefabs* (BOLINGER, 1976);
- unidades lexicais multivocabulares (*multiword lexical units*) (ZGUSTA, 1967);
- expressões feitas (*ready-made utterances*) (LYONS, 1979);
- frases (MELCHUK, 1992);

- unidades lexicais multipalavras (*multi-word lexical units*), pacotes lexicais (*lexical bundles*) (BIBER et al. 1999);
- “tijolinhos” (*building blocks*) (PAWLEY; SYDER, 1983);
- “pedaços” (*chunks*) (NATION, 2001);
- polipalavras (*polywords*), “tijolinhos fixos e por vezes fossilizados” (KJELLMER, 1987) (ORENHA-OTTAIANO, 2004, p. 15)

Aí reside a dificuldade em delimitar a área de atuação da Fraseologia, uma vez que, a depender do caminho percorrido, tem-se o risco de incorrer em questões que fogem da proposta traçada inicialmente.

Por isso, nesta investigação, compreendemos Fraseologia sob ótica de Monteiro-Plantin (2014) que traz a seguinte definição:

Trata-se de uma disciplina independente, relacionada a todos os níveis de análise linguística (do fonético ao discursivo-pragmático), cujo objetivo é o estudo das combinações de unidades léxicas, relativamente estáveis, com certo grau de idiomaticidade, formadas por duas ou mais palavras, que constituem a competência discursiva dos falantes, em língua materna, segunda ou estrangeira, utilizadas convencionalmente em contextos precisos, com objetivos específicos, ainda que, muitas vezes, de forma inconsciente. (p. 33)

Optamos por desta definição, pois, de uma forma direta e objetiva, explana sobre a Fraseologia enquanto disciplina autônoma, cuja área de estudos abarca as combinações de palavras durante a interação entre pessoas. O trecho ‘ainda que, muitas vezes, de forma inconsciente’ ganha destaque nesta pesquisa, tendo em vista que atende ao interesse proposto nas primeiras folhas desta introdução, a saber, analisar o léxico geral e fraseológico de redações de crianças.

Haja vista que as crianças ainda estão no processo de aprendizagem da escrita, as colocações presentes em suas composições são o resgate cultural de sua identidade, atestando o que compreendem da língua portuguesa.

Assim, o percurso discursivo proposto neste tópico sobre a Fraseologia é:

- tratar da fraseologia a partir da historiografia clássica e a contribuição de pesquisadores centrais sobre o tema;
- descrever os fraseologismos, com o recorte das colocações em especial, que serão retomadas posteriormente na análise.

Na sequência, tem-se o primeiro desdobramento deste tema, com a relação dos

diferentes autores que versam sobre as características da Fraseologia, o que nos auxiliará a conceituá-la a partir em seus aspectos inerentes ao escopo deste texto.

Descrever a Fraseologia é uma tarefa que exige bastante esforço e fôlego de quem se dispõe a fazê-lo, segundo também menciona Ruiz Gurillo¹¹ (1997). A intenção destas linhas, contudo, não é esgotar o assunto, mas sim, oferecer o olhar de uma pesquisadora e professora de educação básica sobre o assunto, destacando a importância de o aluno compreender as possíveis construções linguísticas que podem ser arquitetadas pelo falante de língua materna.

Para fins de definição, trataremos no escopo deste trabalho, da Fraseologia enquanto campo de estudos que trata das diversas combinações da língua, ou seja, um recurso linguístico validado graças ao uso por uma dada comunidade de falantes do mesmo idioma. Para efetuar o recorte metodológico das investigações desta pesquisa, limitamos o espaço a uma unidade fraseológica em especial, a saber, as colocações.

O percurso da fraseologia nos remonta a Saussure (1975), responsável por realizar considerações sobre a linguagem fundamentada em um sistema estruturado¹². Logo, as regras que norteiam seu movimento seguem leis internas e dependentes intrinsecamente. Neste âmbito, mesmo que indiretamente, o campo da Fraseologia foi aberto para a interpretação de outros olhares a partir deste autor, haja vista que, no Curso de Linguística Geral, o pesquisador afirma:

Na língua, tudo se reduz a diferenças, mas tudo se reduz também a agrupamentos. Esse mecanismo, que consiste num jogo de termos sucessivos, se assemelha ao funcionamento de uma máquina cujas peças tenham todas uma ação recíproca, se bem que estejam dispostas numa só dimensão (1975, p. 149).

Observa-se aqui o fato de a linguagem ser tratada como um padrão abstrato, subjacente aos atos do discurso. Ao exprimir ideias, a língua adota um conjunto de convenções sociais, permitindo a interação entre seus falantes – como norma, é entendido pelos indivíduos, possibilitando que vivam em sociedade.

¹¹ Para conhecer a fundo as origens da Fraseologia, cito nas referências deste trabalho o segundo capítulo do livro *Aspectos de fraseología teórica española*, intitulado “¿Es La fraseología una disciplina autónoma?”. Ruiz Gurillo (1997) descreve as primeiras descobertas desta vertente linguística, passando pela descoberta dos *refranes* (provérbios) no século XV, privilegiando as pesquisas no idioma espanhol. Contudo, cita autores da academia russa e outros de destaque, como Charles Bally e a revolução que este trouxe ao reconhecer os diversos tipos de fraseologismos existentes, dando início ao guarda-chuva da Fraseologia Moderna.

¹² Apesar da tradução do livro ao qual nos referenciamos ser de 1975, vale informar que Saussure já efetuava tais afirmações muito antes desta data (1894). As informações presentes neste tópico fazem parte das anotações das aulas do genebrino, publicadas postumamente em Curso de Linguística Geral no ano de 1916, pelos seus discípulos Bally e Sechehaye, com a colaboração de Ridlinger.

O termo “agrupamentos” nos indica a ideia seminal da Fraseologia, ainda compreendida dessa forma pela ilustração que o autor faz com o termo “jogo de termos sucessivos”, quer dizer, a sequência pela qual ela se apresenta nos enunciados.

Como estamos no espectro estruturalista, Saussure trata a língua como um sistema dinâmico e operante por si “uma máquina cujas peças tenham todas uma ação recíproca”. Por isso, o linguista também rejeita o contexto em que ocorre a enunciação, pois, a seu ver, não interfere na forma como o falante a utiliza “dispostas numa só dimensão”.

Há um conflito ao aprofundarmos a explanação de Saussure, pois o sistema descrito por ele não condiz com a realização individual desta língua – heterogênea, sujeita aos aspectos exteriores da linguagem. Dessa forma, é necessário recorrer a outros linguistas que considerem a influência do falante para produzir enunciados, a fim de avançarmos na reflexão linguística.

O primeiro autor a nomear estes fenômenos no ocidente e dar status de cientificidade a Fraseologia é justamente discípulo de Saussure e um dos alunos responsáveis pela publicação póstuma do Curso: Bally. Com a publicação de *Traité de Stylistique*, em 1951, prossegue os estudos de seu professor, apresentando as bases dos agrupamentos aos quais já se referia Saussure. A partir da frase: “Nunca poderíamos conservar, nem empregar todas as palavras que sabemos da língua materna, se tivéssemos de aprendê-las separadamente” (1951, p. 67). Bally avança na compreensão de que, para aprendermos e apreendermos verdadeiramente o vernáculo é necessário entender as combinações lexicais existentes. Desse modo, diferencia-se do primeiro autor por conceber a participação ativa do falante neste processo.

Tal envolvimento dá-se por meio dos índices que o autor aponta como a estruturação do campo fraseológico:

- índices exteriores, tratam do aspecto formal dos fraseologismos, como a estabilidade e a não-substituição de seus elementos, relacionados a convencionalidade e fixidez dos itens que a constituem, e
- índices interiores, que estudam a maneira como os usuários captam estas unidades.

Os itens elencados são significativos neste texto por representarem o entendimento do falante do agrupamento em uma expressão com significado próprio, e não mais termos separados.

As “locuções fraseológicas”, conforme Bally (1951) nomeia este fenômeno, se caracterizam por suas unidades “pré-fabricadas”, ou seja, são tratadas como um conjunto de

vocábulo que atua como uma única palavra, devido ao efeito de sentido que possui no discurso do enunciador. Por isso, a fluência de uma pessoa em um dado idioma pode ser mensurada pela habilidade desta em compreender os blocos de palavras, ou *chunks*¹³ da língua com confiabilidade.

Pelo exposto, observa-se que apesar de avançado e complexo, o estudo da Fraseologia pode claramente levar o aluno a ampliar o conhecimento de sua língua materna. Por isso, reiteramos para que esse conteúdo seja implementado nos currículos nacionais, a fim de que, paulatinamente, o estudante brasileiro possa ter condições de otimizar sua linguagem e entendimento sobre as combinatórias existentes.

Ao tratar de fraseologismos, novamente nos deparamos com um tema ao qual são várias as possibilidades de descrição. Se a definição de fraseologia é tratada por diversas vertentes, logicamente a representação de seus ‘afluentes’ é igualmente complexo de ser delimitado e explanado.

Isso ocorre devido à falta de critérios objetivos que definam os aspectos dos fraseologismos, bem como a ausência de limites dos preceitos de frequência, idiomaticidade e fixidez, por exemplo. A variedade em que estes processos ocorrem na língua é difícil de delinear, ocasionando oscilações e divergência entre os teóricos que se aventuram a desbravar ‘estes mares’.

Para ilustrar o exposto acima, trazemos um trecho do quadro de Noimann (2007), cuja síntese dos fraseologismos retrata bem tal impasse:

¹³ Este termo é utilizado em diversas situações de combinatória (como a sequência de teclas necessárias para um musicista extrair os sons melódiosos de um instrumento ao seguir as notações de uma partitura, por exemplo). No contexto linguístico, comparemos as notas musicais aos fonemas e letras, assim, é possível depreender o valor dos *chunks* na fabricação de sintagmas e blocos pré-fabricados de “porções” da língua com sentido independente. Cf. Sinclair (1991) e Wray (2002) para informações adicionais.

Figura 03 – Alguns tipos de Unidades Fraseológicas

PESQUISADOR	TIPOS DE UNIDADES FRASEOLÓGICAS	DESIGNAÇÕES
ZULUAGA (1980)	Locuções e enunciados fraseológicos	Fraseologismo
CROSS (1996)	Locução	Locuções
GURILLO (1997)	Locuções ou modismos	
PENADES MARTINEZ	1. Ditos – 2. Expressões idiomáticas. – 3. Frases – 4. Modismos – 5. Gírias – 6. Idiotismos – 7. Locuções – 8. Modos de dizer – 9. Frases feitas – 10. Refrões – 11. Provérbios – 12. Colocações – 13. Expressões unidades pluriverbais – 14. Unidades léxicas pluriverbais	Fraseologismo
PÉREZ (2000)	1. Combinações de palavras – 2. Provérbios – 3. Refrões – 4. Aforismos – 5. Fórmulas fixas – 6. Frases feitas	Fraseologismo Unidade fraseológica
COLADO (2004)	1. Locuções nominais – 2. Locuções adjetivais – 3. Locuções verbais – 4. Locuções adverbiais – 5. Locuções causais – 6. Locuções preposicionais	Locuções

Fonte: Noimann (2007, p. 28)

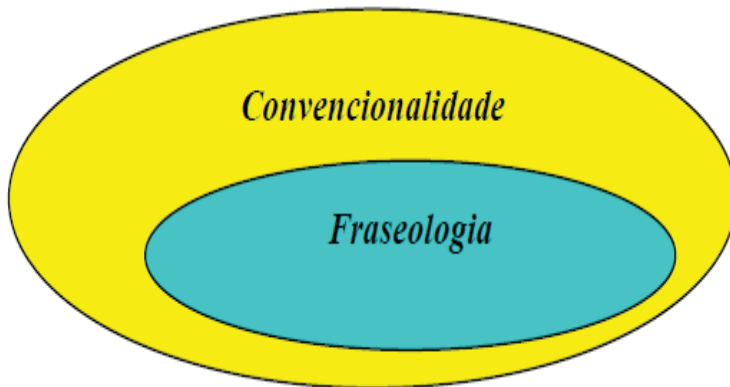
Nele, é possível observar os pesquisadores e o ano da publicação de seus estudos, além das unidades que descrevem e ainda, o termo que utilizam para se referirem aos fraseologismos. Deve-se considerar também o fato de que, sendo uma disciplina dentro da área de humanas, a maioria dos conceitos que envolvem linguagem naturalmente estão sujeitos a várias interpretações, outro fator determinante na profusão de termos e nas inúmeras tentativas de definir tais fenômenos linguísticos.

Adicione às dificuldades iniciais já expostas, o fato de diversos autores cunharem termos díspares para fenômenos semelhantes e o desafio de descrever os fraseologismos estará lançado.

A fim de oferecer uma olhar mais didático a esta questão, pretende-se efetuar o seguinte percurso: primeiramente apontar as características dos fraseologismos; em seguida, realizar uma classificação das unidades fraseológicas (UFs), e por fim, focar a discussão nas colocações.

Corroborando, este entendimento os postulados de Orenha-Ottaiano (2004), que defende o fato da convencionalidade se relacionar à fraseologia, conforme indica abaixo:

Figura 04 – Convencionalidade e Fraseologia



Fonte: Orenha-Ottaiano (2004, p. 13)

Para justificar o fato de a Fraseologia estar contida na convencionalidade, Orenha-Ottaiano (2004) afirma que, enquanto a primeira trata das “combinações estáveis da língua”, a segunda é mais ampla por se responsabilizar pelos usos e costumes utilizados pela sociedade. Ou seja, a primeira lida com aqueles que representam a cultura e memória coletiva de uma dada comunidade de falantes, enquanto a segunda engloba todo e qualquer grupo de agrupamentos linguísticos.

Dentro desta revisão teórica, destacamos o trabalho de Corpas Pastor (1996). A autora traz o conceito de que a Fraseologia é formada por expressões linguísticas concebidas por unidades poliléxicas (grupo de palavras) ou UFs. Estas possuem, por sua vez, idiomaticidade e grau diferentes, podendo existir em díspares aspectos.

Segundo a autora (1996, p. 20-32), há várias características que se destacam nestas unidades, cujo resumo apresenta-se abaixo:

i. Frequência. Diz respeito à quantidade de ocorrências dos elementos que as formam. Para ser considerada uma UF, a frequência conjunta dos itens que a identificam deve ser superior à aparição individual dos mesmos.

ii. Institucionalização. O primeiro item conseqüentemente leva ao segundo, ou seja, a partir da reiteração da frequência mencionada em *i*, ficará comprovado que a taxa de frequência dos dados enquanto expressão é significativa, ou ainda, indicativa que o uso da UF já é consagrado pelo uso por seus falantes.

iii. Fixação. Ao validar a institucionalização de uma unidade fraseológica, o caráter de fixidez também será atestado, legitimando novamente a utilização maciça do termo pela comunidade.

iv. Idiomaticidade. É observado ao verificar que os termos em separado não alcançam o significado desejado, diferentemente do que ocorre ao escrever ou falar o bloco de palavras em conjunto.

v. Variação. É importante lembrar que mesmo possuindo a fixidez descrita em iii., as UFs podem ser variáveis, por questões de adaptação de pessoa em número e gênero, por exemplo. No entanto, a semântica da expressão não pode ser alterada.

vi. Gradação. A gradação ocorre em consequência da variação – pois se a unidade fraseológica admite variação, também haverá uma gradação que indique uma leve mudança na UF original.

Outros autores, como Rodríguez (2004), fazem uma acepção das UFs levando em consideração o grupo de falantes que as enunciam. Neste segmento, vemos que a frequência é o destaque desta análise, conforme atestam as palavras da pesquisadora abaixo:

Se trata de construções anteriores ao ato de fala, produto de processos de repetição na diacronia da língua, que muitas vezes são difíceis de reconhecer na língua materna, posto que pertencem a memória coletiva faz com que eles passem despercebidas. Opostamente, eles são os elementos de interpretação mais caras, e que mais chama a atenção quando enfrentam para a aprendizagem de uma língua estrangeira¹⁴. (p. 9)

Este conceito é interessante, pois considera a ‘memória coletiva’ dos falantes, um fato linguístico que aparece com frequência na redação das crianças. Visto que os alunos participantes desta pesquisa possuem entre sete e nove anos de idade, esse traço pode, nas palavras da autora ‘passar despercebido’, porém é um índice bastante significativo da cultura e léxico de uma dada comunidade linguística.

Assim, o termo ‘diacronia’ é utilizado para indicar a descrição da linguagem das pessoas, o que também sugere a passagem de tempo durante este processo. Fatores como lugar social e uso podem justificar mudanças provenientes de textos de diferentes locais,

¹⁴ *Se trata de construcciones anteriores al acto de habla, producto de procesos de repetición en la diacronía de la lengua, que a menudo es difícil reconocer en la lengua materna, pues su pertenencia a la memoria colectiva hace que pasen inadvertidas. Opuestamente, son los elementos de más costosa interpretación, y que más llaman la atención al enfrentarse al aprendizaje de una lengua extranjera.*

conforme atestam os textos analisados nesta dissertação e que serão estudados nos próximos tópicos.

Visto que a definição das unidades fraseológicas perpassa vários aspectos, a autora acrescenta a fixação (p. 16-18), idiomaticidade (p. 18-20) e o significado pragmático ao rol. Nas palavras de Rodríguez (2004):

Esta é uma propriedade funcional em virtude da qual certas UFS, dotadas de conotações expressivas, influenciam a estrutura geral da conversação, realizando a argumentação com diferentes matizes dependendo do caso (zombaria, ironia, exortação, ameaça, admiração, indignação, etc.)¹⁵ (p. 22)

Por significado pragmático, entendemos que a linguagem adquirida não é somente uma aprendizagem de língua, mas incluem crenças e valores do indivíduo, expressos, igualmente, por meio da linguagem. Da mesma maneira que o ser humano precisa possuir discernimento linguístico para proferir enunciados carregados de carga semântica, o interlocutor também necessita compreender as nuances presentes na mensagem, sob pena de não entender o referente.

Com estes exemplos, queremos apontar que a caracterização das UFs sempre apontará diversos pontos de vista, já que estão intrinsecamente relacionadas entre si. Dessa forma, agruparemos na sequência os fraseologismos a partir do olhar destas duas autoras elencadas acima, visto que trazem à baila tanto os aspectos mais formais das unidades fraseológicas quanto os mais sociais, o que contempla o público alvo deste estudo, a saber, os alunos da escola do centro e os que estudam na área periférica da cidade de Votuporanga, estado de São Paulo.

Após a descrição do panorama geral dos estudos fraseológicos, o enfoque deste penúltimo tópico teórico será o tratamento de alguns dos tipos de fraseologismos existentes, em especial do grupo de palavras que nos é caro neste trabalho, conforme dito em páginas anteriores; as colocações.

¹⁵ *Es ésta una propiedad funcional en virtud de la cual determinadas UFS, dotadas de connotaciones expresivas, tienen una influencia sobre la estructura general de la conversación, realzando la argumentación con diferentes matices según el caso (burla, ironía, exhortación, amenaza, admiración, indignación, etc.)*

1.4 Fraseologismos: as Colocações

Dentro da proposta desta pesquisa, continuamos o percurso teórico com as colocações. Conforme dito sobre a convencionalidade do signo, sabemos que as crianças aprendem a falar desde a mais tenra idade. Retomando Orenha-Ottaiano (2004), observa-se que a importância da aprendizagem dos fraseologismos dá-se desde a infância, justificado em:

Na realidade, aprendemos nossa língua materna dessa maneira, ou seja, em blocos pré-fabricados, como combinatórias prontas que são produzidas de modo automático, sem refletir, de forma inconsciente. Isso se faz necessário pois, se a cada vez que discursássemos tivéssemos de formar novamente uma combinação de palavras, além do tempo que perderíamos para elaborar essa combinação, nosso discurso se tornaria enfadonho. Sob o ponto de vista do ouvinte, como este poderia decodificar um discurso se, a todo momento, o falante criasse novos blocos de palavras? Dessa maneira, o processo de comunicação se tornaria bastante cansativo, tanto por parte do falante quanto do ouvinte. (p. 18)

Dessa forma, é graças às características da convencionalidade, como frequência, institucionalização e fixação, por exemplo, que a criança aprende a articular os blocos pré-fabricados de palavras ao ouvir o discurso daqueles que a cercam. E é na reiteração da escuta atenta, articulada com a fala cotidiana, que este conjunto de vocábulos se constitui numa linguagem natural, aprendida por meio da interação entre pessoas de uma mesma comunidade de falantes.

Por isso, a abordagem dos atributos aplicáveis aos fraseologismos explanados anteriormente é importante, já que nos auxilia a encontrar quais são as delimitações de aplicação dos termos em estudo. Ainda é comum, dentro da sala de aula, nós professores ouvirmos vários questionamentos dos alunos de ensino fundamental I acerca do significado de determinadas palavras e expressões. Isso se dá, muitas vezes, devido ao fato de as crianças ainda desconhecerem certos arranjos existentes no léxico, justamente em se tratando das inúmeras possibilidades de agrupar termos da língua portuguesa do Brasil.

Visto ser necessário realizar um recorte metodológico sobre qual fraseologismo contemplar nas análises dos textos coletados para esta dissertação, a escolha deu-se pelas colocações, já que foi o fraseologismo mais utilizado pelos estudantes dos dois grupos de alunos em estudo.

A seguir, realizamos uma releitura dos principais postulados desta combinatória linguística, desde as diversas definições consagradas das colocações, passando pelas

taxonomias adjacentes e, na medida do possível, tecemos as relações deste tema com o nosso trabalho, contextualizando a discussão.

Uma das primeiras características acerca das colocações a obter destaque é a famosa citação de Firth (1957) sobre a companhia que as palavras mantêm entre si: “Uma palavra é conhecida pela companhia que possui¹⁶”. Essa ‘companhia’ refere-se à forma de articulação das palavras, variável, de acordo com a regularidade com as quais se agrupam. Para um grupo de nativos, a linguagem já é ‘aceita’, sem questionamentos por parte de seus falantes, atestando os critérios de convencionalidade expostos outrora.

Já Lewis (2000a, p. 130) traz a questão da opacidade e transparência nas colocações e expressões idiomáticas. O autor afirma que, se algo é considerado como idiomático ou não, não é decidido por um único fator. A maioria dos linguistas reconhece que há pelo menos dois fatores – sua posição está em um espectro com “fixo” ou “invariável” e “variável” e culmina no segundo espectro, entre a semântica “opacidade” e “transparência”. O autor ainda traz exemplos do *The Oxford Dictionary of Current Idiomatic English* para descrever a gradação de idiomatidade, do maior grau de fixação para o menor:

– Expressão Idiomática Pura: como em *blow the gaff* (revelar um segredo).

São as expressões mais fixas que existem, são quase invariáveis e perdem qualquer sentido literal de interpretação, então são semanticamente opacas.

– Expressão Idiomática Figurativa: como em *catch fire* (de repente, tornar-se interessado ou interessante) e não somente “pegar fogo”; ou a *close shave* (escapar de uma situação por um triz) e não somente “barbear rente”.

Estas últimas são menos opacas, pois, apesar de terem um significado não-literal, podem também ser usadas em seu sentido literal.

Com base nos exemplos acima, observamos que essas expressões possuem outro significado (mais ou menos idiomático), além do literal.

No que tange às colocações, Lewis (2000a, p. 130) as classifica em:

– Colocações Restritivas: ilustrada em *jog someone’s memory* (refrescar a memória de uma pessoa = ajudar uma pessoa a recordar-se de algo).

¹⁶ *You shall know a word by the company it keeps.*

Tais colocações possuem um elemento usado no sentido não literal e outro no sentido denotativo.

– Colocações Abertas: envolve elementos que são (mais ou menos) combinações livres, sendo que cada elemento tem o seu sentido literal.

Para prosseguir na linha de raciocínio proposta nesta pesquisa, utilizamos a acepção de Tagnin (1998), que define colocações como sendo: “combinações lexicais recorrentes, não-idiomáticas, coesas, cujos constituintes são contextualmente restritos e de coocorrência arbitrária” (p. 41).

Essa decisão é justificada pela explicação da autora em manter como ponto comum a conexão intrínseca entre as palavras e o acordo social entre os membros que falam determinado idioma.

Para aprofundar a reflexão sobre a citação acima, retomemos as noções de combinabilidade, trazidas pela autora no nível sintático da convencionalidade: “à faculdade que os elementos linguísticos têm de se combinar. Existem palavras que se associam de forma tão natural que a única explicação possível é de que essa associação tenha sido consagrada pelo uso, isto é, de que ela seja convencional” (TAGNIN, 2013, p. 25-26).

Desta forma, conhecer as palavras que coocorrem com mais frequência torna a comunicação mais fluída, pois não é necessário interrompê-la para explicar o sentido das palavras que o enunciador está empregando. Da mesma maneira, a pessoa que ouve o discurso precisa estar ciente do código e suas articulações, a fim de que o diálogo seja pleno de sentido para ambos.

Analogamente, as colocações também se combinam por convenção, não havendo uma justificativa que as motivem, a não ser o hábito dos falantes em estabelecer essa ordem de reunir as palavras. A fim de caracterizar a taxonomia desta relação, a autora elenca critérios gramaticais para diferenciá-las, o que oferece um respaldo metodológico interessante para a realização das análises no capítulo três desta dissertação.

Segundo Tagnin (2013, p. 64-73), as colocações se dividem em verbais, nominais, adjetivas e adverbiais. Trazemos abaixo uma breve caracterização de cada uma delas, com exemplos retirados do CRAEF:

- Colocações Verbais

Há aquelas que ocorrem com determinados substantivos, como “tomar banho”, “tirar

foto” e “levar picada”, por exemplo. Outro padrão típico das colocações verbais está relacionado àquelas que necessitam de uma preposição: “acender a luz”, “brincar de corrida” e “entrar em greve”.

Mais um formato que a autora destaca são as correspondências naturais que certos verbos possuem ao se combinar com adjetivos. Como exemplificações, temos “viveram felizes para sempre”, “Tá legal” e “ficar doente”.

- Colocações Nominais

Uma das combinatórias decorre da formação de dois substantivos – “igreja matriz” e “concha acústica”.

A outra é formada por meio de preposição, que interliga os termos da colocação nominal: “terra do jardim”, “bolo de milho” e “moradores de rua”.

- Colocações Adjetivas

É a combinação de um adjetivo com um substantivo, como em: “boas ações”, “medo de escuro” e “chuva forte”.

- Colocações Adverbiais

Nesta última, temos duas possibilidades da articulação. Na primeira, o advérbio possui a função de mudar o sentido do adjetivo, presentes em: “bem cara”, “soro meio amarelado” e “estava muito escuro”. Na segunda, o advérbio é responsável por modificar o verbo – “morrendo de medo”, “saiu correndo” e “foi devagarinho no banheiro”.

Para explicar melhor as diferentes disposições deste fraseologismo, apresentamos em formato de quadro os quatro tipos de colocações que Tagnin (2013) elenca, com outros exemplos retirados dos corpora de estudo investigados nesta pesquisa.

Dessa maneira, pretendemos demonstrar que, apesar de se expressarem de maneira simples e coloquial, as crianças são capazes de produzir colocações, relatando, na maioria dos casos, eventos próprios da idade, ou referentes a infância, conforme seguem:

Quadro 02 – Tipologia das colocações

Tipo de colocação	Estrutura	Exemplos
Colocações Verbais	V + S	<i>Empinar pipa, tomar soro e fazer falta</i>
	S + V	<i>A plantinha cresceu e a moeda brilhou</i>
	V + Prep + S	<i>Orar a Deus, andar de skate</i>
Colocações Nominais	V + Adj.	<i>Estar internada e nascer rápido</i>
	S + S	<i>Praça matriz e câmara municipal</i>
	S + Prep + S	<i>Flor na mão, fundo do coração e moeda de ouro</i>
Colocações Adjetivas	S + Adj.	<i>Guarda temporária da criança, sacola plástica e senhorinha velhinha</i>
Colocações Adverbiais	Adv + Adj	<i>Paixão sem fim e muito agradecida</i>
	V + Adv	<i>Ficar muito mau e sair discretamente</i>

Fonte: Autora, relacionando as classificações e formações das colocações expostas por Tagnin (2013) e exemplificando com trechos retirados do CRAEF 1 e CRAEF 2

Nesta categorização, é possível observar que, enquanto um item determina a colocação, o outro é determinado, ou seja, há uma hierarquia entre os termos que formam esta unidade fraseológica – reiterando o formato de unidades pré-fabricadas. Nos dizeres de Orenha-Ottaiano (2012), com base nos preceitos de Hausman (1997): “Aquele que determina é chamado de base, que é o elemento autônomo, enquanto que o outro, o determinado, é chamado de colocado, que somente pode ser interpretado semanticamente quando na colocação” (p. 152). Assim, o termo colocado assume o caráter mais abstrato da expressão, o que justifica a divisão das colocações em verbais, nominais, adjetivas e adverbiais.

Relacionando o estudo das colocações à utilização da tecnologia, temos a ampliação das possibilidades investigativas da linguagem. Neste sentido, o uso de corpora informatizado faz com que tanto a LC quanto a Fraseologia possuam procedimentos bastante peculiares. Gries (2006) é um dos autores que discutem este assunto:

A análise procede baseada na lista de frequência (de palavras, morfemas, padrões gramaticais, etc.), linhas de concordância, em que a palavra de interesse é mostrada em seu contexto natural, e as colocações, ou seja, listas ou tabelas nas quais a palavra de interesse (a mais frequente) aparece

rodeada por todas as palavras que lhe são vizinhas¹⁷. (p. 4)

Graças às linhas de concordâncias geradas, é possível analisarmos, dentre as combinações de palavras, as colocações. Este grupo específico dos fraseologismos revela, por meio da relação sintática e semântica entre os vocábulos que as constituem, a variedade lexical dos idiomas.

O advento do computador, por sua vez, trouxe ferramentas computacionais que possibilitaram um impulso nas análises linguísticas, e em especial nos estudos acerca das colocações, uma vez que identifica itens como frequência, palavras-chave, termos fraseológicos, dados estatísticos, dentre outros aspectos relevantes em uma pesquisa acadêmica.

Neste segmento das abordagens que utilizam colocações em investigações da linguagem, Martelli (2007) descreve duas, a saber: a Abordagem Baseada na Frequência (*Frequency-based approach*) e a Abordagem Fraseológica (*Phraseological Approach*).

A primeira, Abordagem Baseada na Frequência, de cunho firthiano, privilegia, como o nome acusa, a frequência das colocações que co-ocorrem no corpus de estudo. A discussão mais comumente realizada é sobre os dados estatísticos das análises, bem como qual seria a quantidade de palavras a serem analisadas antes e depois das colocações, ou seja, qual a extensão de palavras a serem investigadas, localizadas à esquerda e à direita do termo pesquisado.

Já a Abordagem Fraseológica adentra, por sua vez, no campo da fraseologia e segundo Martelli (2007, p. 16) possui aplicações em lexicografia e pedagogia da linguagem. Outra diferenciação diz respeito à aceitação de outros tipos de combinações de palavras, principalmente expressões idiomáticas e combinações livres.

Embora haja variações na forma como pesquisadores da abordagem fraseológica lidam com a questão de delimitar as colocações, pode-se dizer que o critério da transparência semântica tem sido usado para distinguir as colocações das expressões idiomáticas, enquanto o critério de motivação arbitrária ou semântica tem sido utilizado para diferenciar as colocações das combinações livres.

Neste trabalho, privilegamos a segunda abordagem, já que o foco é analisar como os fraseologismos (colocações) aparecem no CRAEF e qual grupo os utiliza de modo mais frequente, se aquele dos discentes do centro ou da periferia.

¹⁷ (...) - *the analysis proceeds on the basis of frequency lists (of words, morphemes, grammatical patterns, etc.), concordance lines in which the word of interest is shown in its natural context, and collocations, i.e. lists or tables in which for the word of interest the (most frequent) neighboring words are given.*

Outro apontamento interessante trazido por Martelli (2007, p. 33) é o ensino de colocações para aprendizes de segunda língua. Ela afirma que o conhecimento colocacional é desejável, pois o uso correto dessas combinatórias em um contexto adequado de comunicação caracteriza um aluno fluente no idioma alvo.

Fazemos esta breve comparação entre o ensino de idioma estrangeiro e o de língua materna, para demonstrar que, em ambos os casos, os alunos possuem dificuldades em compreender os arranjos entre as palavras de um dado código linguístico. Também destacamos o importante papel do professor em ensinar as colocações às crianças partir da exposição de situações que pedem o uso destes fraseologismos.

Ainda dentro do tema, o conceito de erro colocacional deriva da produção de combinações equivocadas, geralmente produzidas por um aluno que ainda desconheça as sequências lexicais compatíveis com a mensagem a ser transmitida. No entanto, é preciso verificar a razão que leva o discente a elaborar erroneamente tais enunciados, a fim de auxiliá-lo a se comunicar melhor.

Neste sentido, Martelli (p. 34) cita Corder (1981, p. 11) e os três motivos pelos quais é importante investigar as prováveis causas dos desvios linguísticos desta natureza:

- i. observar o que o aluno assimilou e o que falta para o mesmo aprender;
- ii. com o olhar de pesquisador, constatar como a língua estrangeira é compreendida pelo aprendiz;
- iii. ver o erro como uma pista das táticas empregadas pelos discentes para testar as hipóteses que possuem da língua alvo.

De posse dessas informações, o professor de idiomas passa a enxergar o erro como uma ocorrência normal dentro do processo de aprendizagem do aluno, que naturalmente precisa perpassar por estas etapas de desenvolvimento linguístico. Corrobora este entendimento o próprio Corder (1981, p. 10) em: “o conceito chave [...] é a de que o aluno está usando um determinado sistema de linguagem em cada ponto do seu desenvolvimento [...]. Os erros do aluno são a evidência deste sistema e eles próprios são sistemáticos¹⁸”.

Analogamente, Tagnin (2002) segue na mesma direção ao retomar o termo “falante ingênuo”, outrora abordado em Fillmore (1979). Apesar de utilizá-lo citando a função do tradutor, a autora traça um paralelo entre os dois, afirmando que ambos não têm consciência dos blocos pré-fabricados da língua, e acrescenta:

¹⁸ *the key concept [...] is that the learner is using a definite system of language at every point in his development [...]. The learner's errors are evidence of this system and are themselves systematic.*

Observe-se também que, dependendo da situação, podemos todos ser falantes ingênuos em nossa própria língua materna. Como poderia um leigo conhecer os termos técnicos (na maioria, colocações) de certas profissões como medicina ou direito? Ou, como saberíamos o que dizer (usar as fórmulas corretas) em situações desconhecidas como, por exemplo, um velório, se jamais fomos a um? (2002, p. 192)

Essas indagações são pertinentes para esta pesquisa, visto que a criança, ainda em fase de aprendizagem do letramento, também se encontra na posição de “falante ingênuo”. Dessa forma, paralelamente ao aprendiz de segunda língua ou ao tradutor iniciante, o aluno de ensino fundamental I também precisa ser incentivado a conhecer as combinatórias de palavras, a fim de alcançar uma proficiência em sua própria língua, no que tange à fala, à escrita e à compreensão das colocações nas atividades cotidianas.

Pedagogicamente, para Lewis (2000a, p.132) as colocações coocorrem naturalmente na língua e, exatamente por esse motivo, os alunos precisam aprender como “colocar as palavras juntas”, de modo a produzir colocações comumente empregadas na língua. De acordo com o autor, se os vocábulos formam uma combinação única, estas são diferentes de compreender o funcionamento das mesmas em separado. É consideravelmente mais difícil para o aprendiz unir as palavras a fim de formar colocações.

Neste sentido, Lewis (2000a) afirma:

Qualquer análise de fala ou escrita dos alunos revela uma falta de competência colocacional. A falta desta competência nesta área compele os alunos a cometer erros gramaticais, pois criam enunciados mais longos devido ao fato de não conhecerem as colocações que expressam exatamente o que querem dizer. Os professores, muitas vezes, focam na correção de erros gramaticais, falhando em perceber que isso não fará nenhuma diferença – os desvios não são causados por culpa da gramática, mas pela falta das colocações¹⁹. (p. 49)

Uma prova desta afirmação é a ocorrência das fórmulas de rotina nos textos dos alunos. Essa propriedade é mais observada nos enunciados fraseológicos, descritos por Corpas Pastor (1996). Remetem-se aos enunciados fraseológicos, se caracterizando por aparecerem no discurso falado, pertencendo, portanto, ao arcabouço social da linguagem de uma dada comunidade que partilha de uma mesma identidade cultural e linguística. A estrutura é fixa,

¹⁹ *Any analysis of students' speech or writing shows a lack of this collocational competence. Lack of competence in this area forces students into grammatical mistakes because they create longer utterances because they do not know the collocations which express precisely what they want to say. Teachers often then focus on correcting the grammar mistakes, failing to realize that it will make no difference – the mistakes are not made because of faulty grammar but a lack of collocations.*

apresentando forma estruturada e conteúdo autônomo, embora precise se relacionar ao contexto para ser entendida.

Notadamente usados no discurso oral, estes itens estão muito marcados aqui, nas produções escritas coletadas. É válido ressaltar o caráter pragmático destas unidades, haja vista que buscam marcar uma identificação maior com o interlocutor, conforme exemplifica, Orenha-Ottaiano (2004):

É por isso que, quando esbarramos em alguém, por exemplo, dizemos “desculpe”. Essa expressão faz parte de um costume social consagrado e previsível; não fazer uso dela, nessa situação, implicaria falta de educação por parte daquele que cometeu o ato de esbarrar, pois uma pessoa considerada educada, de acordo com as regras sociais estabelecidas, pediria desculpas automaticamente. (p. 9)

Nesta ilustração, a palavra “desculpe” se caracteriza como uma fórmula de rotina, intrinsecamente ligada às normas sociais e o que se espera em termos de uma comunicação saudável entre falantes. Em outras palavras, as nuances identificadas entre diversos grupos de falantes de um mesmo vernáculo, nos levam a crer na existência de um acordo tácito entre membros, que legitimam a língua conforme o uso de determinadas combinações linguísticas consolidadas no cotidiano.

Essa definição é corroborada por Rodríguez (2004, p. 28), que se refere a este fraseologismo como *las fórmulas rutinarias*. Para a autora, as fórmulas de rotina também se caracterizam pela interação social, ou seja, só ocorrem em determinadas situações que exijam certa convenção entre pares. Essa característica esteve presente nos dois corpora coletados, conforme seguem alguns exemplos do CRAEF:

Fórmulas de abertura – visam facilitar o discurso, iniciando a interação ou proporcionando a fluência durante a conversa com o interlocutor: “Oi amigo, tudo bem?”.

Fórmulas de encerramento – como o nome sugere, encerra o diálogo, de forma respeitosa e sem deixar dúvidas sobre o fim da interação: “De nada”; “Tá bom”.

Fórmulas de Transição – novamente relacionadas às questões verbais entre indivíduos, este item é responsável por regular a conversa. Podem oferecer espaço para o interlocutor expressar sua opinião (“Em que posso ajudar?”); iniciar outro tópico da conversação ou testar o canal de comunicação: “Você é mudo? Me responde!”.

Já as fórmulas psicossociais têm o objetivo de expressar os sentimentos do emissor do discurso, marcando a carga semântica envolvida nesta interação, diferenciando-as em:

Fórmulas expressivas – indicando sentimentos ou atitudes do emissor: “Oba!”; “Caramba!”.

Fórmulas comissivas – remetem a algum tipo de “comprometimento” firmado com o interlocutor, geralmente empregados no tempo futuro, tais como promessas e ameaças: “Vamos brincar!”; “Seu surdo, vai ficar deitado”.

Fórmulas diretivas – visam a persuadir o enunciatário, a partir dos indícios de exortação, informação e ânimo durante o discurso: “Quando você pode ir na minha casa?”.

Fórmulas de rituais de saudação – divididas em saudações, empregadas no início da conversação e despedidas, ligadas às fórmulas de abertura e fechamento, já elencadas acima: “Bom dia!”; “Feliz Natal!”.

Uma justificativa para esse uso diz respeito ao fato de os alunos, ainda em processo de alfabetização, apresentar dificuldades de compreender as diferenças entre a linguagem oral e a escrita. Assim, ao redigir textos, ocasionalmente produzem fraseologismos típicos da oralidade, como as fórmulas de rotina, que não deveriam aparecer em uma modalidade escrita. Fazemos uma pequena ressalva para falar sobre o gênero proposto, visto que em uma história em quadrinhos, é desejável que transpareçam marcas orais, posto que a interação entre os personagens pressupõe este tipo de registro ou em um diálogo, numa história fictícia. Contudo, em um relatório sobre um filme ou na descrição sobre a cidade, a linguagem esperada é mais formal, e portanto, dispensa o uso de fórmulas de rotina.

Retomando a aprendizagem de primeiro idioma, testificamos a importância do professor de língua portuguesa em atentar para o ensino de fraseologismos na escrita de seus alunos.

Por isso, destacamos a importância de tratarmos deste tema da produção textual com a coleta de corpus de aprendizes infantil, pois é preciso observar como os discentes do ensino fundamental I compreendem as colocações. A partir desta investigação, traçar estratégias para oportunizar a ampliação da competência lexical, como meio de melhorar a redação de textos, de modo a adquirir um léxico mais rico e significativo em língua portuguesa do Brasil.

Neste contexto, o último tópico da fundamentação teórica versa sobre a aplicação da Aprendizagem Movidada por Dados na elaboração de atividades que objetivam ampliar a competência lexical dos estudantes de ensino fundamental I, conforme segue.

1.5 Aprendizagem Movidada por dados para a elaboração de atividades baseadas em corpora

Embora a abordagem AMD já fora mencionada nesta dissertação no tópico da LC direcionada ao ensino, acreditamos ser necessário aprofundar a exploração deste tema. Realizamos esta assertiva, pois, a partir das análises efetuadas, constatamos que, de modo geral, as crianças de ambas as escolas produziram poucas colocações em suas produções textuais.

Tal fato nos inquietou a ponto de pensarmos em como auxiliar os alunos a perceberem que as colocações já estão presentes no seu dia a dia, em diversos contextos e situações comunicativas. O resultado das reflexões nos levou a aplicar os conceitos explanados pela AMD, já que, dos três segmentos de ensino abordados na fundamentação teórica, este é o que mais se aproximou dos nossos objetivos metodológicos.

Retomando a explanação feita em páginas anteriores, a AMD visa a destacar o papel do aluno enquanto responsável pela sua aprendizagem, enquanto utiliza exemplos da língua para compreender melhor o funcionamento da mesma. Para caracterizá-la, Kennedy (1991, p. 110) não menospreza as informações trazidas em gramáticas e dicionários, mas afirma que “a AMD oferece uma descrição linguística por meio das estatísticas que servem tanto para apoiar quanto para contradizer as intuições dos alunos sobre o uso da língua²⁰”.

Ou seja, embora os livros didáticos tenham a sua representatividade no ensino, de uma forma geral, apresentam o conteúdo de uma forma bastante formal, o que pode gerar um entendimento de que a linguagem só exista por meio de regras normativas e estáticas.

Isso ocorre, pois, a escolha em descrever a estrutura da língua nem sempre permitirá uma abertura para o uso de determinadas construções linguísticas pelos falantes; até porque, um material didático privilegia um ou outro campo de estudos, sendo impossível abarcar todas as vertentes de um idioma em um único volume.

A utilização de corpora, ao contrário, mostra as possibilidades de combinar a língua,

²⁰ *the DDL provides a statistical aspect of linguistic description to both support and contradict the learners' intuitions about the use of language.*

nem sempre discriminadas nos manuais tradicionais. O próximo passo nesta empreita é atingir o ponto destacado por Teixeira e Chishman (2008, p. 248): “É no corpus que os alunos esclarecem as suas dúvidas quanto às especificidades da língua, e não mais simplesmente fazendo questionamentos ao professor”. Aos poucos, os alunos serão aptos a compreender melhor o funcionamento das colocações, como combinatórias de palavras recorrentes.

A partir da proposta de que o próprio aluno interprete os dados existentes nos corpora coletados, os discentes seriam capazes de procurar, questionar e refletir sobre os padrões léxico-gramaticais, o que o auxiliaria sobremaneira na ampliação destes conteúdos.

Por isso, os procedimentos introdutórios de apresentação dos dados explicitados anteriormente são importantes, a fim de garantir a autonomia dos alunos frente aos dados. Ao evitar que o entendimento das informações com corpora se restrinja a um baixo número de estudantes, a turma aprende a identificar as colocações empregadas pelos falantes, reverberando uma aprendizagem mais profícua do idioma materno.

A AMD também é vantajosa devido à ligação que proporciona entre teoria e prática no ensino da língua portuguesa do Brasil, ainda mais se considerarmos o contexto real da produção de enunciados. Aliada ao uso de recursos tecnológicos, a interação entre homem e máquina pode favorecer o interesse na busca do saber, não só da gramática, mas também, da ampliação do léxico e a compreensão das combinatórias existentes em sua língua.

Contudo, o docente que optar por essa forma de atuação, não deve perder de vista os procedimentos metodológicos mais apropriados para utilizar com a sua turma de alunos. De acordo com Boulton (2011):

Em muitos casos, é provável que no início, o professor continue a ser o centro das atenções, decidindo os pontos de estudo da língua, os corpora e ferramentas a utilizar, as atividades, a progressão, e assim por diante [...]. Isso também pode ser uma forma de tornar as informações mais acessíveis como uma alternativa para simplificar o corpus propriamente dito, como as linhas de concordância podem ser selecionadas, classificadas e organizadas com antecedência. (p. 574)²¹

Com efeito, nesta trajetória rumo ao conhecimento, não basta simplesmente coletar dados. Antes, as informações linguísticas devem estar em concordância com os propósitos do observador. Por meio de uma linguagem autêntica, é possível observar quais combinações são mais recorrentes e utilizadas com mais frequência pela comunidade de falantes de um dado

²¹ *In many cases, the teacher is likely to remain the centre of attention at the start, deciding the language points to study, the corpora and tools to use, the activities, progression, and so on [...]. This can also be a way to make the data more accessible as an alternative to simplifying the corpus itself, as concordance lines can be selected, graded and organised in advance.*

idioma.

Cobb e Boulton (2015, p. 481) por sua vez, afirmam “que tal atividade pode ser muito exigente com o aluno, que provavelmente necessitará de um treinamento intensivo [...] a fim de colher todos os benefícios do corpus de consulta²²”. Assim, o objetivo é tornar o aluno pesquisador da língua a qual estuda, enquanto o docente se torna um intermediador ou facilitador tanto no uso das ferramentas computacionais quanto na compreensão das informações geradas.

Esta é a forma usada pelos pesquisadores para investigar o léxico, o que nos leva ao método empírico-indutivo da AMD. Este trata as informações do particular para o geral, ou seja, a partir do corpus, o pesquisador chega a conclusões empíricas acerca do tema.

Contextualizando o exposto para o Ensino Fundamental I, não podemos deixar de destacar ainda mais o papel do professor, pois será ele que fará a intermediação entre os corpora e os alunos. Também será este profissional que elaborará as atividades pedagógicas com a utilização de recursos tecnológicos da LC, que proporcionarão aos alunos descobrirem os padrões léxico-gramaticais por meio de linhas de concordância. Porém, para viabilizar estes procedimentos, é necessária a união de uma série de fatores, que incluem:

- um profissional com conhecimentos teóricos sobre a LC e suas possibilidades didáticas voltadas ao ensino, a fim de realizar a compilação dos dados e transformá-los em exercícios de fácil compreensão para o aluno;
- a existência, na unidade escolar, de um laboratório de informática composto por um número de computadores suficiente para atender a demanda. Além disso, é desejável que este local tenha acesso à *Internet* de alta conectividade para a realização de pesquisas via *online*.

Para tanto, retomamos o exposto na introdução desta pesquisa, sobre a falta de formação destes profissionais, não só na área de Letras, mas também nos cursos de licenciatura de pedagogia. Ora, se o principal objetivo do pedagogo é alfabetizar os alunos, porque não investir na capacitação deste futuro professor em abordagens modernas e proficuas como a AMD e demais segmentos?

Deixamos em aberto este questionamento, haja vista que, para respondê-lo, teremos que adentrar em esferas que, infelizmente, fogem de nossa alçada de pesquisadora, cuja resolução não se encontra em nossas mãos.

²² *such activity can be quite demanding on the learner, who is likely to need intensive training [...] in order to reap the full benefits of corpus consultation.*

Nesta dissertação, a linguagem autêntica das quais falam a AMD são as próprias redações elaboradas pelos alunos. Após os procedimentos de coleta, descritos a seguir no capítulo metodológico, observamos as colocações empregadas pelos discentes para recortar quais seriam mais adequadas na elaboração de atividades baseada em corpus. Em seguida, utilizamos o Corpus do Português²³ para adicionar outras e, assim, propor a ampliação da competência colocacional dos alunos, descrita em detalhes no capítulo 4 deste texto.

Portanto, a AMD foi a forma mais profícua para assegurar que o foco das atividades fossem as colocações, graças aos recursos existentes em seu bojo. Fazemos tal afirmação, pois a aprendizagem de um idioma, seja em língua materna ou estrangeira, pressupõe que o estudante articule o léxico para se comunicar, seja oralmente ou por meio escrito, como é caso desta pesquisa.

Dessa maneira, é necessário fazer o uso coerente das palavras, bem como saber como relacioná-las no cotidiano, frente aos diferentes contextos discursivos. Por isso, a escola ainda se configura como o ambiente ideal para a aprendizagem deste conhecimento tão necessário a sua comunicação.

Contudo, Monteiro-Plantin (2011) ressalta que, apesar de o Brasil ter avançado na divulgação de trabalhos acadêmicos e congressos sobre fraseologia, ainda há poucos estudos voltados para o português do Brasil. Os fatores apontados para essa ocorrência vão desde a falta de material específico para esse público-alvo até a orientação de não trabalhar com as unidades fraseológicas, pelo entendimento errôneo de que estas demonstram falta de ‘criatividade’ do aluno.

Mesmo sendo os dizeres da autora datados do ano de 2011, enquanto observamos um número crescentes de pesquisas voltadas à fraseologia, ainda são escassas as pesquisas que enfocam os fraseologismos no âmbito infantil, principalmente no que concerne às colocações e sua importância na aprendizagem desde o ensino fundamental.

Haja vista o número de alunos que não conseguem expressar suas ideias até mesmo na língua materna, é lamentável que o estudo e, conseqüentemente, o ensino de fraseologia não sejam mais incentivados em nosso país. Afinal, com a exploração dos padrões léxico-gramaticais da língua portuguesa, os índices de alunos letrados e que se expressam satisfatoriamente em sua escrita textual provavelmente seriam maiores em relação aos que possuímos atualmente.

Com o encerramento deste tópico, terminamos a revisão teórica dos fundamentos

²³ É um corpus linguístico resultante da compilação de textos da língua portuguesa, que compreendia (até 2017), 1 bilhão de palavras.

principais do nosso trabalho. Iniciamos com as motivações pessoais para a escolha do tema, perpassando pelas bases da LC, com ênfase nas pesquisas com corpora de aprendizes, em especial nas pesquisas com redações de crianças. Por fim, realizamos incursões sobre a fraseologia, bem como sobre os fraseologismos, com ênfase nas colocações.

Segue o capítulo sobre Metodologia, que traz dados referentes aos textos que compõem os corpora de estudo e a descrição dos dois grupos de estudantes participantes deste trabalho e o posterior tratamento das informações coletadas.

2. METODOLOGIA

O enfoque deste tópico será a caracterização dos Corpora de Redações de Aprendizês do Ensino Fundamental (CRAEF 1 e 2) quanto à natureza de seus dados. Estão elencadas aqui a descrição dos dois grupos de pesquisa; dos procedimentos metodológicos aplicados às redações, bem como a maneira pela qual os dados em estudo foram coletados e sistematizados para a realização das análises.

Por meio de teóricos como Granger (1998) e Viana (2010), trazemos a descrição sistemática dos dois corpora de estudo. Estas informações são fundamentais para que o leitor compreenda as particularidades da faixa etária investigada nesta pesquisa e obtenha informações complementares sobre as duas turmas de estudantes.

Posteriormente, descreveremos os dois corpora, no que tange à maneira como os dados foram compilados, além das temáticas das redações coletadas. Tal item é significativo, pois a partir destes dados é possível observarmos as diferentes propostas oferecidas aos discentes, e que serão discutidas propriamente no terceiro capítulo.

Finalizando esta etapa, abordaremos pormenorizadamente a manipulação das informações, desde a coleta dos textos até as ferramentas computacionais utilizadas no *software* por meio do qual serão processadas as redações. Neste trajeto, também explanamos as fases de digitalização das produções textuais, identificação dos dados, referenciais usados no programa *WordSmith Tools*, dentre outras informações correlatas.

Assim, delinearemos os contornos metodológicos desta pesquisa, para que tanto os agentes participantes da pesquisa quanto os corpora coletados estejam contextualizados frente às temáticas teóricas já apresentadas no capítulo teórico anterior.

2.1 Caracterização dos Corpora de Redações de Aprendizês do Ensino Fundamental (CRAEF 1 e 2)

Para apresentar as principais peculiaridades dos nossos corpora, traremos as contribuições de dois teóricos que versam sobre este tópico: Viana (2010) e Granger (1998) Dessa forma, oferecemos um olhar mais pormenorizado sobre os aspectos do CRAEF 1 e CRAEF 2.

A partir da taxonomia de corpora, descrita por Viana (2010), o esquema representado

auxilia a compreensão do corpus, possibilitando um olhar mais pormenorizado sobre os itens apresentados.

Figura 05 – Taxonomia de Corpora

Critérios		Corpus		
Abrangência		Geral Especializado		
Meio		Oral Escrito		
Tempo	Número de períodos históricos	Sincrônico Diacrônico		
	Relação com a atualidade	Contemporâneo Histórico		
Renovação		Dinâmico Estático		
Línguas	Número	Monolíngue	Paralelos	Alinhados Não alinhados
		Multilíngue		
	Produtores	De primeira língua (L1) De segunda língua (L2) De língua estrangeira (LE)		
Emprego	Estudo Referência			

Fonte: Viana (2010, p. 30)

Seguindo a classificação do autor, o CRAEF 1 e 2 se enquadram na abrangência geral, pois os textos redigidos pelos alunos não são de nenhuma área específica do conhecimento, principalmente pelo fato de terem sido escritos por crianças de 7 a 9 anos de idade. No que diz respeito ao meio pelo qual foram coletados, trata-se de textos escritos, elaborados pelos próprios alunos nas salas de aulas citadas, sem interferência do meio externo.

Com relação ao item tempo, os corpora de aprendizes CRAEF 1 e 2 foram compilados a partir de textos do mesmo aluno ao longo de um ano (em cinco momentos distintos em 2015), em ambas as escolas citadas. No escopo deste trabalho, a investigação será longitudinal, visto que analisamos a escrita dos dois grupos ao longo do ano letivo citado. Ainda nesse critério, o rol de textos é de caráter contemporâneo, pois foram coletados no ano de 2015.

O CRAEF 1 e o CRAEF 2 também se caracterizam pela dinamicidade, conforme explicitado no parágrafo anterior, bem como por ser monolíngue, visto que utiliza um único

idioma – seus produtores ainda estão no processo de aprendizagem da língua materna. Por fim, o CRAEF 1 e o CRAEF 2 são corpora de estudo, não de referência da língua portuguesa, uma vez que nos fornecerão dados de pesquisa.

A segunda pesquisadora acima citada, Granger (1998), elenca dois critérios que devem ser observados na compilação de um corpus: aspectos relacionados à língua (escrito ou falado; o gênero utilizado; o tema da atividade e quais as condições para a produção deste material) e aqueles que indicam as características do aprendiz (idade, gênero e língua materna, por exemplo).

O tamanho de um corpus (ou corpora) de estudo também é tema recorrente no campo da LC. Haja vista que ainda há uma grande discussão sobre a validade ou não de pesquisas com corpora de tamanho reduzido, acreditamos ser importante realizar uma breve reflexão deste tema. Pelo fato desta pesquisa coletar textos de crianças, o número total dos dados é pequeno em relação a outros trabalhos com corpora de maior volume. Contudo, o tamanho não delimita o valor deste estudo, pois é representativo de um determinado grupo social, posto que os dados analisados são compostos de material linguístico autêntico e real.

Conforme indicam Bowker e Pearson (2002, p. 45-46), apesar de nem sempre ser possível realizar pesquisas com um grande número de dados, é possível encontrar em pequenos corpora muitas informações úteis. Tais investigações e estudos corroboram as afirmações de Trask (2004, p. 68), “a partir de corpora, podem-se fazer observações precisas sobre o real comportamento linguístico de falantes reais, proporcionando informações altamente confiáveis e isentas de opiniões e de julgamentos prévios sobre os fatos de uma língua”.

Outra vantagem é o fato de pequenos corpora possibilitarem que o pesquisador tenha a sua disposição material específico de certas áreas do conhecimento, abrindo caminho para a preparação de materiais de ensino diferenciados. Corrobora para este entendimento o fato do uso de corpora proporcionar, por meio da análise de palavras-chave e linhas de concordâncias geradas, a compreensão sobre os padrões do uso autêntico da língua.

Ainda no âmbito pedagógico, temos a contribuição de Flowerdew (2001, p. 369-376) sobre os usos de pequenos corpora, eficazes na educação para atingir os objetivos de ensinar aos educandos de língua estrangeira, no caso de sua pesquisa, padrões colocacionais, apropriação pragmática e recursos discursivos.

Neste sentido, Koester (2010), traz à baila um dos motivos principais dos pesquisadores utilizarem um corpus de estudo reduzido: a ligação entre o corpus e o contexto onde foi produzido, acrescentando:

Onde corpora muito grandes, por meio de sua descontextualização, apresentam *insights* sobre padrões léxico-gramaticais na língua como um todo, corpora especializados menores oferecem ideias sobre os padrões de uso da língua em cenários específicos. Com um pequeno corpus, o compilador muitas vezes é o próprio analista, e, por conseguinte, normalmente possui um alto grau de familiaridade com o contexto²⁴. (p. 67)

Por meio das redações de dois grupos sociais distintos, como é o caso desta pesquisa, é possível verificar como o léxico se relaciona a cada um, fazendo parte da identidade destes brasileiros. Ragan (2001) também colabora para este entendimento, quando afirma que o tamanho da amostra é menos importante do que a preparação e adaptação do produto linguístico, considerando o perfil de um indivíduo ou grupo no uso da língua (p. 211).

Após efetuar estas considerações, prossigamos com o exposto, elencando os temas das produções dos alunos e também informações sobre as condições de coleta e compilação dos dados no próximo tópico.

2.2 Caracterização das escolas 1 e 2 (Centro e Periferia)

A fim de contextualizar a realidade social existente sobre os agentes participantes da pesquisa, caracterizamos, na sequência, o perfil das escolas onde os textos dos alunos foram coletados.

As redações de estudantes do terceiro ano do Ensino Fundamental que compõem o Corpora de Redações de Aprendizes do Ensino Fundamental (CRAEF) são oriundas de duas escolas do município de Votuporanga. A escola 1, CRAEF 1, localiza-se na zona oeste do município, em um bairro de classe média alta, atendendo, no entanto, alunos de muitos bairros circundantes e também da Zona Rural. Assim, a instituição oferece atendimento a estudantes de diferentes padrões socioeconômicos, desde educandos carentes até aqueles de poder aquisitivo elevado. Desse modo, os textos coletados nesta escola deram origem ao CRAEF 1.

Outro dado interessante deste ambiente refere-se ao grau de escolaridade dos pais ou responsáveis. Com base nos questionários respondidos pelos familiares no momento da

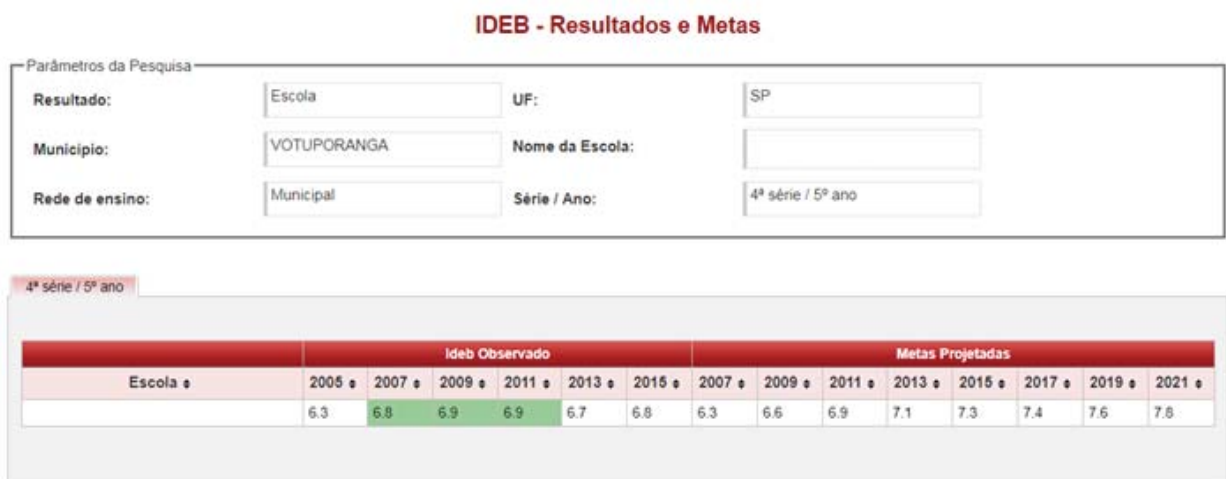
²⁴ *Where very large corpora, through their de-contextualisation, give insights into lexico-grammatical patterns in the language as a whole, smaller specialized corpora give insights into patterns of language use in particular settings. With a small corpus, the corpus compiler is often also the analyst, and therefore usually has a high degree of familiarity with the context.*

matrícula escolar, compreende-se que: 18% dos pais (ou responsáveis) possuem nível superior; 58% cursaram até o Ensino Médio; 15% possuem o Ensino Fundamental Completo; 9% não completaram o Ensino Fundamental e 0% não frequentou a escola.

Em relação à aprendizagem dos alunos, os últimos dados são da avaliação realizada pelo Ministério da Educação (MEC) denominada PROVA BRASIL²⁵. Realizada com os alunos do 5º ano, este procedimento visa a aplicar uma mesma prova a todos os alunos da federação. Dessa forma, é garantido o princípio da isonomia, ao submeter os estudantes ao mesmo processo avaliativo.

Os últimos resultados, inclusos na plataforma do dia 05/09/2016, apontam os seguintes números:

Figura 06 – IDEB da escola do Centro



Obs:

* Número de participantes na Prova Brasil insuficiente para que os resultados sejam divulgados.

** Sem média na Prova Brasil 2015: Não participou ou não atendeu os requisitos necessários para ter o desempenho calculado.

*** Calculado a partir da proficiência média dos alunos nas avaliações estaduais, em decorrência do extravio de provas e impossibilidade do cálculo da proficiência para a Prova Brasil.

Os resultados marcados em verde referem-se ao Ideb que atingiu a meta.

Fonte: Plataforma do IDEB

Disponível em: <<http://ideb.inep.gov.br/resultado/resultado/resultado.seam?cid=3125006>>

Acesso em: 03/09/2017

Pelos índices, observamos que a meta proposta pelo IDEB foi atingida nos anos de 2007, 2009 e 2011 (indicados pelos resultados em verde), fazendo com que a meta projetada aumente progressivamente para os outros anos.

Contudo, é pertinente lembrar que o desempenho dos alunos pode ser maior ou menor

²⁵ O IDEB (Índice de Desenvolvimento da Educação Básica) assume um valor calculado a partir do desempenho obtido pelos alunos que participaram da Prova Brasil/Saeb 2015 e das taxas de aprovação, mensuradas com base nas informações prestadas ao Censo Escolar 2015. Dessa forma, cada uma dessas unidades de agregação tem seu próprio Ideb e metas estabelecidas ao longo do horizonte do PDE, ou seja, até 2021.

por inúmeras variáveis, como: assiduidade dos alunos (e também do professor responsável pela turma naquele ano); comprometimento dos estudantes e também de seus pais ou responsáveis, ao acompanhar as tarefas escolares, procurarem profissionais de saúde, caso o aluno apresente dificuldades neurológicas, psicológicas ou sociais, que interferem diretamente no rendimento escolar, dentre outros fatores externos.

Assim, nem sempre o índice apresentado em um ano é compatível com o outro, como é observado no ano de 2013 – o alcançado pelos estudantes do 5º ano foi de 6.7 pontos, ao contrário da meta de 7.1, por exemplo.

Após elencarmos algumas características da escola 1, verificamos como é a realidade social e educacional do próximo centro educacional municipal. Reiteramos a necessidade de traçar o perfil das escolas, bem como o da comunidade que a rodeia, para compreendermos o léxico representado pelos discentes, e que, de uma forma geral, descrevem a sua relação com o bairro e o entorno onde residem.

A unidade escolar 2 está situada na zona sudoeste do município, próxima à rodovia Péricles Belini, em uma vizinhança de classe econômica baixa. Atende, também, alunos de um bairro construído pela administração municipal, para atender pessoas que se encontravam em moradias de risco na cidade (Residencial Noroeste Monte Verde), além da Zona Rural. Tanto os alunos do Residencial quanto os sitiantes possuem transporte gratuito para frequentarem a Unidade Escolar. Dessa maneira, os textos coletados nesta escola deram origem ao CRAEF 2.

A clientela é itinerante, devido à grande procura de famílias vindas do nordeste para trabalhar em serviços temporários. O fluxo de alunos matriculados e que se transferem para outros bairros ou cidades é grande ao longo do ano. O nível escolar destes pais geralmente é de Ensino Fundamental completo; poucos com Ensino Médio e ou Ensino Superior. As profissões mais encontradas entre os familiares dos alunos são: funcionários de indústrias, usinas, comércio e autônomos.

Assim como outrora indicado, trataremos agora da aprendizagem da Escola 2. Os dados abaixo se referem a mesma avaliação já descrita e as informações possuem a mesma fonte, a saber, a *Website* do IDEB:

Figura 07 – IDEB da escola da Periferia

IDEB - Resultados e Metas

Parâmetros da Pesquisa

Resultado:	Escola	UF:	SP
Município:	VOTUPORANGA	Nome da Escola:	
Rede de ensino:	Municipal	Série / Ano:	4ª série / 5º ano

4ª série / 5º ano

Escola	Ideb Observado					Metas Projetadas								
	2005	2007	2009	2011	2013	2015	2007	2009	2011	2013	2015	2017	2019	2021
	5.4	6.4	6.2	6.2	5.7	5.9	5.4	5.7	6.1	6.3	6.5	6.8	7.0	7.2

Fonte: Plataforma do IDEB

Disponível em: <<http://ideb.inep.gov.br/resultado/resultado/resultado.seam?cid=3125360>> Acesso em: 03/09/201

Verificamos que a meta de partida estipulada é menor do que a apresentada na escola do Centro. Esta se justifica pelos critérios elencados na nota de rodapé número 13 – a nota da última avaliação. Paralelamente à Escola 1, notamos que a escola da periferia também se manteve acima da média nos anos de 2007, 2009 e 2011; vindo a decair, contudo, a partir do ano de 2013.

Os dados apresentados neste tópico são importantes, uma vez que demonstram que as escolas possuem características diferentes, não só de localização, mas também em questões sócio-econômicas que influenciam na aprendizagem dos alunos. De posse destas informações, passamos a descrever como os textos dos alunos foram coletados, bem como as motivações da escrita e, ainda, o tratamento metodológico realizado com as composições.

2.3 Procedimentos para coleta dos dados e temáticas para a construção dos corpora de aprendizes

Visto que o corpus foi produzido por crianças, primeiramente foram feitas reuniões com os pais de ambas as escolas, elucidando a relevância da pesquisa. Depois, solicitamos que os responsáveis pelos discentes autorizassem o uso dos textos para a análise, explicando que os mesmos não seriam identificados, preservando, portanto, a identidade dos discentes

envolvidos (Anexo 1).

De posse das autorizações, foram reunidas 224 redações ao longo do ano de 2015, de estudantes do 3º ano do Ensino Fundamental I (123 produções do CRAEF 1, escola do centro da cidade e ainda 101 textos do CRAEF 2, unidade de ensino localizada na periferia de Votuporanga). Cada turma inicialmente era composta por 30 alunos, com idades entre 7 a 9 anos.

Contudo, devido a fatores não esperados, como transferência de alunos das duas unidades escolares, o número de alunos ao final da pesquisa foi alterado, pois alguns mudaram de cidade ou bairro, não realizando as últimas propostas textuais. Outra situação não prevista foi o recebimento de novos estudantes em ambas as escolas, que também interferiu no total de estudantes atendidos no final do ano.

Independente destes fatores, todas as redações foram digitadas e analisadas, salvo aquelas dos alunos não alfabetizados (quatro discentes da escola da periferia e uma estudante da unidade do centro). Esse posicionamento foi necessário, pois, infelizmente, não foi possível recolher os itens lexicais na escrita destas crianças para proceder às investigações propostas.

Para a elaboração da pesquisa, a professora solicitou que os discentes redigissem uma redação inicial, com fins diagnósticos. Os estudantes dos dois grupos CRAEF 1 e CRAEF 2 escreveram a redação com o tema proposto, sem interferências acerca de questões ortográficas, regências nominais e verbais, para preservar a integridade do corpus.

Uma vez por bimestre, ambas as classes redigiram uma produção escrita para compor o corpus de pesquisa, gerando dessa forma, um banco com aproximadamente 60 redações por amostragem. Desse modo, os estudantes elaboraram suas composições com gêneros variados ao longo do ano, seguindo a orientação proposta, e sem a interferência das professoras.

A seguir, apresentamos as explanações sobre as cinco propostas de redações coletadas ao longo do ano de 2015:

REDAÇÃO 1: Escrita sobre animais do zoológico (escola da periferia) – 02/2015

Reescrita do filme os *Croods* – 02/2015

Apesar do planejamento do projeto propor que todas as propostas de tessitura fossem idênticas para ambas as turmas, houve um imprevisto na primeira amostra. A princípio, o tema da reescrita de uma história do filme *Os Croods* em vídeo seria utilizado também pelos alunos da periferia, contudo, uma forte chuva ocorreu no início daquela semana, prejudicando

o funcionamento da Lousa Digital.

Fomos orientados pela direção escolar a não utilizar o aparelho até a manutenção do mesmo. Sem a estipulação do prazo ao qual o equipamento seria consertado, decidimos por alterar o tema, a fim de não prejudicar o cronograma, e principalmente, não perder os primeiros dias de aula. Esses são importantes para a pesquisa, pois os discentes ainda não receberam muitas interferências pedagógicas dos docentes sobre as normas de escrita das redações, o que proporciona uma amostra mais original e autêntica das crianças.

Assim, a realização da escrita espontânea dos alunos da periferia foi sobre os animais que conheciam ou que já tinham conhecimento via Internet, televisão ou experiência vivenciada, com o tema, animais do zoológico.

REDAÇÃO 02: “Como é ser criança em Votuporanga” – Concurso Cultural promovido pela Secretaria da Educação: 05/2015

No município citado, desde 2006 há o festival literário denominado FLIV (Festival Literário de Votuporanga). De acordo com informações da *Web Site* oficial:

A tradicional Feira do Livro de Votuporanga, deu espaço, em 2011, ao Festival Literário de Votuporanga – Fliv, um evento multicultural que reúne em nove dias diversas atividades ligadas à literatura. O acesso é gratuito. A missão é promover o hábito à leitura, incentivando a população, desde a infância, com iniciativas que democratizem o acesso aos livros e à cultura.

Próximo a data do evento, que ocorre no 2º semestre, uma vez ao ano, a Secretaria Municipal de Educação (SME) solicita às escolas atendidas pelo município que recolham redações e desenhos dos alunos acerca das impressões que possuem da cidade.

A fim de que a autenticidade dos textos seja a maior possível, esta atividade é realizada na sala de aula. Dentro do prazo estipulado, normalmente de duas a três semanas, as professoras propõem aos discentes que redijam o texto de forma livre e autônoma. De posse dos textos de todas as escolas do Ensino Fundamental I, a equipe da Secretaria Municipal da Educação de Votuporanga (SME) elege aqueles que mais se destacam.

Como premiação, expõem as ilustrações e textos das crianças no evento por meio de banners ou publicações. Apresentam, além das produções, o nome completo da criança, a série e escola em que estuda e ainda o nome do docente responsável pela turma.

REDAÇÃO 3: Proposta da semente mágica – 06/2015

Partindo da premissa de uma semente mágica, foi pedido aos alunos que imaginassem o que poderia ser plantado com tal item. Dessa forma, abre-se espaço para o desenvolvimento da criatividade, concomitante ao vocabulário próprio que o tema exige.

REDAÇÃO 4: Escrita de uma narrativa folclórica – 08/2015

A escolha deste tema se deu pelo interesse em verificar o conhecimento dos alunos sobre um tema nacional, ou seja, o folclore, a cultura do nosso povo. Solicitamos aos discentes que redigissem sua escrita pautados nessa premissa: o encontro do personagem Chico Bento com um ser fantástico, representado pela figura lendária da Mula sem cabeça.

REDAÇÃO 05: “Como posso ajudar as pessoas que vivem perto de mim?” – 11/2015

Neste último segmento de redações coletadas, foi proposta uma produção que tratasse do tema “Solidariedade”. Para tanto, cogitamos a oportunidade de os alunos escolherem um planejamento discursivo dentre duas variáveis vertentes.

Dessa forma, foi possível investigar como os aprendizes expressam, em palavras, o seu entendimento sobre o contexto que envolve crianças em tratamento médico ou doação de brinquedos para um menino carente.

Após elencar as estratégias realizadas para a coleta dos dados, a seguir passamos a tratar sobre os procedimentos metodológicos realizados nas redações em estudo. Estes dados são importantes para proporcionar uma análise cuidadosa do léxico e das colocações presentes nos textos dos alunos.

2.4 Tratamento de informação adotado nos corpora de aprendizes

Após a coleta de dados, as composições do CRAEF 1 e do CRAEF 2 foram digitadas e armazenadas no computador, como dados digitalizados em formato txt (texto sem formatação), no programa Bloco de Notas, pertencente ao rol de *softwares* acessórios do Pacote *Office do Windows*.

Em relação à identificação dos textos, foi criado um código, indicando: escola a qual

pertencem, número do aluno em ordem alfabética e o número da redação coletada. Portanto, os textos estão organizados por siglas, como: “A_03_01”, leia-se Escola A – Centro, aluno número 03 da sala, Tema 01.

Nesta etapa do trabalho, observamos que a escrita de palavras como “brincano”, ao invés de “brincando”, por exemplo, gerava um novo vocábulo, alterando, assim, o número de *types* e *tokens* registrado, induzindo a uma leitura equivocada das informações, já que ambas as palavras se referem ao ato de brincar, e não a ações diferentes. Desse modo, foi preciso retomar o passo anterior, repassando todas as palavras digitadas, a fim de corrigir as trocas ortográficas (“caza” no lugar de “casa”) e demais formas grafadas de forma errônea pelas crianças.

Além disso, o fato das questões de pesquisa se relacionarem diretamente à quantidade e variedade lexical dos estudantes, também justifica a alteração dos dados originais, haja vista que não faz parte do escopo desta dissertação, a verificação de erros ortográficos ou, ainda, de hiper e hipossegmentação entre palavras.

Com a resolução deste impasse metodológico, utilizamos o *software WordSmith Tools*, versão 6.0, para manusear os dados coletados.

Figura 08 – Interface do *Software WordSmith Tools* e as ferramentas utilizadas para as análises



Fonte: *Print* da tela do programa citado

No primeiro momento, a ferramenta *Wordlist* foi usada para listarmos as palavras mais utilizadas pelos estudantes. Geralmente atestam para a existência de termos predominantemente gramaticais (expressos pelas preposições, artigos e pronomes), haja vista se tratar de uma lista de vocábulos mais frequentes. Outra informação de destaque encontrada por este instrumento são os dados estatísticos sobre as redações.

Dentre as indicações numéricas, é válido observar a densidade lexical a partir do número total de palavras (*tokens*) e o número de palavras distintas (*types*). A partir destes

dados, é possível identificar o item *type-token ratio*²⁶, descrito por Berber Sardinha (1999).

Para o autor, mesmo que o texto possua um alto número de vocábulos no total, a variedade lexical será baixa se essas se repetirem com frequência. O inverso também ocorre na avaliação linguística, ou seja, se a redação não possuir muitas reiterações de palavras, o índice do *type-token ratio* será maior, indicando um texto “mais rico ou variado”.

Em seguida, o acessório *KeyWords*²⁷ também foi utilizado, já que disponibiliza as palavras-chave que mais se destacaram em comparação com o corpus de referência – representativo da língua portuguesa. Este rol de palavras mais significativas foi contrastado com o corpus de referência *Folha Wordlist*²⁸.

O item a seguir abordará os critérios para a realização das análises, e, posteriormente, os resultados sobre:

- i. comparação dos dois corpora CRAEF 1 e CRAEF 2;
- ii. análise do léxico dos alunos; e
- iii. o levantamento das colocações encontradas nas redações das salas de aula já mencionadas.

²⁶ A relação *Type-token ratio* demonstra a riqueza lexical de um texto. É a razão entre o número de palavras total (*tokens*) e os vocábulos que não se repetem (*types*).

²⁷ De acordo com Berber Sardinha (1999, p. 7) “Esta ferramenta permite a seleção de itens de uma lista de palavras (ou mais) por meio da comparação de suas frequências com uma lista de referências. *KeyWords* contrasta uma lista de palavras (ou mais de uma) de um corpus de estudo com uma lista de palavras de um corpus de referência. O resultado do contraste é uma lista de palavras-chave, ou palavras cujas frequências são estatisticamente diferentes no corpus de estudo e no corpus de referência.”

²⁸ É considerado um dos corpus mais representativo da língua portuguesa, trazendo as palavras de maior relevância do idioma.

3. ANÁLISE DOS RESULTADOS

De acordo com os objetivos propostos nesta dissertação, inicialmente a análise baseada no CRAEF 1 e CRAEF 2 possuía duas vertentes: a primeira voltada aos aspectos lexicais, enquanto que a segunda fase tratava das colocações presentes nas produções escritas dos aprendizes, ambas comparando as redações dos dois grupos CRAEF 1 e CRAEF 2.

Contudo, conforme já apontado, durante as análises dos fraseologismos, era esperado um número maior de colocações do que o verificado. Após tal constatação, outra hipótese foi levantada dentro desta dissertação: enquanto professora de educação básica (PEB I), como poderia proceder para tentar melhorar o léxico destes estudantes?

Em discussões com a orientadora Profa. Dra. Adriane Orenha-Ottaiano, concordamos em ampliar os objetivos específicos do trabalho. Dessa maneira, depois de tratar do desenvolvimento dos tópicos que compararão os dois corpora CRAEF 1 e CRAEF 2, seguido da análise do léxico dos discentes e posteriormente do levantamento e discussão das colocações presentes nas produções, traremos algumas propostas de atividades, a fim de melhorar a competência colocacional de estudantes do ensino fundamental I.

Para esta investigação, não efetuamos a aplicação das atividades e tampouco a avaliação das turmas neste quesito empírico, haja vista que vários discentes saíram de suas respectivas salas, inviabilizando a verificação dos temas em voga neste estudo. Dessa forma, em trabalhos posteriores, pretendemos realizar esta tarefa e descrevê-la em detalhes na tese de doutorado a ser iniciada em 2018.

Apresentamos, na sequência deste capítulo, os dados coletados. Isso se faz necessário, pois, conforme relatado, a proposta inicial era coletar sessenta redações em cada amostragem do CRAEF 1 e 2, o que resultaria em trezentos textos para análise. Porém, os fatores citados no tópico 2.3, acerca de transferências, recebimentos de alunos, dentre outros motivos, mostraram-se uma causa que reverberou negativamente no total de composições coletadas nas duas escolas.

Por isso, o próximo tópico visa a delinear os resultados obtidos após a digitação das redações, e que nos ajudará a traçar um perfil mais completo das salas de aula em estudo.

3.1 Análise de dados do CRAEF

Com o final da etapa de coleta e digitalização das redações, teve início a fase de análise das mesmas. A diferença entre o total de textos planejado para a coleta de dados, considerando cinco produções por aluno e a quantidade de redações efetivamente digitadas se justifica pelos fatores já descritos sucintamente na metodologia. Para esclarecer os motivos de alguns textos não terem sido coletados, segue a legenda com a explanação das siglas:

Figura 09 – Legenda das siglas

LEGENDA	
RD	REDAÇÃO DIGITADA
NF	NÃO FEZ (A CRIANÇA FALTOU DURANTE A SEMANA DE APLICAÇÃO)
AR	ALUNO RECEBIDO APÓS A COLETA DA REDAÇÃO
TR	ALUNO TRANSFERIDO PARA OUTRA UNIDADE DE ENSINO, CIDADE OU ESTADO
PS	PRÉ-SILÁBICA: CRIANÇA NÃO ALFABETIZADA

De fato, a tabela abaixo mostra o número final das redações coletadas da escola da área central de Votuporanga tabuladas (CRAEF 1), já com as siglas relacionadas a cada caso:

Tabela 01 – Redações do CRAEF 1

CRAEF 1 (ESCOLA DO CENTRO)													
ALUNOS	REDAÇÕES						ALUNOS	REDAÇÕES					
	1	2	3	4	5	TOTAL		1	2	3	4	5	TOTAL
1	NF	RD	RD	RD	RD	4	17	NF	NF	RD	NF	RD	2
2	RD	RD	RD	RD	RD	5	18	RD	RD	RD	RD	RD	5
3	RD	RD	RD	RD	RD	5	19	RD	RD	RD	RD	TR	4
4	NF	RD	RD	RD	RD	4	20	RD	RD	RD	RD	RD	5
5	RD	RD	RD	RD	RD	5	21	RD	RD	NF	RD	RD	4
6	RD	RD	RD	RD	RD	5	22	RD	NF	RD	RD	RD	4
7	RD	RD	RD	RD	RD	5	23	RD	RD	RD	RD	RD	5
8	RD	RD	RD	RD	RD	5	24	RD	RD	RD	RD	RD	5
9	RD	RD	RD	RD	RD	5	25	AR	RD	RD	RD	RD	4
10	RD	RD	RD	RD	RD	5	26	AR	RD	RD	RD	RD	4
11	RD	RD	RD	RD	RD	5	27	AR	AR	AR	RD	RD	2
12	RD	NF	NF	RD	RD	3	28	AR	AR	AR	RD	RD	2
13	RD	RD	RD	RD	RD	5	29	TR	TR	TR	TR	TR	0
14	RD	RD	RD	RD	RD	5	30	RD	TR	TR	TR	TR	1
15	RD	RD	RD	RD	RD	5	31	PS	PS	PS	PS	PS	0
16	RD	RD	RD	RD	RD	5	TOTAL	21	23	25	27	27	123

Ao todo, vinte e sete redações não foram coletadas e cinco não puderam ser utilizadas em nossa análise pelas seguintes razões:

- seis alunos faltaram durante a semana de aplicação da atividade (incluindo 1 que passou por procedimento cirúrgico com complicações, impedindo que frequentasse as aulas durante várias semanas), fato que diminuiu o número de redações coletadas em nove;
- a ausência de três crianças, cujas transferências foram requisitadas pelos pais, também alterou a quantidade pretendida, reduzindo-as em dez composições (o aluno número 29 entrou na escola no início do ano letivo, porém foi transferido na primeira semana de aula, por isso, não realizou nenhuma proposta de redação);
- Outro fator que influenciou no total de textos diz respeito aos alunos recebidos ao longo de 2015: quatro ao todo. Estes chegaram após o início das atividades regulares, também modificando o montante de proposições esperadas;
- Por fim, há um estudante no nível pré-silábico de escrita na sala de aula. Nesta hipótese, a criança ainda está formulando o seu entendimento sobre a língua e ao redigir palavras, frases e textos, possui muitas dificuldades em realizar esta tarefa com autonomia. Apesar de ter realizado todas as propostas, não é possível utilizar os textos do aluno 31, reduzindo em cinco o rol de tessituras.

Após a catalogação dos textos, com a compilação dos textos para processamento no programa computacional *WordSmith Tools* (SCOTT, 2012), versão 6.0, a ferramenta *Wordlist* computou dados eletrônicos sobre as produções desta unidade escolar.

Na aba *Statistics* do referido *software*, é possível visualizar os *tokens* e *types* de todas as tessituras digitalizadas, assim como constatar a relação *type-token ratio*, já descrita anteriormente no item 2.4 do presente trabalho. Para facilitar a apresentação dos números, dispomos as informações em formato de tabela, conforme segue:

Tabela 02 – Rol de redações coletadas (CRAEF 01)

Total de Textos Coletados	129
Redações Digitalizadas	123
Tokens (número de palavras total)	12.436
Types (número de palavras distintas)	2.200
Relação Type-token ratio(TTR)	18

Haja vista que os alunos escreveram sobre os mesmos cinco temas, não houve um número muito expressivo de vocábulos distintos nas redações. Consequentemente, a relação que denota a “riqueza lexical” já descrita em páginas anteriores é bastante baixa, sendo necessário, assim, apresentar os dados estatísticos que as comprovem:

Figura 10 – Estatísticas do CRAEF 1

N	Overall	1	2	3	4	5	6	7	8	9	10	11	12
text file	Overall	an.txt	an.txt	an.txt	an.txt	dro.txt	dro.txt	dro.txt	dro.txt	dro.txt	ue.txt	ue.txt	ue.txt
file size	65.010	255	356	304	408	457	587	333	317	533	514	1.141	925
tokens (running words) in text	12.527	52	66	56	72	90	111	65	61	106	95	227	186
tokens used for word list	12.436	51	66	56	72	89	107	63	61	105	84	227	186
types (distinct words)	2.200	41	50	39	49	64	74	44	42	52	54	103	91
type/token ratio (TTR)	18	80	76	70	68	72	69	70	69	50	64	45	49

Vale mostrar, contudo, a individualidade de cada criança no momento de redigir sua escrita, já que há bastante variação entre a quantidade de palavras redigida pelos discentes deste mesmo grupo estudantil, o que também altera a proporção do *Type/token ratio*. Nota-se que, enquanto há estudantes que redigem 1.141 palavras em uma única composição, outros escrevem apenas 255 vocábulos.

Por esse motivo, ao longo deste capítulo, apresentamos alguns textos das crianças, de modo a reforçar a singularidade destes discentes, atestando para o que foi afirmado na fundamentação teórica sobre a competência colocacional. É necessário desenvolvê-la desde os primeiros anos da infância, a fim de garantir que, na vida adulta, possam se expressar com autonomia e segurança, articulando, com propriedade, sua comunicação com os demais indivíduos.

Paralelamente à escola do Centro, elencamos abaixo as informações sobre o CRAEF 2, a escola localizada na periferia da cidade de Votuporanga, após a transposição das redações físicas em digitalizações no *software* citado.

Assim como na escola 1, a unidade educacional da periferia também apresentou dissonância entre a quantidade esperada daquela efetivamente realizada. A seguir, apresentamos a Tabela 03 com a coleta sistematizada sem a legenda, já que as informações presentes na Figura 09 também se enquadram nesta.

Tabela 03 – Redações do CRAEF 02

CRAEF 2 (ESCOLA DA PERIFERIA)													
ALUNOS	REDAÇÕES						ALUNOS	REDAÇÕES					
	1	2	3	4	5	TOTAL		1	2	3	4	5	TOTAL
1	RD	RD	RD	NF	RD	4	19	RD	TR	TR	RD	RD	3
2	NF	RD	NF	RD	RD	3	20	RD	RD	RD	NF	RD	4
3	RD	RD	RD	RD	RD	5	21	RD	RD	RD	RD	RD	5
4	RD	RD	RD	RD	RD	5	22	RD	RD	NF	RD	RD	4
5	RD	RD	RD	RD	RD	5	23	NF	RD	TR	TR	TR	1
6	NF	RD	RD	RD	RD	4	24	RD	RD	NF	RD	RD	4
7	RD	RD	NF	NF	NF	2	25	AR	RD	TR	TR	TR	1
8	NF	RD	RD	RD	RD	4	26	AR	RD	NF	RD	RD	3
9	RD	RD	RD	RD	RD	5	27	AR	AR	AR	RD	TR	1
10	RD	RD	RD	RD	RD	5	28	AR	AR	AR	RD	RD	2
11	RD	RD	RD	RD	RD	5	29	AR	AR	AR	AR	RD	1
12	NF	RD	NF	RD	TR	2	30	PS	PS	PS	PS	PS	0
13	RD	RD	NF	NF	RD	3	31	PS	PS	PS	PS	PS	0
14	RD	RD	RD	RD	RD	5	32	PS	PS	PS	PS	PS	0
15	NF	NF	NF	RD	RD	2	33	PS	PS	PS	PS	PS	0
16	RD	RD	RD	RD	RD	5	34	PS	PS	PS	PS	PS	0
17	NF	RD	RD	RD	RD	4	TOTAL	17	24	15	21	24	101
18	RD	RD	RD	NF	RD	4							

Conforme aponta a Tabela acima, 44 textos não foram coletados e 15 não utilizados, pelos motivos relacionados abaixo:

- quinze alunos faltaram durante a semana de aplicação da atividade, fato que diminuiu o número de redações coletadas em vinte e duas;
- a ausência de cinco crianças, cujas transferências foram requisitadas pelos pais, também alterou a quantidade pretendida, reduzindo-as em dez composições (o aluno número 19 foi transferido da escola por questões familiares, porém retornou à mesma unidade de ensino após o recesso escolar, quando os pais fizeram as pazes e voltaram a morar juntos);
- Outro fator que influenciou no total de textos diz respeito aos alunos recebidos ao longo de 2015: cinco ao total. Estes chegaram após o início das atividades regulares, também modificando o montante de proposições esperadas.

Esta diferença numérica é, portanto, significativa. A descrição das duas unidades escolares no capítulo metodológico aponta as causas do êxodo das famílias: a rotatividade de emprego em áreas rurais e usineiras da região noroeste paulista. Por isso, há muito mais alunos transferidos e recebidos na escola da periferia do que na unidade central, cujos pais aparentemente se estabelecem por mais tempo em um mesmo bairro.

Seguindo a análise, os resultados numéricos na ferramenta *Wordlist* apresentam:

Tabela 04 – Dados das Redações do CRAEF 2

Total de Textos Coletados	126
Redações Digitadas	101
<i>Tokens</i> (número de palavras total)	7.999
<i>Types</i> (número de palavras distintas)	1.376
Relação <i>Type-token ratio</i> (TTR)	17

Visto que há cinco discentes pré-silábicos nesta sala de aula, não foi possível realizar as análises em textos cujos autores se encontram nas primeiras hipóteses alfabéticas. O número de textos digitados, portanto, diminuiu em vinte e cinco, conforme apontado pela tabela.

A partir dos dados da *Wordlist*, também efetuamos as estatísticas de escrita desta turma, dispostos a seguir:

Figura 11 – Estatísticas do CRAEF 2

N	Overall	1	2	3	4	5	6	7	8	9	10	11	12
text file	Overall	ely.txt	ely.txt	ely.txt	ely.txt	los.txt	los.txt	los.txt	los.txt	los.txt	s..txt	s..txt	s..txt
file size	41.589	310	607	375	408	394	533	273	989	365	38	319	212
tokens (running words) in text	8.054	54	108	69	73	89	102	53	202	70	7	63	43
tokens used for word list	7.999	54	107	69	73	79	101	53	202	69	7	62	43
types (distinct words)	1.376	40	71	50	49	41	68	38	99	46	6	48	22
type/token ratio (TTR)	17	74	66	72	67	52	67	72	49	67	86	77	51

Conforme explicitado anteriormente, reiteramos a necessidade de analisar o rendimento lexical individualizado, pois, nas duas turmas, há discentes que escrevem mais de duzentas palavras enquanto outros redigem apenas seis. Esta diferença indica que, em ambas as escolas, há uma variedade considerável de desempenho linguístico entre crianças do mesmo ambiente educacional.

A reiteração desta ocorrência indica que, nas duas escolas, há provavelmente a influência de fatores extralinguísticos, por exemplo, alternância de vocabulário familiar, desenvolvimento discrepante da oralidade e diferença no acesso à cultura via plataformas como cinema, livros, revistas, *Internet*. Essas causas, dentre outras, podem ser os motivos que justifiquem tal flutuação na produção das palavras dos estudantes.

Terminada esta etapa, realizamos uma comparação entre as escolas, no que tange ao número de redações coletadas e efetivamente utilizada para a análise. De posse da somatória dos textos, relacionamos as informações da seguinte forma:

Tabela 05 – Particularidades das redações

COMPARATIVO DAS REDAÇÕES		
TEXTOS COLETADOS POR ESTUDANTE	CRAEF 1 – número de alunos	CRAEF 2 – número de alunos
5 PROPOSTAS	17	9
4 PROPOSTAS	7	8
3 PROPOSTAS	1	4
2 PROPOSTAS	3	4
1 PROPOSTA	2	4
0 PROPOSTAS (REDAÇÕES COLETADAS, MAS NÃO DIGITALIZADAS)	1 (5 textos)	5 (25 textos)
PARTICIPANTES	31	34

Nesta tabela, um dado que chama a atenção diz respeito aos alunos ainda não alfabetizados: um estudante da unidade escolar do centro e cinco crianças da escola da periferia. Infelizmente, devido aos fatores já elencados outrora, estas redações não puderam ser digitalizadas nesta dissertação. Dessa forma, cinco textos da escola do centro e vinte e cinco da unidade educacional da periferia não pertencem ao rol analisado no CRAEF. Essa quantidade de proposições textuais foi determinante na dissonância da coleta entre as escolas, porém não foi o único fator que ocasionou essa diferença.

Os números também demonstram que, apesar de a escola 2, localizada na periferia, contar com um número maior de participantes ao longo do ano, somente nove alunos realizaram todas as cinco propostas de redações, contabilizando quarenta e cinco coletas completas. Enquanto isso, na escola 1, aquela que representa os alunos do centro da cidade, dezessete crianças redigiram os cinco textos, resultando em oitenta e cinco proposições inteiras.

A discrepância numérica entre as escolas também prosseguiu na contagem de quatro redações por discente, sendo sete alunos do CRAEF 1 e oito do CRAEF 2, totalizando,

respectivamente, vinte e oito e trinta e duas composições. A diferença continua nas demais linhas em menor quantidade de coletas, conforme disposto na tabela.

As causas que apontam as ocorrências adversas da coleta dos textos são:

Tabela 06 – Quantidade de textos não coletados

COMPARATIVO DAS REDAÇÕES NÃO COLETADAS		
MOTIVO	CRAEF 1	CRAEF 2
FALTA DOS ALUNOS SUPERIOR A 3 DIAS NA SEMANA DA COLETA	9	22
TRANSFERÊNCIA PARA OUTRA UNIDADE ESCOLAR	10	10
ALUNOS RECEBIDOS DURANTE O ANO LETIVO	8	12
TOTAL DE TEXTOS NÃO COLETADOS	27	44

Observa-se na tabela as causas que resultaram em um baixo número de textos coletados, principalmente na escola da periferia. Dentre os motivos apontados, a falta dos alunos na semana da coleta justificou a não coletânea de vinte e duas redações, um número bastante alto se comparado com a escola do centro da cidade, de apenas nove. Esse dado nos indica que o alto índice de ausência destes alunos da escola 2 na sala de aula ao longo do ano.

Já os discentes que foram transferidos para outra unidade escolar, seja por mudanças de endereço para outro bairro ou até mesmo de cidade, foi a causa da falta de vinte coletas textuais, dez de cada escola. Os alunos que entraram nas suas respectivas salas após o início das aulas regulares também afetou a contagem final, sendo oito redações para a escola 1 e doze para a escola 2.

Estes itens resultaram em vinte e sete textos a menos para o CRAEF 1 e quarenta e quatro para o CRAEF 2. Acrescidos das propostas dos alunos pré-silábicos já contabilizados, chegamos ao seguinte número: cento e vinte e três textos digitalizados na escola do centro e cento e uma redações analisadas na escola da periferia. Assim, a diferença de vinte e dois textos entre as escolas influenciou também na quantidade de palavras escritas, um total de 12.436 palavras no CRAEF 1 em oposição a 7.999 vocábulos no CRAEF 2.

De posse dos dados da escola do centro e da periferia, realizamos outra análise no *software* já mencionado, a fim de compor todas as informações do CRAEF:

Tabela 07 – Somatória do CRAEF 1 e 2

Total de Textos Coletados no CRAEF	255
Redações Digitadas no CRAEF	224
<i>Tokens</i> (número de palavras total)	20.435
<i>Types</i> (número de palavras distintas)	4.116
Relação <i>Type-token ratio</i> (TTR)	18

Conforme dito anteriormente na etapa metodológica, é possível analisar a linguagem dos indivíduos mesmo em um corpus de poucas palavras. Endossamos tal afirmação, pois, embora pequenos, estes corpora são significativos por representarem a linguagem de crianças que estudam no terceiro ano do ensino fundamental I, na cidade de Votuporanga, localizada a noroeste do estado de São Paulo. Além disso, atestamos que estes dados são suficientes para responder as nossas perguntas de pesquisa desta dissertação, sobre a maneira como cada sala de aula trata da variação lexical, também acerca das colocações empregadas e, ainda, de colocações que poderão ser aprendidas por meio de atividades com vistas ao desenvolvimento da competência colocacional destes aprendizes.

Finalizando a fase de descrição do CRAEF, gostaríamos de apontar alguns traços distintivos entre as escolas por nós observados, enquanto professora responsável pelas duas turmas investigadas empiricamente no ano letivo de 2015. A seguir, trazemos algumas ocorrências, alheias ao ensino sistematizado oferecido nas instituições escolares, presentes com mais veemência nos alunos da periferia, tais como:

- mais faltas do que os alunos que estudam no centro da cidade, denotando a falta de incentivo dos pais e familiares para estudar, haja vista que a grande parte das faltas ocorre nas segundas e sextas-feiras. Ao questionar os pais sobre o motivo dessas ausências, estes justificam as faltas por viagens a passeio, além de alegarem que a criança dorme tarde e não se sente disposta para estudar no período da manhã.

- recebimento de alunos no meio do ano letivo, ocasionado pelo fluxo de trabalho em usinas da cidade e empresas adjacentes, que contratam funcionários em período de colheita da cana-de-açúcar. Estes empregos sazonais também interferem na vida escolar dos filhos, pois quando o contrato termina, por volta de três, quatro meses depois, essas crianças também migram com os pais para outras cidades e estados, resultando em várias transferências e recebimento de estudantes ao longo do ano escolar.

- alunos não alfabetizados. Esse tema não será desenvolvido nesta dissertação, pois demanda um espaço maior e várias reflexões teóricas para uma abordagem mais completa sobre esta ocorrência.

Essas ocorrências não são exclusivas da periferia, mas, em comparação com os discentes do centro, as crianças do CRAEF 2 demonstram vivenciar com uma maior frequência esses tipos de infortúnios em seu cotidiano.

Por todo o exposto, seguimos para o próximo tópico, que realiza análises no tocante aos aspectos lexicais e comparativos entre o CRAEF 1 e CRAEF 2, bem como às colocações encontradas nestes corpora, relacionadas em seguida.

3.2 O léxico infantil presente no CRAEF 1 e 2

A proposta deste tópico é responder a pergunta de pesquisa referente a hipótese se há ou não diferenciação lexical entre textos escritos por alunos do centro da cidade e redações redigidas por crianças da periferia. Se a hipótese se confirmar, será necessário justificar de que forma ocorre esta mudança.

Para atingir tal objetivo, voltamos aos temas das propostas textuais já explanados na metodologia, e realizamos a análise pelo programa computacional *WordSmith Tools*, versão 6.0, observando a relação entre a lista de palavras gerada, as palavras-chave e os itens lexicais que obtiveram destaque nos textos do CRAEF 1 e 2. Ao longo deste tópico, apresentamos ainda alguns trechos das escritas das crianças, afim de corroborar as discussões e reflexões concebidas neste capítulo analítico.

Antes de prosseguirmos, é coerente fazer uma observação metodológica dos dados digitalizados, posto que os textos que compõem o CRAEF foram produzidos por crianças de sete a oito anos; vale lembrar que estes estudantes do terceiro ano do ensino fundamental I ainda estão no processo de consolidação da alfabetização e letramento. Assim, já era esperado que os alunos de ambas as escolas apresentassem uma escrita com desvios ortográficos. E de fato, tal característica ocorreu em todas as etapas de coletas, em ambas as unidades escolares.

Assim, apesar de as imagens dos textos mostrados na sequência apresentarem a escrita original, da forma como foram redigidas inicialmente, foi necessário fazermos a correção

destes termos na etapa da digitalização dos dados. Isso se dá em razão de o *software* interpretar cada termo redigido erroneamente como uma “nova palavra” (por exemplo, “cabeza”, em vez de “cabeça”) e, caso as escritas fossem mantidas no formato inicial, a análise das palavras contabilizadas sofreria alterações, resultando em dados incorretos.

Outro detalhe digno de nota diz respeito aos valores numéricos presentes nas análises do *software WordSmith Tools* (2012). Conforme já observado em tópicos anteriores, priorizaremos nesta dissertação o estudo qualitativo da escrita das crianças. Contudo, por não se tratar de dados tão extensos, manteremos a observação quantitativa.

É importante retomar, por fim, o fato de que cada proposta teve um objetivo específico de atuação, conforme será explicado em detalhes nas páginas seguintes, em referência aos Temas 01, 02, 03, 04 e 05.

Primeiramente, faremos uma breve análise do Tema 01.

Tema 01 – Relato do filme *Os Croods* (escola do centro) – 02/2015

Escrita sobre animais do zoológico (escola da periferia) – 02/2015

Em razão de a primeira proposta ser aplicada logo no início do ano, precisamente na primeira semana de aula, acrescido do fato de as crianças serem recém saídas do segundo ano do ensino fundamental I, decidimos por sugerir inicialmente um tema lúdico, como um relato de filme.

Desse modo, na escola do centro, começamos a abordagem apresentando o filme *Os Croods*, voltado ao público infantil, com a história de uma família que vive no período pré-histórico e enfrenta várias aventuras neste ambiente inóspito. A película possui várias cenas engraçadas e que chamam a atenção dos alunos. Na aula seguinte, oferecemos a atividade de escrita deste filme, que foi prontamente aceita pelos educandos e realizada com afinco.

Já, na escola da periferia, não foi possível coletar a primeira proposta da mesma forma que a outra unidade escolar. Retomando o que está disposto no capítulo metodológico, por motivos inerentes ao arbítrio da pesquisadora, à questão da energia elétrica impossibilitou a exibição do filme. Com a ocorrência de chuvas, raios e relâmpagos na cidade de Votuporanga durante a primeira semana de aula no mês de fevereiro de 2015, a gestão escolar sugeriu que as professoras evitassem utilizar a lousa digital durante o período de duas semanas, a fim de evitar perdas ao patrimônio.

Dessa maneira, ao invés de assistirem aos “*Os Croods*”, realizamos duas rodas de conversa sobre o tema zoológico e animais que vivem neste espaço, ilustrando a atividade

com fotos de animais retirados de revistas e livros educativos. Realizamos a escolha deste tema, pois observamos que as crianças gostam bastante deste assunto, além de demonstrarem um interesse genuíno na abordagem deste tema, confirmado na aceitação de redigirem o texto escrito, elencando itens ditos na abordagem oral.

Para efetuar as análises deste primeiro bloco de textos, apresentaremos a *Wordlist* referente à listagem de palavras do CRAEF 1 e 2:

Tabela 08 – Comparação do CRAEF 1 e 2 (1º Tema)

DADOS ESTATÍSTICOS DO 1º TEMA DE REDAÇÃO			
CRAEF 1 – CENTRO		CRAEF 2 – PERIFERIA	
Redações digitadas	21	Redações digitadas	17
<i>Tokens</i> (número total de palavras)	1.796	<i>Tokens</i> (número total de palavras)	823
<i>Types</i> (número de palavras distintas)	485	<i>Types</i> (número de palavras distintas)	254

Apesar de o total de textos coletados do CRAEF 1 e 2 serem próximos, díspares apenas por quatro composições, a diferença entre a quantidade de palavras produzidas pelas duas classes foi extremamente grande, visto que o número feito pelos alunos do CRAEF 1 foi mais do que o dobro do CRAEF 2, resultando em uma diferença de 973 vocábulos. Consequentemente, as palavras distintas, ou seja, o número total de *types*, também possui um valor interessante, pelo fato das crianças participantes terem de sete a nove anos e escreverem muitos vocábulos diferentes.

Por meio da ferramenta *Keyword* do programa *WordSmith Tools*, utilizamos também o Folha Wordlist, como corpus de referência do Português. Dessa forma, foi possível obter o rol das palavras-chave, ou seja, aquelas mais significativas presente nos textos dos alunos deste primeiro tema. Seguem na próxima página as primeiras quinze palavras:

Tabela 09 – KeyWords gerada pelo software na análise do 1º Tema – CRAEF 1 e 2

KEYWORDS DO TEMA 1					
CRAEF 1			CRAEF 2		
N	Key Word	Freq.	N	Key Word	Freq.
1	CAVERNA	23	1	GIRAFÁ	12
2	MINHA	43	2	EU	30
3	PEDRA	22	3	VIMOS	12
4	FAMÍLIA	32	4	E	82
5	BANHO	20	5	ZOOLÓGICO	9
6	NÓS	29	6	VI	12
7	ESTIMAÇÃO	12	7	LEÃO	11
8	EU	32	8	ELEFANTE	8
9	MEU	24	9	ZEBRA	6
10	ELES	29	10	MÃE	11
11	PAI	18	11	MINHA	12
12	TINHA	20	12	MONTE	4
13	E	109	13	FAMÍLIA	4
14	ERA	25	14	HIPOPÓTAMO	4
15	ENERGIA	8	15	GOLFINHO	5

A sala que escreveu sobre o filme “*Os Croods*”, pertencente ao CRAEF 1, apontou elementos existentes neste universo, como caverna, atingindo a frequência de 23 repetições, respectivamente. Outra palavra do mesmo campo semântico é pedra, com 22 incidências, relacionado ao referente “idade da pedra”, ou usando “pedra” enquanto material usado neste ambiente, como em: “nós pegamos uma pedra e esfrega as duas pedras junta e faz fogo no escuro”; “desenhavam em uma pedra” e “fechamos a caverna com uma pedra e não com uma porta”.

Visto que essa turma efetuou comparações entre a vida do homem da caverna com a sua própria, a palavra família apareceu 32 vezes, comprovando o alto índice de equiparações feitas neste âmbito: “eu tenho uma família” e “eu não queria que fosse a minha família eu não queria esta lá porque lá dava terremotos”, foram algumas das sentenças produzidas pelos estudantes.

Outro termo que merece destaque diz respeito à incidência da palavra “energia“, com 8 ocorrências. Esta foi usada para designar a principal dificuldade que teriam, caso morassem naquela época: “a energia era o sol”; “na idade da pedra não existia energia tv celular e computador” e “Não tem energia e para nós tomar banho nós toma banho no rio”.

Já, no CRAEF 2, os vocábulos que mais se destacaram foram os nomes de animais: “girafa” – 12 ocorrências; “leão” – 11; “zebra” – 6 e “hipopótamo” – 4, por exemplo. Neste contexto, os alunos elencaram tanto os animais típicos do zoológico, quanto aqueles que

gostariam de conhecer ou ver “de perto”, como “golfinho”, que apareceu 5 vezes no corpus e “elefante” – 8.

Também é digna de nota a ocorrência da palavra “mãe”, 11 reincidências, enquanto a família teve 4 citações. Haja vista que a figura paterna nem sempre é presente em lares da periferia, observa-se que, ao falar da progenitora, a criança já se refere à família como um todo: “eu já fui no zoológico com a minha mãe” e “depois pedi pra minha mãe pra visitar a onça”.

Apesar do tema desta primeira proposta ser diferente para as duas escolas, trouxemos o *print* de alguns textos, para tecer comentários mais pontuais sobre a escrita das crianças nesta primeira coleta. Haja vista que o número de palavras foi bastante baixo, para o tema 01, não efetuaremos a análise por *Keywords* neste bloco de textos, pois algumas escritas não apontaram nenhuma palavra de conteúdo na análise por palavras-chave. Nestes casos, indicamos a ausência das palavras-chave ou das frequências por meio do traço, conforme aponta a próxima tabela:

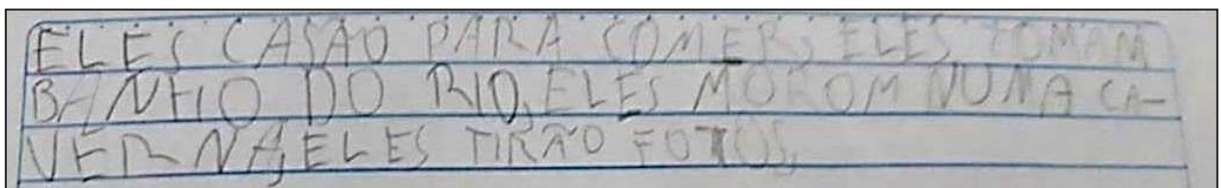
Tabela 10 – *KeyWords* das Amostras do 1º Tema de redação

CRAEF 1 – CENTRO				CRAEF 2 – PERIFERIA			
ALUNO 18		ALUNO 30		ALUNO 03		ALUNO 07	
KEY WORD	FREQ.	KEY WORD	FREQ.	KEY WORD	FREQ.	KEY WORD	FREQ.
Eles	04	Caverna	02	Eu	05	-	-
-	-	Felizes	02	Vi	03	-	-

Provavelmente por se tratar de uma primeira redação, efetuada logo após as férias escolares, muitas redações demonstraram este padrão, dos dois grupos escolares. Por isso, não foi possível efetuar a análise dos dados de *Tokens*, *types* e *type-token ratio*, com resultados relevantes a este estudo. Somente os Temas 02, 03, 04 e 05 trarão a referida análise.

Inicialmente, trazemos duas amostras da escola do centro e outras duas da periferia, conforme segue:

Figura 12 – Amostra da Escola 1 (Aluno 18)

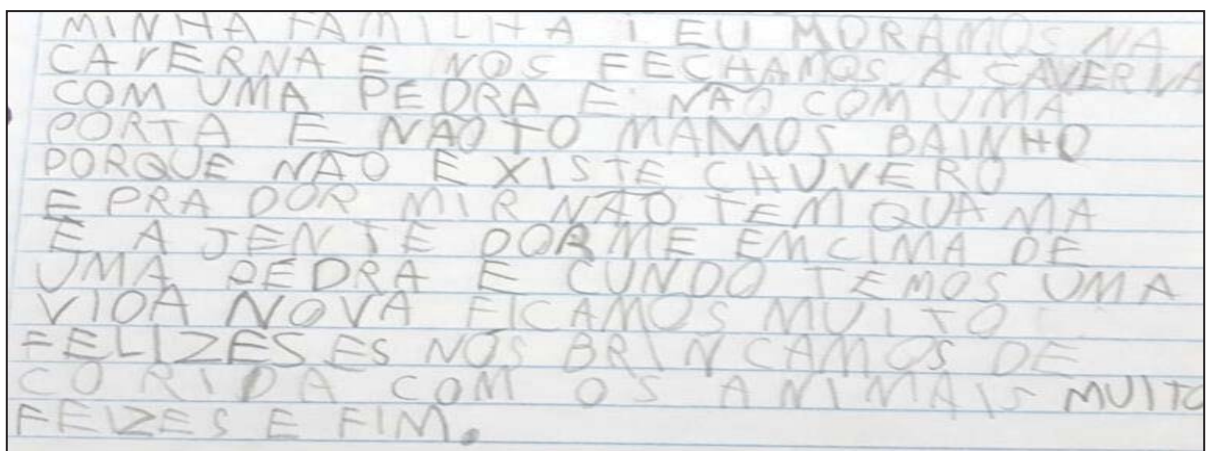


Nesta amostra, a aluna iniciou a descrição da família apresentada no filme, quando

abruptamente parou de escrever e começou a chorar. Este foi um dos poucos casos em que a criança se recusou a continuar o texto, pois disse que não sabia como iria agir, caso sua família tivesse que morar em “um lugar como aquele”.

Assim, apesar do número global de palavras da escola do centro ser superior em relação ao número da periferia, chamamos atenção para o fato de haver alunos do CRAEF 1 que também redigem poucas palavras ou não demonstram interesse em produzir muitas frases, caso não se identifiquem com a proposta oferecida.

Figura 13 – Amostra da Escola 1 (Aluno 30)

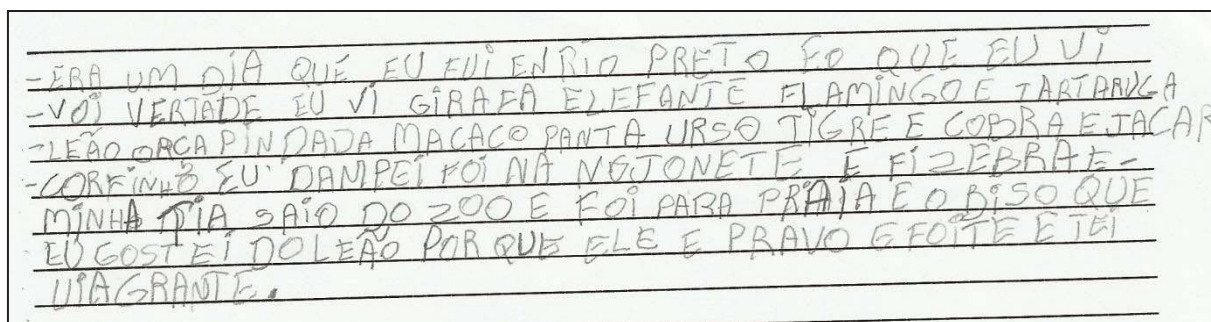


Ao contrário da anterior, esta estudante se divertiu bastante assistindo ao filme. Isso pode ser observado em seu texto, em que assumiu seu local na história, usando as marcas de primeira pessoa do plural ao elencar itens do cotidiano da família que vivia na idade da pedra.

É destaque também a forma como a criança resumiu os principais aspectos do filme, com a descrição das cenas iniciais, perpassando pela mudança de residência, aqui tratado por “quando temos uma vida nova”, e o relato da cena final, em que a família mostra-se adaptada ao novo ambiente – “nós brincamos de corrida com os animais”.

Nas redações do CRAEF 2, o tema dos animais foi retratado de forma interessante pelos alunos, uma vez que, enquanto vários lamentavam ainda não conhecer um zoológico, outros afirmavam que já haviam ido a esse local, como atestam os textos a seguir:

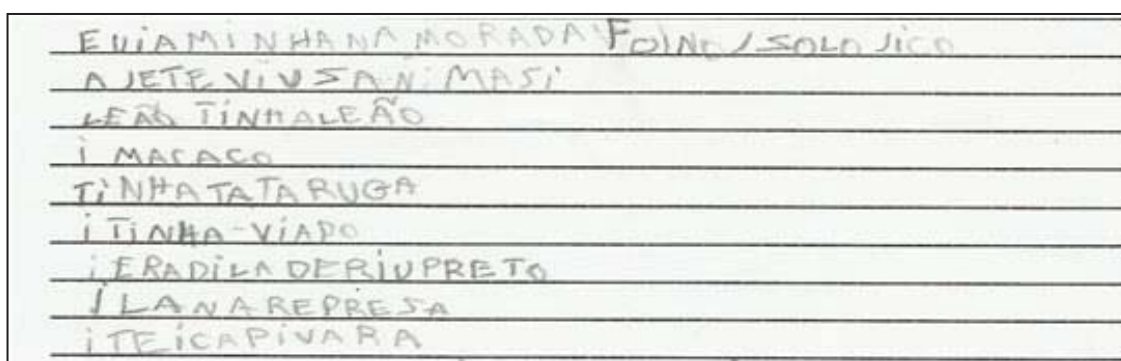
Figura 14 – Amostra da Escola 2 (Aluno 03)



Nesta composição, o discente trouxe a cidade de “Rio Preto”, referindo-se a São José do Rio Preto como o local onde pôde ver os animais, chamado de zoo: “leão”, “macaco”, “cobra”, “jacaré” e “tartaruga”, dentre outros. No entanto, ele lista uma série de outros que não estão presentes no zoológico desta cidade, visto que são próprios de outros ambientes; aquáticos, tal como o “golfinho” ou, ainda, climáticos, como o “panda”, por exemplo.

Outra referência interessante é a opinião do aluno sobre qual o animal de que ele mais gostou: “Eu gostei do leão porque ele é bravo e forte e tem unha grande”. Aqui há uma opinião do senso comum, pois, da maneira como ele se apresenta no zoológico, geralmente deitado e inexpressivo, não há como verificar a ferocidade e a força deste animal.

Figura 15 – Amostra da Escola 2 (Aluno 07)



Apesar de morar em Votuporanga, este estudante fez um comentário pertinente ao tema, apontando um animal que aparece “com relevante frequência” na represa de São José do Rio Preto, “a capivara”, mostrando que independentemente do local social, as crianças fazem inferências significativas do assunto em voga.

Neste texto, o discente também traz vários animais, porém, em formato de lista, o que destoa daquilo que se espera de uma produção textual. Por esse motivo, retomamos a importância de a criança realizar este tipo de atividade com frequência regular, a fim de que, a

cada vez que entrar em contato com a escrita de um texto, ela compreenda sistematicamente a maneira mais adequada de proceder esta importante tarefa.

Realizamos este comentário, posto que essa é uma das duas únicas redações coletadas deste aluno, pois, nas demais, ele não estava presente durante a etapa da coleta. Infelizmente, a falta de uma fiscalização mais pormenorizada dos órgãos competentes, bem como de uma ação mais contundente aos responsáveis pelos discentes faltosos, é uma das causas que contribui para ocorrências como essas citadas aqui.

Observando as produções de maneira geral, os discentes da unidade escolar 1, pertencentes ao centro, buscaram relacionar a vida dos habitantes das cavernas ao seu próprio cotidiano, destacando as diferenças de hábitos destes ambientes. Acreditamos que a motivação para este padrão seja a identificação que tiveram entre os personagens da história vista no vídeo e a de suas próprias famílias, pois a grande maioria dos alunos fez comparações, por exemplo, em relação à falta de tecnologia e de como seria difícil viver naquela época.

Já com o segundo grupo de estudantes, os textos apontaram outra realidade: por propor um tema leve e agradável, trabalhado em sala por meio do diálogo e de imagens, esperávamos que as crianças da periferia se sentissem motivadas a produzir várias sentenças. Os alunos, contudo, escreveram poucas palavras, enquanto outros apenas se limitaram a descrever os nomes dos animais de que mais gostam ou ainda quais gostariam de conhecer, contrariando a expectativa da pesquisadora.

Por ser uma primeira coleta, estes resultados só confirmam a necessidade de trabalhar o léxico, bem como o léxico-fraseológico em sala de aula, a fim de proporcionar uma oportunidade de aprendizagem desde as séries iniciais, conforme já pontuado em páginas anteriores.

Tendo traçado algumas características e análise do Tema 01, passamos, a seguir, ao Tema 02.

Tema 02 – Sobre a cidade de Votuporanga – 05/2015

Visto que o tópico acerca do concurso cultural já foi abordado na metodologia, limitaremos este espaço à análise lexical dos itens de destaque neste ambiente.

O intuito desta segunda redação foi observar o que os alunos relatam sobre o local onde moram e quais as suas impressões sobre esse lugar.

Iniciamos com a *Wordlist*, que apresenta os dados estatísticos das duas escolas em

estudo neste trabalho, como foco no Tema 02:

Tabela 11 – Comparação do CRAEF 1 e 2 (Tema 02)

DADOS ESTATÍSTICOS DO 2º TEMA DE REDAÇÃO			
CRAEF 1 – CENTRO		CRAEF 2 – PERIFERIA	
Redações digitadas	23	Redações digitadas	24
<i>Tokens</i> (número total de palavras)	2.132	<i>Tokens</i> (número total de palavras)	1.633
<i>Types</i> (número de palavras distintas)	422	<i>Types</i> (número de palavras distintas)	317

Esta foi a única proposta em que a coleta de redações foi maior na periferia do que na escola do centro. Além disso, é significativo observar o aumento de 810 palavras em relação ao Tema 01, da escrita sobre animais, denotando que o tema sobre a cidade de Votuporanga proporcionou às crianças uma identificação maior, visto que puderam falar sobre o seu cotidiano e suas preferências, bem como a opinião que possuem do município.

Contudo, apesar desta melhora linguística, os discentes do CRAEF 2 ainda redigiram menos vocábulos em comparação aos alunos do CRAEF 1: há precisamente 499 palavras (*tokens*) de diferença entre as duas turmas.

Novamente, comprovamos a afirmação feita sobre o grupo anterior de que este tema, ou seja, a escrita sobre o local onde vivem, despertou a atenção e o interesse dos estudantes em escrever. Assim, ressaltamos a necessidade de o professor oferecer diversos argumentos textuais à sua turma, a fim de oferecer ao aluno a chance de se expressar sobre outros assuntos, principalmente àqueles que permitam à criança tecer sua opinião, com o posicionamento frente a algum tema relevante para as mesmas.

Paralelamente ao primeiro rol de textos, neste tema também efetuamos a análise por meio da *Keyword* para observar a lista das dez palavras mais representativas desta proposta, conforme segue:

Tabela 12 – *KeyWords* gerada pelo software na análise do 2º Tema – CRAEF 1 e 2

KEYWORDS DO TEMA 2					
CRAEF 1			CRAEF 2		
N	Key Word	Freq.	N	Word	Freq.
1	EU	132	1	DIVIRTO	6
2	VOTUPORANGA	44	2	PRACINHA	5
3	VOU	74	3	VÓ	4
4	MORO	26	4	VÔ	3
5	PRA	35	5	AMO	4
6	DIVERTIR	24	6	TIA	4
7	MINHA	43	7	IRMÃOS	5
8	FAMÍLIA	25	8	PIPA	3
9	ESCOLA	32	9	PEGA	4
10	PRACINHA	12	10	IR	6

Para os alunos do centro da cidade, observamos a incidência de verbos, como “ir” (vou), “morar” (moro) e “divertir”, além dos locais: “escola”, “pracinha” e “Votuporanga”, denotando uma variedade lexical significativa para crianças que ainda cursam o terceiro ano do ensino fundamental I. Já, os alunos da periferia, apontaram outros verbos, como “ir” e “divertir” (divirto). Outro dado interessante corroborou uma verificação da redação anterior (Tema 01), em que as crianças deste ambiente nem sempre denominam “família” a formação de pessoas que moram com eles. Neste caso, o tema de Votuporanga levou os alunos a elencarem outros familiares com os quais residem, tais como “vó”, “vô”, “tia” e “irmãos”. Aqui também aparecem vários elementos lúdicos, como: “pipa”, “pega-pega” e “pracinha”.

Prosseguindo a análise desta proposta, baseadas em quatro redações escolhidas aleatoriamente, sendo duas do CRAEF 1 e duas do CRAEF 2, elaboramos a lista de palavras-chave, por meio da ferramenta *KeyWords*. Estes quatro textos foram contrastados com o corpus de referência *Folha Wordlist*:

Tabela 13 – *KeyWords* das Amostras do 2º Tema de redação

CRAEF 1 – CENTRO				CRAEF 2 – PERIFERIA			
ALUNO 10		ALUNO 24		ALUNO 06		ALUNO 17	
KEY WORD	FREQ.	KEY WORD	FREQ.	KEY WORD	FREQ.	KEY WORD	FREQ.
Votuporanga	02	Eu	07	Votuporanga	02	Pracinha	02
Moro	02	Aqui	05	-	-	Votuporanga	03
Eu	04	Votuporanga	02	-	-	Máximo	02
-	-	Muito	04	-	-	Eu	04

Visto que o tema da redação era sobre a cidade de Votuporanga, é natural que este topônimo tenha se destacado em todas as produções. O pronome pessoal do caso reto “Eu”

também obteve destaque, já que os estudantes redigem em primeira pessoa suas opiniões.

Em seguida, temos a presença do verbo “moro” e do advérbio de lugar “aqui”, reforçando o léxico acerca da cidade. Entre o rol de palavras-chave apresentados, o vocábulo “máximo” se destaca, por denotar uma opinião pessoal sobre o prefeito que, de acordo com o ponto de vista do autor do texto, “colocou” os brinquedos na praça onde ele frequenta.

A seguir, apresentamos imagens com trechos das redações descritas acima, já que a partir da escrita, é possível observar os diferentes olhares que possuem sobre a mesma cidade.

Figura 16 – Amostra da Escola 1 (Aluno 10)

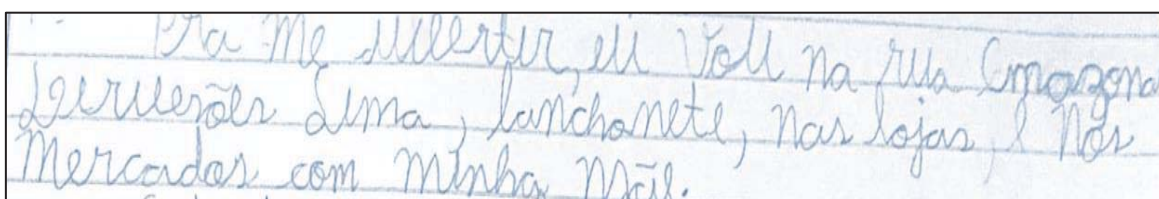
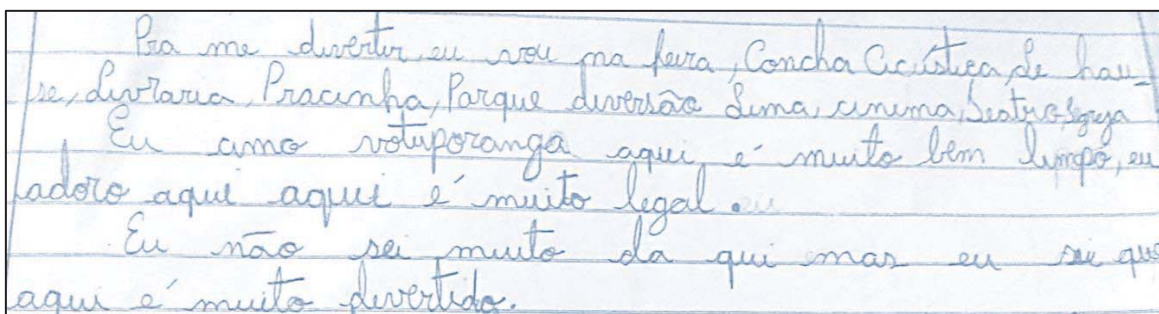


Figura 17 – Amostra da Escola 1 (Aluno 24)



No rol de escritas da escola 1, localizada no centro da cidade, a descrição de lugares particulares foi um tema que se destacou bastante. Podemos comprovar a afirmativa pela citação de “Diversões Lima”, local de diversões infantis com diversos brinquedos; “Lan House”; “livraria”; “lanchonete”, todos locais cuja entrada é condicionada a pagamento e ou é necessário ir com dinheiro para consumir os produtos ali existentes.

Vejamos como este tema é articulado no CRAEF 2, a partir do ponto de vista e posicionamento dos estudantes da periferia:

Figura 18 – Amostra da Escola 2 (Aluno 17)

PRA ME DIVERTIR EU VOU NA PRAÇINHA
 EU GOSTO DA CIDADE DE VOTUPORANGA,
 PORQUE TODO QUASE DIA EU VOU
 NA PRAÇINHA BALANÇAR O
 PRAFEITO É O MAXIMO PORQUE
 ELE COLOCOU OS BRINQUEDOS
 GANGORRA BALANÇO ESCORREGADIA

Figura 19 – Amostra da Escola 2 (Aluno 24)

PRA ME DIVERTIR EU VOU NA CASA DO MEU
 AMIGO, LINDO DE BICICLETA ETC...
 EU GOSTO DA CIDADE DE
 VOTUPORANGA, PORQUE DOME IR SHOPPING, NO
 CINEMA, NO JOGO DE FUTEBOL, NO
 ZOOLOGICO PORQUE É MUITO LEGAL A
 Z BOM TAMBÉM VIAJA E I NA
 IGREJA BOM ISSO FOI TUDO TIAU.

Diferentemente do material analisado na escola do centro, as produções do CRAEF 2, situada na periferia de Votuporanga, demonstram uma predileção por locais comunitários, tais como: “pracinha”, “jogo de futebol no campinho”, “igreja”, dentre outros. É interessante observar o relato acerca das melhorias feitas em relação aos brinquedos da praça: “gangorra, balanço, escorregador”, o que nos leva a crer que os alunos percebem as mudanças ocorridas nos lugares públicos. Já o trecho “quase todo dia eu vou na pracinha balançar” nos dá um indicativo de que este local é muito frequentado pelas crianças para diversão.

Também houve a descrição de locais que não existem em Votuporanga como, por exemplo, o zoológico e o shopping, este último em construção na cidade desde o ano de 2015, caracterizando o desconhecimento que vários discentes possuem do espaço em que vivem. Tanto o zoológico quanto o shopping mais próximo de Votuporanga estão localizados na cidade de São José do Rio Preto, aproximadamente a 88 quilômetros de distância do município, denotando ainda uma hipótese de confusão geográfica do aluno em descrever o espaço em que habita.

Tabela 14 – *Tokens, types e type-token ratio* das amostras do Tema 02

CRAEF 1 – CENTRO				CRAEF 2 – PERIFERIA			
ALUNO 10		ALUNO 24		ALUNO 06		ALUNO 24	
<i>Tokens</i>	107	<i>Tokens</i>	90	<i>Tokens</i>	68	<i>Tokens</i>	85
<i>Types</i>	74	<i>Types</i>	68	<i>Types</i>	56	<i>Types</i>	62
<i>Type-token ratio (ttr)</i>	69	<i>Type-token ratio (ttr)</i>	76	<i>Type-token ratio (ttr)</i>	82	<i>Type-token ratio (ttr)</i>	73

Conforme explicado na metodologia, a variedade lexical é verificada pela relação entre *tokens* e *types*, ou seja, número total de palavras e número de vocábulos não repetidos. Observamos que tanto no CRAEF 1 quanto no CRAEF 2, os discentes possuem um alto índice de itens sem repetição, se compararmos aos *tokens*. A variação gera o dado *Type-token ratio* (TTR), que, de uma forma geral, mostrou um valor alto para alunos de Ensino Fundamental I.

Neste contexto, a concepção de comunidade e local social nortearam as escritas, em concordância com as produções escolhidas dos corpora. Esta análise demonstrou que, apesar de redigirem sobre o mesmo tema, crianças provenientes de diferentes contextos refletem seu local social em sua escrita. A descrição deste ambiente provavelmente é um dos fatores determinantes para que os alunos diversifiquem o léxico, trazendo nomes de lugares e suas impressões pessoais sobre os mesmos. Como resultado, tem-se uma linguagem viva, dinâmica e variada (do ponto de vista lexical) do município de Votuporanga.

Os alunos da Escola 1 (CRAEF 1), do centro da cidade, tendem a abordar aspectos positivos da cidade, destacando os locais particulares; enquanto os discentes da Escola 2 (CRAEF 2), situados na periferia de Votuporanga elencam situações ocorridas em sua realidade social, citando os lugares públicos que frequentam para se divertirem.

Da mesma forma que as crianças que habitam na periferia, aquelas que residem no centro da cidade também estão sujeitas a vivenciarem obstáculos em suas vidas. Percebemos, contudo, que, em um local mais próximo ao centro da cidade, com mais oportunidades de trabalho, lazer, cultura e diversão, há mais possibilidades desse público ser mais bem assistido frente às suas necessidades e, portanto, viver melhor sem precisar se deslocar para outros locais mais afastados de suas casas.

Prosseguindo a análise, tratamos ao Tema 03 para verificar como o léxico foi redigido pelos dois grupos.

Tema 03 – Proposta da semente mágica – 06/2015

Para encerrar o primeiro semestre, o objetivo do terceiro tema foi verificar o grau de criatividade das crianças frente a uma proposta que despertasse aos alunos a oportunidade de criação a partir do argumento da semente mágica.

Neste sentido, as crianças foram instigadas a ultrapassar os limites do reino vegetal, podendo tratar a semente como uma infinita possibilidade de realizações semânticas.

A seguir, tem-se os dados estatísticos da escrita dos alunos:

Tabela 15 – Comparação do CRAEF 1 e 2 (Tema 03)

WORDLIST DO 3ª TEMA DE REDAÇÃO			
CRAEF 1 – CENTRO		CRAEF 2 – PERIFERIA	
Redações digitadas	25	Redações digitadas	15
<i>Tokens</i> (número total de palavras)	2.539	<i>Tokens</i> (número total de palavras)	1.131
<i>Types</i> (número de palavras distintas)	593	<i>Types</i> (número de palavras distintas)	373

Esta foi a produção que teve a menor coleta de textos dos alunos da periferia, somente 15 participantes. Visto que foi realizada antes do recesso escolar do mês de julho, observa-se a questão, já relatada anteriormente sobre este público, que se refere ao êxodo urbano. O fluxo de trabalho dos pais nas usinas da região reverbera também na escola, com a transferência e recebimento de crianças, a depender da época do plantio e colheita da cana-de-açúcar.

Acerca do número total de palavras, observa-se que os alunos do CRAEF 1 continuam seguindo em uma escala crescente de escrita, redigindo 407 palavras a mais do que a anterior. Outro item importante diz respeito às palavras distintas, ou seja, aquelas que só aparecem uma vez no corpus. Sobre essa informação, ambas as escolas aumentaram este escore, sendo o CRAEF 1 em 171 palavras e o CRAEF 2 em 56 vocábulos.

Com base na análise acima e nos dados das redações seguintes, as quais devido a questões de delimitação não foram apresentadas nesta dissertação, observamos que conforme avançam em sua aprendizagem sobre a língua portuguesa, as crianças ampliam gradativamente a competência colocacional. Dessa maneira, elas possuem condições cognitivas para elaborar sentenças cada vez mais completas e complexas.

Nesta proposta também observamos o uso de palavras próprias do plantio, elencadas na tabela das *KeyWords*:

Tabela 16 – *KeyWords* gerada pelo software na análise do 3º Tema – CRAEF 1 e 2

KEYWORDS DO TEMA 3					
CRAEF 1			CRAEF 2		
N	Key Word	Freq.	N	Key Word	Freq.
1	SEMENTE	60	1	SEMENTE	39
2	EU	76	2	ELA	25
3	FEIJÃO	33	3	MÁGICA	11
4	ELA	64	4	E	78
5	MÁGICA	29	5	ENTÃO	14
6	PLANTAR	24	6	PLANTAR	15
7	NASCER	11	7	EU	16
8	FALOU	22	8	CRESCEU	4
9	CRESCEU	9	9	ÁRVORE	6
10	ERA	36	10	PASSARAM	4

Neste rol, juntamente com as palavras chave “semente” e “mágica”, próprias da atividade sugerida, apareceram “plantar”, contabilizadas 24 vezes no CRAEF 1 e 15 vezes no CRAEF 2; “nascer”, com 11 ocorrências na escola do centro; “cresceu”, também citado pelos dois corpora e ainda “passaram”, com 4 registros na escola da periferia.

Este dado nos aponta que, apesar de os alunos não conviverem no meio rural, conhecem o vocabulário próprio do plantio de uma semente, pelo conhecimento empírico que possuem deste contexto. Partindo desta hipótese, os estudantes também estariam aptos a desenvolver outras tarefas que envolvam o léxico, e que os auxiliariam a desenvolver ainda mais a expressão do vocabulário sobre a língua portuguesa do Brasil.

De forma análoga à proposta anterior, utilizamos quatro textos, dois de cada corpus, para efetuar comentários acerca do léxico desses alunos.

Tabela 17 – *KeyWords* das Amostras do 3º Tema de redação

CRAEF 1 – CENTRO				CRAEF 2 – PERIFERIA			
ALUNO 1		ALUNO 26		ALUNO 04		ALUNO 08	
KEY WORD	FREQ.	KEY WORD	FREQ.	KEY WORD	FREQ.	KEY WORD	FREQ.
SEMENTE	2	PORISSO	4	ELA	8	EU	6
PLANTAR	2	MÁGICA	3	VELHINHA	3	IA	3
VOU	3	COMER	3	DIAMANTES	2	BASTANTE	3
MINHA	3	COMIDA	3	ESPEROU	2	DINHEIRO	4
-	-	FOME	3	-	-	COISA	3
-	-	PRA	3	-	-	MINHA	3

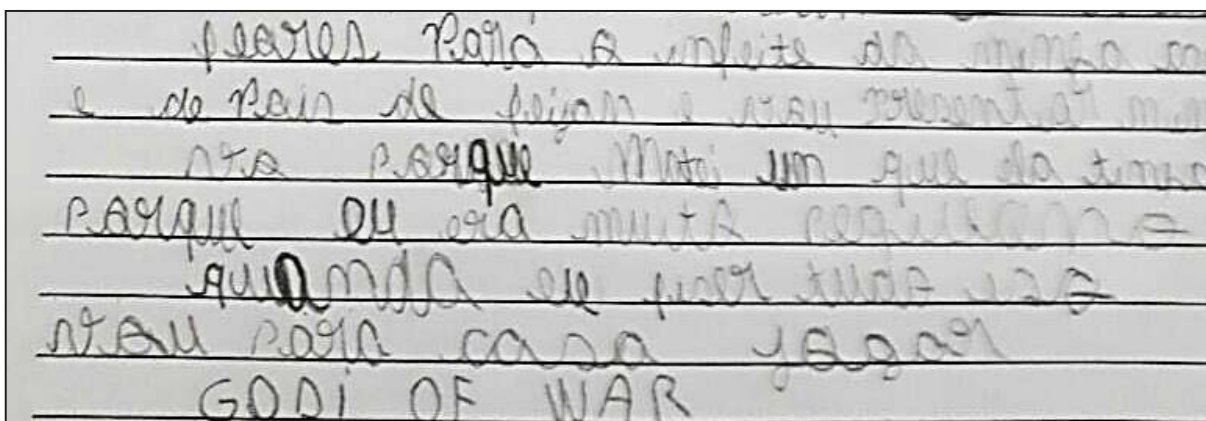
Conforme citado acima, as palavras do campo semântico ganharam destaque na maioria dos textos, como é o caso do aluno 01, do CRAEF 1, comprovados pelos termos

“semente”, com 2 ocorrências, e “plantar”, também com 2 recorrências. Já o aluno 26, ainda da escola do centro, trouxe outros termos que pertencem à questão alimentar, indiretamente relacionadas ao tema plantar: “comida”, apareceu 3 vezes no corpus, e “fome” – 3 repetições.

Já em redações que não enfocaram no plantio de verduras e legumes, transpareceram outros temas alheios a uma plantação tradicional. A título de exemplificação, seguem dois textos do CRAEF 2: o estudante 04 apresenta termos “velhinha”, 3 ocorrências, e “diamantes”, com 2 ocorrências. Outra criança, a de número 08, foi o protagonista da história, indicado pela frequência da palavra “eu” (6). Outras palavras que se destacam neste contexto são “bastante”, 3 vezes, e “dinheiro”, com 4 ocorrências.

Seguem a descrição destas produções:

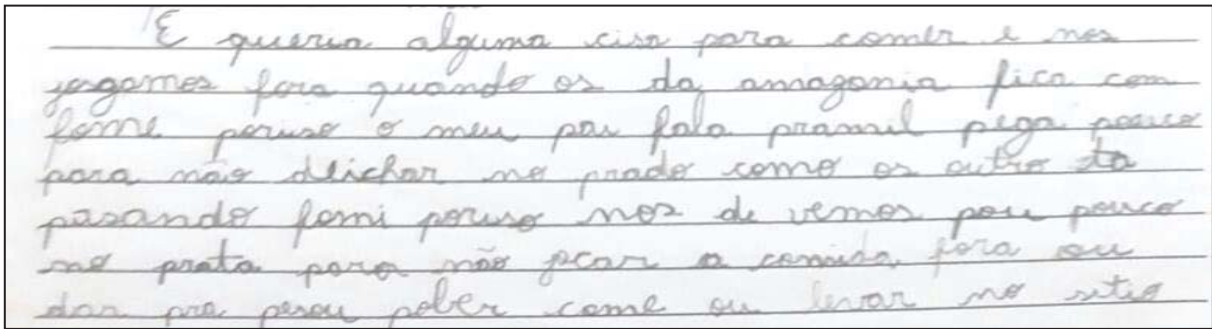
Figura 20 – Amostra da Escola 1 (Aluno 01)



A escolha desta redação deu-se pela espontaneidade do aluno em escrever enquanto relembra fatos pessoais, comprovando a teoria de que desde crianças, todos nós temos um repertório lexical próprio, que pode ser acessado a qualquer momento. Ao afirmar que irá “presentear minha vó porque matei um que ela tinha porque eu era muito pequeno”, o discente escreve “matei” referindo-se a flor da avó, resgatando da memória lembranças sobre o tópico dado, o que o auxilia a produzir um texto único e com identidade.

Outro fato que comprova isso é a projeção do futuro a curto prazo, algo realizado também mentalmente pelas crianças, e que aparece na escrita quando o estudante afirma “quando eu fizer tudo isso vou para casa jogar GOD OF WAR.” Aqui há o planejamento de uma atividade que será efetuada após ele realizar as tarefas que programou no início do texto.

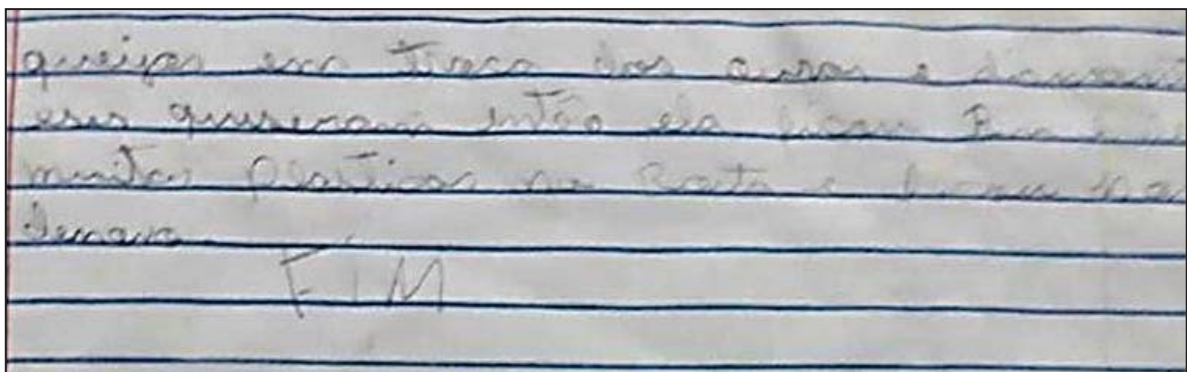
A próxima imagem ilustra outra redação, também proveniente do CRAEF 1:

Figura 21 – Amostra da Escola 1 (Aluno 26)

A aluna 26 apresenta uma preocupação genuína com o próximo, representada por vários trechos ao longo do texto: “por isso o meu pai fala pra mim pega pouco pra não deixar no prato como os outros tá passando fome”; “por isso nós devemos por pouco no prato para não jogar a comida fora ou dar pra pessoa pobre comer ou levar no sítio para as galinhas ou ficar em uma sacola plástica”.

Esta é uma outra característica inerente ao vocabulário infantil, reproduzir o discurso dos adultos quando estes fazem alguma exortação ou conselho. Pela forma como intercala a fala do pai com as reflexões que ela própria realiza: “porque tem gente que não tem comida e fica doente e tem pessoas que morre de fome ou fica muito mau”, observamos que a discente já demonstra características de respeito e preocupação com os demais, valores estes passados pela família da menina, e já faz parte do repertório linguístico desta criança.

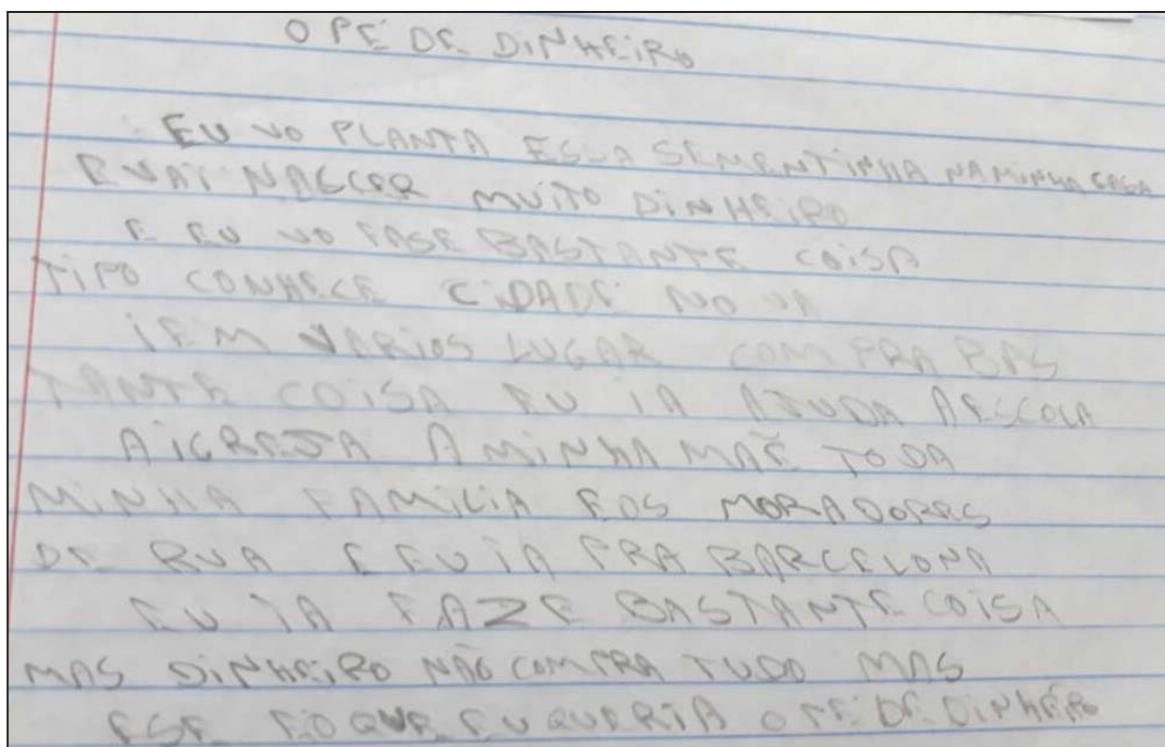
Do CRAEF 2, trazemos outros dois textos, conforme seguem:

Figura 22 – Amostra da Escola 2 (Aluno 04)

Aqui tivemos uma reinvenção da história do João e o pé de feijão, porém com a mudança de vários itens. Nesta escrita, a criança faz uma narrativa bem humorada, retratando uma senhora bastante ativa, que “sobe na árvore”, “troca queijos com os ratos” em troca de diamantes e faz “cirurgias plásticas”.

A inventividade infantil é muito bem-vinda dentro da sala de aula, pois indica a capacidade que os estudantes têm em desenvolver novos mundos e realidades, possível dentro de um texto que permita essa criação. Mais uma vez chamamos a atenção para o desenvolvimento de atividades que propiciem a reflexão lexical, a fim de possibilitar mais textos criativos como este apresentado. Outra escrita segue, proveniente do mesmo corpus:

Figura 23 – Amostra da Escola 2 (Aluno 08)



Nesta composição, o aluno protagoniza a história, pois assinala o que fará com o dinheiro que brotará da sua semente mágica. Dentre os planos, estão: “ajudar a escola a igreja a minha mãe toda minha família e os moradores de rua e eu ia pra Barcelona”. Assim, as crianças demonstram sua pureza, seja em exteriorizar o desejo íntimo que possuem, seja por citarem em suas produções, pessoas e locais que estão no seu cotidiano. Outra frase de destaque neste texto é a afirmativa “mas dinheiro não compra tudo”, sinalizando o início de uma reflexão, que não foi desenvolvida pelo discente.

Em seguida, apresentamos o índice das palavras totais (*tokens*), em contraste com as palavras que não se repetem no texto (*types*):

Tabela 18 – *Tokens, types e type-token ratio* das amostras do Tema 03

CRAEF 1 – CENTRO				CRAEF 2 – PERIFERIA			
ALUNO 1		ALUNO 26		ALUNO 04		ALUNO 8	
<i>tokens used for word list</i>	66	<i>tokens used for word list</i>	144	<i>tokens used for word list</i>	132	<i>tokens used for word list</i>	79
<i>types (distinct words)</i>	50	<i>types (distinct words)</i>	86	<i>types (distinct words)</i>	85	<i>types (distinct words)</i>	49
type-token ratio (TTR)	76	type-token ratio (TTR)	60	type-token ratio (TTR)	64	type-token ratio (TTR)	62

Os números indicam uma variação grande entre as redações, posto que cada indivíduo produz de acordo com a sua capacidade linguística, dentro dos itens já elencados aqui, como conhecimento prévio e identificação com o assunto, repertório lexical, dentre outros aspectos.

Assim, constatamos que, dentro de uma mesma sala, há crianças com perfis díspares de escrita. No CRAEF 1, por exemplo, o aluno 1 só redigiu 66 palavras totais, contudo, 50 foram vocábulos, resultando em uma alta razão *type-token ratio* entre esses dados, 76. Na mesma sala, a estudante 26 escreveu 144 palavras, contudo, mesmo tendo elaborado 86 termos sem repetições, seu escore lexical foi menor – 60.

Nas amostras do CRAEF 2, ocorreu algo diferente com os alunos 4 e 8. Enquanto o primeiro produziu 132 palavras, destes 85 vocábulos, resultaram no valor 64. Já o segundo, apesar de ter escrito menos, 79 palavras, o valor de 49 termos ocasionou uma proporção de 62. Assim, as análises mostram que a variedade lexical dessas crianças é relevante, considerando que ainda estão no ensino fundamental I.

De forma geral, ambas as turmas demonstraram um grande interesse em redigir o tema concebido e o fizeram com afinco. Após as análises desta terceira proposta, verificamos que os alunos do centro tiveram a seguinte atitude frente ao tema “semente mágica”:

- 8 crianças reescreveram a história João e o pé de Feijão, alguns da forma tradicional, e outros fizeram algumas alterações, como o nome do personagem principal, do gigante e demais itens;
- 10 discentes plantaram frutas, verduras e legumes, como cenoura, pimentão, morango e maçã, por exemplo;
- 4 redigiram um texto cuja semente mágica nasceria brinquedos: bola de handebol e bola de futebol, com três ocorrências;
- 1 menina descreve animadamente sobre as fadas que saíram de sua plantação;
- 1 aluno plantou uma semente que deu origem a uma árvore de dinheiro e

- 1 aluna colheu diamantes da plantação de suas sementes mágicas.

Já com os discentes do CRAEF 2, obtivemos o seguinte resultado:

- 8 estudantes cultivaram algum item relacionado ao reino vegetal, como árvore, tomate, acerola, milho, maracujá, dentre outros;
- 2 crianças elaboraram uma narrativa com a descrição dos brinquedos que brotaram das sementes mágicas, bolas de futebol;
- 2 discentes colheram dinheiro, oriundos do plantio das sementes;
- 1 aluna recolheu pedras preciosas de sua plantação e intercalou com alguns elementos da história do João e o pé de feijão;
- 1 aluno relatou que tinha a semente da vida, que curava quem a tinha consigo e
- 1 aluno escreveu que gostaria de plantar “no fundo do meu coração, amor, atenção tudo o que há de bom”.

É pertinente afirmar que a professora regular das salas constantemente lê livros para as crianças e, durante tais atividades orais, a história citada sobre o pé de feijão já foi narrada aos dois grupos, e, portanto, todos conheciam este tema e estavam livres para reescrever a narrativa, se assim o desejassem.

Dessa maneira, observamos que os textos mais criativos vieram da escola da periferia, seja pela variedade de “plantações” ou, ainda, pela originalidade em criar uma semente da vida, além do simples, mas tocante desejo de plantar bons sentimentos.

Por isso, os docentes de ensino fundamental I não podem rotular seus educandos e efetuar julgamentos a partir do local social em que se encontra. As crianças são seres em formação e naturalmente criativos e inovadores, como acabamos de observar nesta amostra.

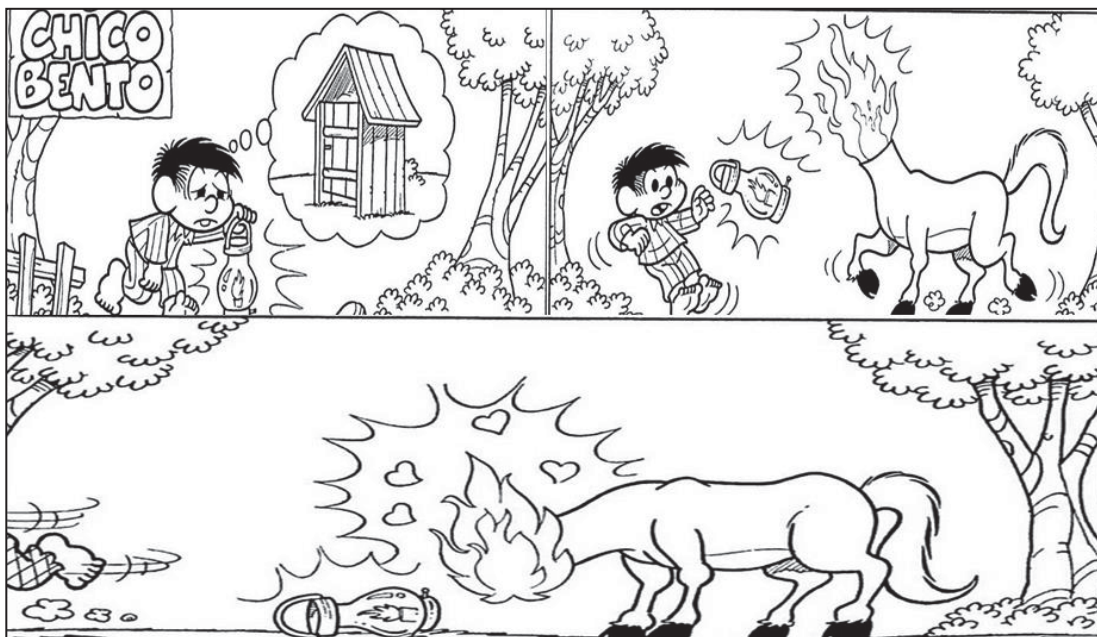
Após as devidas explicações, segue a análise lexical do Tema 04.

Tema 04 – Escrita de uma narrativa folclórica – 08/2015

Após as férias escolares, o penúltimo bloco de redações teve o propósito de tratar do folclore, tema já conhecido pelos discentes e frequentemente trabalhado pelos professores. Assim, o enfoque desta proposta foi possibilitar que as crianças pudessem dissertar sobre um tema que já conhecem.

Assim, para oferecer um estímulo a mais, adicionamos uma tirinha da Turma da Mônica com o personagem Chico Bento, morador da Vila Abobrinha, localidade rural em que ocorrem várias situações com teor popular e folclórico, como indica a figura:

Figura 24 – Imagem utilizada como apoio ao Tema 04



Na figura acima, temos o mote da história: Chico Bento sai durante a noite para ir ao banheiro, também chamado de “casinha” pelos habitantes da área rural, e repentinamente se encontra com a mula sem cabeça. Assustado, o menino foge, deixando para trás o lampião.

Em seguida, a personagem folclórica mostra-se apaixonada pelo item deixado pelo menino.

Paralelamente aos temas anteriores, a tabela abaixo inicia as discussões do Tema 04:

Tabela 19 – Comparação do CRAEF 1 e 2 (Tema 04)

WORDLIST DA 4º TEMA DE REDAÇÃO			
CRAEF 1 – CENTRO		CRAEF 2 – PERIFERIA	
Redações digitadas	27	Redações digitadas	21
<i>Tokens</i> (número total de palavras)	3.132	<i>Tokens</i> (número total de palavras)	2.464
<i>Types</i> (número de palavras distintas)	576	<i>Types</i> (número de palavras distintas)	479

Como se vê, as duas salas continuam avançando no que concerne à quantidade de palavras utilizadas nas produções. No CRAEF 1, foi produzido 593 vocábulos a mais do que na proposta anterior; enquanto o CRAEF 2 redigiu 1.333 mais palavras em relação ao texto 3,

visto que nesta coleta tivemos seis crianças redigindo a mais neste grupo. Consequentemente, no que se refere aos *tokens*, o número de palavras não repetidas também aumentou, nos dois grupos de estudantes.

No Tema 04, apresentamos também as *Keywords*:

Tabela 20 – *KeyWords* gerada pelo software na análise do 4º Tema – CRAEF 1 e 2

CRAEF 1			CRAEF 2		
N	Key Word	Freq.	N	Key Word	Freq.
1	MULA	99	1	MULA	105
2	CHICO	92	2	CHICO	74
3	LAMPIÃO	48	3	CABEÇA	67
4	CABEÇA	74	4	BENTO	38
5	BANHEIRO	52	5	LAMPIÃO	32
6	BENTO	50	6	BANHEIRO	32
7	AI	33	7	SEM	67
8	SEM	73	8	LAMPARINA	15
9	ELE	85	9	ELE	63
10	LAMPARINA	17	10	E	173

É interessante observar o paralelismo dos dados nos dois grupos de estudantes, já que, com exceção de “aí” e “e”, todas as demais oito palavras apareceram nos dois corpora, sendo a primeira, segunda e nona de forma idêntica, só variando a quantidade de ocorrências: “Mula” (99 vezes no CRAEF 1 e 105 no CRAEF 2); “Chico” (92 ocorrências na escola do centro e 74 na escola da periferia) e “ele” (com 85 repetições na escola 1 e 63 na escola 2). Por essas informações, é possível inferir que a presença das figuras influenciou muito a escolha lexical dos alunos, posto que limitou a variedade de palavras outrora apresentada nas demais propostas.

Posto os itens estatísticos, passemos à apreciação lexical desta proposta, com dois exemplos de cada corpus coletado. Inicialmente, analisamos as palavras-chave destas produções, conforme atesta a listagem:

Tabela 21 – KeyWords das Amostras do 4º Tema de redação

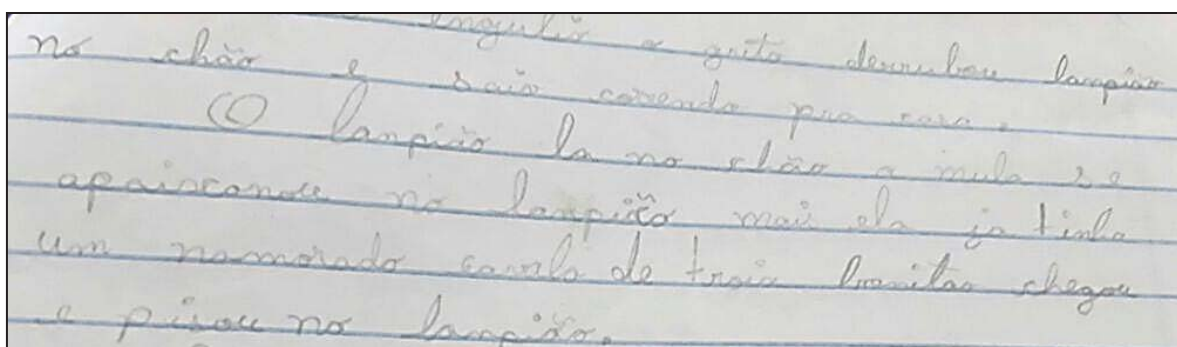
CRAEF 1 – CENTRO				CRAEF 2 – PERIFERIA			
ALUNO 6		ALUNO 22		ALUNO 05		ALUNO 19	
KEY WORD	FREQ	KEY WORD	FREQ	KEY WORD	FREQ.	KEY WORD	FREQ
MULA	9	MULA	9	MULA	3	LAMPARINA	8
CHICO	13	CHICO	10	LAMPARINA	2	MULA	5
LAMPIÃO	8	CURUPIRA	3	CABEÇA	3	ACORDOU	3
BENTO	3	-	-	OLHANDO	2	BENTO	3
NAMORADO	3	-	-	BENTO	2	CHICO	3
CABEÇA	4	-	-	-	-	FOGO	3
ESTAVA	5	-	-	-	-	SAIU	3
CHÃO	3	-	-	-	-	FLORESTA	2
FALOU	3	-	-	-	-	BANHEIRO	2

Na escola do Centro, como era de se esperar, “mula”, “Chico”, “Bento” e “cabeça” dominam os primeiros lugares, pois estão intrinsecamente ligados à proposta 04, respectivamente com 9, 13, 3 e 4 ocorrências. Também tivemos a presença de verbos, como “estar” (estava), com frequência 5 e “falar” (falou), com 3 repetências. Outra palavra de destaque neste contexto foi “lâmpião”, com 8 aparições. Um personagem que originalmente não estava previsto era “Curupira”, presente no texto da discente 22 do CRAEF 1, com 3 repetições.

Na unidade escolar da periferia, os estudantes 5 e 19 também repetiram os itens padrão nesta proposição textual, citando “mula”, com 3 e 5 citações; “cabeça”, com 3 recorrências; “Chico”, com 3 reiterações e “Bento”, que consta 2 vezes no texto da aluna 5 e 3 vezes no aluno 19.

Outras palavras de destaque foram “floresta” e “banheiro”, ambas com 2 repetições, além dos verbos “olhar” (olhando), “acordar” (acordou) e “sair” (saiu), respectivamente, com 2, 3 e 3 ocorrências.

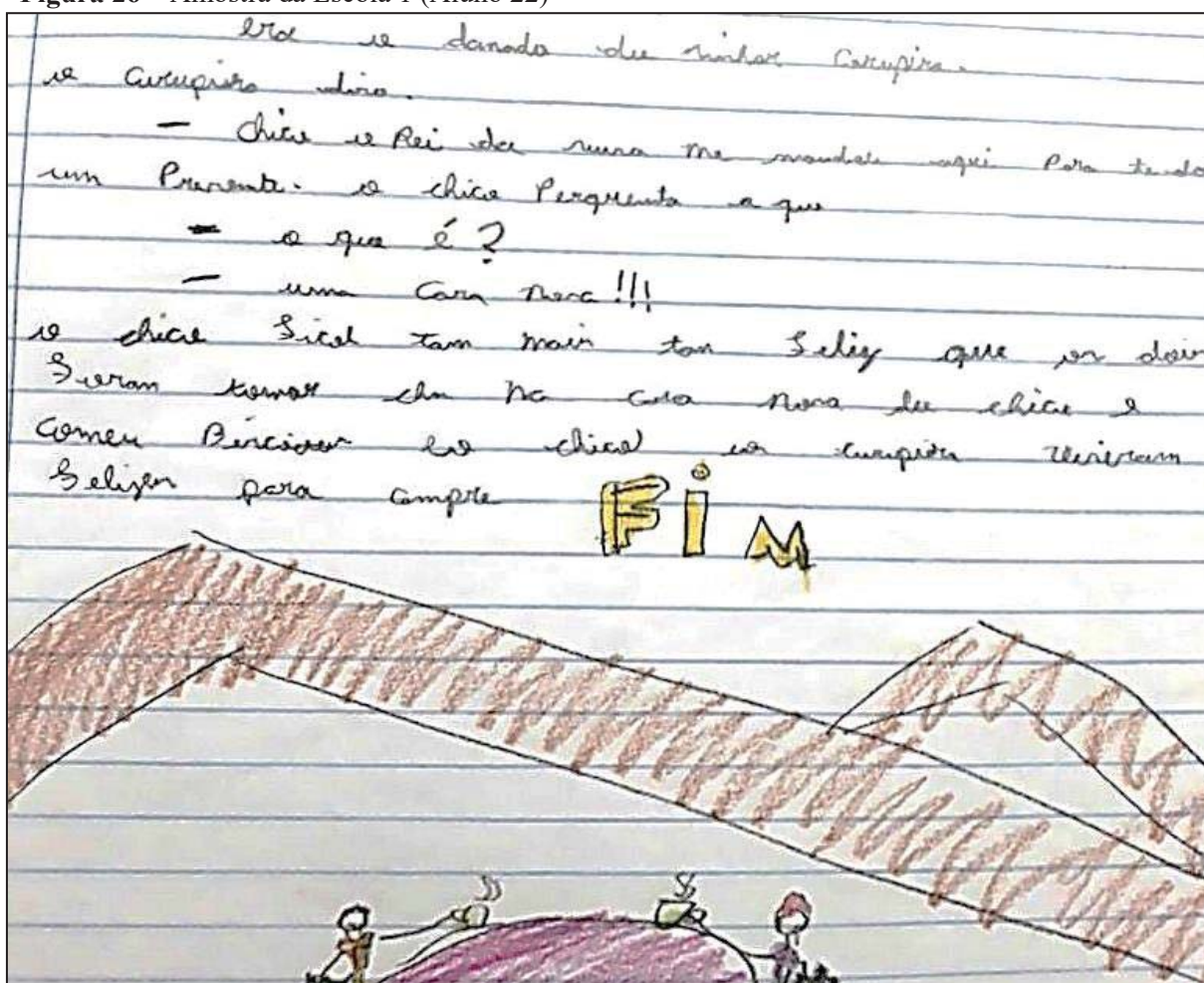
De posse destes dados, verificamos mais pontualmente o conteúdo destas produções.

Figura 25 – Amostra da Escola 1 (Aluno 06)

Escolhemos este texto em específico, por observar que o aluno tentou reproduzir o modo de falar do Chico Bento “e deu um grito: - AA!!! e falou: - eita só”, mostrando que conhece a variedade linguística das pessoas que moram na área rural. Outra situação interessante atesta para a espontaneidade do discurso infantil, que traz a onomatopeia de sons, e toda a sorte de expressão que eles conseguem transformar em linguagem.

No âmbito lexical, o aluno inovou, ao trazer outro personagem na narrativa: “o cavalo de tróia”. No bojo desta proposta, que traz a mula sem cabeça, esta figura é pertinente e oferece uma novidade inusitada ao texto. Neste contexto, a caracterização também chama a atenção: “ela já tinha um namorado cavalo de troia bonitão chegou e pisou no lampião”. Posto que a arma usada pelos gregos é conhecida pela opulência e majestade, a descrição do discente é condizente com o argumento, atestando para a importância do conhecimento de mundo e aporte lexical, já tratado aqui em páginas anteriores.

Figura 26 – Amostra da Escola 1 (Aluno 22)

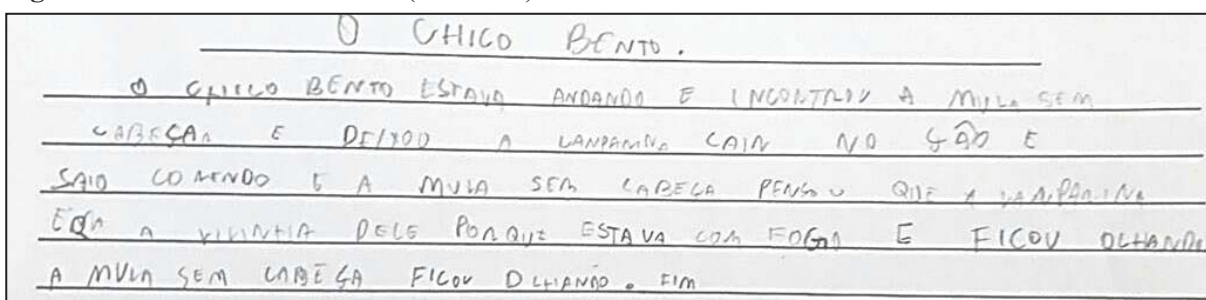


Esta escrita aponta um enredo consistente, pois, desde as primeiras linhas, indica que o

personagem Vitor, cujo apelido é Chico Bento, é pobre e não possui casa. A criança em análise foi a única aluna do centro que não citou o banheiro em seu texto, criando outra situação que motivasse o encontro do personagem com a mula sem cabeça. Por fim, a mula se transforma no curupira e realiza o sonho do menino de ter uma casa. A narrativa termina com os personagens lanchando na casa nova do Chico Bento. Apesar da simplicidade do texto, é notório o encadeamento de ideias e a coerência criada pela discente para encaixar a imagem sugerida com a história imaginada e descrita.

Segue a análise de e uma das amostras da periferia, Escola 2, Aluno 5:

Figura 27 – Amostra da Escola 2 (Aluno 05)

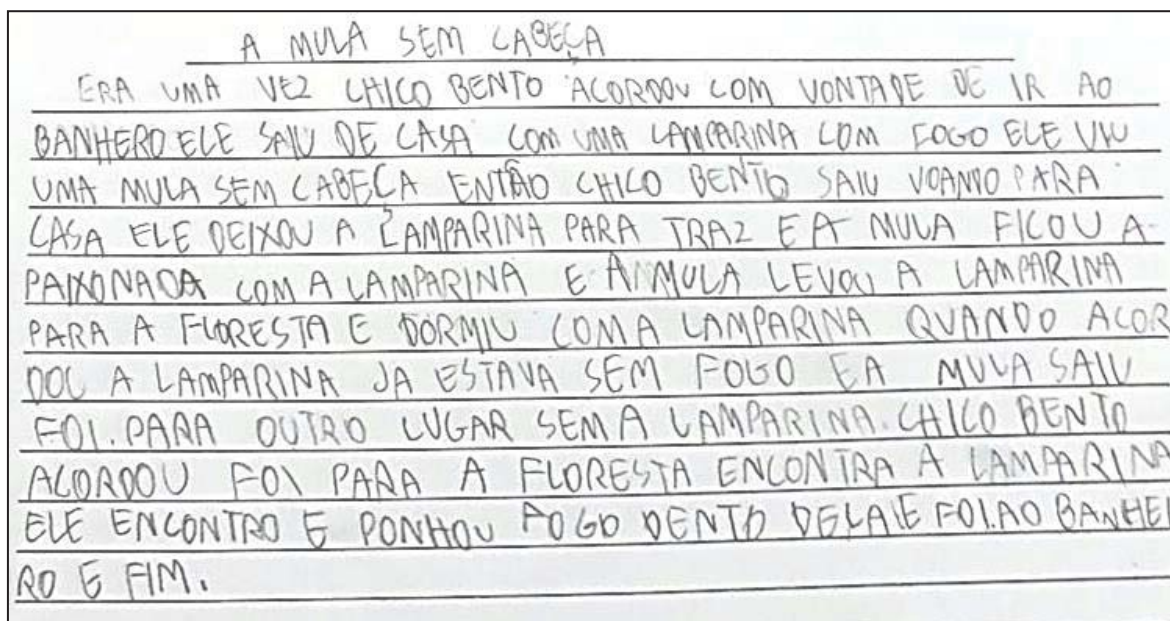


Por ter elaborado uma única oração, a criança não usa conectivos, o que torna essa narrativa um exemplo de texto que só descreve os elementos das figuras. Assim, refletimos sobre a necessidade dos educadores trabalharem atividades que desenvolvam a linguagem oral, já que há vários itens das imagens que não foram abordados pela criança, como: a floresta, a casinha, ou banheiro, o fato do Chico Bento estar sonolento, dentre outros.

Fazemos essa afirmação, pois, uma das funções do letramento é oferecer condições para o aluno expor sua opinião, fazendo-o pensar sobre o assunto tratado em sala de aula. Assim, nos momentos de escrita, a criança fará esse processo de refletir novamente acerca de um tema, geralmente relacionando-o ao seu repertório lexical e cultural, para ter mais condições de escrever acerca das propostas efetuadas pelo professor.

Na sequência, apresentamos a amostra da Escola 2, do aluno 19:

Figura 28 – Amostra da Escola 2 (Aluno 19)



Neste último texto, destacamos a informalidade, outra característica da escrita infantil, presentes em “então Chico Bento saiu voando para casa” e “ele encontrou e pôs fogo”, por exemplo. Esse traço é muito comum nas redações, pois demonstra a presença de traços provenientes da linguagem oral no texto escrito. Outra marca proveniente desta natureza diz respeito ao uso do pronome “ele”, que teve 4 ocorrências, e da conjunção aditiva “e”, com 7 repetições. Nota-se que o discente não segmenta as frases, mas as une por esses elementos, tornando mais difícil a leitura da produção textual e, conseqüentemente, o seu entendimento.

Neste sentido, uma opção de intervenção pedagógica a ser realizada pelo docente seria a reescrita deste texto coletivamente, com toda a sala. Após a leitura do professor, é possível realizar a escrita em conjunto com o grupo de estudantes, que decidem como refazer o texto, pontuando-o e inserindo os conectivos da maneira mais adequada.

Após as análises do léxico apresentado, discorre-se a observação acerca do número total de palavras, bem como das que não se repetem e a razão *type-token ratio* dos mesmos:

Tabela 22 – Tokens, types e type-token ratio das amostras do Tema 04

CRAEF 1 – CENTRO				CRAEF 2 – PERIFERIA			
ALUNO 6		ALUNO 22		ALUNO 05		ALUNO 19	
<i>tokens used for word list</i>	247	<i>tokens used for word list</i>	208	<i>tokens used for word list</i>	51	<i>tokens used for word list</i>	108
<i>types (distinct words)</i>	125	<i>types (distinct words)</i>	109	<i>types (distinct words)</i>	29	<i>types (distinct words)</i>	45
<i>type-token ratio(TTR)</i>	51	<i>type-token ratio(TTR)</i>	52	<i>type-token ratio(TTR)</i>	57	<i>type-token ratio(TTR)</i>	42

Nesta tabela, observamos que os dois discentes do centro redigiram mais de 200 palavras, 247 e 208 respectivamente, e também notamos um alto índice de palavras distintas, 125 e 109. Portanto, o valor que indica a riqueza lexical pontuou a escrita desses discentes em 51 e 52.

Os alunos que representam a periferia, por sua vez, escreveram bem menos vocábulos, 51 e 108. Em compensação, o número de itens distintos, a saber 29 e 45, possibilitou um escore no *Type-token ratio* de 57 e 42.

Finalizando esta proposta 4, observamos nos CRAEF 1 e 2 quais textos somente descreveram as imagens dadas na proposta e quais foram além da linguagem não verbal. Segue o resultado:

- No CRAEF 1, 9 (nove) alunos fizeram o relato das figuras, enquanto 18 (dezoito) trouxeram elementos novos, como os pais do Chico Bento, outros personagens da turma da Mônica e do folclore;
- No CRAEF 2, 7 (sete) crianças reproduziram por escrito o que havia nas gravuras, e 14 (quatorze) usaram a criatividade para inovar, seja pela participação de personagens do folclore ou ainda explorando o ambiente do sítio, adicionando cavernas, rios e floresta ao cenário.

Assim, é possível afirmar que sim, a presença do pictórico influenciou a escrita das crianças, contudo, este fato não impossibilitou o emprego da criatividade, visto que, para alguns discentes, as figuras serviram de inspiração para ir além da proposta, conforme observado nas análises efetuadas.

Finalizando o tópico, passemos a descrição lexical do Tema 05 e os procedimentos de análise na sequência.

Tema 05 – “Como eu posso ajudar as pessoas que vivem perto de mim?”

A última proposta de redações teve como foco oferecer a oportunidade das crianças redigirem sobre a solidariedade. Dessa maneira, decidimos ilustrar o tema com figuras que despertassem nos estudantes a empatia, a capacidade de se colocar no lugar do outro e perceber as suas necessidades:

Figura 29 – Ilustração do Tema 05



Nas imagens, é possível verificar duas situações diferentes: enquanto a primeira nos remete a um leito de hospital, com uma criança visitando a outra que está doente, a outra ilustração mostra uma família doando brinquedos a um menino carente. Os estudantes foram convidados a escrever sobre uma das duas situações dispostas nos desenhos.

Seguindo o padrão apresentado nas propostas anteriores, tem-se a tabela com os dados estatísticos do CRAEF:

Tabela 23 – Comparação do CRAEF 1 e 2 (Tema 05)

WORDLIST DA 5ª TEMA DE REDAÇÃO			
CRAEF 1 – CENTRO		CRAEF 2 – PERIFERIA	
Redações digitadas	27	Redações digitadas	24
<i>Tokens</i> (número de palavras total)	2.899	<i>Tokens</i> (número de palavras total)	1.817
<i>Types</i> (número de palavras distintas)	601	<i>Types</i> (número de palavras distintas)	475

Contrariando as expectativas, ambas as salas apresentaram queda na escrita da última proposta, sendo que o CRAEF 1 apresentou 233 palavras a menos do que a quarta proposta e o CRAEF 2 reduziu sua produção em 647 vocábulos. Apesar disso, o número de palavras diferentes foi levemente maior em relação ao texto folclórico: 25 para os estudantes do centro e 4 aos discentes da periferia.

Tabela 24 – *KeyWords* gerada pelo software na análise do 5º Tema – CRAEF 1 e 2

KEYWORDS DO TEMA 5					
CRAEF 1			CRAEF 2		
N	Key Word	Freq.	N	Key Word	Freq.
1	FALOU	36	1	ELA	50
2	ELA	61	2	MENINO	21
3	BRINQUEDOS	31	3	BRINQUEDOS	15
4	MENINO	28	4	FELIZ	18
5	VOCE	43	5	IRMÃ	16
6	AI	18	6	VISITAR	10
7	EU	46	7	DIA	35
8	DOENTE	19	8	HOSPITAL	15
9	MENINA	21	9	DOENTE	9
10	DOAR	15	10	MÃE	16

No último rol de textos, houve muitas semelhanças entre as palavras-chave das duas escolas, pois em ambas, os alunos se dividiram entre as propostas. No CRAEF 1, essa afirmação é comprovada pelas palavras “doar”, “brinquedos”, “meninos”, referente a doação (tema do segundo quadrinho) e os vocábulos “doente” e “menina”, relacionados a primeira ilustração). Já no CRAEF 2 ainda apareceram os termos “visitar” e “hospital”, também ligados ao argumento da criança doente e os itens “menino” e “brinquedos” – apontados como característicos do segundo projeto.

Além da elaboração da lista de palavras-chave, oriundos da manipulação dos dados na ferramenta *KeyWords*, aproveitamos a oportunidade para exemplificar a análise por meio da comparação entre *types*, *tokens* e *type-token ratio* de duas propostas, uma de cada corpus.

Dessa forma, além da análise quantitativa, é possível oferecer um olhar qualitativo às investigações linguísticas, propiciando uma visão mais abrangente do léxico em foco.

Tabela 25 – *KeyWords* das Amostras do 5º Tema de redação

CRAEF 1 – CENTRO		CRAEF 2 – PERIFERIA	
ALUNO 03		ALUNO 04	
KEY WORDS	FREQ.	KEY WORDS	FREQ.
MENINO	9	LEOM	2
TINHA	11	FAMILIA	2
VIZINHO	7	BRINQUEDOS	3
EU	10	ABRIGO	2
FALADO	4	-	-
AI	3	-	-

Neste tema acerca da solidariedade, as palavras de maior chavicidade que se destacam

na amostra da escola do Centro, são “menino”, “eu” e “falado”; enquanto a amostra da Periferia traz o substantivo próprio “Leon”, uma criança de rua – personagem central da história redigida pelo aluno 04, “família” e “abrigo”, indicando uma diferença lexical significativa entre os textos, que será explorada mais adiante, acerca do *type-token ratio* (TTR).

Tem-se, na sequência, as amostras de onde retiramos as palavras-chave de cada corpus, a fim de contextualizar a análise do léxico, proposto a seguir:

Figura 30 – Amostra da Escola 1 (Aluno 03)

- Eu tenho vários brinquedos que eu não
brinco mais, eu vou perguntar pros meus pais
se eles deixam eu doar pra você.
O menino tinha falado pros pais dele
e os pais dele tinha deixado si ele doar
pra vizinho, e a vizinho tinha agrade-
cido.

Figura 31 – Amostra da Escola 2 (Aluno 04)

em um dia chueva um mendigo
estava sem abrigo e com fome enrolado no
papelão. ele foi de casa a casa pedindo comida
mais apenas uma deu água, comida, abrigo e
até Brinquedos para ele. a família pediu
a a juiz a guarda temporária da criança e
que se chamava Leon. eo juiz deu a

De acordo com o léxico apresentado, o aluno da Escola 1 (Centro) conta sua história a partir do ponto de vista do doador do brinquedo, assumindo ser este o personagem principal de sua redação. As orações comprovam esta escolha em: “Eu tenho vários brinquedos”; “eu vou perguntar (...) se eles deixam eu doar pra você”.

O doador atua como ator principal deste processo, conforme atesta a presença exacerbada do pronome pessoal do caso reto “Eu”. Além disso, o favorecido pela ação é citado apenas duas vezes – sempre na posição de coadjuvante, enquanto agente passivo.

Já, o discente da Escola 2 (Periferia), assume outra perspectiva. Ao contar a narrativa sob o ponto de vista de um mendigo, seu léxico descreve de que forma esta pessoa vive (“sem abrigo e com fome enrolado no papelão”).

No desenrolar da história, aparecem os termos próprios do Conselho Tutelar – “abrigo”; “juiz” e “guarda temporária da criança”. Tem-se a descrição de outra realidade

social, mostrando o cotidiano daquele que necessita receber atenção e ajuda da sociedade.

Após serem manuseados pelo programa *WordSmith Tool* (SCOTT, 2012), observamos também aspectos relativos à densidade lexical dos textos em questão, tomando como ponto de partida o número total de palavras (*tokens*) e também o número de palavras diferentes (*types*) empregadas pelos aprendizes. A partir da elaboração da *Wordlist* com as redações dos corpora de aprendizes, CRAEF 1 e CRAEF 2, elaboramos uma tabela que contempla os itens citados:

Tabela 26 – *Tokens, types e type-token ratio* das amostras do Tema 05

ESCOLA 1 – ALUNO 03		ESCOLA 2 – ALUNO 04	
Tokens	183	Tokens	79
Types	70	Types	55
Type-token ratio (TTR)	38	Type-token ratio (TTR)	69.62

Os resultados apontam várias particularidades nas escritas dos discentes em questão. Apesar de o Aluno 03 ter escrito mais palavras em sua redação (183), muitas foram repetidas – resultando em um baixo escore no item *type-token ratio* (38). Como visto na digitalização deste texto, a palavra “Eu” ocorre quatro vezes no mesmo parágrafo, reafirmando os resultados apresentados pelo *software*.

Já, o Aluno 04, redigiu menos da metade de vocábulos (79) da redação anterior analisada (183), contudo, o número de palavras não repetidas, ou seja, o número de *types* (55) foi comparativamente maior que o aluno 03 (70), se considerarmos o número total de palavras empregadas. Por esse motivo, o programa demonstra que a razão *type-token ratio* está muito mais elevada (69.62) no aluno 04.

A análise demonstrou que, crianças oriundas de realidades econômicas diferentes, possuem percepções divergentes sobre uma mesma situação social. Ao se colocar no lugar do doador de brinquedos (Escola 1 – Centro), o aluno 03 expressa a realidade e o lugar em que vive. No exemplo apresentado, esta prerrogativa é confirmada pela reiteração da palavra “eu” (quatro vezes em um mesmo parágrafo).

O inverso também foi mostrado na redação do discente da periferia (Aluno 04). A maneira como descreve a realidade social do “mendingo”, apontado posteriormente como uma criança de rua, chama a atenção por se tratar de um texto escrito por uma estudante de 07 anos. A capacidade que a discente possui, desde a infância, em se colocar empiricamente no lugar do outro para abordar esta triste realidade, foi um dos grandes destaques no CRAEF 2.

Portanto, utilizamos o suporte da LC e do programa *WordSmith Tools* para analisar as cinco propostas de produção textual oferecidas aos dois grupos, CRAEF 1 e CRAEF 2. Após

o exposto, concluímos que ocorreram muitas variações na quantidade de palavras dos textos coletados. Estas alternâncias ocorreram em decorrência de vários fatores, como faltas, transferências e recebimentos de alunos, além de crianças no estágio pré-silábico de escrita, situações verificadas em ambas as turmas, principalmente na escola da periferia.

Apesar desses imprevistos, foi possível constatar que, ao longo das coletas, houve um aumento substancial dos vocábulos redigidos pelas crianças, bem como o número de palavras distintas, escritas uma única vez nas composições. Este dado, entretanto, está atrelado ao número de textos coletados, pois, mesmo que um aluno produza vários enunciados, é preciso um montante significativo de redações para justificar um corpus com mais vocábulos.

Sobre o léxico, a idade dos participantes, de 7 a 9 anos, justificaria por si só a simplicidade do vocabulário observado e os desvios de escrita, visto que ainda estão cursando as séries iniciais. Porém, é preciso chamar a atenção para o engenho e a inventividade que alguns alunos demonstraram em suas produções, fruto do empenho e individualidade dos estudantes, comprovados nos dois corpora.

Assim, ser discente do centro ou da periferia não influenciou o desempenho lexical das crianças. Retomamos às afirmações anteriores sobre o local social em que os alunos se encontram e a variedade lexical. Independentemente do bairro e cidade em que moram, os traços de personalidade, empatia e expressão estão, em maior ou menor grau, presentes em todo ser humano, desde a infância.

Por isso, o professor deve oferecer temáticas diferenciadas aos alunos, para que estes se sintam confortáveis em escrever sobre assuntos diversos, ao mesmo tempo em que desenvolve um repertório lexical rico. Por fim, é desejável que o próprio educador tenha um olhar sem pré-conceitos e julgamentos, afinal, via de regra, as crianças mostram-se interessadas e dispostas a aprender os conteúdos ensinados.

A próxima etapa contempla os fraseologismos, finalizando as análises sobre o CRAEF 1 e CRAEF 2, ao apresentar a terceira parte deste capítulo avaliativo, que versa sobre as colocações encontradas nas produções textuais.

3.3 Análise fraseológica do CRAEF 1 e 2

O objetivo deste subitem é identificar as colocações mais comumente usadas pelos alunos a partir das redações que compõem o CRAEF 1 e CRAEF 2. Haja vista que

enfaticamente várias vezes neste trabalho a importância do desenvolvimento da competência lexical e colocacional nas séries iniciais, esta análise é coerente, uma vez que aponta empiricamente para a necessidade de o professor refletir sobre o tema, possibilitando que pensem em atividades e situações para o aperfeiçoamento de seus alunos.

Além disso, será com base nas colocações apresentadas que desenvolvemos uma das estratégias pensadas para aprimorar os conhecimentos dos estudantes frente a este conteúdo linguístico tão importante: a proposta de atividades que sucede este tópico.

Primeiramente, é preciso esclarecer que a classificação das colocações será feita segundo a classificação de Tagnin (2013), já apresentada na fundamentação teórica. Relembrando a classificação sugerida pelos autores, as colocações são subdivididas em verbais, nominais, adjetivas e adverbiais. Estas serão relacionadas às cinco propostas textuais, de forma análoga à análise lexical realizada anteriormente.

Além disso, uma informação pertinente diz respeito à variedade de tempos verbais identificada, alternando entre colocações verbais indicando passado, presente e futuro. Desse modo, para facilitar a visualização dos dados, optamos por apresentá-los, nesta análise, no modo infinitivo.

Outro esclarecimento digno de nota é sobre a frequência das colocações. Apesar de o CRAEF ser composto por 224 textos e um total de 20.435 palavras, o corpus tornou-se pequeno, devido à linguagem simples das crianças participantes, que possuem de sete a nove anos de idade. Neste sentido, o enfoque fraseológico não será pautado somente nas estatísticas, mas principalmente, no aspecto linguístico. Desse modo, cabe lembrar que todas as colocações apresentadas neste tópico foram redigidas pelas crianças, embora saibamos que não estejam adequadas a convencionalidade da língua portuguesa do Brasil. Por esta razão, no intuito de auxiliá-las na produção de colocações que sejam frequentes e que estejam de acordo com a convencionalidade da língua, incluímos o capítulo 4 nesta dissertação, conforme será explicado em detalhes nas próximas páginas.

Feitos os esclarecimentos de cunho metodológico, iniciamos a discussão pelas palavras-chave geradas, por meio das ferramentas *Keywords* e *Corcord*, do programa *WordSmith Tools*. Graças às referidas ferramentas, constatamos o emprego de cento e uma colocações no CRAEF 1 oitenta e oito no CRAEF 2.

Visto que cada redação possui uma temática diferente, iniciamos pela 1ª proposta, Tema 01, conforme segue.

Escrita sobre animais do zoológico (escola da periferia) – 02/2015

Neste bloco de produções textuais, as colocações registradas pelo CRAEF 1 discorreram sobre o relato do filme “*Os Croods*”, resultando em dezesseis colocações. Estas refletiram o vocabulário próprio do cotidiano dos “homens das cavernas”, além das comparações que as crianças efetuaram das diferenças entre este ambiente e o seu próprio. Assim, as palavras-chave que proporcionaram o maior número de colocações foram “banho”, “pedra”, “comida” e “dentes”. A seguir, apresentamos as linhas de concordância do item lexical “banho”, formando a colocação verbal “tomar banho”:

Figura 32 – Tela das concordâncias com a palavra “BANHO”

N	Concordance
1	a gema do ovo duas pessoas iam banho na chuva a luz ia se foca não
2	comida e quando nós quer toma banho nós tirava o pó e nós dormia
3	e para nós 1 tomar banho nós toma banho no rio pra nós 2 ficar seguro
4	e fim. Minha família 1 pra eles toma Banho Bate na Roupa 2 deles com o
5	eles caçam para comer, eles tomam banho do rio, eles moram numa
6	e não com uma porta e não tomamos banho porque não existe chuveiro e
7	Maria Clara Bom nós para tomar banho nós procura "água" ! Para
8	chegou a noite e é a hora de tomar banho a gente já tomou banho agora
9	tem que ser com pedra e para tomar banho e a caverna é pequena e não
10	os homem das caverna faz para tomar banho as vezes eles usam pedra para
11	pedra não tinha - chuveiro para tomar banho eu ia no - rio nessa época não
12	gente tinha 4 que caçar e para tomar banho que 5 minha mãe me
13	pra casa e todos foram tomar banho na chuva e ficou limpos e ficou
14	Não tem energia e para nós 1 tomar banho nós toma banho no rio pra nós
15	no lugar dos Croods eu poderia tomar Banho no Mar Como eu Poderia
16	das Cavernas Como Eu Poderia Tomar Banho Se eu Tivesse no lugar dos
17	e nós comemos o ovo e nós tomava banho assim minha mãe batia o pau
18	moravam em cavernas eles tomavam banho com pó nem existia iluminação
19	em cavernas para eles tomavam banho tirando poeira da roupa e os
20	hora de tomar banho a gente já tomou banho agora é hora de escovar os

Fonte: *PrintScreen* da tela do programa *WordSmith Tools*

Foram identificadas vinte ocorrências da palavra-chave “banho”, a partir da qual extraímos colocações. Neste sentido, a expressa maioria denota a colocação verbal “tomar banho”, cujo verbo foi conjugado em diversos tempos verbais; presente: “toma”, “tomam”, “tomamos” e “passado”: “tomou”, “tomava”. Além das formas compostas, como: “foram tomar banho”, “quer tomar banho” e “poderia tomar banho”.

Também houve uma ocorrência do termo junto ao verbo ir: “iam banho”. Este dado parece-nos mostrar o desconhecimento do aluno em como combinar este substantivo com o verbo adequado. Contudo, é preciso contextualizar o fato de que o conteúdo das formas

verbais só é trabalhado nas escolas municipais de Votuporanga a partir do 4^a ano, o que justifica a pouca variedade empregada pelos alunos.

Dessa maneira, a preocupação que temos demonstrado ao longo deste texto com a necessidade de se trabalhar os fraseologismos nas salas de aula desde o ensino fundamental I é válida, visto que também auxiliaria os docentes no ensino de outros conceitos e colocações da língua portuguesa do Brasil.

Na sequência temos as concordâncias geradas a partir de “pedra”, o que levou a formação da colocação adjetiva “idade da pedra”:

Figura 33 – Tela das concordâncias com a palavra “PEDRA”

N	Concordance
1	no coco e colocamos uma pedra e a 7 pedra bati e nós para ficar a marca 8
2	para fazer fogo nós pegamos uma 9 pedra e esfrega as duas pedra junta
3	e para tirar foto tem que ser com pedra e para tomar banho e a
4	essa é minha Família A vida da pedra Bom vamos começar eu me
5	. se minha família tivesse a idade da pedra eles no saberiam o que era um
6	eu e a minha família - na idade da pedra não tinha - televisão e nem
7	80 anos a minha família na idade da pedra não existia energia tv celular e
8	felizes e fim. minha família na idade da pedra minha família dorme no chão
9	eu caçava animais vivos - na idade da pedra não tinha - chuveiro para tomar
10	uma 9 pedra e esfrega as duas pedra junta e faz 10 fogo no escuro
11	inteiro e batiam a cara e o corpo na pedra e na pedra ficava a marca e
12	a cara e o corpo na pedra e na pedra ficava a marca e para aqueles
13	bati e nós para ficar a marca 8 na pedra para fazer fogo nós pegamos
14	e também do nó cego que trombou na pedra A família das Cavernas Como
15	foca não ia ter tu para tira foto é uma pedra que ia bater na gente ia mora
16	e nos fechamos a caverna com uma pedra e não com uma porta e não
17	cama e a gente dorme encima de uma pedra e quando temos uma vida nova
18	a terra e desenhavam em uma pedra. Como somos feliz Minha
19	6 areia no coco e colocamos uma pedra e a 7 pedra bati e nós para
20	em uma caverna e colocamos 3 uma pedra na frente da caverna para nós
21	e animais eles tiraram foto em uma pedra as roupas é tom de animais ele
22	para tomar banho as vezes eles usam pedra para tirar suas sujeiras e não

Fonte: PrintScreen da tela do programa WordSmith Tools

Com base nos dados acima, identificamos a presença da colocação adjetiva “idade da pedra”, com 5 ocorrências. Além disso, apareceram 17 vezes o uso material deste elemento para realizar diversas atividades, mas que não formam colocações: “eles usam pedra para tirar suas sujeiras”, “fechamos a caverna com uma pedra” e “desenhavam em uma pedra”, “a gente dorme encima de uma pedra”, por exemplo, denotando a utilização de uma ferramenta padrão para aquela época.

Ainda, relatamos uma combinação lexical não recorrente em língua portuguesa,

gerando uma colocação inadequada, conforme convenção linguística nesta língua: “a vida na pedra”. O emprego de uma combinação lexical inadequada pode atestar o que já foi exposto na página anterior sobre a preocupação que o professor deve ter acerca da competência colocacional dos estudantes.

A próxima palavra neste rol lexical é “comida”, que resultou em diversas combinatórias verbais, como “caçar comida”, por exemplo:

Figura 34 – Tela das concordâncias com a palavra “COMIDA”

N	Concordance
1	com fome e a família fomos caçar a comida fomos voltamos pra casa e
2	sai da caverna e lá fora é difícil acha comida e para comer tem que ser
3	e noite e quando saímos é para caçar comida chegou a noite e é a hora de
4	nós ficava com fome nós caçava comida e quando nós quer toma
5	não come o que a gente e não come comida eles comem uns bichos muito
6	eu tinha Que caçar Ate encontrar Comida Mas Que dó Meu Pai sempre

Fonte: PrintScreen da tela do programa WordSmith Tools

Ao todo, registramos 6 frequências com a palavra-chave comida. Há algumas variações encontradas neste contexto, como as colocações verbais “caçar comida”, “encontrar comida”, “comer comida”, próprios da linguagem utilizada para os homens primatas. Visto que as refeições são os animais, os mesmos precisam ser “caçados”, “encontrados” e “comidos” neste ambiente inóspito.

Para encerrar a lista, destacamos a palavra “dentes”, cujas colocações verbais trazemos na sequência:

Figura 35 – Tela das concordâncias com a palavra “DENTES”

N	Concordance
1	vamos um graveto e limpávamos os dentes e finalmente a hora de
2	banho agora é hora de escovar os dentes não tinha escova de dente

Fonte: PrintScreen da tela do programa WordSmith Tools

Os dois apontamentos com a forma dentes estão adequados neste contexto, gerando as

colocações verbais “limpar os dentes” e “escovar os dentes”. Estes fraseologismos já estão consagrados em nosso meio, devido ao uso destas combinações na linguagem diária dos brasileiros.

Novamente citamos a necessidade metodológica de realizar um recorte no relato das palavras que foram base ou colocados para as colocações existentes no CRAEF 1 e 2. Por isso, escolhemos somente algumas mais representativas para análise. Para finalizar, apresentamos todas as combinações que ocorreram na escola do centro, iniciando pela primeira proposta de redações.

A listagem completa das colocações apresentadas pelos estudantes encontra-se no quadro 03:

Quadro 03 – Colocações do CRAEF 1 (Tema 01)

Tipo de colocação	Estrutura	Exemplos
Colocações Verbais	V + S / V + Prep. + S	<i>Caçar comida, tomar banho, tirar foto, fazer retrato, contar história, fazer fogo, acender a luz, escovar os dentes, limpar os dentes e brincar de corrida.</i>
Colocações Nominais	S + Prep + S	<i>Pelo de animais.</i>
Colocações Adjetivas	S + Adj.	<i>Bichos de estimação, noite escura, jovem menina e idade da pedra.</i>
Colocação Adverbial	V + Adv.	<i>Correr até cansar.</i>

Conforme observado, houve uma predominância das colocações de cunho verbal, dez ao todo. Depois, temos as combinações nominais, com uma ocorrência no total. Por fim, também observam-se quatro colocações adjetivas e apenas uma no formato adverbial.

Dois fraseologismos chamam a atenção nesta lista: “tirar foto” e “fazer retrato”, ambos com 1 ocorrência cada no CRAEF 1. Para verificar a frequência destas colocações, recorreremos a ferramenta *Sketch Engine*²⁹ (KILGARRIFF et al, 2004), que apresentou a ocorrência de 3.245 resultados para “tirar foto” e somente 36 recorrências de “fazer retrato”, atestando para a convenção do uso da primeira combinação em detrimento da segunda.

Mais uma vez destacamos aqui a utilidade desta pesquisa, em oferecer respaldo teórico e prático aos professores de língua materna para que possam facilitar o entendimento das colocações às crianças e, conseqüentemente, ampliar a competência colocacional de seus

²⁹ Gerenciador de corpus e *software* que analisa uma grande quantidade de textos. Pelas suas especificidades, atende o interesse daqueles que lidam com a linguagem, como lexicógrafos, tradutores ou estudantes de línguas, além de pesquisadores de um modo geral.

estudantes.

Agora vejamos como o CRAEF 2 apresentou as colocações na primeira proposta de redações. Neste primeiro bloco de textos, observamos nove colocações, relacionadas, em sua maioria, às condições climáticas e a especificidades do mundo animal, apontadas nas produções dos alunos. Semelhantemente à escola do centro, observamos que os vocábulos que mais chamaram a atenção nesta primeira coleta foram “chuva”, “forte” e “pescoço”, relacionadas a seguir. Iniciemos com a palavra “chuva”:

Figura 36 – Tela das concordâncias com a palavra “CHUVA”

N	Concordance
1	na praia ele foi nada e começou a chuva e a mãe do menino disse -
2	TODAS AS ESPÉCIES. o menino e a chuva era uma vez um menino estava
3	nado na piscina na chuva (ilegível) cai chuva de Deus que ta gostoso manda
4	ZOOLÓGICO. eu nado na piscina na chuva (ilegível) cai chuva de Deus que
5	CRIANÇAS NO CARRO ENTÃO VEIO CHUVA FORTE E VEIO O

Fonte: PrintScreen da tela do programa *WordSmith Tools*

Os dados apresentam 5 ocorrências com “chuva”, destes, três se relacionam com verbos, como “começar a chuva”, “cair chuva” e “vir chuva”. Devido ao uso destas formas estarem cristalizadas pelos falantes do português, tais colocações verbais são representativas no corpus dos alunos da periferia.

Entre as linhas de concordâncias, chama a atenção a expressão “chuva de Deus”, pois Votuporanga é uma cidade bastante quente, com temperaturas próximas a 40° Celsius. Assim, o tempo abafado ocasiona chuvas, que ajudam a diminuir o calor proveniente desta climática tropical.

A próxima palavra é “forte”, que adjetiva tanto verbos quanto substantivos, conforme segue:

Figura 37 – Tela das concordâncias com a palavra “FORTE”

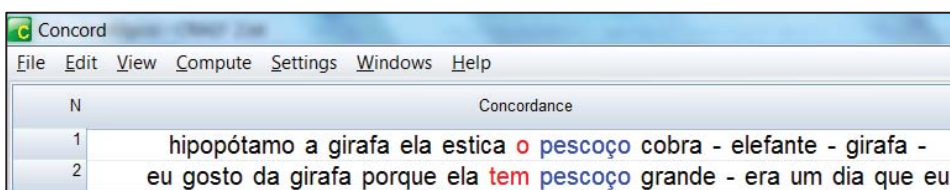
N	Concordance
1	NO CARRO ENTÃO VEIO CHUVA FORTE E VEIO O RELÂMPAGO FIM
2	gostei foi do leão porque ele é bravo e forte e tem unha grande. EU E
3	e começou a trovejar e relampear forte e um menino estava mexendo no

Fonte: PrintScreen da tela do programa *WordSmith Tools*

Dentro das três linhas apresentadas, 2 itens se referem a colocações, uma adjetiva “chuva forte” e a outra verbal “relampear forte”. Visto que o tema era animais, seria esperado que mais alunos destacassem as características dos mesmos. Esse caso ocorreu somente uma vez, mas da forma como foi redigido não gerou colocação.

Já a palavra “pescoço” ocorreu na caracterização da girafa, descrita na sequência:

Figura 38 – Tela das concordâncias com a palavra “PESCOÇO”



Fonte: PrintScreen da tela do programa *WordSmith Tools*

Oriundas desta palavra temos uma colocação verbal “esticar o pescoço” e outra adjetiva “pescoço grande”, termos representativos da linguagem natural das crianças em descrever uma das características mais marcante da girafa.

Na sequência, paralelamente a escola do centro, segue o escopo de colocações encontradas no CRAEF 2 desta proposta:

Quadro 04 – Colocações do CRAEF 2 (Tema 01)

Tipo de colocação	Estrutura	Exemplos
Colocações Verbais	V + S / V + Prep. + S	<i>Cair chuva, começar chuva, vir chuva, mexer no celular, cair um raio e esticar o pescoço.</i>
Colocações Adjetivas	S + Adj.	<i>Pescoço grande, zebra listrada e chuva forte.</i>
Colocação Adverbial	V + Adv.	<i>Relampear forte.</i>

Semelhantemente ao CRAEF 1, no texto das crianças da periferia também predominou o uso de colocações verbais, seis ao todo. Ocorreram ainda, três colocações adjetivas, usadas para indicar atributos ou características aos seres e uma adverbial. Contudo, nesta proposta não obtivemos combinatórias nenhuma combinatória nominal.

Um adendo interessante é a respeito da colocação “zebra listrada”. Apesar de aparecer somente 1 vez no CRAEF 2, as pesquisas via *Web* apontaram 57.000 recorrências, a maioria delas mais comumente usada em histórias infantis ou se referindo a pintura, adesivo ou outro

ornamento nesta configuração.

Uma outra contribuição desta pesquisa, portanto, é indicar que o professor promova discussões em sala de aula, para que a criança, desde cedo, consiga identificar padrões e atentar para o uso destes em sua comunicação, de acordo com o contexto da enunciação.

Neste primeiro rol de textos, observamos que a escola do centro produziu mais colocações, tanto em quantidade quanto em uma maior variedade de combinatórias do que a periferia. Como já relatado anteriormente, não foi possível oferecer o mesmo tema para ambas as escolas, o que de uma certa maneira, prejudicou a comparação das colocações entre as turmas.

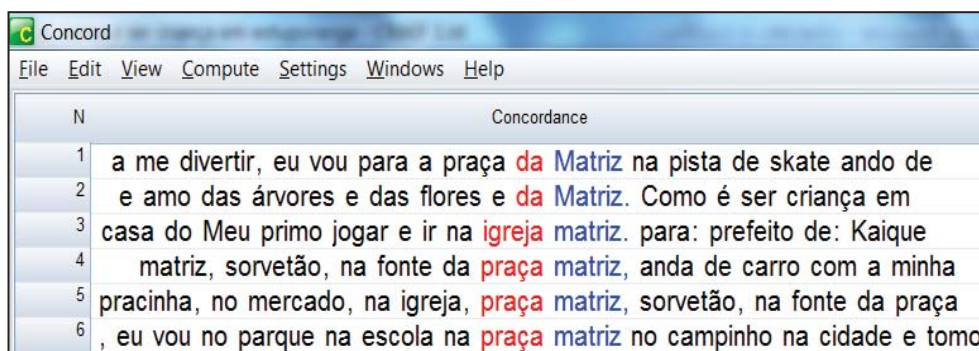
Feitas as considerações sobre o Tema 01, passemos ao segundo grupo de redações, a respeito da cidade de Votuporanga.

Tema 02 – “Como é ser criança em Votuporanga” – Concurso Cultural promovido pela Secretaria da Educação: 05/2015

Prosseguindo a análise, as colocações elencadas pelo CRAEF 1 e CRAEF 2 na segunda proposta tiveram como argumento a cidade de Votuporanga. Iniciando as reflexões pela escola do centro, compilamos os textos no programa *WordSmith Tools*. Após esta etapa, obtivemos ao todo vinte colocações, que apontaram, de uma forma geral, alguns locais pertencentes a cidade, as impressões que os discentes possuem sobre a mesma, além dos tipos de brincadeiras e atividades que realizam neste local.

Dessa forma, as palavras-chave escolhidas para discussão foram “matriz”, “Deus”, “greve” e “jogar”. Para começar, apresentamos as linhas de concordância de “matriz”, responsável pela formação das colocações nominais “igreja matriz” e “praça matriz”:

Figura 39 – Tela das concordâncias com a palavra “MATRIZ”



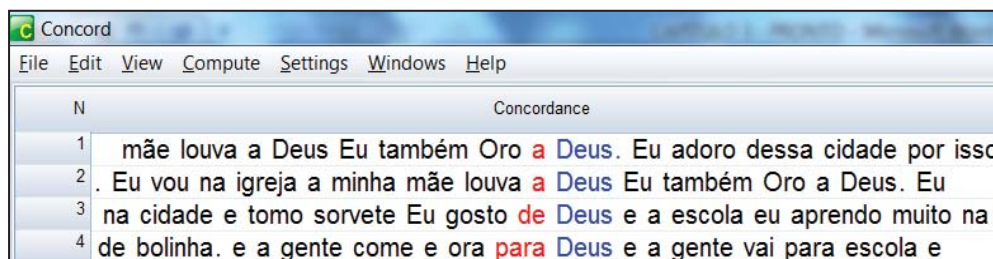
N	Concordance
1	a me divertir, eu vou para a praça da Matriz na pista de skate ando de
2	e amo das árvores e das flores e da Matriz . Como é ser criança em
3	casa do Meu primo jogar e ir na igreja matriz . para: prefeito de: Kaique
4	matriz, sorvetão, na fonte da praça matriz , anda de carro com a minha
5	pracinha, no mercado, na igreja, praça matriz , sorvetão, na fonte da praça
6	, eu vou no parque na escola na praça matriz no campinho na cidade e tomo

Fonte: *PrintScreen* da tela do programa *WordSmith Tools*

O *software* retornou 6 reiteraões, dentre as quais quatro colocações nominais, 1 com a combinação “igreja matriz” e três com a formação “praça matriz”. Haja vista que há outras praças na cidade, também tivemos a ocorrência de “praça da Matriz” e somente do substantivo “Matriz”. Devido a localização privilegiada, tanto a praça, quanto a igreja, se localizam exatamente no marco zero da cidade, sendo um local bastante conhecido dos estudantes, que moram nos arredores deste bairro. A escola do centro, por exemplo, se situa a leste, distante a alguns quarteirões da praça citada.

Na sequência, temos os retornos a palavra-chave “Deus”, item lexical também abordado pelos alunos:

Figura 40 – Tela das concordâncias com a palavra “DEUS”

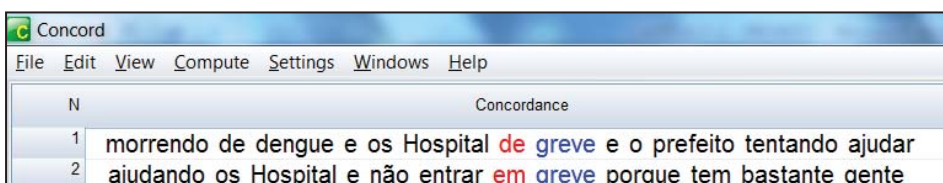


Fonte: *PrintScreen* da tela do programa *WordSmith Tools*

A partir deste vocábulo, obtivemos 4 colocações, todas verbais acrescidas de preposição, “a”, que ocorreu duas vezes, “de” e “para”, formando as combinações “oro a Deus”, “louva a Deus”, “gosto de Deus” e “ora para Deus”. Apesar de não estar relacionada diretamente ao tema da cidade de Votuporanga, a palavra-chave “Deus” se destaca neste contexto por abordar, mesmo que indiretamente, uma atividade realizada em conjunto com a família, geralmente na igreja, palavra mencionada 11 vezes no CRAEF 1.

A seguir temos a abordagem do item “greve”, com o objetivo de mostrar uma crítica social:

Figura 41 – Tela das concordâncias com a palavra “GREVE”



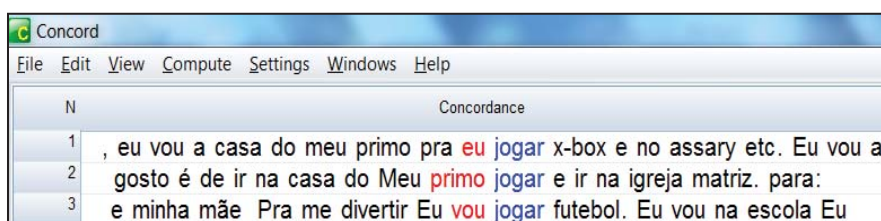
Fonte: *PrintScreen* da tela do programa *WordSmith Tools*

Com somente 2 menções no CRAEF 1, esta palavra gerou duas colocações, uma adjetiva “hospital de greve” e outra verbal “entrar em greve”. Aqui tivemos uma abordagem atípica para o linguajar infantil, embora seja extremamente pertinente dentro do tema cidade.

Ao citar as “pessoas morrendo de dengue” e “gente doente”, as duas redações chamam a atenção para um problema de saúde pública, tão presente e necessário a ser discutido na sala de aula. Reconhecemos a simplicidade da abordagem nos textos dessas crianças, cuja crítica não foi aprofundada nas redações, mas deixamos o registro da sensibilidade de alunos em reconhecer a falta de suporte para um setor de atendimento tão importante para o bem estar da população.

Por último, trouxemos um dos temas lúdicos elencados pelos alunos da escola do centro, “jogar”:

Figura 42 – Tela das concordâncias com a palavra “JOGAR”



N	Concordance
1	, eu vou a casa do meu primo pra eu jogar x-box e no assary etc. Eu vou a
2	gosto é de ir na casa do Meu primo jogar e ir na igreja matriz. para:
3	e minha mãe Pra me divertir Eu vou jogar futebol. Eu vou na escola Eu

Fonte: *PrintScreen* da tela do programa *WordSmith Tools*

Embora só com 3 linhas de concordância, observamos que os discentes do CRAEF 1 destacaram mais os locais de recreação do que a atividade realizada em si, já explanado anteriormente no item 3.2 da presente dissertação. Portanto, há poucas referências as brincadeiras, presentes aqui nas duas colocações verbais “jogar x-box” e “jogar futebol”.

Além dessas, também ocorreram três formações verbais de andar + de, resultando em “andar de bicicleta”, “ando de patinete” e “andar de carro com o meu pai”, na referência de atividades diárias que realizam com frequência. A listagem completa das colocações apresentadas pelos estudantes encontra-se no quadro 05:

Quadro 05 – Colocações do CRAEF 1 (Tema 02)

Tipo de colocação	Estrutura	Exemplos
Colocações Verbais	V + S / V + Prep. + S	<i>Tomar banho, assistir filme, jogar x-box, tomar sorvete, louvar a Deus, orar a Deus, andar de bicicleta, fazer piquenique, andar de patinete, arrumar o cabelo, andar de carro e entrar em greve.</i>
Colocações Nominais	S + S / S + Prep. + S	<i>Igreja matriz, parque de diversões, futebol no campinho, dia de Natal, pista de skate e fonte da praça.</i>
Colocações Adjetivas	S + Adj.	<i>Câmara municipal e hospital de greve.</i>

Conforme observado, houve uma predominância das colocações de cunho verbal, doze ao todo, seguida de sete combinações nominais e duas adjetivas. Neste tema 02 não foram encontradas nenhuma colocação adverbial no CRAEF 1, o que atesta para a dificuldade que as crianças têm em caracterizar com detalhes os locais da cidade ou as atividades que executam neste espaço.

Já na escola da periferia, a análise das colocações chegou ao total de nove fraseologismos. As combinatórias, ao contrário do primeiro grupo, não focaram em lugares específicos de Votuporanga, ou estes foram citados em sintagmas simples, “pracinha”, com 5 ocorrências ou “campinho”, 2 citações, de maneira que não formam colocações.

O assunto mais elencado pelos estudantes foram as atividades lúdicas que realizam no ambiente, caracterizando uma abordagem mais lúdica da cidade. Dessa maneira, as palavras-chave escolhidas para discussão foram “pipa”, “andar” e “coração”.

Introduzindo as discussões, temos as linhas de concordância do item lexical “pipa”, responsável pela composição de colocações verbais deste universo infantil:

Figura 43 – Tela das concordâncias com a palavra “PIPA”

N	Concordance
1	ME DIVERTIR, EU VOU EMPINAR PIPA EU GOSTO DA CIDADE DE
2	cidade de Votuporanga, porque erguer pipa COMO É SER CRIANÇA EM
3	pra pracinha eu vou pra feira ou solto pipa ou jogo futebol. eu gosto da

Fonte: PrintScreen da tela do programa WordSmith Tools

Embora possua somente três linhas de concordância, o destaque deste item lexical foram as possibilidades empregadas pelos alunos na estruturação das colocações verbais: “empinar pipa”, “erguer pipa” e “soltar pipa”. Tal brincadeira é bastante recorrente das crianças da periferia, seja por ser um brinquedo de fácil execução, as próprias crianças podem construir uma pipa em suas casas, seja pela periferia dispor de um trânsito menos movimentado do que o centro da cidade, e, portanto, oferecer condições para que brinquem mais livremente.

Estas colocações verbais serão uma das combinações que farão parte da lista de tarefas baseada em corpora a ser explorada no próximo capítulo, com base no tipo de colocado a ser utilizado para formar colocações. Devido a esse fator, a discussão desta ocorrência verbal será tratada em detalhes em páginas posteriores. Segue a análise com o item lexical “andar”:

Figura 44 – Tela das concordâncias com a palavra “ANDAR”

Concordance	
N	
1	porque dá para brincar e dá para andar de bicicleta em Votuporanga é
2	de Votuporanga porque aqui eu posso andar de bicicleta e também de
3	cidade de Votuporanga porque posso andar de bicicleta É muito legal
4	, IRMÃ PRA ME DIVERTIR, EU VOU: ANDAR DE BICICLETA EU GOSTO
Concordance	
N	
1	amigo e é muito legal porque eu fico andando de bicicleta com meu amigo
2	DIVIRTO MESMO ASSIM SOZINHO ANDANDO DE BICICLETA OU
Concordance	
N	
1	EU VOU NA CASA DO MEU AMIGO, ANDO DE BICICLETA ETC... EU
2	Votuporanga porque eu brinco na rua e ando de bicicleta foi muito legal

Fonte: PrintScreen das telas do programa *WordSmith Tools*

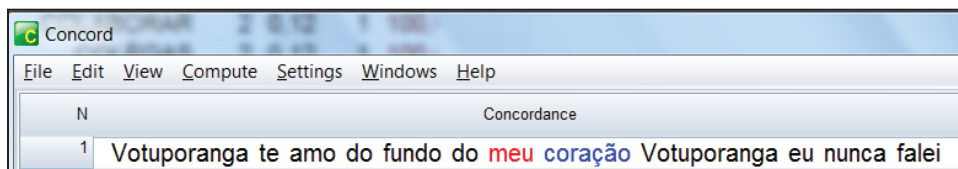
Devido à existência de várias conjugações do verbo “andar”, resolvemos adicionar às formas infinitivas, aquelas apresentadas no gerúndio e também as de primeira pessoa, a fim de justificar as 8 ocorrências existentes desta colocação verbal. Chama a atenção o fato de todas tratarem do ato de “andar de bicicleta”, algo muito comum na periferia da cidade, não só entre as crianças, mas também para os adultos que residem nestes locais.

Neste sentido, temos mais uma visão particularizada deste ambiente, em que os alunos descrevem o que acontece de mais significativo em seu cotidiano, não só como forma de divertimento, mas também como meio de transporte que podem utilizar longe da área central da cidade. Estas informações são comprovadas nas linhas de concordância: “aqui eu posso

andar de bicicleta” e “é muito legal porque eu fico andando de bicicleta com o meu amigo”.

Segue a última palavra a ser analisada no Tema 02, “coração”:

Figura 45 – Tela da concordância com a palavra “CORAÇÃO”



Fonte: PrintScreen da tela do programa *WordSmith Tools*

Justificamos a presença deste item lexical devido a ocorrência da única colocação adverbial do tema sobre Votuporanga, “amar do fundo do meu coração”. Apesar de não termos a presença do advérbio que marcaria tal combinatória, destacamos o uso da expressão por ser uma utilização criativa da língua, aqui empregada para destacar o amor do discente pela cidade.

Já a lista das colocações empregadas neste bloco de composições textuais está na sequência:

Quadro 06 – Colocações do CRAEF 2 (Tema 02)

Tipo de colocação	Estrutura	Exemplos
Colocações Verbais	V + S / V + Prep + S	<i>Soltar pipa, erguer pipa, empinar pipa, passar carro, trazer felicidade, jogar videogame, andar de bicicleta e brincar de corda.</i>
Colocação Adverbial	V + Adv. / Adv + Adj	<i>Amar do fundo do meu coração.</i>

Novamente, temos a presença quase absoluta de colocações verbais, com a presença ou não de preposição, resultando em oito combinatórias, somente um fraseologismo adverbial, mas nenhuma nominal ou adjetiva.

Uma ocorrência verbal digna de nota é “trazer felicidade”, com uma frequência. Esta também aponta para o teor de inventividade dos alunos já apontado anteriormente e que coopera para a criação de enunciados imaginativos e originais, propiciado pela manipulação da língua portuguesa do Brasil.

Em comparação com as redações do CRAEF 1, o grupo escolar do centro redigiu mais do que o dobro de colocações do que a sala de aula da periferia, embora ainda utilizasse duas formas de combinatória sintática (verbais e nominais). Estas trataram de locais presentes nos

espaços urbanos, como “igreja matriz”, “praça matriz” e também elencou rapidamente sobre um problema social existente na cidade, como a saúde pública “hospital em greve”; reverberando para o entendimento de situações que são passíveis de ocorrer neste ambiente.

A periferia, por sua vez, trouxe para a discussão itens lexicais próprios da infância, como as brincadeiras “empinar pipa” e “andar de bicicleta”, por exemplo, típicos da linguagem de criança. Estas também se configuram como algo representativo do seu local de origem, resultando, dessa forma, em maneiras distintas de descrever o município a partir do seu ponto de vista.

Em seguida, prosseguimos a análise, agora na esfera de um tema que visava atingir o imaginário infantil.

Tema 03 – Proposta da semente mágica: 06/2015

Na terceira coleta de textos, os dois grupos tiveram o desafio de discorrer sobre o que brotaria de uma semente mágica. Neste sentido, o tema propiciou um vocabulário voltado principalmente as ações sobre o plantio, além de vários nomes de frutas, verduras e legumes.

Os alunos do CRAEF 1, por exemplo, redigiram vinte e cinco colocações diversas, o maior número de combinatórias de todas as propostas de redações. Baseadas neste número apresentamos os itens lexicais “pé”, “rápido”, “bem” como palavras-chave a serem explanadas. A seguir, apresentamos as linhas de concordância do item lexical “pé”, que culminou na apresentação de grande número de colocações nominais deste tópico:

Figura 46 – Tela das concordâncias com a palavra “PÉ”

N	Concordance
1	para a mãe: - Mãe porque meu pé de feijão gigante em? disse a mãe
2	ele ficou muito grande e ele escalou o pé de feijão. O feijão de banana Era
3	o diamante e todos ficaram feliz. o pé de feijão mágico um dia Fábio foi
4	tarde Dora lia as histórias do João e o pé de feijão. e um dia ela foi na casa
5	mágico. Até que depois nasceu o pé de feijão mágico e o fazendeiro
6	embora para casa e dormir. Fim. O pé de feijão mágico Era uma vez um
7	semente e ele ficou muito feliz com o pé de maçã. Mas ele comeu todas
8	cresceu agora vou comer. Fim João o pé de Feijão Era uma vez um menino
9	feijão. durante de alguns dias o seu pé de feijão. ele ficou muito grande e
10	mágica A semente mágica era um pé de morango ela ia nascer em um
11	mágico um dia Fábio foi plantar um pé de feijão e o Fábio viu um mágico

Fonte: *PrintScreen* da tela do programa *WordSmith Tools*

Conforme afirmado anteriormente, o item lexical pé resultou em 11 colocações nominais, relacionadas ao tema 03: “pé de feijão”, com nove ocorrências, pé de morango, com uma citação e pé de maçã, também com uma referência.

Conforme dito anteriormente, pelo fato do tema pressupor a plantação de uma semente, várias crianças do CRAEF 1 se basearam na narrativa do João e o pé de feijão para redigirem o texto, e este dado foi comprovado nesta análise. A seguir, temos a palavra-chave “rápido”, cuja presença do adjetivo propiciou a formação de colocações verbais e adverbiais:

Figura 47 – Tela das concordâncias com a palavra “RÁPIDO”

N	Concordance
1	for para a escola para ela nascer bem rápido que nem o da escola ai eu vou
2	hora que ele dormiu a planta cresceu rápido e na hora que o garoto
3	acordou ele foi ver a planta cresceu rápido e o moleque falou - que
4	faço na minha casa para crescer mais rápido e eu vou fazer - tudo que deve
5	e ela salvou seus amigos e fugiu muito rápido e tudo ficou feliz mais a história
6	as minhas plantas cresceu muito rápido... - Eu não acredito das
7	das ruas ETC. Fim O feijão que nasce rápido Eu vou plantar a semente em

Fonte: *PrintScreen* da tela do programa *WordSmith Tools*

Dentre as sete linhas de concordâncias geradas, temos três colocações verbais, “crescer rápido”, que ocorre duas vezes e “nascer rápido”, com uma ocorrência. Já a combinatória “nascer bem rápido”, que aparece uma vez, forma a primeira colocação adverbial desta proposta, seguida de “crescer mais rápido”, com uma reiteração e “muito rápido”, presente duas vezes no CRAEF 1 ao lado dos verbos “fugiu” e “cresceu”, resultando em quatro colocações adverbiais.

A próxima palavra-chave é “bem”, que registrou sete colocações:

Figura 48 – Tela das concordâncias com a palavra “BEM”

The screenshot shows a window titled 'Concord' with a menu bar (File, Edit, View, Compute, Settings, Windows, Help) and a table of concordance results. The table has two columns: 'N' (line number) and 'Concordance' (text snippet). The results are as follows:

N	Concordance
1	. Amanhã uma de alface para cuidar bem da minha família então irei
2	semente e mágica e a semente ficou bem grande e sabia ela jogou bola de
3	e mesmo assim eu cuidei muito bem dela e eu e a semente mágica
4	Isabela ficou muito feliz e cuidou muito bem das fadas. A semente mágica
5	eu for para a escola para ela nascer bem rápido que nem o da escola aí
6	A maçã mágica Era uma vez semente bem brilhante que ficava num canteiro
7	Fábio não fala com estranho tá bem você não quer acreditar porque
8	de morango ela ia nascer em um vaso bem bonito e terra mágica e ia ficar

Fonte: *PrintScreen* da tela do programa *WordSmith Tools*

Das oito linhas geradas por este item lexical, sete são colocações adverbiais, “cuidar bem”, com três frequências, “ficar bem grande”, com uma frequência e “nascer bem rápido”, que já foi abordado há pouco, com uma referência. As outras duas também são combinações adjetivas “semente bem brilhante” e “vaso bem bonito”, visto que intensificam o sentido exposto pelos adjetivos. Já a linha 7, “tá bem”, não formou colocação.

Finalizando este tópico, o rol geral das colocações do CRAEF 1:

Quadro 07 – Colocações do CRAEF 1 (Tema 03)

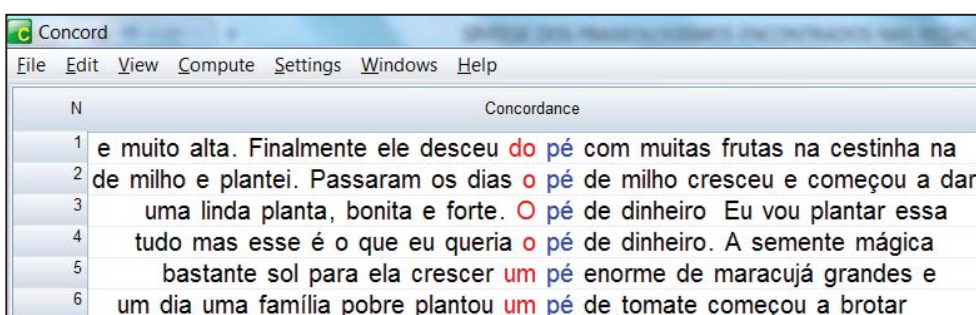
Tipo de colocação	Estrutura	Exemplos
Colocações Verbais	V + S / V + Prep + S	<i>Tomar cuidado, gastar dinheiro, passar fome, aguardar a florzinha, escalar o pé de feijão, regar as plantas, gritar de felicidade, cair na ideia dele e dar um jeito.</i>
Colocações Nominais	S + Prep + S	<i>Pé de feijão, pé de maçã, pé de morango, bola de futebol, bola de handebol, semente de feijão e banho de sol.</i>
Colocações Adjetivas	S + Adj.	<i>Sacola plástica, feijão da sorte, semente bem brilhante e vaso bem bonito.</i>
Colocação Adverbial	V + Adv. / Adv + Adj	<i>Cuidar bem, ficar bem grande, cuidar muito bem, morrer de fome e bicho em apuros.</i>

O Tema 03 gerou nove colocações verbais, sete de ordem nominais, quatro combinações adjetivas e cinco adverbiais. Novamente atribuímos a quantidade e variedade de colocações ao tópico proposto, visto que os alunos se motivaram a descrever com mais ênfase os processos de regar e o crescimento natural da planta, além do tipo de árvore que foi

plantada e como ocorreu este processo.

Após tais constatações, prossigamos com os fraseologismos descritos pelo CRAEF 2. Semelhantemente ao grupo de alunos do centro, as redações do Tema 03 foram as que mais apresentaram colocações, resultando em trinta e duas combinações ao todo. A título de análises, estudamos as palavras-chave “pé”, “milho” e “muito”. Na sequência, temos as linhas de concordância do item lexical “pé”, a fim de traçarmos um paralelo com o que foi efetuado pela escola do centro:

Figura 49 – Tela das concordâncias com a palavra “PÉ”



N	Concordance
1	e muito alta. Finalmente ele desceu do pé com muitas frutas na cestinha na
2	de milho e plantei. Passaram os dias o pé de milho cresceu e começou a dar
3	uma linda planta, bonita e forte. O pé de dinheiro Eu vou plantar essa
4	tudo mas esse é o que eu queria o pé de dinheiro. A semente mágica
5	bastante sol para ela crescer um pé enorme de maracujá grandes e
6	um dia uma família pobre plantou um pé de tomate começou a brotar

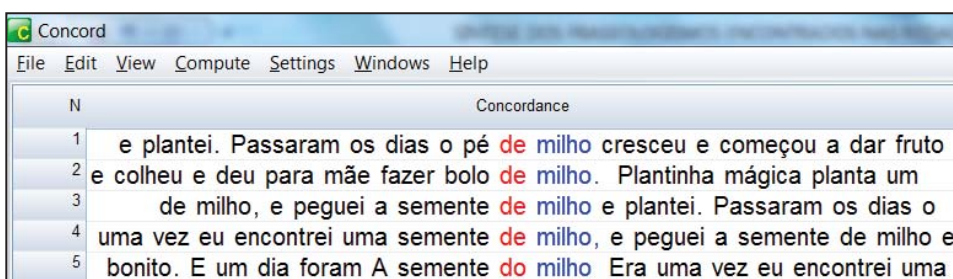
Fonte: PrintScreen da tela do programa *WordSmith Tools*

A análise mostrou 6 linhas de concordância, ao invés das 11 apontadas pelo CRAEF 1. No entanto, não houve menções à história de João o pé de feijão pelas crianças da periferia, o que resultou em combinações diferentes. Temos quatro colocações nominais, “pé de dinheiro”, com duas ocorrências, “pé de milho” com uma frequência e “pé de tomate”, que também apareceu somente uma vez. Ainda na linha 1 ocorreu uma combinação verbal “descer do pé” e a linha 5 traz uma colocação adjetiva, “pé enorme de maracujá”.

Assim, apesar de terem produzido menos colocações com a palavra-chave “pé”, as ocorrências descritas no CRAEF 2 foram mais inventivas e originais do que as do centro.

Em seguida, temos as linhas geradas pelo item lexical “milho”:

Figura 50 – Tela das concordâncias com a palavra “MILHO”



N	Concordance
1	e plantei. Passaram os dias o pé de milho cresceu e começou a dar fruto
2	e colheu e deu para mãe fazer bolo de milho. Plantinha mágica planta um
3	de milho, e peguei a semente de milho e plantei. Passaram os dias o
4	uma vez eu encontrei uma semente de milho, e peguei a semente de milho e
5	bonito. E um dia foram A semente do milho Era uma vez eu encontrei uma

Fonte: PrintScreen da tela do programa *WordSmith Tools*

As cinco linhas de concordância geradas se referem a colocações nominais: “semente de milho”, com duas ocorrências, “semente do milho”, apareceu somente uma vez no corpus, “pé de milho”, ocorreu uma vez e ainda “bolo de milho”, que também possui somente uma citação.

A fim de verificar qual dos dois formatos é mais recorrente, pesquisamos no *Sketch Engine* as frequências de “semente de milho” e “semente do milho”: enquanto a primeira obteve 1.657 resultados, a segunda teve 59 retornos, atestando para o uso consagrado da combinatória “semente de milho”.

De acordo com o exposto, retomamos a importância de tratar deste tema em sala de aula, para proporcionar condições dos alunos se comunicarem melhor em língua materna. Finalizando esta etapa, passemos a abordar o uso da palavra-chave “muito”:

Figura 51 – Tela das concordâncias com a palavra “MUITO”

N	Concordance
1	ela foi crescendo muito rápido e brotou muito mais maracujá ficou gostosos e
2	. De repente ela foi crescendo muito rápido e brotou muito mais
3	então ele subiu na árvore grande e muito alta. Finalmente ele desceu do
4	a plantinha cresceu e a Maria ficou muito feliz e chamou os seus pais
5	para ver, os pais de Maria gostam muito da Ex sementinha. A semente
6	ela acordou tinha uma árvore grandona muito mas muito grande com tomates
7	Era uma vez uma semente mágica muito alegre, e tinha uma casa, perto
8	tinha uma árvore grandona muito mas muito grande com tomates de ouro e
9	sementinha na minha casa e vai nascer muito dinheiro. E eu vou fazer
10	sementinha Era uma vez uma semente muito bonitinha. Até que um belo dia
11	de tomate começou a brotar tomates muito grandes eles comeu alguns e

Fonte: *PrintScreen* da tela do programa *WordSmith Tools*

O item lexical “muito” obteve onze linhas de concordância no CRAEF 2. Contudo, apresenta-se como colocação nas demais ocorrências, gerando, assim, seis combinações adverbiais da seguinte forma: três colocações oriundas da formação advérbio + verbo, como em: “gostam muito”, “crescendo muito rápido” e “nascer muito dinheiro”, todos com uma ocorrência cada um. Por fim, há três colocações resultantes de advérbio + adjetivo, verificado em: “muito alegre”, “semente muito bonitinha” e “tomates muito grandes”, também com a frequência de uma repetição por colocação mencionada.

A única combinatória que apareceu duas vezes foi “muito, mas muito grande”, que

retomou o termo em estudo novamente na mesma linha de concordância. Neste contexto, observamos a tendência que as crianças têm em redigir expressões próprias da oralidade em um modalidade escrita, já mencionado em páginas anteriores.

Finalizando esta etapa, temos o rol das colocações empregadas pelos discentes do CRAEF 2:

Quadro 08 – Colocações do CRAEF 2 (Tema 03)

Tipo de colocação	Estrutura	Exemplos
Colocações Verbais	V + S / V + Prep + S	<i>Fazer plásticas, realizar desejos, regar a semente, brotar tomates, passar o dia, aguardar a planta, andar de moto, dar água, crescer um pé e descer do pé.</i>
Colocações Nominais	S + Prep + S	<i>Sementinha de tomate, pé de tomate, plantação de cenouras, suco de maracujá, bolo de milho, dor nas costas, terra do jardim, pé de dinheiro, moradores de rua, e moeda de ouro.</i>
Colocações Adjetivas	S + Adj.	<i>Senhorinha velhinha, árvore grandona, dia lindo, boas ações, árvore grande e dia chuvoso.</i>
Colocação Adverbial	V + Adv. / Adv + Adj	<i>Nascer muito dinheiro, gostar muito, semente muito bonitinha, bem gostosa, muito grande e muito rápido.</i>

Neste tema, os discentes utilizaram as quatro variedades de colocações, sendo: dez verbais, dez nominais, seis adjetivas e por fim, seis adverbiais, que resultaram em trinta e duas colocações criativas para o assunto, além da caracterização em detalhes da planta cultivada.

Após observar o desempenho de ambas as salas de aula, verificamos que os alunos da escola da periferia redigiram sete colocações a mais do que o grupo do centro. Dessa forma, corroboramos o entendimento de que determinados temas requerem uma maior descrição, como é o caso desta proposta sobre o plantio da semente. Conforme os alunos escrevem suas composições, é possível notar o desenvolvimento do grão, o que possibilita uma utilização mais desenvolvida do léxico-fraseológico.

Por isso, reafirmamos a importância do docente em variar os temas sugeridos, para que o aluno possa utilizar o seu repertório lexical com maior propriedade. A próxima análise versa sobre o Tema 04, acerca de um personagem folclórico brasileiro.

Tema 04 – Escrita de uma narrativa folclórica: 08/2015

A quarta proposta textual, acerca da lenda da mula sem cabeça, apresentou menos colocações do que o tema anterior, em ambas as escolas. Os alunos do CRAEF 1 redigiram vinte colocações, nesta etapa. De uma forma geral, muitas empregaram termos que remetem a noite, como “ficar dia e noite sem dormir” e “morrer de sono”, ou ainda relacionados ao universo rural: “animais venenosos”, “levar picada” e “subir na árvore”, bem como descreveram os eventos fantásticos ocorridos na área rural, “chorar fogo” e “mula falante”.

Da mesma forma como realizamos as análises anteriores, seguimos com as palavras-chave do Tema 04, que são “cavalo” e “fogo”:

Figura 52 – Tela das concordâncias com a palavra “CAVALO”

N	Concordance
1	ficou a noite ela ouviu uma pata de cavalo e ela ficou com medo e
2	, um escutei um barulho de pata de cavalo mais você não falou que não
3	quarteirão ai ele ouviu um barulho de cavalo e pensou que era um ladrão
4	lampião mais ela já tinha um namorado cavalo de troia bonito chegou e pisou
5	na mão dele ai, apareceu um cavalo com fogo em vez da cabeça.

Fonte: *PrintScreen* da tela do programa *WordSmith Tools*

O programa computacional retornou cinco resultados para este item lexical, que gerou 4 colocações nominais, a saber: “pata de cavalo”, com duas ocorrências, “barulho de cavalo”, recorrente uma vez e “cavalo de troia”, uma vez no CRAEF 1. Acreditamos que as combinatórias “barulho de cavalo” e “pata de cavalo” referem-se ao som do andar deste animal, parecido com o cavalgar característico do cavalo.

Escolhemos esta palavra-chave, por verificarmos a associação feita pelos discentes entre os dois animais, mula e cavalo. Talvez pelo primeiro não ser um mamífero muito conhecido dos alunos, estes decidiram atribuir ao ser folclórico atributos do equino, um animal que eles já estão mais acostumados a descrever.

A próxima palavra é “fogo”, responsável por formar colocações verbais e nominais desta proposta número 04:

Figura 53 – Tela das concordâncias com a palavra “FOGO”

N	Concordance
1	mais perto e o chico falou: - aquele fogo anda? ou é impressão minha. Foi
2	mula brigou com o namorado e chorou fogo. Mas Chico comprou outro
3	mão dele ai, apareceu um cavalo com fogo em vez da cabeça. eles dois,
4	era um bandido com uma flecha de fogo então era o que não era tudo
5	a mula sem cabeça sentiu o cheiro de fogo e se apaixonou. os apaixonados
6	que era bandido com uma flecha de fogo mas não era. ai a mãe dele
7	dela ele ficou tão assustado com esse fogo que ele não ficou preguiçoso que
8	apareceu e falou: - Eu quero esse fogo! - O gente toma o fofo fogo! Ela
9	esse fogo! - O gente toma o fofo fogo! Ela falou: - Cadê o bebezinho
10	na mão caiu no chão e a mula veio o Fogo e Ficou apaixonada e o Chico
11	Estava muito escuro e viu um fogo e o fogo estava chegando perto mais
12	perto dela e o fogo foi mantendo o fogo da cabeça dela ele ficou tão
13	e ela estava chegando perto dela e o fogo foi mantendo o fogo da cabeça
14	ficou a apaixonado pelo lampião e pelo fogo lindo que ela tinha igual o dela e
15	pelo lampião porque o lampião tinha fogo igual a mula sem cabeça E a
16	lâmparina e dentro da lâmparina tinha fogo e a mula sem Cabeça não
17	chegou mais Perto viu que tinha um fogo que nem ela e ficou apaixonado
18	Estava muito escuro e viu um fogo e o fogo estava chegando perto

Fonte: PrintScreen da tela do programa *WordSmith Tools*

Apesar deste item lexical gerar 18 linhas de concordâncias, a maioria delas não produziram colocações, como atestam as linhas 1, 3, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17 e 18. As demais que possuem o substantivo “fogo” como protagonista formaram cinco combinatórias, sendo: uma colocação verbal “chorou fogo”, com uma ocorrência, três colocações nominais “flecha de fogo”, duas vezes e “cheiro de fogo”, com uma frequência, além de uma adjetiva “fogo lindo”.

As colocações “chorou fogo” e “fogo lindo” não são recorrentes na língua portuguesa do Brasil³⁰. Porém, dentro do contexto da história de uma entidade fantástica que possui chamas flamejantes ao invés de uma cabeça, tal brecha abre espaço para a criação de combinatórias que descrevam esta narrativa.

Na página subsequente, temos o quadro com as colocações elencadas pelos estudantes do centro da cidade:

³⁰ Após pesquisas no *Sketch Engine*, observamos que a combinatória “chorar fogo” obteve 442 resultados, enquanto “fogo lindo” teve 24.900 recorrências.

Quadro 09 – Colocações do CRAEF 1 (Tema 04)

Tipo de colocação	Estrutura	Exemplos
Colocações Verbais	V + S / V + Prep + S	<i>Chorar fogo, levar picada, subir na árvore, dar um cascudo, sentir o cheiro, apagar o lampião e dar de cara.</i>
Colocações Nominais	S + Prep + S	<i>Mula sem cabeça, cavalo de troia, flecha de fogo, tanque de guerra, bocado de armas, dor de garganta e pata de cavalo.</i>
Colocações Adjetivas	S + Adj.	<i>Mula falante e animais venenosos.</i>
Colocação Adverbial	V + Adv.	<i>Sair discretamente, aparecer do nada, ficar dia e noite sem dormir e morrer de sono.</i>

Houve a predominância de colocações verbais nos textos sobre a mula sem cabeça, sete ao todo; além de sete da ordem nominal; duas combinatórias adjetivas e outras quatro adverbiais. Apesar de terem representado todas as variedades de colocações, esse bloco de produções retornou menos fraseologismos do que o Tema 03. Uma das possibilidades levantadas para esse resultado diz respeito às ilustrações que acompanharam essa proposta, reflexão já discutida em páginas anteriores, na análise lexical.

Em relação ao Tema 03, o CRAEF 2 também diminuiu a quantidade de fraseologismos. No entanto, o padrão das colocações foi descrever as reações do personagem sitiante frente aos acontecimentos, comprovados por “mijar na calça”, “sair correndo” e “morrer de medo”, totalizando dezenove combinatórias. Na sequência, temos o debate dos itens “sair” e “fogo”, iniciando pelo verbo da 3ª conjugação “sair”:

Figura 54 – Tela das concordâncias com a palavra “SAIR”

The screenshot shows a window titled 'Concord' with a menu bar (File, Edit, View, Compute, Settings, Windows, Help) and a table of concordance results. The table has two columns: 'N' (line number) and 'Concordance' (text snippet). The text is from a story about Chico Bento and a headless mule. The word 'SAIR' is highlighted in red in the original image, and the words 'SAIU' and 'SAIU' are highlighted in blue.

N	Concordance
1	chama mula sem cabeça e o Chico Bento saiu correndo e esqueceu a lanterna e a
2	uma mula sem cabeça então Chico Bento saiu voando para casa ele deixou a
3	PELO LAMPIÃO O CHICO SAIU CORRENDO PARA PEGAR UM
4	só que ele viu a mula sem cabeça e saiu correndo aí a mula sem cabeça se
5	e deixou a lamparina cair no chão e saiu correndo e a mula sem cabeça
6	CARA COM A MULA SEM CABEÇA E SAIU CORRENDO E NÃO VOLTOU
7	DA MULA SEM CABEÇA ELE SAIU CORRENDO DE MEDO E
8	uma toca e foi no banheiro assim que ele saiu do banheiro ele foi para a toca e
9	ele encontrou a mula sem cabeça e ele saiu correndo desesperado e largou o
10	PARA PEGAR. MAS QUANDO ELE SAIU DE CASA DEU DE CARA COM A
11	com vontade de ir ao banheiro ele saiu de casa com uma lamparina com
12	a lamparina já estava sem fogo e a mula saiu foi para outro lugar sem a
13	voltou e encontrou o banheiro e outra saiu e viu a mula sem cabeça e nunca

Fonte: PrintScreen da tela do programa *WordSmith Tools*

Das treze linhas de concordância, destacamos a presença de oito colocações, todas relacionadas com a ação de sair de algum local em alta velocidade. Dessa forma, a colocação verbal “sair correndo”, apareceu seis vezes, “sair voando”, outra combinação verbal com uma ocorrência e uma colocação adverbial “sair correndo desesperado”.

Conforme apontado nos parágrafos anteriores, reafirmamos o destaque para as ações e reações do Chico Bento, descritas em detalhes pelas crianças. Atribuímos esse padrão pelo fato de o personagem ser uma criança, e assim, despertar a empatia e a identificação dos alunos.

A próxima palavra-chave é “fogo”, descrita aqui como forma de comparar a escrita dos discentes do centro e da periferia na abordagem desta palavra:

Figura 55 – Tela das concordâncias com a palavra “FOGO”

N	Concordance
1	e deu beijo na lamparina e falou ti amo fogo porque a mula não tem cabeça a
2	toda história acabou o céu abalou com fogo e a banda tocou e estava
3	era a velinha dele porque estava com fogo e ficou olhando a mula sem
4	saiu de casa com uma lamparina com fogo ele viu uma mula sem cabeça
5	mula não tem cabeça a da mula é de fogo para pagar o Chico ele tem
6	FORÁ COM UM LAMPIÃO CHEIO DE FOGO E ENCONTROU A MULA SEM
7	a mula ela sentiu um cheiro de fogo ela ficou apaixonada do fogo e o
8	cheiro de fogo ela ficou apaixonada do fogo e o Chico Bento correu com
9	não quebrou a mula sentiu o cheiro do fogo ela pegou a lamparina e ficou
10	ELA TEM CABEÇA DE MUITO FOGO E AS VEZES ELA FICA
11	sem cabeça e pensou que mataram o fogo e foi embora e o Chico boto
12	pelo fogo e - quebrou o lampião e o fogo sumiu e a mula sem cabeça e
13	cama dela dormir no mato e voltou o fogo na cabeça a mula ficou contente
14	Chico tacou água na mula apagou o fogo da mula foi no rio tomar banho
15	atrás da mula então a mula esperou o Fogo então nunca mais Fez isto e Fim
16	da cuca feio não é olha a panela o fogo é coisa feia nossa olha a sala
17	um filho dela e ficou apaixonado pelo fogo e - quebrou o lampião e o fogo
18	a lamparina ele encontrou e ponhou fogo dentro dela e foi ao banheiro e
19	acordou a lamparina já estava sem fogo e a mula saiu foi para outro lugar
20	que linda ficou impressionado você tem fogo eu também mais você não fala
21	que assustou o Chico Bento tinha fogo na cabeça e corpo de cavalo no

Fonte: PrintScreen da tela do programa WordSmith Tools

As vinte e uma linhas de concordância resultaram em seis colocações, do tipo verbais e nominais. Das combinatórias verbais, obtivemos “apagar o fogo” e “ponhar fogo”, com uma ocorrência cada. Já as colocações nominais foram “lamparina com fogo” e “fogo na cabeça”, com um retorno cada e “cheio de fogo”, que apareceu duas vezes no CRAEF 2.

Novamente, chamamos a atenção para a existência de combinatórias que não são usuais na língua escrita, como “ponhar fogo”, por exemplo, segundo as normas cultas da língua portuguesa. Proveniente do linguajar oral, esta forma de expressar o verbo “pôr” não é indicada para uso em uma produção textual, corroborando a necessidade de ensinar a criança a se expressar de outras formas, dentro do que for mais adequado para o contexto.

Finalizando esta etapa, temos o rol de todas as colocações existentes no Tema 04, conforme segue:

Quadro 10 – Colocações do CRAEF 2 (Tema 04)

Tipo de colocação	Estrutura	Exemplos
Colocações Verbais	V + S / V + Prep + S	<i>Tomar banho, cavalgar com a mula, dar de cara, mijar na calça, dar surra nele, grudar com o lampião, encher a barriga e apagar o fogo.</i>
Colocações Nominais	S + Prep + S	<i>Lamparina com fogo, cheio de fogo, cheiro de fogo, fogo na cabeça e leite de vaca.</i>
Colocações Adjetivas	S + Adj.	<i>Cara de pau.</i>
Colocação Adverbial	V + Adv. ou Adv + Adj	<i>Ir devagarinho, morrer de medo, sair correndo, sair correndo desesperado e chegar logo em casa.</i>

De posse dos dados, apontamos para o alto índice de colocações verbais redigidas pelos alunos da periferia, nove ao todo; seguido das colocações nominais e adverbiais, com cinco ocorrências em cada variedade. Por fim, obtivemos somente uma colocação adjetiva.

Em comparação ao CRAEF 1, ambas as salas diminuíram o número de fraseologismos descritas neste tema; embora não deixassem de expressar com colocações a linguagem do mundo fantástico do folclore, com combinações expressivas e descritivas deste ambiente.

Novamente, os alunos do centro produziram mais fraseologismos do que a escola do centro, com uma colocação de diferença. Além da influência da imagem na folha de redação oferecida aos alunos, ainda creditamos este número menor de fraseologismos redigido nos dois grupos devido a um provável desconhecimento de certos detalhes da narrativa. Como exemplos, apontamos o fato de nenhum aluno ter citado que a mula sem cabeça é a “punição folclórica” para a mulher que se apaixona por um padre, ou então de que forma “quebrar o feitiço” para que a mesma volte ao normal.

Por isso, além da variação de temas e gêneros textuais a serem escritos pelas crianças, também é preciso explorar mais a fundo as lendas abordadas, para que o estudante, efetivamente, sinta-se seguro no relato de suas impressões sobre o assunto dado. Finalizando esta etapa das colocações, divulgamos o próximo e último bloco da análise de textos do ponto de vista fraseológico, acerca da solidariedade.

Tema 05 – “Como posso ajudar as pessoas que vivem perto de mim?”: 11/2015

Analogamente às demais propostas, trazemos neste último segmento as discussões sobre as colocações encontradas acerca desta temática. Visto esta propositura abordar a

solidariedade, um substantivo abstrato, ilustramos o ato de bondade por meio de duas figuras, uma no contexto hospitalar e outra praticando a doação de brinquedos. Após realizar a compilação das redações, constatamos a presença de vinte colocações, que procuraram descrever as situações citadas, ora referindo-se ao primeiro exemplo de ação solidária, como em “soro meio amarelado” e “buquê de flores”, ora fazendo alusão ao segundo desenho, notado nas combinatórias “doador de brinquedos” e “carrinho de controle remoto”.

Por isso, as palavras-chave destacadas neste corpus são “natal” e “boa”. Na sequência, apresentamos as linhas de concordância do item lexical “natal”:

Figura 56 – Tela das concordâncias com a palavra “NATAL”

The screenshot shows a window titled 'Concord' with a menu bar (File, Edit, View, Compute, Settings, Windows, Help) and a table of concordance results. The table has two columns: 'N' (line number) and 'Concordance' (text snippet). The word 'natal' is highlighted in red in the original image, and 'Natal' is highlighted in blue.

N	Concordance
1	felizes para sempre. Fim quando chega natal eu ganho e doo brinquedos tenho
2	sempre. Boa ação. Era um belo dia de natal . O menino Cleiton acordou todo
3	todo ansioso para ver na árvore de natal se tinha algum presente. A o chegar
4	. Um dia ela acordou e era véspera de natal e recebeu varias visitas e a ultima
5	presente. A o chegar perto da árvore de natal viu um monte de presentes e
6	perguntar o que ele tinha ganhado de natal . O menino perguntou: - O que você
7	pra você Um dia era véspera de natal e os pais de um menino tinha
8	e a irmã foi levar um presente de natal e levou um buquê de flores ai os
9	sarou!- E as duas ficaram felizes e era Natal elas passaram o Natal inteiro juntas
10	pais dela levaram ela até o hospital e era natal e a melhor amiga dela foi visitar a
11	a amiga. - Toma essas flores e Feliz natal todo mundo está com saudades de
12	para Ele: Que vou comprar Brinquedo no natal E no final Do ano também ele falou:
13	menino que tinha ganhado um presente no Natal e o brinquedo que ele não brinca
14	felizes e era Natal elas passaram o Natal inteiro juntas e junto com a família
15	Quero ver a minha amiga e tá chegando o natal . Eu vou dar umas flores para ela. -
16	uma menina triste e doente, ela amava o Natal que é de Jesus. Essa menina
17	quem não tem. Chegou no Dia seguinte O natal filho não vai dar para te dar
18	crianças pedindo brinquedos. No outro natal eu ganhei o que eu mais queria um

Fonte: PrintScreen da tela do programa WordSmith Tools

No total, obtivemos 18 linhas de concordância para a palavra “natal”, que gerou nove colocações: uma de ordem verbal “chega natal”, com uma ocorrência; sete nominais, como “dia de natal”, “presente de natal” e “presente no natal”, apresentadas uma vez e “árvore de natal” e “véspera de natal”, com duas recorrências cada e, por fim, uma combinatória adjetiva, “feliz natal”. Provavelmente, o fato de escreverem esta redação ao final do mês de novembro, portanto próximo as férias escolares, colaborou para o entendimento de que já estávamos às vésperas desta data comemorativa.

A seguir, temos o adjetivo “boa”, destacado pelo contexto da colocação “boa ação”, disposto na sequência:

Figura 57 – Tela das concordâncias com a palavra “BOA”

N	Concordance
1	para pessoas pobres e humildes. A Boa ação Numa linda noite uma
2	deles e viveram felizes para sempre. Boa ação. Era um belo dia de natal.

Fonte: PrintScreen da tela do programa *WordSmith Tools*

As duas colocações adjetivas são passíveis de registro, pois, demonstram que apesar de terem de sete a nove anos de idade, as crianças conseguem retratar itens coerentes com o tema exposto, solidariedade. Além disso, este tema reforça o desejo de paz e boas atitudes ao próximo, o que justifica a escrita deste fraseologismo e de outros, como por exemplo: “receber visitas” e “doação de brinquedos”.

Já a classificação das colocações apresentadas encontra-se no Quadro 11:

Quadro 11 – Colocações do CRAEF 1 (Tema 05)

Tipo de colocação	Estrutura	Exemplos
Colocações Verbais	V + S / V + Prep + S	<i>Andar de bicicleta, atravessar a rua, tomar um soro, dar um atestado, tomar banho, receber visitas e quebrar a perna.</i>
Colocações Nominais	S + S ou S + Prep + S	<i>Menino doador, doação de brinquedos, carrinho de controle remoto, freio da bicicleta, véspera de natal, doador de brinquedos, buquê de flores, presente de natal e árvore de natal</i>
Colocações Adjetivas	S + Adj.	<i>Soro meio amarelado, feliz natal e boa ação.</i>
Colocação Adverbial	V + Adv.	<i>Passar mal.</i>

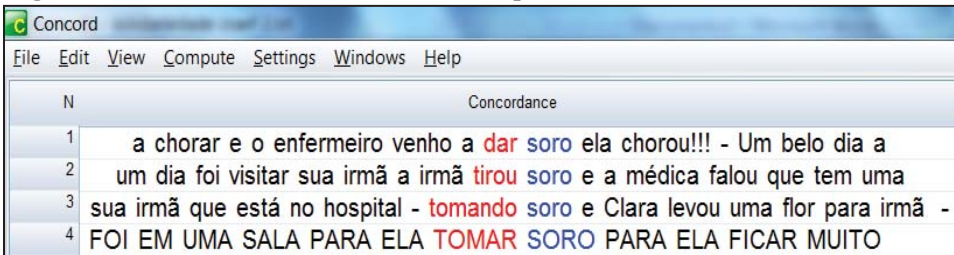
Nesta seleção, tivemos a inclusão de sete colocações verbais, nove combinações nominais, três de ordem adjetivas e uma adverbial. Ao privilegiar as ações e nomear vários substantivos diferentes, notamos que os alunos realizaram menos caracterizações, indicada pelos poucos adjetivos e advérbios presentes nas combinações acima.

Passemos agora a descrição dos fraseologismos pelo segundo grupo, composto por discentes da periferia, cuja soma de combinatórias encontradas foi de dezenove colocações na última proposta textual. Paralelamente à turma de estudantes do centro, o CRAEF 2 também utilizou diversas classes gramaticais para expressar o tema da solidariedade, comprovado pelo

uso de colocações como “(receber de) braços abertos”, “sentir sua falta” no relato da proposta sobre as amigas no hospital, além das combinações “sem abrigo” e “ajudar o próximo” observados na escrita acerca da doação de brinquedos.

Visto que a maioria dos discentes deste grupo destacou a primeira ilustração, trazemos as análises dos itens lexicais “soro” e “câncer”, na discussão deste último tema, iniciando com as linhas de concordância para “soro”, conforme segue:

Figura 58 – Tela das concordâncias com a palavra “SORO”



N	Concordance
1	a chorar e o enfermeiro venho a dar soro ela chorou!!! - Um belo dia a
2	um dia foi visitar sua irmã a irmã tirou soro e a médica falou que tem uma
3	sua irmã que está no hospital - tomando soro e Clara levou uma flor para irmã -
4	FOI EM UMA SALA PARA ELA TOMAR SORO PARA ELA FICAR MUITO

Fonte: PrintScreen da tela do programa *WordSmith Tools*

Ao todo, foram identificadas quatro colocações verbais, com a predominância de “tirar soro”, ocorrida três vezes e “dar soro”, com uma frequência. Visto a aparição desta combinatória “dar soro”, realizamos uma pesquisa no *Sketch Engine* para verificar a frequência desta combinatória, na qual encontramos somente cinco ocorrências, todas resultantes da transcrição oral de reportagens sobre saúde. Já “tirar soro” obteve 1.730 resultados, que variam entre “soro de creme de leite”, “soro de queijo”, “soro do braço”; enquanto “tomar soro” ocorreu 6.637 vezes, com a maioria dos casos no contexto mencionado.

Dessa forma, atestamos para a importância de oferecer respaldo aos alunos no tocante a compreensão dos fraseologismos, devido à especificidade da linguagem e das circunstâncias em que são enunciados os discursos. De posse destas informações, a probabilidade de o aluno se comunicar de maneira mais eficaz na escrita será significativamente maior.

O último item lexical a ser estudando diz respeito a “câncer”, que obteve quatro linhas de concordâncias:

Figura 59 – Tela das concordâncias com a palavra “CÂNCER”

The screenshot shows a window titled 'Concord' with a menu bar (File, Edit, View, Compute, Settings, Windows, Help). Below the menu is a table with two columns: 'N' and 'Concordance'. The table contains four rows of text where the word 'câncer' is highlighted in blue and 'DE CÂNCER' is highlighted in red.

N	Concordance
1	ela no hospital quando ela estava com câncer quando foi o outro dia sua
2	com câncer Um dia a Carol ficou com câncer e os cabelos caíram e ela
3	FELIZ. Carol e Ana a menina com câncer Um dia a Carol ficou com
4	PARA DOAR PARA O HOSPITAL DE CÂNCER PARA AJUDAR CEIA DE

Fonte: PrintScreen da tela do programa WordSmith Tools

Ao todo, quatro composições trouxeram esta doença para retratar a menina internada no hospital. Nelas, relatamos duas colocações verbais “estar com câncer” e “ficar com câncer”, com uma frequência cada uma e uma combinatória nominal “hospital de câncer”.

Neste rol ainda é possível ver o conhecimento que as crianças possuem sobre esta doença, presente nas colocações “cair os cabelos”, “raspar a cabeça” e “ficar muito mal”. Tal constatação denota que, independente de termos usados, a imagem de uma criança doente na proposta desta redação, denotando um assunto mais “denso” a ser escrito por estudantes de sete a nove anos, os discentes conseguem se expressar e apontar elementos que se enquadram na temática sugerida.

Quadro 12 – Colocações do CRAEF 2 (Tema 05)

Tipo de colocação	Estrutura	Exemplos
Colocações Verbais	V + S / V + Prep + S	<i>Tomar soro, sentir sua falta, chegar um filho, tocar a campainha, raspar a cabeça, ajudar o próximo, andar de bicicleta e limpar a casa.</i>
Colocações Nominais	S + Prep + S	<i>Picada de marimondo, hospital de câncer, ceia de natal e hospital de saúde.</i>
Colocações Adjetivas	S + Adj.	<i>Braços abertos, anjo da guarda e guarda temporária.</i>
Colocação Adverbial	V + Adv. ou Adv + Adj	<i>Ficar muito mal, acabar tudo bem, feliz para sempre e paixão sem fim.</i>

Neste tema, os alunos do CRAEF 2 escreveram oito colocações verbais, quatro nominais, três adjetivas e quatro adverbiais. De uma forma geral, as combinatórias foram bem distribuídas, com a costumeira predominância dos fraseologismos verbais em detrimento aos outros.

De acordo com o disposto das produções textuais número 5, foi possível verificar que,

em ambas as turmas, CRAEF 1 e CRAEF 2, o léxico-fraseológico ilustrou o tema proposto, com a descrição da solidariedade por meio de uma linguagem significativa e empática com as imagens dispostas na folha de redação. Neste contexto, a presença das ilustrações pode ter sido benéfica aos alunos, pois permitiu que explorassem com propriedade as situações vivenciadas por crianças que estão internadas em hospitais ou então que não possuem brinquedos para se divertir.

Ainda no trato das colocações, os estudantes do centro novamente se sobressaíram na quantidade de colocações redigidas, repetindo a diferença exposta no tema anterior, com apenas uma combinatória a mais do que a escola da periferia.

Concluindo, por meio das fundamentações teóricas já dispostas acerca da fraseologia e das colocações, foi possível analisar as redações coletadas das crianças das duas salas de aula em estudo. Para realizar das investigações empíricas, foi necessário adicionar os textos no *software WordSmith Tools* a fim de encontrar mais facilmente as palavras-chave que geraram as colocações destacadas nas páginas anteriores.

Devido à quantidade de itens lexicais encontrados, não foi possível realizar a análise pormenorizada de todas as colocações. Contudo, ao longo do item 3.3, efetuamos reflexões sobre algumas, procurando apontar variedades e especificidades da linguagem infantil. Dessa forma, apesar dos fatores já tratados nas primeiras páginas desta análise, os alunos da periferia produziram quase a mesma quantidade de colocações do que os discentes do centro, apenas treze combinatórias a menos do que os estudantes do CRAEF 1.

Assim, mesmo oriundos de locais sociais diferentes, reiteramos a qualidade e a criatividade de ambas as turmas, que souberam expressar suas impressões e individualidade nas linhas redigidas. Não houve, portanto, uma diferença expressiva na maneira de relatar as colocações, visto que houve ocorrências de combinatórias não usuais nos dois corpora de estudo.

Contudo, por termos coletado o montante de 224 textos, esperávamos uma quantidade expressiva de fraseologismos. Após a realização das análises, contabilizamos 189 colocações, dos dois conjuntos de redações coletadas. Enquanto algumas produções geraram quatro ou até cinco combinatórias, várias outras não produziram nenhuma que pudéssemos utilizar na dissertação.

Por esse motivo, decidimos acrescentar mais um objetivo específico neste trabalho, a fim de contribuir para o ensino das colocações aos alunos do Ensino Fundamental I. Para tanto, o próximo capítulo fará um detalhamento das estratégias empregadas na produção das atividades e, na sequência deste relato, apresenta efetivamente os dez exercícios elaborados

para ampliar a competência colocacional dos estudantes de língua materna, a saber, a língua portuguesa do Brasil.

4. PROPOSTAS DE ATIVIDADES PARA O DESENVOLVIMENTO DA COMPETÊNCIA COLOCACIONAL DOS ALUNOS DO ENSINO FUNDAMENTAL I

Conforme observado ao longo desta análise, os alunos efetuaram vários arranjos de palavras, observados por meio da análise das produções textuais, resultando no uso de combinatórias lexicais que nem sempre estão de acordo com o uso consagrados pelos falantes de língua portuguesa do Brasil, tais como “irem banho” ou a “vida na pedra”, ao invés de “tomar banho” e “idade da pedra”.

Tal fato linguístico demonstra que não basta somente alfabetizar, ou ensinar a criança como articular fonemas e letras para escrever. O estudante também necessita adquirir a habilidade de usar as combinatórias de palavras em sua própria língua materna com propriedade.

Por isso, ao final de todas as análises acerca das produções textuais coletadas do CRAEF, sentimos a necessidade de refletir a partir de um olhar pedagógico sobre os resultados obtidos, o que nos levou a pensar de que maneira poderíamos contribuir para que os alunos possam elaborar uma escrita mais rica e variada, no que tange ao uso das colocações.

Neste sentido, decidimos propiciar situações em sala de aula, por meio de atividades didáticas, a fim de que o professor auxilie no desenvolvimento da competência colocacional dos alunos.

Essa decisão se explica em razão de minha prática docente atuante nas séries iniciais há 11 anos, desde que assumi o cargo público de Professora de Educação Básica (PEB I) na prefeitura de Votuporanga, em abril de 2007. Tal prerrogativa nos permite, juntamente com a professora orientadora desta dissertação, elaborar exercícios de cunho prático que atendam à lacuna observada a respeito do uso dos fraseologismos nas redações destes estudantes.

Assim, a inserção de colocações em atividades didáticas voltadas às crianças é pertinente à proposta desta pesquisa, pois visa a atender os seguintes objetivos:

- ampliar o vocabulário dos discentes do ensino fundamental I, ao passo em que consolidam a etapa da alfabetização da língua materna, a partir de tarefas com o uso de uma linguagem autêntica, partindo do cotidiano do aluno;
- propiciar condições pedagógicas para os alunos melhorarem a escrita de redações, ao empregar as colocações em suas produções textuais de forma natural e de acordo com o

gênero proposto;

- possibilitar o entendimento de que há formas próprias de articulação da linguagem, consagradas pelo uso da comunidade de falantes da língua portuguesa do Brasil;
- prover o desenvolvimento da consciência fraseológica acerca das combinações lexicais comumente usadas na sua língua materna, auxiliando o discente a empregá-las no seu dia a dia, tanto na comunicação oral quanto escrita;
- proporcionar, a longo prazo, ao aperfeiçoamento da compreensão e interpretação de textos de qualquer espécie, reverberando na melhora significativa da alfabetização como um todo.

A fim de atingir os objetivos propostos acima, a finalidade deste último capítulo como mencionamos, é sugerir atividades didáticas baseadas em corpus que ofereçam estes subsídios ao professor das séries iniciais. Vale lembrar que no escopo desta pesquisa, a aplicação dessas atividades serão etapas a serem realizadas no doutorado da autora deste trabalho. Desse modo, este capítulo objetiva descrever a metodologia empregada para a elaboração e a seleção das colocações a serem empregadas nessas atividades, para melhor atender as especificidades do público infantil.

Partindo da premissa de que as crianças necessitam sentir segurança nas ações que executam com autonomia, as atividades possuem níveis de dificuldade variados, desde o grau fácil, perpassando pelo intermediário e culminando em propostas mais complexas e desafiadoras. Caso contrário, os alunos podem se sentir desmotivados por realizarem somente exercícios muito fáceis ou extremamente difíceis.

Nomeamos de atividades fáceis, aquelas que utilizam as colocações presentes no CRAEF, ou seja, trazem as combinatórias elencadas pelas crianças, sendo representativas do interesse dos mesmos, pertencendo à realidade de mundo dos alunos. Estes exercícios possuem o propósito de introduzir o tema, apresentando o assunto aos discentes que não produziram nenhum fraseologismo em suas produções textuais.

Visto que estas tarefas foram produzidas por meio da seleção de colocações redigidas pelos próprios estudantes do ensino fundamental, temos certeza de que serão facilmente compreendidas por esse público. Além disso, após sua análise, tais colocações também foram selecionadas em razão de as consideramos relevantes para comporem o conjunto lexical dos alunos dessa faixa etária, segundo nossa experiência como professora.

Assim, para prosseguirem nas próximas etapas de aprendizagem, extraímos colocações do universo infantil dentro das propostas de texto apresentadas, com base nas palavras-chave do CRAEF. Em seguida, a partir dessas palavras, buscamos no Corpus do

Português, novas colocações para incluir nos exercícios que possuem um nível de dificuldade maior, instigando um desafio maior aos estudantes. As atividades de grau intermediário e difícil são aquelas compostas por um vocabulário não listado no CRAEF. Enquanto as primeiras são compostas por colocações mais próximas da linguagem infantil, as últimas representam combinatórias mais elaboradas, oriundas de um aprofundamento maior no léxico das propostas.

Mais uma preocupação de ordem metodológica, oriunda da nossa formação no CEFAM (Centro Específico de Formação e Aperfeiçoamento do Magistério) e no curso superior em Pedagogia foi como apresentar o conteúdo fraseológico de forma lúdica para o discente. Por isso, a concepção destas atividades buscou formas atraentes, a fim de chamar a atenção do discente, tanto por meio de ilustrações quanto pelas frases de incentivo à sua realização.

4.1 Procedimentos metodológicos para a seleção das colocações presentes nas atividades

Antes de apresentarmos as atividades, faremos uma introdução para explicar quais os parâmetros utilizados na seleção das colocações que compõem os exercícios.

Inicialmente, verificamos as colocações empregadas nos cinco temas das redações coletadas. Visto que cada proposta privilegiava campos semânticos diferentes, examinamos quais combinatórias relacionavam-se efetivamente aos temas propostos e quais não fizeram parte do rol elencado pelas crianças.

Neste contexto, elaboramos uma tabela com as colocações empregadas pelos alunos que possuam relação com cada tema sugerido. No entanto, a nossa experiência como docente nas séries iniciais nos indica que, por estarem na faixa etária de sete a nove anos, os alunos já poderiam ter utilizado outras combinatórias além daquelas que aparecem no CRAEF.

Dessa forma, inserimos algumas palavras-chave baseados no Corpus do Português na coluna “Colocações possíveis, mas não citadas”, a fim de compor os exercícios de categoria intermediária e complexa. A intenção é possibilitar o uso de colocações variadas já nesta etapa de ensino, e não postergar este conteúdo somente para o Ensino Fundamental II. Por isso, reiteramos a importância destas atividades, posto que serão úteis para os discentes desenvolverem a competência colocacional desde o início de sua escolarização.

Visto que as atividades foram elaboradas por meio de corpus para a utilização em

escola pública, consideramos a necessidade de o docente levar os discentes ao laboratório de informática para pesquisarem a frequência das colocações. Dessa maneira, de forma prática, eles compreenderão que a convenção de uma combinatória ocorre pelo uso frequente das colocações pela comunidade de falantes.

No entanto, sabemos das inúmeras limitações existentes neste âmbito, tais como: poucos computadores disponíveis para uma grande quantidade de alunos, conectividade baixa e falta de recursos financeiros que permitam comprar a licença de uso de *sites* especializados como o *Sketch Engine*, por exemplo.

Por estas razões, sugerimos ao educador que examine as colocações no site de busca do *Google*³¹, pois, além de gratuito, este recurso possui uma *interface* que o estudante domina e usa constantemente em seu cotidiano. Ensinando aos estudantes como realizar as pesquisas e observar os resultados apresentados, ele será capaz de identificar padrões recorrentes na língua e detectar outros que não sejam tão usuais da fala e escrita da língua portuguesa do Brasil.

Uma última explicação necessária para a elaboração deste capítulo das atividades, diz respeito à nomenclatura “colocação” ser usada em algumas atividades. Nesses casos, pressupomos que o professor fará uma breve, porém necessária, explanação sobre o significado deste fraseologismo, para que o aluno sintam-se seguros na realização dos exercícios.

Novamente, cabe ao professor a incumbência de demonstrar quais as alternativas mais empregadas pelos falantes nativos e apresentá-las às crianças de forma didática e clara. Feitos os esclarecimentos de cunho metodológico, iniciamos a descrição da escolha das colocações. Do primeiro rol de textos, temos:

Quadro 13 – Campo semântico das colocações do Tema 01

Tema 1 CRAEF 1: Relato do filme “Os Croods”		Tema 2 CRAEF 1: Redação sobre animais	
COLOCAÇÕES REDIGIDAS PELOS ALUNOS	COLOCAÇÕES POSSÍVEIS, MAS NÃO CITADAS	COLOCAÇÕES REDIGIDAS PELOS ALUNOS	COLOCAÇÕES POSSÍVEIS, MAS NÃO CITADAS
Caçar/ achar/ comida	Homem das cavernas	<i>Pescoço grande</i>	Animais em extinção
Idade da pedra	Pintura rupestre	Estica o pescoço	Habitat natural
Pelo de animais	Pré-história	Zebra muito listrada	Animais selvagens / domésticos

Conforme apontado pelo quadro acima, durante a primeira proposta de redações do CRAEF 1, os alunos trouxeram as colocações relacionadas a “caçar comida” 5 vezes; “idade da pedra” teve 5 ocorrências; e “pelo de animais” só teve 1 caso registrado no corpus.

³¹ Pesquisas realizadas na *website* <www.google.com.br>, nos dias 23 e 24 de janeiro de 2018.

Conforme apontado no último tópico, houve a frequência 1 para “a vida da pedra”, apontando desconhecimento de alguns alunos sobre a convenção adequada da colocação “idade da pedra”.

É possível perceber que, em muitos casos, os alunos utilizam termos pouco recorrentes na língua portuguesa do Brasil, enquanto desconhecem os mais usuais a respeito do período histórico referido do filme “Os Croods”. Fazemos essa afirmação, pois as colocações empregadas pelos discentes do CRAEF 1, possuem, respectivamente a seguinte frequência no google: “caçar comida”, com 7.890 resultados; “idade da pedra”, com 409.000 retornos; e “pelo de animais”, com 186.000 repetições.

Há outros fraseologismos significativos deste ambiente, tais como “homens das cavernas”, “pintura rupestre”, também chamada de “arte rupestre” ou ainda “gravura rupestre” e “pré-história”, que não foram citados nenhuma vez no corpus da escola do centro. Este fato nos fez procurar a frequência dos itens abordados acima, cujo resultado apontam que as colocações sugeridas acima são mais comumente usadas pelos falantes da língua portuguesa do Brasil do que aquelas empregadas pelos alunos.

Fazemos essa afirmação, pois a frequência registrada no Google para o uso de “caçar comida” e “pelo de animais”, por exemplo, é bem menor do que “homens das cavernas”, que possui 347.000 resultados, “pintura rupestre”, com frequência de 422.000 vezes e “pré-história”, com 5.190.000 ocorrências. Apesar de serem combinações próprias do universo do filme, chamamos a atenção para a função do professor de língua portuguesa oferecer outras possibilidades linguísticas, que ajudem o discente a se comunicar com mais propriedade, ampliando o léxico com o acréscimo de mais colocações, como as sugeridas neste tópico.

Os discentes do CRAEF 2, por sua vez, elencaram traços distintivos mais pontuais dos animais, como o tamanho característico do pescoço da girafa, 2 ocorrências ou as listras da zebra, descrita 1 vez no corpus.

Analogamente ao efetuado anteriormente na escola do centro, também averiguamos no CRAEF 2 os dados relacionados a constância das colocações apresentadas no site de busca outrora citado, e chegamos no seguinte índice numérico: “pescoço grande” obteve um escore de 77.900 ocorrências, “estica o pescoço” foi relacionado em 31.200 resultados e “zebra muito listrada” aparece somente 2 vezes no site Google.

Neste rol temático, outras possibilidades inexploradas nas redações das crianças da periferia seriam as colocações “animais em extinção”, “habitat natural” e “animais selvagens ou domésticos”, haja vista que também pertencem ao campo semântico proposto.

Já as colocações nomeadas acima tiveram as seguintes frequências na *website*:

“animais em extinção” 275.000 vezes, “habitat natural” foi indicado 545.000 vezes, “animais selvagens” apareceu em 3.130.000 resultados e “animais domésticos” obteve 2.050.000 retornos. Estes números comprovam a afirmação feita anteriormente sobre o CRAEF 1, acerca da baixa frequência das colocações trazidas pelos alunos, reiterando em um possível desconhecimento de formas mais usuais, mais elaboradas ou, ainda, mais específicas do português do Brasil.

Portanto, no primeiro grupo de textos coletados, ambos os corpora em estudo CRAEF 1 e CRAEF 2 ignoram ou omitem as colocações mais usuais ou mais específicas sobre o assunto proposto nas redações, resultando em combinatórias linguísticas menos representativas ou menos sofisticadas da língua portuguesa do Brasil.

Por isso, a apresentação das colocações não mencionadas pelos discentes é o mote da atividade 1 (página 170), que pretendem relacionar as possibilidades linguísticas destes temas. Objetiva-se, dessa forma, aliar as colocações que os alunos já possuem em seu repertório lexical àquelas que ainda desconhecem ou que não estão habituadas a utilizar em seus textos.

Já, no segundo argumento de redações, seguem as colocações elencadas das duas salas de aula em estudo, relacionadas à cidade de Votuporanga:

Quadro 14 – Campo semântico das colocações do Tema 02

Tema 2– Redação sobre a cidade de Votuporanga		
COLOCAÇÕES PRESENTES NO CRAEF 1	COLOCAÇÕES PRESENTES NO CRAEF 2	COLOCAÇÕES POSSÍVEIS, MAS NÃO CITADAS NO CRAEF
Igreja / praça matriz	Erguer / soltar/ empinar pipa	Área urbana e área rural
Câmara municipal	Andar de bicicleta	Densidade demográfica
Concha Acústica		Agência dos correios

Neste quadro, as colocações nominais predominaram no CRAEF 1, comprovados pelos termos igreja/prça matriz, que obtiveram, respectivamente, 1 e 3 repetições. Outro local citado foi a Câmara Municipal, com 1 ocorrência apenas, além da Concha Acústica, que também teve somente 1 citação, indicando lugares recorrentes para as crianças daquele local, já que estas moram mais perto do centro da cidade.

Aprofundando o uso das combinações de palavras no site de busca mencionado, relatamos que os alunos do CRAEF 1 realizaram uma escrita com colocações bastante representativas da língua portuguesa, atestados por: “igreja matriz”, que apresentou 1.030.000 resultados no *Google*; “praça matriz”, com 459.000 retornos, seguidos de “câmara municipal”, possuindo 22.400.000 citações e, “concha acústica”, retratado 624.000 vezes.

Já, os habitantes da periferia, por morarem em um bairro afastado da região central de

Votuporanga, não elencaram nenhum dos lugares descritos pelo outro grupo, preferindo expor o que fazem nas praças, calçadas e ruas da cidade. Neste sentido, predominaram as colocações verbais, confirmadas pelas ações desempenhadas ludicamente. Dessa forma, temos as colocações com a base “pipa”: “erguer pipa” – 1 ocorrência; “soltar pipa” – 1 vez; e “empinar pipa” – 1 frequência; além da base “bicicleta”: combinada com o verbo andar, em 8 ocorrências.

Pesquisando as colocações no Google, enquanto a colocação “andar de bicicleta” apareceu 417.000 vezes, encontramos as três variáveis apontadas pelo CRAEF 2 para indicar a brincadeira com pipa, aferida da seguinte forma: “erguer pipa”, com 1.070 ocorrências; “soltar pipa” apresentou 146.000 resultados; e “empinar pipa” 70.900 relatos. A forma “erguer” nos indica um uso menos usado desta combinação, enquanto as demais são mais frequentes pelos falantes.

De maneira geral, ambas as turmas trataram das brincadeiras que fazem, o que é característico do vocabulário infantil, embora as crianças do CRAEF 1 retratam locais pertencentes ao centro da cidade, conforme outrora mencionado.

Sobre as colocações não utilizadas, deixamos claro que reconhecemos o fato de os agentes participantes da pesquisa serem crianças e, assim, demonstrarem um vocabulário um pouco limitado devido à idade. Este resultado indica a necessidade de ampliar o léxico-fraseológico destes alunos, pois defendemos o ensino das colocações no ensino fundamental como forma de auxiliá-los a compreender melhor como se articulam as combinatórias de palavras na língua portuguesa do Brasil.

Contudo, esperávamos que descrevessem mais itens relativos à “área urbana” ou “área rural” do município, visto que:

- são conceitos tratados largamente no conteúdo de geografia e fáceis de identificar nos espaços existentes da cidade e
- estão presentes no vocabulário da criança, posto que as crianças diferenciam um ambiente do outro, atestado pelas descrições que fazem dos passeios aos finais de semana: “Ontem fui no sítio da minha tia e andei de cavalo”, “Na fazenda onde meu avô trabalha, a gente pesca e nada no rio”, dentre outros relatos característicos destes locais.

Esses fraseologismos sugeridos também variam bastante quanto à sua ocorrência em língua portuguesa, indicados por: “área urbana” – 3.180.000 recorrências e “área rural” com frequência de 853.000 resultados. Defendemos a exploração destas colocações, pois, se as

mesmas fossem empregadas pelos estudantes, poderiam proporcionar um texto mais rico lexicalmente falando, visto que indicariam mais um elemento relacionado à identidade dos locais em que moram.

Ainda neste contexto sobre município, o trabalho com fraseologismos que expressem locais diferentes uns dos outros também é útil. Representado pelas colocações nominais “área urbana”, “área rural” e demais combinatórias provenientes da investigação deste campo semântico, estes tópicos serão observados na atividade 2 (página 171).

Reitera-se, dessa forma, a importância do objetivo elencado no início deste capítulo acerca de facilitar o contato da criança com colocações linguísticas díspares daquelas que ela utiliza em seu cotidiano, já que há muitas combinatórias possibilitadas por essa temática não foram abordadas tanto pelos alunos do centro quanto pelos discentes da periferia.

Prosseguindo, o terceiro rol de produções escritas, acerca da semente mágica, propiciou a apreciação das colocações dispostas na sequência:

Quadro 15 – Campo semântico das colocações do Tema 03

Tema 3 – Redação sobre a “Semente Mágica”		
COLOCAÇÕES PRESENTES NO CRAEF 1	COLOCAÇÕES PRESENTES NO CRAEF 2	COLOCAÇÕES POSSÍVEIS, MAS NÃO CITADAS NO CRAEF
Banho de sol	Terra do jardim / plantação de cenouras	Adubar o solo / Cultivar a semente /
Aguar a florzinha / regar as plantas /	Regar a semente / aguar a planta / dar água	Colher os frutos
	Brotar tomates / crescer um pé de maracujá / Virar uma linda planta	Solo arenoso / solo argiloso / solo calcário

No CRAEF 1, as colocações que se relacionam ao tema sugerido obtiveram as seguintes frequências: “banho de sol” – 1; “aguar a florzinha” – 3; e “regar as plantas” – 6. A busca destas colocações na *Internet* mostrou que a maioria delas também é convencional na comunidade linguística de falantes do português brasileiro, como provam “banho de sol”, com 586.000 resultados e “regar as plantas”, com 150.000. Neste quesito, somente a combinação “aguar a florzinha” obteve pouca frequência, com 5 retornos, visto que se trata de uma combinação presente mais na oralidade do que na escrita.

Em contrapartida, no CRAEF 2, observamos a presença das seguintes colocações nominais: “terra do jardim”, com 1 frequência e “plantação de cenouras”, também com 1 ocorrência; além das verbais “regar a semente”, 1 referência; “aguar a planta”, com 4 resultados; “dar água”, 1 forma; “brotar tomates”, 2 frequências, “crescer um pé de (alguma planta)”, 7 vezes e “virar uma planta”, 1 vez.

Os resultados encontrados na *website* Google apontam que “terra do jardim” é bastante utilizada pelos falantes do português brasileiro, com 207.000 recorrências, bem como “regar a semente”, com 28.200 ocorrências e “plantação de cenouras”, com 5.900 repetições. Já as combinatórias “crescer um pé de” teve 3.820 ocorrências, “virar uma linda planta” foi usada 1.310 vezes, “aguar a planta” retornou 1.820 dados, ao passo que “brotar tomates” obteve 43 resultados e “dar água” apareceu somente 8 vezes na pesquisa.

Nestes números, chama a atenção a fato de “aguar a planta” ser menos frequente no corpus da Internet do que “regar a semente”, enquanto o inverso ocorreu no corpus da periferia. Ou seja, os alunos usam a expressão menos utilizada pela convencionalidade “aguar” do que àquela considerada padrão “regar”. Depreende-se a influência da oralidade no texto escrito, visto que a primeira é uma expressão mais informal do ato de molhar a planta, enquanto a segunda refere-se a uma maneira mais formal de descrever a mesma ação.

Este é um indício da necessidade latente do professor apresentar outras maneiras de expressão linguística, conforme sugerido a seguir. A partir do exposto, sinalizamos a inclusão de outras combinatórias que também tratem do plantio e que são mais recorrentes do que aquelas apresentadas pelos discentes do CRAEF 1 e 2.

Após a realização da pesquisa *on-line*, tem-se: “adubar o solo”, com 22.600 retornos, “cultivar a semente”, obtendo 39.100 resultados, “germinar as sementes”, que aparece 12.600 vezes, “colher os frutos” possui 226.000 recorrências, “desabrochar das flores”, 15.700. Além destas colocações, indicamos outras um pouco mais complexas, como: “solo arenoso”, com 148.000 resultados, “solo argiloso”, com 98.000 retornos e “solo calcário”, com 35.300 reiterações.

Por isso, a fim de ampliar o rol de colocações sobre o tema do plantio, trazemos, ainda na atividade 3 (página 172), um vocabulário mais aprofundado, por meio de exercícios que explicam didaticamente os conceitos dispostos neste tópico.

O quarto conjunto de redações coletadas, sobre o folclore, resultou nas seguintes combinatórias:

Quadro 16 – Campo semântico das colocações do Tema 04

Tema 4 – Redação sobre a “Semente Mágica”		
COLOCAÇÕES PRESENTES NO CRAEF 1	COLOCAÇÕES PRESENTES NO CRAEF 2	COLOCAÇÕES POSSÍVEIS, MAS NÃO CITADAS NO CRAEF
Mula sem cabeça	Mula sem cabeça / múmia sem cabeça	Relinchar
Dar um cascudo	Cavalgar com a mula	Quebrar o feitiço
Sentir o cheiro de fogo	Deu um coice	

Sobre as colocações empregadas pela escola do centro, houve poucas relacionadas ao contexto da história, como “mula sem cabeça”, com 68 ocorrências, “dar um cascudo”, com 1 repetição e “sentir cheiro de fogo”, com 1 reiteração.

Na *website* já citada, mula sem cabeça obteve 208.000 resultados, “dar um cascudo”, obteve 16.800 recorrências e “sentir cheiro de fogo” apareceu 4 vezes, denotando que a sinestesia não costuma ser muito empregada pelos falantes nativos do português brasileiro.

Já, na periferia, tivemos a recorrência de “mula sem cabeça” 64 vezes no CRAEF 1, além da presença da combinação “múmia sem cabeça”, com frequência igual a 3. Além dessas, podemos relacionar ao tema “deu um coice”, ocorrida uma única vez 1 e “cavalgar com a mula”, também escrita 1 vez.

No Google, nos surpreendemos com a aparição da “múmia sem cabeça”, referida 519 vezes pelos usuários. Porém, as referências foram utilizadas no sentido egípcio da palavra, e não como a retomada da lenda brasileira, configurando uma maneira equivocada de expressão do aluno. Já, a combinatória “deu um coice” mostrou-se recorrente, com 42.100 citações, enquanto “cavalgar com a mula” teve uma frequência igual ao CRAEF 2, ou seja, somente 1 recorrência.

Haja vista oferecer outras possibilidades dos alunos utilizarem as colocações, aquelas dignas de nota seriam “relinchar” e “quebrar o feitiço”, visto que estão relacionadas a Mula sem cabeça: quando transformada em animal, a entidade fantástica não “fala”, mas seu som seria o “relinchar” e para voltar ao normal, é necessário “quebrar o feitiço”, retirando o freio de ferro existente no pescoço. As pesquisas efetuadas na Internet apontam a frequência de, “relinchar” 136.000 casos e “quebrar o feitiço” 123.000 ocorrências, confirmando o alto índice destas combinatórias e, conseqüentemente, justificando sua inclusão nas atividades.

Pelo baixo número de fraseologismos apresentados nesta proposta, os exercícios versam sobre colocações verbais na Atividade 4 (página 173), oferece a oportunidade de o estudante saber detalhes pormenorizados da lenda da mula sem cabeça, enquanto reflete sobre qual colocação utilizar no decorrer das frases acerca da história citada.

Finalizando a última proposta de textos, que versam sobre a solidariedade, elencamos os itens a seguir:

Quadro 17 – Campo semântico das colocações do Tema 05

Tema 5 – Redação sobre a “Solidariedade”		
COLOCAÇÕES PRESENTES NO CRAEF 1	COLOCAÇÕES PRESENTES NO CRAEF 2	COLOCAÇÕES POSSÍVEIS, MAS NÃO CITADAS NO CRAEF
Doação de brinquedos	Guarda temporária da criança	Ajudar o próximo
Tomar um soro	Tomar soro	Tratar dos animais
Dar atestado	Anjo da guarda	Dividir o lanche

Inicialmente, acerca do disposto no CRAEF 1, destacamos aquelas mais representativas do tema, como “doação de brinquedos”, com 11 recorrências, “tomar um soro”, com 4 ocorrências e “dar atestado”, com 1 resultado. Com a averiguação na *website* do Google, temos: “doação de brinquedos” com 326.000 repetições, “tomar um soro” aparece “14.400” vezes e “dar atestado” ocorre 17.700 vezes. Estes itens apontam que a linguagem empregada pelos alunos do centro está dentro da convencionalidade linguística no português do Brasil.

Já, no CRAEF 2, representando a escola da periferia, a combinatória “tomar soro” também ocorre 4 vezes. Neste contexto, destacamos a colocação adjetiva “guarda temporária da criança”, com 1 resultado e “anjo da guarda”, também com 1 repetição. Nas pesquisas oriundas do site de busca, encontramos 1.610 resultados para “guarda temporária da criança”, enquanto “anjo da guarda” ocorreu 450.000 vezes. Dentro do tema da solidariedade, estes itens também foram representativos positivamente, auxiliando por promover textos de qualidade em expressão e criatividade, conforme expresso anteriormente na análise lexical do tópico 3.2.

Dessa forma, finalizamos as sugestões de ampliação do léxico-fraseológico, incorporando o tema solidariedade, como forma de sensibilizar os alunos frente a um tema atual e necessário a ser abordado em sala de aula. Nas pesquisas pelo Google, “ajudar o próximo” obteve 512.000 resultados, enquanto “tratar dos animais” retornou 50.500 vezes e “dividir o lanche” apareceu 16.800 vezes.

Ilustrando o exposto, a atividade número 5 (página 174) solicita ao aluno que troque um bloco de palavras pelas colocações mais convenientes em cada caso.

Após analisar as cinco propostas textuais, sentimos a necessidade de ampliar o rol de possibilidades pedagógicas deste tema, mostrando de que forma o docente pode apresentar os fraseologismos a sua turma em diversos contextos. Seja em um simples jogo da memória, mote da atividade 6 (página 175), ou no desafio de resolver cruzadinhas e caça palavras mostradas por meio das folhas 7 e 8 (páginas 176 e 177), é possível tratar das colocações

verbais, nominais, adjetivas e adverbiais relacionadas ao universo da criança, incluindo datas comemorativas e históricas.

Finalizando esta etapa, disponibilizando algumas sugestões de como trabalhar atividades baseadas em corpora no laboratório de informática. Para tanto, solicitamos conferir as tarefas de número 9 e 10 (páginas 178 e 179), conforme atestam as páginas citadas desta dissertação.

4.2 Atividades baseadas em corpora para o Desenvolvimento da Competência Colocacional dos estudantes de Ensino Fundamental I

De acordo com o exposto anteriormente, as primeiras tarefas são baseadas nas temáticas das redações do CRAEF. Estas utilizam tanto as colocações empregadas pelos estudantes e que compõem o CRAEF 1 e 2 quanto aquelas pertencentes ao léxico das propostas baseadas em corpus, mesclando o grau de dificuldade para atender à necessidade dos alunos, conforme já explicado no subcapítulo anterior:

ATIVIDADE 01 – UGA UGA

ATIVIDADE 02 – PAPO DE GENTE GRANDE

ATIVIDADE 03 – SEGREDOS DA TERRA

ATIVIDADE 04 – SAIBA MAIS SOBRE A LENDA DA MULA SEM CABEÇA

ATIVIDADE 05 – REESCREVENDO A FRASE COM OUTRAS PALAVRAS...

Após cumprir o protocolo das propostas, importantes por contextualizar as colocações possíveis de serem ensinadas dentro dos campos semânticos sugeridos, elaboramos exercícios extras, com outras temáticas. Estes também visam ampliar o vocabulário infantil, dentro da concepção didática argumentada desde o início deste tópico.

Para tanto, acrescentamos ainda a proposta de alguns temas infantis, além de rotinas escolares e datas comemorativas nacionais, como tema que ofereça condições de as crianças refletirem sobre as colocações mais representativas deste universo. Posto que os discentes já estudam estes conteúdos com o professor em sala de aula nas tarefas interdisciplinares, consideramos incluí-las no rol de exercícios apresentados.

Outra justificativa digna de nota é o fato de levar os estudantes a compreenderem que

as colocações estão presentes em seu cotidiano, já que representam a sua identidade e cultura do país, conforme elencados abaixo:

ATIVIDADE 06 – JOGO DA MEMÓRIA PARTE 2 (COLOCAÇÕES ADJETIVAS)

ATIVIDADE 07 – CRUZADINHA DA HIGIENE BUCAL

ATIVIDADE 08 – CAÇA PALAVRAS DA PÁSCOA

Haja vista que destacamos o papel da AMD no escopo deste trabalho, consideramos pertinente adicionar mais dois exercícios voltados à pesquisa de colocações em língua portuguesa. Estes visam a verificação de como as palavras se combinam na articulação da linguagem. Para realizar as investigações, sugerimos a exploração do site Corpus do Português.

Comparando os dados com esse corpus, enfatizamos que estas últimas atividades são as que mais refletem a abordagem da AMD. Contudo, reconhecemos que o uso de recursos como; laboratório de informática com o número de computadores suficiente para a turma de alunos e que possuam conexão com a Internet adequada, ainda seja um dos entraves que possa dificultar a plena realização desta tarefa em escolas públicas do Brasil.

Neste sentido, dentro da realidade escolar de cada unidade de ensino, o professor pode realizar adaptações, como apresentar aos alunos as linhas de concordância projetadas em tela para a visualização da sala toda, ou, em último caso, realizar as pesquisas com corpora em casa e entregar aos alunos os resultados já impressos para as reflexões sobre os padrões apresentados:




ATIVIDADE 09 – ADJETIVANDO AS BRINCADEIRAS!

ATIVIDADE 10 – APRENDENDO MAIS SOBRE OS CÔMODOS DA CASA

Baseada na proposta 1 (CRAEF 2) – Campo lexical do tema ANIMAIS trazido pelos alunos e introdução às colocações nominais

ATIVIDADE 01 – TAREFA ANIMAL!

Cada figura traz uma característica bastante marcante destes animais. Observe as figuras com atenção para descobrir o segredo e aproveite para escrever cada colocação ao lado da imagem que o representa!

		COMPRIDO
		REI
		PESCOÇO
		MUITO LISTRADA
		DA SELVA
		ZEBRA

As próximas dicas completarão as frases sobre o reino animal! Não deixe o bicho te pegar, hein!!

HABITAT	ANIMAIS	EM EXTINÇÃO	NATURAL
----------------	----------------	--------------------	----------------

ANIMAIS DOMÉSTICOS E SELVAGENS

Os animais domésticos são aqueles que podem conviver pacificamente entre os seres humanos, como cachorros e gatos, por exemplo. No entanto, há outro grupo de bichinhos que são selvagens, pois possuem hábitos alimentares e comportamento agitado que podem ser prejudiciais aos demais animais ou até aos homens. O _____ destes seres vivos costuma ser a floresta, matas e outros ambientes. Infelizmente, devido a fatores como o crescimento das cidades, caça e queimadas, alguns correm o risco de não existir mais, e podem se tornar _____, como a ararajuba e o cervo-do-Pantanal.

GABARITO:
 PESCOÇO COMPRIDO. ZEBRA MUITO LISTRADA. REI DA SELVA
 HABITAT NATURAL / ANIMAIS EM EXTINÇÃO

Baseada na 2ª proposta de redações (CRAEF), sobre a cidade de Votuporanga
APRESENTAÇÃO DE NOVAS COLOCAÇÕES NOMINAIS

ATIVIDADE 02 – PAPO DE GENTE GRANDE³²

Faça como os repórteres da televisão e entreviste três pessoas que moram na sua rua, pedindo que respondam a opinião sobre os itens do quadro. Para cada linha, faça um X que demonstra o quão satisfeitas estão com os serviços públicos oferecidos pelo município. Depois, traga a folha respondida de volta para a escola, e junto com a professora, montem um gráfico para expor os resultados coletados.

Serviços Oferecidos	Entrevistado 1			Entrevistado 2			Entrevistado 3		
	Satisfeito	Pouco Satisfeito	Insatisfeito	Satisfeito	Pouco Satisfeito	Insatisfeito	Satisfeito	Pouco Satisfeito	Insatisfeito
Água tratada									
Rede de esgoto									
Coleta de lixo									
Iluminação pública									
Postos de saúde									
Escolas públicas									
Praças Públicas									
Lazer e Cultura									
Transporte Coletivo									

SUGESTÃO AO PROFESSOR:

Após a coleta das informações, discutir com os alunos se todos conhecem o significado de cada colocação nominal existente nesta atividade, água tratada, rede de esgoto, coleta de lixo, iluminação pública, postos de saúde, escolas públicas, praças públicas, lazer e cultura e transporte coletivo. Assim, os discentes compreenderão melhor a especificidade destas combinatórias relacionadas ao tema CIDADE.

GABARITO:

VISTO QUE A PROPOSTA É UMA ENTREVISTA, AS RESPOSTAS SERÃO DE CUNHO PESSOAL DE CADA PARTICIPANTE.

³² Atividade retirada do site educativo: Loucos por Tecnologia. Disponível em: <<https://loucosportecnologias.blogspot.com.br/2014/01/zona-rural-e-zona-urbana-meios-de.html>>. Acesso em: 24 jan 2018.

Baseada na 3ª proposta de redações (CRAEF), sobre a Semente Mágica
APRESENTAÇÃO DE COLOCAÇÕES NOMINAIS E ADJETIVAS

ATIVIDADE 03 – SEGREDOS DA TERRA!³³

A seguir, estão os nomes de três tipos de solo. Para descobrir quais são você deve desembaralhar as letras. Para ajudar, abaixo das plaquinhas há informações sobre as características destas formações terrestres. Vamos lá?

OSLO OSONERA

_____ : possui grande quantidade de areia, por isso é muito permeável, pois a água infiltra facilmente pelos espaços formados entre os grãos de areia. Normalmente ele é pobre em nutrientes.

OSLO CARIOLÁC

_____ : com pouco nutriente e grande quantidade de partículas rochosas em sua composição, é inadequado para o cultivo de plantas. É típico de regiões desérticas.

OSLO IGOSRAOL

_____ : é formado por grãos pequenos e compactos, sendo impermeável e apresentando grande quantidade de nutrientes. É bom para a prática da atividade agrícola.

Há vários procedimentos que favorecem o cultivo da terra. Relacione as colocações abaixo com a descrição de cada um e conheça mais sobre eles.

1. CURVAS DE NÍVEL

a) Plantar alternadamente, diversas culturas no mesmo solo, com ao menos uma leguminosa: feijão ou soja.

**2. ADUBAÇÃO
ORGÂNICA**

b) A adição de matéria orgânica ao solo, como esterco, bagaço de cana, ossos moídos, dentre outros elementos.

**3. ROTAÇÃO DE
CULTURA**

c) Sulcos feitos no terreno onde há passagem de água, a fim de prevenir a erosão.

GABARITO:
SOLO ARENOSO. SOLO CALCÁRIO. SOLO ARGILOSO.
1-C / 2-B / 3-A

³³ As informações acerca de solo presentes nesta atividade foram retiradas da palestra “O solo, mais do que simplesmente a camada superficial da Terra”, disponível em: <<http://slideplayer.com.br/slide/10155758/>>. Acesso em: 29 jan 2018.

Baseada na 4ª proposta de redações (CRAEF), sobre a Mula sem cabeça
COLOCAÇÕES DIVERSAS

ATIVIDADE 04 – SAIBA MAIS SOBRE A LENDA DA MULA SEM CABEÇA

Você conhece bem a lenda da mula sem cabeça? Complete as frases abaixo com a opção entre parênteses mais apropriada para cada caso e fique sabendo mais detalhes sobre essa história sobrenatural...

1. Existem muitas versões da lenda da _____,
(*múmia-sem-cabeça / mula-sem-cabeça*)

que _____ da mulher que se apaixona por um padre.
(*conta a história / retrata a historieta*)

2. Como punição por tentar viver um _____, ela se transforma
(*impossível amor / amor impossível*)

na mula-sem-cabeça, cuja sina é _____ pelo nariz e boca.
(*soltar fogo / acender fogo*)

3. Nas noites de _____, ela assusta os moradores dos sítios e das fazendas,
(*lua minguante / lua cheia*)

_____ até o dia amanhecer.
(*gritando muito grande / relinchando bem alto*)

4. Para _____ da mula-sem-cabeça é preciso ter
(*abrir o encanto / quebrar o feitiço*)

coragem! Pois a mula só volta a ser uma mulher novamente quando alguém tira o

_____ do pescoço dela.
(*freio de ferro / pedal do freio*)

----- SUGESTÃO AO PROFESSOR:

Após a realização da atividade, discutir com a turma o porquê realizamos determinadas escolhas lexicais em detrimento a outras, baseados em critérios como gramaticalidade, ordem e combinabilidade das colocações, relacionadas a convencionalidade semântica. Aqui cabe novamente uma pesquisa comparativa entre as duas variáveis apresentadas, afim de ilustrar com mais propriedade esta importante ocorrência da língua portuguesa do Brasil.

GABARITO:

1. MULA-SEM-CABEÇA / CONTA A HISTÓRIA
2. AMOR IMPOSSÍVEL / SOLTAR FOGO
3. LUA CHEIA / RELINCHANDO BEM ALTO.
4. QUEBRAR O FEITIÇO / FREIO DE FERRO.

Baseada na 5ª proposta de redações (CRAEF), sobre a Solidariedade
COLOCAÇÕES VERBAIS, NOMINAIS E ADJETIVAS

ATIVIDADE 05 – REESCREVENDO A FRASE COM OUTRAS PALAVRAS...

Reescreva as frases a seguir, trocando a expressão em negrito por uma das colocações a seguir:

DOAÇÃO DE BRINQUEDOS – A GUARDA TEMPORÁRIA DA CRIANÇA – ANJO DA GUARDA – DAR ATESTADO

1. Vou faltar da escola para ir ao consultório da oftalmologista hoje. Minha mãe vai pedir para a médica **escrever um papel** comprovando que estive lá no horário da tarde.

2. O professor percebeu que ontem eu estava chateado e veio conversar comigo. Depois falou pro meu pai que a nossa família tem que ser mais unida e me dar mais atenção em casa. Ele foi um verdadeiro **amigo quando precisei de ajuda**.

3. A diretora da escola falou pra gente que por causa da crise econômica, tem muita criança que vai passar o natal sem presente. Por isso, ela vai fazer uma campanha de **juntar todos os brinquedos que os alunos não usam mais e dar** para as crianças mais pobres.

4. Quando um menor de idade é maltratado pelos pais, ele não precisa sofrer pra sempre. Se encontrar uma família que tenha condições financeiras pra cuidar dele, estes podem pedir para o juiz **trocá-lo de casa, pelo menos por um tempo**, até a justiça decidir com quem ela deve ficar.

GABARITO:

1. DAR ATESTADO.
2. ANJO DA GUARDA.
3. DOAÇÃO DE BRINQUEDOS.
4. A GUARDA TEMPORÁRIA DA CRIANÇA.

Exercícios extras, utilizando temas do vocabulário infantil
COLOCAÇÕES ADJETIVAS

ATIVIDADE 06 – JOGO DA MEMÓRIA: PARTE 2 (COLOCAÇÕES ADJETIVAS)

Após ver onde está cada carta, identifique as colocações adjetivas e sua imagem adjacente.
 Boa diversão!

<p>PESSOA DOENTE</p>		<p>CAFÉ PRETO</p>	
	<p>SINAL VERDE</p>		<p>CHICLETE TUTTI FRUTTI</p>
<p>CEBOLA ROXA</p>		<p>CABELO RUIVO</p>	

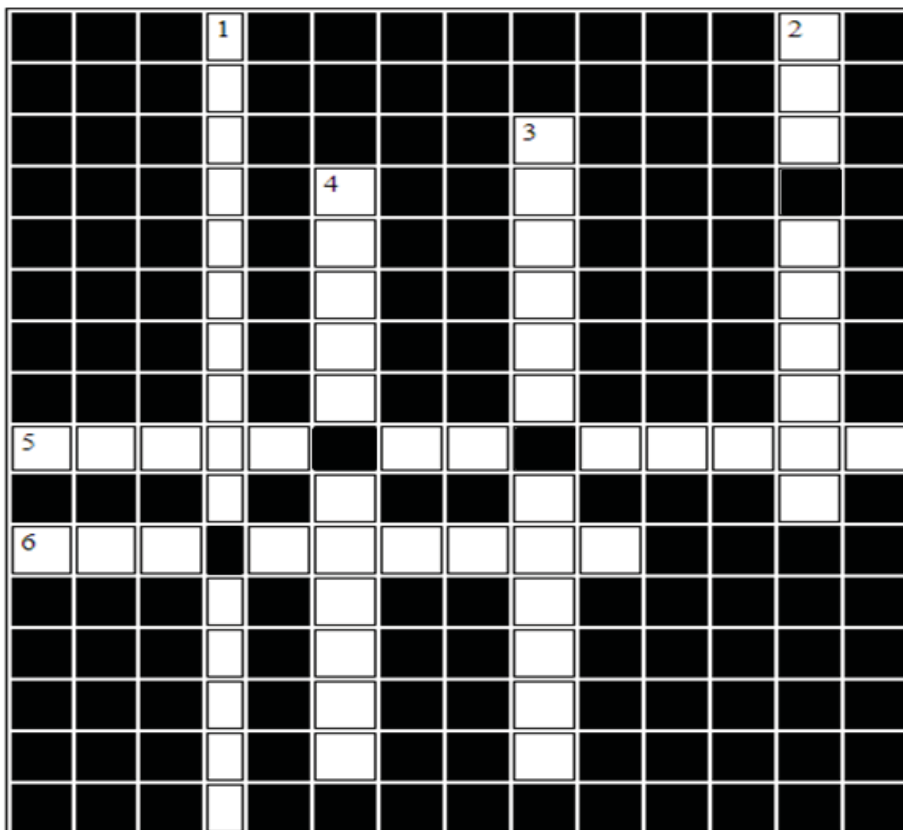
SUGESTÃO AO PROFESSOR:

Após a atividade lúdica, o docente pode pedir aos alunos para escrever frases sobre as colocações presentes no jogo, auxiliando-o a compreender em quais contextos elas são enunciadas, como “sinal verde”, que se relaciona a ideia de consentimento. Outra possibilidade é explorar com o grupo algumas noções de prosódia semântica, vistas na carta “dor nas costas”, haja vista que “doente” sempre vai ser empregado como algo negativo.

GABARITO:
 ATIVIDADE LÚDICA.

ATIVIDADE 07 – CRUZADINHA DA HIGIENE BUCAL

Leia as dicas e resolva este passatempo super divertido com colocações!



Horizontal

5. OUTRO NOME PARA CREME DENTAL.
6. MAIS UM ITEM USADO NA HIGIENE BUCAL, PARECIDO COM LINHA DE COSTURA.

Vertical

1. PRODUTO USADO PARA FAZER BOCHECHO.
2. UM PROBLEMA DE SAÚDE BUCAL CAUSADO PELA FALTA DE HIGIENE.
3. TODA ESCOVA DE DENTE DEVE TER PARA NÃO MACHUCAR A GENGIVA.
4. COLOCAMOS NA ESCOVA DE DENTES PARA FAZER A LIMPEZA DIÁRIA DA BOCA.

GABARITO:

HORIZONTAL:

5. PASTA DE DENTE
6. FIO DENTAL

VERTICAL:

1. ENXAGUANTE BUCAL
2. MAU HÁLITO
3. CERDAS MACIAS
4. CREME DENTAL

Exercícios extras, utilizando temas do cotidiano infantil
COLOCAÇÕES NOMINAIS E ADJETIVAS

ATIVIDADE 08 – CAÇA PALAVRAS DA PÁSCOA

Procure e marque no diagrama de letras 6 colocações nominais e adjetivas:

COELHINHO DA PÁSCOA
 ÚLTIMA CEIA

OVOS DE CHOCOLATE
 SEMANA SANTA

RESSURREIÇÃO DE JESUS
 COLOMBA PASCAL

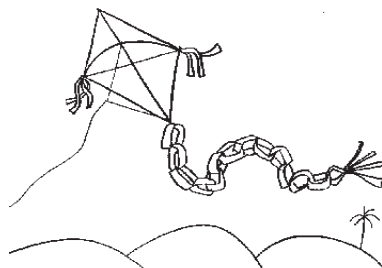


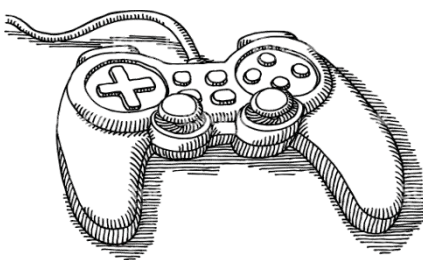
Exercícios baseados em AMD (Aprendizagem Movida por Dados)
COLOCAÇÕES ADJETIVAS

ATIVIDADE 09 – ADJETIVANDO AS BRINCADEIRAS!

Na atividade 03, você relacionou os verbos com cada brinquedo. Agora o jogo mudou!! Acesse o site Corpus do Português³⁴ para pesquisar e escrever nos espaços quais os adjetivos que mais ocorrem mais frequentemente com as brincadeiras a seguir:







GABARITO:
CONSERVADA, ELÉTRICA, LEVE
CORTANTE, COLORIDA, PERDIDA.
NOVO, PORTÁTIL, VIOLENTO (HÁ OUTRAS POSSIBILIDADES)

³⁴ Disponível em <<https://www.wordandphrase.info/port/>>, acesso em 1. fev. 2018.

Exercícios baseados em AMD (Aprendizagem Movidada por Dados)
COLOCAÇÕES ADJETIVAS

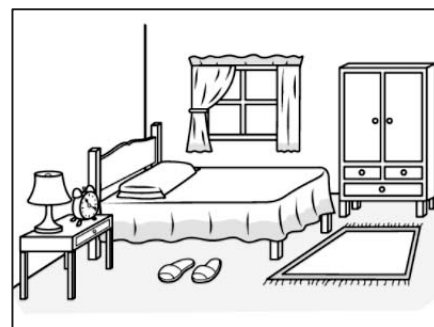
ATIVIDADE 10 – APRENDENDO MAIS SOBRE OS CÔMODOS DA CASA

Já que estávamos falando agora a pouco sobre ajudar a mamãe em casa, aproveite para verificar quais os adjetivos que mais combinam com cada ambiente doméstico. Para ajudá-lo nesta missão, retorne ao site Corpus do Português para verificar as palavras que se harmonizam com os cômodos a seguir:

VEGETARIANA E
MODERNA

DUPLO E
CONFORTÁVEL

FEMININO E
HIGIÊNICO



SUGESTÃO AO PROFESSOR:

Para proporcionar uma aula com um maior aproveitamento por parte da turma, sugerimos que antes de efetivamente encaminhar os alunos ao laboratório de informática, o professor elabore uma aula introdutória sobre o Corpus do Português, a fim de descrever as ferramentas existentes nesta *website*, bem como a maneira de pesquisar as palavras-chave no mesmo.

GABARITO:

BANHEIRO: FEMININO E HIGIÊNICO.

COZINHA: VEGETARIANA E MODERNA.

QUARTO: DUPLO E CONFORTÁVEL

É relevante lembrar que, a princípio, este capítulo não estava previsto na dissertação, mas foi elaborado após a análise das produções textuais compiladas. O exame aprofundado das composições, por meio do *software WordSmith Tools*, nos possibilitou verificar que, tanto no CRAEF 1 quanto no CRAEF 2 não obtivemos muitas colocações empregadas pelos alunos. Por isso, utilizando nossa experiência como professora de séries iniciais, o intuito dessas páginas foi apresentar sugestões para ampliar o léxico-fraseológico das crianças de ensino fundamental, a fim de que pudessem melhorar o entendimento sobre as combinatórias de palavras em língua portuguesa.

Para cumprir este propósito, primeiramente elencamos os objetivos a serem atingidos pelas atividades, seguidos dos percursos metodológicos realizados, incluindo as estratégias pedagógicas empregadas para atrair a atenção das crianças e os critérios didáticos para a escolha das colocações a serem utilizadas. Posteriormente, apresentamos os dez exercícios que compõem a segunda parte deste capítulo, sendo os cinco primeiros oriundos da exploração das colocações relacionadas às cinco propostas das redações coletadas; os próximos três relacionados a exercícios extras, com combinatórias voltadas ao léxico infantil e feriados nacionais e os últimos dois às propostas que visam a pesquisa com o uso da Aprendizagem Movida por Dados, a serem efetuadas, preferencialmente, no laboratório de informática, a fim de que o aluno entre em contato com as ferramentas computacionais de busca e consulta a dados autênticos, baseados em corpora.

Logo, acreditamos que diferentes métodos de tarefas aumentem a probabilidade de uma maior aceitação dos alunos em relação ao conteúdo colocacional. Por isso, os exercícios são variados, seja por meio de tarefas mais acessíveis de cunho introdutório à temática das colocações, seja pelas atividades mais difíceis, que são mais desafiadoras para o grupo de estudantes.

Após identificar e compreender as propostas mais fáceis acreditamos ser 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 oportunidade ao aluno para ampliar as habilidades colocacionais da língua portuguesa do Brasil.

Com o final deste último capítulo, o próximo tópico versa sobre as considerações finais da dissertação, fechando as discussões levantadas ao longo de todo o texto. Além de propor a reflexão acerca dos resultados apontados nas análises, dispõe ainda dos encaminhamentos futuros desta investigação, conforme apontam as próximas páginas na sequência.

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-ROSKOWSKI, 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.

ANEXO

Autorização dos pais



Prezados Senhores

Eu, Elaine Cristina Ferreira de Oliveira, RG. 34.384.230-0 e CPF. 217.843.678-18; estou elaborando o trabalho acadêmico na Universidade Estadual Paulista (UNESP) como tese de mestrado em Estudos Linguísticos.

O trabalho constará de REDAÇÕES DOS ALUNOS que servirão de fonte de pesquisa para estudos na área ensino de PORTUGUÊS.

Um dos principais objetivos das redações é constituir-se como uma fonte de pesquisa para compreender a aprendizagem do aluno e melhorar o ensino em nosso país.

Sua colaboração é importante para este projeto. Se concordarem com a utilização da redação de seu filho (sua filha) no trabalho acima referido, peço a gentileza de assinar esta autorização na linha abaixo.

Caso desejarem mais informações, coloco-me a sua inteira disposição para quaisquer esclarecimentos que se façam necessários.

Este trabalho não possui fins lucrativos e não constará o nome do aluno.

Votuporanga, ____ / ____ / _____

Assinatura do Responsável

Nome do aluno (da aluna)



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

Prezados Senhores

Eu, Elaine Cristina Ferreira de Oliveira, RG. 34.384.230-0 e CPF. 217.843.678-18; estou elaborando o trabalho acadêmico na Universidade Estadual Paulista (UNESP) como tese de mestrado em Estudos Linguísticos.

O trabalho constará de REDAÇÕES DOS ALUNOS que servirão de fonte de pesquisa para estudos na área ensino de PORTUGUÊS.


Um dos principais objetivos das redações é constituir-se como uma fonte de pesquisa para compreender a aprendizagem do aluno e melhorar o ensino em nosso país.

Sua colaboração é importante para este projeto. Se concordarem com a utilização da redação de seu filho (sua filha) no trabalho acima referido, peço a gentileza de assinar esta autorização na linha abaixo.

Caso desejarem mais informações, coloco-me a sua inteira disposição para quaisquer esclarecimentos que se façam necessários.

Este trabalho não possui fins lucrativos e não constará o nome do aluno.

Votuporanga, 13/09/2015


Assinatura do Responsável

Victor William de Souza Maurício
Nome do aluno (da aluna)



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

Prezados Senhores

Eu, Elaine Cristina Ferreira de Oliveira, RG. 34.384.230-0 e CPF. 217.843.678-18; estou elaborando o trabalho acadêmico na Universidade Estadual Paulista (UNESP) como tese de mestrado em Estudos Linguísticos.

O trabalho constará de REDAÇÕES DOS ALUNOS que servirão de fonte de pesquisa para estudos na área ensino de PORTUGUÊS.

Um dos principais objetivos das redações é constituir-se como uma fonte de pesquisa para compreender a aprendizagem do aluno e melhorar o ensino em nosso país.

Sua colaboração é importante para este projeto. Se concordarem com a utilização da redação de seu filho (sua filha) no trabalho acima referido, peço a gentileza de assinar esta autorização na linha abaixo.

Caso desejarem mais informações, coloco-me a sua inteira disposição para quaisquer esclarecimentos que se façam necessários.

Este trabalho não possui fins lucrativos e não constará o nome do aluno.

Votuporanga, 04/11/2015

Glivanete R.F. do Nascimento

Assinatura do Responsável

Mez Steyves Ribério Cavalcini

Nome do aluno (da aluna)