

UNIVERSIDADE ESTADUAL DE GOIÁS
CÂMPUS CORA CORALINA
PROGRAMA DE PÓS-GRADUAÇÃO EM LÍNGUA, LITERATURA E
INTERCULTURALIDADE

Jeane Cardoso Costa

**ANÁLISE DE REDAÇÕES EM LÍNGUA PORTUGUESA A PARTIR DO PERFIL
LEXICAL: CONTRIBUIÇÕES DO PROCESSAMENTO LINGUÍSTICO-
COMPUTACIONAL**

GOIÁS
2021

JEANE CARDOSO COSTA

**ANÁLISE DE REDAÇÕES EM LÍNGUA PORTUGUESA A PARTIR DO PERFIL
LEXICAL: CONTRIBUIÇÕES DO PROCESSAMENTO LINGUÍSTICO-
COMPUTACIONAL**

Dissertação apresentada ao Programa de Pós-Graduação *Stricto Sensu* em Língua, Literatura e Interculturalidade da Universidade Estadual de Goiás/Câmpus Cora Coralina, como requisito parcial para a obtenção do título de Mestre em Língua, Literatura e Interculturalidade.

Linha de Pesquisa: Estudos de Língua e Interculturalidade.

Orientador: Prof. D.r Eduardo Batista da Silva.

GOIÁS
2021

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UEG
com os dados fornecidos pelo(a) autor(a)

CC837a	<p>Cardoso Costa, Jeane A ANÁLISE DE REDAÇÕES EM LÍNGUA PORTUGUESA A PARTIR DO PERFIL LEXICAL: CONTRIBUIÇÕES DO PROCESSAMENTO LINGUÍSTICO-COMPUTACIONAL / Jeane Cardoso Costa; Orientador: Eduardo Batista da Silva. – Goiás, 2021. 177 p.</p> <p>Mestrado – Língua, Literatura e Interculturalidade – Câmpus Cora Coralina, Universidade Estadual de Goiás, 2021.</p> <p>1. Lexicologia. 2. Linguística de Córpus. 3. Redações. 4. Perfil Lexical. 5. Processamento Linguístico-Computacional. I. Batista da Silva, Eduardo, orient. II. Título.</p>
--------	---

DISSERTAÇÃO DO MESTRADO EM LÍNGUA, LITERATURA E
INTERCULTURALIDADE

FOLHA DE APROVAÇÃO

JEANE CARDOSO COSTA

ANÁLISE DE REDAÇÕES EM LÍNGUA PORTUGUESA A PARTIR DO PERFIL
LEXICAL: CONTRIBUIÇÕES DO PROCESSAMENTO LINGUÍSTICO-
COMPUTACIONAL

Trabalho aprovado em 04 de fevereiro de 2021 pela Banca Examinadora constituída
pelos seguintes professores:

BANCA EXAMINADORA

Prof. D.r Eduardo Batista da Silva
Universidade Estadual de Goiás
(Presidente)

Prof^a. Dr.a Darcília M. P. Simões
Universidade Estadual de Goiás/Universidade Estadual do Rio de Janeiro
(Membro interno)

Prof^a. Dr.a Maria José Bocorny Finatto
Universidade Federal do Rio Grande do Sul
(Membro externo)

Dedico essa dissertação a Deus, autor e consumidor da minha fé. Dedico também àqueles que seguraram firme em minhas mãos sempre acreditando que seria possível.

AGRADECIMENTOS

Agradeço a Deus porque até aqui me ajudou. Nos momentos de crises, cansaços e instabilidades, nos quais não sabemos se conseguiremos continuar, Ele chega como uma brisa suave e nos lembra de que em todo tempo está presente e jamais se ausentará.

Agradeço ao meu orientador Eduardo Batista da Silva por tudo o que acrescentou à minha jornada acadêmica. Suas orientações e a nossa caminhada enquanto aluna e professor foram um divisor de águas na minha vida. Carregarei esse crescimento para sempre e sou muito grata por isso.

Agradeço à banca examinadora que de modo tão gentil e competente se dispôs a avaliar este trabalho. Minhas sinceras considerações a cada professor(a).

Agradeço à minha querida mãe, Joane, por cada cuidado materno, por ser o meu esteio, pelo seu amor, disposição e compreensão para comigo sempre. Obrigada por ser a minha mãe.

Agradeço ao meu pai, Luis Carlos, por sempre acreditar em mim, sendo presente, solícito e muito bondoso. Obrigada por ser o meu pai.

Agradeço a minha irmã, Jordane, amiga de todas as horas, conselheira, companheira e cheia de amor. Obrigada por tanto me acolher em todos os momentos me trazendo tanta segurança.

Agradeço aos meus quatro cachorros, Sansão, Sol, Feijão e minha pequena Salomé. Como é bom ter vocês por perto, meus bichinhos.

*Porque dEle e por Ele, e
para Ele, são todas as
coisas; glória, pois, a Ele
eternamente. Amém.
(BÍBLIA, Romanos, 11,
36)*

RESUMO

A escrita de textos depende do conhecimento do vocabulário que será utilizado, bem como de sua variabilidade quanto às faixas de frequência. O presente trabalho trata da avaliação de textos escritos por alunos de Ensino Médio considerando análises quantitativas lexicais e a correção feita por professores de redação e do processamento computacional. Nesse contexto, o objetivo geral deste trabalho é demonstrar a qualidade do vocabulário contido em redações mediante o perfil lexical. Com relação aos objetivos específicos, a pesquisa pretende: 1) constituir um *cópus* de redações; 2) traçar um perfil lexical a partir do processamento do *cópus* e 3) verificar a influência do vocabulário mais formal nas avaliações dos textos escritos. Para embasar a fundamentação teórica, essa pesquisa recorrerá fundamentalmente aos trabalhos desenvolvidos no âmbito da Lexicologia (ANTUNES, 2012; BARBOSA; SOUZA 2016; BIDERMAN, 1999, 2006; GUERRA; ANDRADE, 2012; GENOUVRIER; PEYTARD, 1985; ROCCO, 1981), da Linguística de *Cópus* (BERBER SARDINHA, 2004, 2010, 2012; BIBER; REPPEN, 2015; HALLIDAY, 1993; SINCLAIR, 2004) e da Interculturalidade (BARBOSA, 2009; BARBOSA; SOUZA, 2016). Com relação ao material, foram necessários um *cópus* de redações de Ensino Médio, um formulário eletrônico para avaliação de redações, um *software* de processamento linguístico-estatístico: *Wordsmith Tools*, uma lista de palavras do português, o dicionário eletrônico *Houaiss*, a impressora virtual Doro Writer, uma calculadora estatística *online* e o cálculo amostral. Quanto aos procedimentos, foi realizado um estudo do perfil lexical das redações e à correção dos textos realizada por dez professores de redação e Língua Portuguesa. Por perfil lexical, entendemos a somatória relativa à duas porcentagens complementares: a de vocabulário fundamental e aquela de vocabulário não fundamental (vocabulário mais formal). O vocabulário mais formal variou de 6,42% a 25,29%. Dos 10 professores participantes, 7 foram influenciados pela presença de vocabulário mais formal nas redações, o que indica uma correlação positiva de moderada a forte na avaliação dos corretores. Os resultados evidenciam a necessidade de que os professores de Língua Portuguesa e de redação atentem-se ao ensino de vocabulário tendo em vista o direcionamento das escolhas vocabulares e o aumento de repertório lexical.

ABSTRACT

The writing of texts depends on the knowledge of the vocabulary that will be used, and the writing quality is referred to the ability to use various words in the productions. This work deals with the evaluation of texts written by high school students considering lexical analyses based on human correction and computational processing. Thus, the general objective of this work is to demonstrate the quality of the vocabulary contained in compositions through the lexical-grammatical profile drawn and the comparison between the human and linguistic-computational corrections made by software. With regard to specific objectives, the research aims to: constitute a corpus of high school compositions; propose a methodology for identifying the vocabulary quality of texts; draw a lexical-grammatical profile from the processing of the corpus; propose analyses based on readability metrics; to establish comparisons between the human correction of texts and the linguistic-computational. To support the theoretical foundation, this research will mainly use the works developed within the scope of Lexicology (ANTUNES, 2012; BARBOSA E SOUZA 2016; BIDERMAN, 1999, 2006; GUERRA; ANDRADE, 2012; GENOUVRIER, PEYTARD, 1985; ROCCO, 1981), From Corpus Linguistics (BERBER SARDINHA, 2004, 2010, 2012; BIBER; REPPEN, 2015; HALLIDAY, 1993; SINCLAIR, 2004) and of the Interculturality (BARBOSA,2009; BARBOSA; SOUZA,2016). On relation with the material, it was necessary a essay corpus from high school, a electronic form for the evaluation of the essays, A software that processes static linguistic: Wordsmith tools, a list of Portuguese words, the electronic dictionary Houaiss, a printer virtual Doro writer, a online static calculator and the sample of the calculation, regarding the procedures we proceed with a lexical way study on the essays and the correction of the texts with 10 Portuguese writing teachers. Using the lexical way of study, we understand the relative sum of two percentage complements: the one with a fundamental vocabulary and the one with out a fundamental vocabulary (a more formal vocabulary). The most formal vocabulary has a variation of 6.42% to 25.29%. From the 10 teachers participating, 7 was influenced by the presence of a more formal vocabulary on the essays. That indicates a strong positive correlation on the evaluation on who's correcting it. The results shows that the Portuguese teachers and writing teachers needs to pay attention on the teaching of vocabulary having in view the vocabulary choices and the increased of lexical repertoire.

LISTAS DE ILUSTRAÇÕES

Figura 1 - Sumário presente no manual de correção do Enem.....	31
Figura 2 - Critérios de avaliação de redações da Vunesp	32
Figura 3 - Seção 1 do formulário de correção das redações.....	83
Figura 4 - Seção 2 do formulário com as correções digitadas.....	84
Figura 5 - Ferramenta WordList do software WordSmith Tools	85
Figura 6 - Lista bruta da frequência de palavras na CorPop	86
Figura 7 - Dicionário eletrônico Houaiss – Conjugações	88
Figura 8 - Aplicativo Doro – Impressora Virtual	89
Figura 9 - Calculadora estatística online.....	90
Figura 10 - Cálculo amostral	91
Figura 11 - Lista de palavras deslematizada e ampliada.....	94
Figura 12 - Redação com menor mediana.....	99
Figura 13 - Redação com maior mediana.....	101

LISTAS DE TABELAS

Tabela 1 - Escala de correlação	97
Tabela 2 - Notas atribuídas às redações e suas respectivas medianas	98
Tabela 3 - Correlação entre as notas dos corretores e os types da redação.....	102
Tabela 4 - Perfil lexical das redações.....	104

LISTAS DE ABREVIATURAS E SIGLAS

AVL	Lista de vocabulário acadêmico
BNC	<i>British National Córpus</i>
CEPRIL	Centro de Pesquisa, Recursos e Informação em Linguagem
CN	Córpus NILC
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico.
COCA	<i>Córpus of Contemporary American</i>
COMET	Córpus multilíngue de aprendizagens e <u>córpus</u> de tradução.
CONELA	Congresso Nordeste de Linguística Aplicada
CORPOBRAS	Córpus do Português do Brasil
CORPOP	Córpus do Português Popular Brasileiro
CORPOPAM	Córpus do Português Popular Brasileiro Ampliada
CORPOPEX	Córpus do Português Popular Brasileiro Extra
DDL	<i>Data-Driven Learning</i>
ENEM	Exame Nacional do Ensino Médio
EXTEX	Extração de Textos de Ficheiros Formatados
GSL	<i>New General Service List</i>
GT	Grupo de Trabalho
INEP	Instituto Nacional de Estudos e Pesquisas
LC	Linguística de Córpus
LOB	<i>Lancaster-Oslo/Bergen Córpus</i>
MICASE	<i>Michigan Córpus of Academic Spoken English</i>
NILC	Núcleo Interinstitucional de Linguística Computacional
PDF	<i>Portable Document Format</i>
PLN	Processamento de Língua Natural
SELL	Simpósio Internacional de Estudos Linguísticos e Literários
SIELLI	Simpósio Internacional sobre Língua, Literatura e Interculturalidade
TERMEX	Extração Automática de Termos e Elaboração Colaborativa de Terminologias para Intercâmbio e Difusão de Conhecimento Especializado
TERMNEO	Observatório de Neologismos Científicos e Técnicos do Português Contemporâneo do Brasil
TOEFL	<i>Test of English as a Foreign Language</i>

UFSCAR Universidade Federal de São Carlos
USP Universidade de São Paulo
VUNESP Fundação responsável pelo vestibular da Unesp
WST *WordSmith Tools*

SUMÁRIO

1 INTRODUÇÃO	15
2 FUNDAMENTAÇÃO TEÓRICO-METODOLÓGICA	22
2.1 Lexicologia	22
2.1.1 <i>O ensino do léxico e o uso do vocabulário</i>	<i>25</i>
2.1.2 <i>A atenção dispensada ao vocabulário no Enem e na Vunesp.....</i>	<i>30</i>
2.1.3 <i>Conceitos de Lexicultura e Carga Cultural Compartilhada.....</i>	<i>33</i>
2.2 Linguística de Córpus.....	36
2.2.1 <i>A questão da extensão.....</i>	<i>51</i>
2.2.2 <i>A questão da representatividade</i>	<i>52</i>
2.2.3 <i>Tipos de córpus.....</i>	<i>52</i>
2.2.4 <i>Visão panorâmica de projetos e definições de córpus.....</i>	<i>53</i>
2.2.5 <i>Linguística de Córpus e Ensino.....</i>	<i>59</i>
2.3 Lexicologia e Linguística de Córpus	76
3 MATERIAL E MÉTODO.....	80
3.1 Material.....	81
3.1.1 <i>O córpus de redações.....</i>	<i>81</i>
3.1.2 <i>O formulário de avaliação de redações.....</i>	<i>83</i>
3.1.3 <i>Wordsmith Tools.....</i>	<i>84</i>
3.1.4 <i>A lista CORPOP</i>	<i>86</i>
3.1.5 <i>Dicionário Houaiss.....</i>	<i>87</i>
3.1.6 <i>Aplicativo Doro PDF Writer 2. 15</i>	<i>88</i>
3.1.7 <i>Calculadora estatística online.....</i>	<i>89</i>
3.1.8 <i>Cálculo amostral e coeficiente de correlação de Pearson</i>	<i>90</i>
3.2 Método	91
4 RESULTADOS E ANÁLISE DOS DADOS	98
5 CONSIDERAÇÕES FINAIS	107
REFERÊNCIAS	111
APÊNDICE A – Redações da Vunesp 2020.....	118
APÊNDICE B – Dados quantitativos das redações.....	142
APÊNDICE C – Palavras não presentes na lista ampliada	153

1 INTRODUÇÃO

O contato com estudos lexicológicos iniciados na especialização e desenvolvidos ao longo do curso de mestrado em Língua, Literatura e Interculturalidade no Programa de Pós-graduação da Universidade Estadual de Goiás/Câmpus Cora Coralina foi o ponto de partida principal para a presente pesquisa, que também recorre à Linguística de Córpus¹. A partir de conversas com o professor orientador e devido à minha participação em eventos como o VII Simpósio Internacional de Estudos Linguísticos e Literários (SELL), no Grupo de Trabalho de Estudos de Córpus, realizado pela Universidade Federal do Triângulo Mineiro e o XIII Colóquio de Estudos Lexicais, promovido pelo projeto TermNeo, realizado pela Universidade de São Paulo, surgiu a ideia de desenvolver um trabalho relacionado à avaliação de vocabulário em redações. No ano de 2020, a minha participação no Congresso Nordestino de Linguística Aplicada (Conela) e no Simpósio Internacional sobre Língua, Literatura e Interculturalidade (Sielli), eventos nos quais compartilhei recortes dessa pesquisa, foi importante para aprofundamentos nos estudos lexicológicos e de Linguística de Córpus.

Frente ao exposto, nossa pesquisa investiga o perfil lexical de redações em língua portuguesa escritas por candidatos ao vestibular da Universidade Estadual Paulista (Unesp) no ano de 2020. Os arquivos das redações foram gentilmente cedidos pela Fundação responsável pelo vestibular da Unesp, conhecida como Vunesp.

Por perfil lexical, entendemos a somatória relativa à duas porcentagens complementares: a de vocabulário fundamental (as palavras mais comuns) e aquela de vocabulário não fundamental (que chamaremos, doravante, de vocabulário mais formal).

Trabalhos anteriores já trataram do léxico e práticas de redação. Iniciamos nossa exposição com a pesquisa de Rocco (1981). A autora examinou 1.500 redações de candidatos a vestibulares (1978) obtidas da Fuvest. Seu objetivo era caracterizar a linguagem escrita dos vestibulandos e a existência de uma crise na linguagem escrita, particularmente desses indivíduos. As redações foram escolhidas pela oportunidade de obtenção de um corpus homogêneo. A pesquisa revelou alguns problemas na escrita desses textos, dentre eles, a carência de nexos, de continuidade e quantidade de informações e a ausência de originalidade.

Ferreira e Vieira (2013), por exemplo, desenvolveram uma pesquisa acerca de atividades envolvendo o léxico em sala de aula e os desafios para o ensino de língua materna.

¹ Optamos pela grafia “córpus”, uma vez que se trata de termo paroxítono terminado em “us”, devendo ser acentuado, de acordo com as regras ortográficas da língua portuguesa. Para a forma no plural, adotamos igualmente “córpus”.

As autoras salientam que o estudo do vocabulário/léxico é notadamente voltado para o ensino de língua estrangeira, o que torna escassas as pesquisas sobre o ensino do vocabulário/léxico nas aulas de língua materna. Segundo elas, o trabalho com o léxico demanda critérios que extrapolam a dimensão linguística, uma vez que o uso das palavras constitui em uma atividade discursiva. Desse modo, o estudo do vocabulário/léxico contempla apenas itens lexicais como a palavra, grupo de palavras e expressões fixas. Para as autoras, ficou evidenciado que o uso do dicionário como recurso metodológico exige uma preparação por parte do professor para que a utilização desse material seja realizada de forma profícua.

Finatto, Spagnolo e Bergmann (2018) realizaram uma pesquisa a partir da lista Oxford, que oferece um conjunto básico das palavras mais frequentes na escrita no inglês britânico para todos os tipos de textos. As palavras dessa lista foram selecionadas por especialistas em línguas e ensino de línguas para estrangeiros, das palavras mais frequentes no BNC *British National Corpus*. Ela é usada como referência para os dicionários da editora Oxford para aprendizes de inglês como língua estrangeira e representa um vocabulário básico com as palavras mais úteis e importantes para que o estudante tenha em mente. Esse vocabulário marca a escolha de palavras que fazem parte de uma definição em um verbete de dicionário Oxford, originando uma forma de “vocabulário controlado” com a proposta de ser um conjunto de palavras mais acessível para o entendimento do aprendiz (BNC), um acervo enorme de textos em inglês que está disponível na internet. Pelo fato de não terem encontrado uma lista semelhante para o português brasileiro (PB), as pesquisadoras realizaram um experimento inicial e exploratório na tentativa de identificar um conjunto de palavras correspondentes a um vocabulário básico do PB, empregado na modalidade escrita de modo mais frequente.

Existem alguns desafios na prática de ensino de Língua Portuguesa no que tange às atividades de leitura, interpretação textual e produção textual. Dentre esses desafios, encontra-se a missão de despertar nos alunos o gosto pela leitura e envolvê-los em atividades de prática de escrita que ativem o reconhecimento de vocabulário direta ou indiretamente. Pensando nisso, Silva e Silva (2020) realizaram uma pesquisa acerca do repertório vocabular limitado, ou seja, o reduzido conhecimento lexical dos estudantes, especialmente nas primeiras séries da segunda fase do Ensino Fundamental. Para o desenvolvimento da pesquisa supracitada os autores recorreram aos contos literários em língua portuguesa por possuírem várias palavras que se repetem e que podem ser usadas como estratégia pelo professor na aula de Língua Portuguesa para a aquisição de vocabulário. A pesquisa preocupou-se com a recepção do texto escrito em língua portuguesa e objetivou ampliar o vocabulário de alunos de uma escola pública da rede estadual de Itumbiara, na região sul-goiana, por meio da leitura de trechos de contos literários

e da aplicação de testes (questionários) aos alunos. Os resultados da pesquisa evidenciaram que de modo geral o nível de conhecimento vocabular dos alunos, com relação às palavras do teste, era baixo. Confirmaram que houve aumento lexical de modo geral no pós-teste.

Souza e Xavier (2019) também nos trazem uma relevante investigação acerca da Lexicologia e da Língua Aplicada e suas implicações no ensino de língua materna. As autoras buscaram traçar inter-relações entre a Língua Aplicada (LA) e a proposta de ensino do léxico, objetivando mostrar a importância dele em todos os níveis de ensino. As autoras afirmaram que apesar do léxico possuir um papel relevante no ensino de língua, nem sempre as discussões a ele atinentes têm chegado até os alunos dos mais variados níveis educacionais. Buscaram investigar por revisão bibliográfica as dificuldades encontradas nas abordagens dos conteúdos lexicais, como por exemplo, a falta de formação adequada dos professores de Língua Portuguesa para tal tarefa. Analisaram a grade curricular dos cursos de Letras oferecidos pela Universidade Federal de Goiás da cidade de Catalão/2018, a fim de conhecerem as disciplinas na área do léxico que estavam sendo ofertadas, bem como a natureza delas, se eram obrigatórias ou optativas. A pretensão era entender se realmente os professores de Língua Portuguesa concluem a graduação estando preparados a ensinar sobre o léxico em suas aulas. As pesquisadoras concluíram que a necessidade de inserção dos estudos lexicais na formação de professores de Língua Portuguesa é urgente e que a Língua Aplicada surge como auxílio para um olhar teórico e metodológico sobre a necessidade de estudos de se difundir estudos lexicais nas instituições em todos os níveis de educação. As autoras averiguaram que a maioria dos problemas apontados nas pesquisas analisadas esbarraram na formação do professor no que diz respeito às disciplinas sobre léxico ofertadas no período de formação superior.

Costa e Silva (2020) desenvolveram uma investigação sobre a riqueza lexical de redações escritas em língua portuguesa a partir da razão *type-token ratio* (TTR): a razão forma/item (ou vocábulo/ocorrência) expressa em porcentagem. De acordo com os resultados, verificou-se que o aumento de insumo lexical que o aprendiz tem pode impactar em seu desempenho em uma redação, por exemplo – a depender da exposição e emprego consciente do léxico.

Numa perspectiva lexicológica, apesar de também tratar da díade léxico-redação, nossa pesquisa difere dos trabalhos anteriores por três motivos principais: 1) Recorre a uma lista de palavras frequentes na língua portuguesa, obtida a partir de um corpus; 2) Processa uma amostra de redações de um grande vestibular brasileiro; 3) Preocupa-se com o estudo do vocabulário fundamental e do vocabulário mais formal, destacando a relevância do uso consciente do léxico

para a escrita de textos – porquanto, em uma redação, para além de questões ortográficas, gramaticais e discursivas, o repertório lexical deve ser considerado e 4) Discute o impacto do perfil lexical na avaliação das redações.

Assim, argumentamos a favor de um ensino de práticas de redação em que se promova o uso consciente de vocabulário, de modo que seja refletido na escrita de textos, culminando eventualmente na ampliação do repertório lexical por parte do alunado e consequente êxito no que se refere à qualidade do texto.

É útil ressaltar que o texto (a redação), não deve ser vista como um produto de consumo pronto e acabado, mas como uma prática social realizada por sujeitos em um contexto comunicativo, discursivo. O texto, assim, está em constante transformação, é sempre inconcluso, marcado pela incompletude. Considerando isso, o ensino de redação na escola pressupõe a escrita como processo/trabalho.

Pensando, pois, no trabalho com a linguagem que o sujeito realiza em suas interações, no contexto da sala de aula essa concepção contribui para que o sujeito aprendiz assuma um papel mais protagonista na ampliação da discussão e na introdução de outras falas, outras vozes. Isso abre espaço para que sentidos e aprendizados sejam construídos de forma mais autônoma por esses sujeitos. Na verdade, tanto quem fala/escreve como aquele para quem se fala/escreve são vistos como sujeitos ativos que dialogicamente constroem sentidos (BAKHTIN, 2006) do texto, visto assim como um produto inacabado que exige de ambos diferentes conhecimentos e escolhas (linguísticas, textuais, discursivas e conceituais etc).

Procuramos trazer contribuições aos estudos lexicais e de Linguística de Córpus a partir da percepção de um fenômeno importante e atual. Uma vez que não são muitos os trabalhos acerca desse tema, uma dissertação de mestrado que explore essa temática tende a ser útil a pesquisadores, estudantes e professores que desejem compreender ou aprofundar conhecimentos nessas áreas.

Partindo, ainda, do pressuposto de que nas correções de redações, além dos quesitos tradicionais como a adequação ao tema e gênero, a capacidade de argumentação, a coesão e coerência, uso de conectivos, surge a necessidade de explorar a qualidade do léxico nessas correções, há uma lacuna a ser preenchida. Em outras palavras, é preciso que maior atenção seja dada ao nível do vocabulário dos alunos ao empregar a língua escrita nos contextos de produção.

Outra questão preponderante na produção escrita dos alunos são os usos vocabulares como reflexos de mudanças sociais e culturais que ocorrem; isso se torna possível por ser o léxico um sistema aberto que pode ser expandido e alterado. Por conta disso, as modificações

multiculturais que influenciam diretamente no valor lexical não podem de modo algum serem ignoradas no ensino de língua.

Pelos aspectos supracitados, recorreremos à Lexicultura por pensarmos que o maior patrimônio de um povo é a herança linguística. Novos sentidos vão surgindo na língua de acordo com as diferentes mudanças, acontecimentos, e diversidades culturais, o que resulta, algumas vezes, no desuso de determinados itens lexicais. Inversamente, alguns termos podem voltar ao uso, porém com conotações distintas. Vê-se, assim, que cada palavra demonstra uma realidade, um valor específico; o léxico torna-se enriquecido por novos vocábulos ou novas significações.

É importante ressaltar esforços feitos no sentido de contribuir com essa ampliação lexical e com o uso de vocabulário. Isso tem sido feito por meio de ferramentas interativas de divulgação que podem ser acessadas por grande parte dos alunos. No âmbito federal, o Instituto Nacional de Estudos e Pesquisas (INEP), por exemplo, em sua página *online* oficial tem compartilhado ideias acerca das descobertas vocabulares por parte dos alunos, chamando a atenção à necessidade de ampliação do insumo lexical para a escrita e interpretação de textos.

Nosso estudo, que envolve a correção realizada por professores de redação e o processamento linguístico-computacional, continua uma tradição, uma vez que a comunicação escrita é uma habilidade necessária entre os alunos. Dado o fato de que o ensino é um empreendimento complexo com variáveis distintas, as contribuições decorrentes das análises realizadas aqui fornecerão uma perspectiva mais ampla sobre a formação lexical que deve estar presente no ensino. Não possuímos a preocupação de caracterizar erros, tampouco destacar deficiências, mas, sim, angariar esforços e apresentar aos profissionais que lidam com produção textual possibilidades pautadas no processamento linguístico-computacional, considerando aspectos lexicais presentes nos textos. Seguimos a orientação de que a Lexicologia configura-se como um rico campo a ser estudado, à medida que trata do léxico por excelência. Enquanto área do conhecimento que trata diretamente da língua, pode contribuir para a melhoria da qualidade do ensino de língua.

Frente ao exposto, delinhamos as seguintes questões de pesquisa: 1) Qual o perfil lexical das redações? e 2) Qual o vocabulário mais formal empregado?

Na tentativa de tratar dos problemas mencionados, o objetivo geral desse trabalho é demonstrar a qualidade do vocabulário contido em redações mediante o perfil lexical. Com relação aos objetivos específicos, a pesquisa pretende: 1) constituir um *corpus* de redações; 2) traçar um perfil lexical a partir do processamento do *corpus* e 3) verificar a influência do vocabulário mais formal nas avaliações dos textos escritos.

Lançando luz sobre as características para a avaliação de textos, que por sua vez, englobam quesitos específicos, como por exemplo, o respeito às características do gênero escolhido e aos padrões da linguagem escrita, o uso consciente da escrita alfabética, a interferência ou não da oralidade na escrita, a ausência de troca de fonemas, o domínio das regras básicas de concordância nominal e verbal da língua, a segmentação e disposição textual, dentre outros; considerando tais aspectos, buscamos investigar numa perspectiva lexical o grau de importância de um variado repertório de palavras que ao ser utilizado no texto interfere positiva ou negativamente em sua avaliação. Frente a isto, pudemos verificar que as evidências coletadas demonstram que o vocabulário empregado na escrita de textos exerce significativa influência sobre as correções dos professores (ver seção de Resultados).

No que diz respeito à estrutura deste trabalho, foi organizado em partes independentes, porém, sequenciais. No primeiro capítulo, as linhas teóricas que fundamentarão as discussões deste trabalho serão apresentadas. A fundamentação teórico-metodológica será embasada fundamentalmente nos trabalhos desenvolvidos no âmbito da Lexicologia (ANTUNES, 2012; BARBOSA; SOUZA, 2016; BIDERMAN, 1999, 2006; GUERRA; ANDRADE, 2012; GENOUVRIER; PEYTARD, 1985; ROCCO, 1981) e da Linguística de Córpus (BERBER SARDINHA, 2004, 2010, 2012; BIBER; REPPEN, 2015; HALLIDAY, 1993; SINCLAIR, 2004).

No segundo capítulo, os materiais e os procedimentos metodológicos serão descritos. Serão caracterizados o córpus de redações que foi utilizado, bem como, o formulário de avaliação enviado aos professores. A ferramenta linguístico-estatística utilizada para o processamento dos dados e análises foi o *software* WordSmith Tools, versão 8.0. Os demais materiais compreendem o córpus de redações, o formulário de avaliação, a lista de palavras ampliada no *software*, o perfil lexical (objeto de pesquisa) que foi traçado, o dicionário eletrônico Houaiss, a impressora virtual Doro PDF Writer, o cálculo amostral, o coeficiente de Pearson e a calculadora estatística online. De modo a subsidiar a análise dos dados quantitativos obtidos pelo processamento linguístico-computacional, contamos com a participação voluntária de professores de redação atuantes em diferentes cidades brasileiras. Ainda no segundo capítulo, delinearemos os procedimentos metodológicos. A coleta de dados, a conversão para o formato texto simples, o processamento no software, a tabulação parcial dos dados, a correção dos professores de redação e a tabulação final dos dados serão, do mesmo modo, aprofundados no decorrer do capítulo.

Os capítulos seguintes trarão os resultados e a análise dos dados obtidos acerca do vocabulário utilizado nas redações, do perfil lexical, do modo como o corpus foi distribuído e processado e de como se deu a avaliação dos textos.

No capítulo cinco, que trata das considerações finais, retomamos os capítulos anteriores e o cerne de cada um, recapitulando a sequência teórica delineada no decorrer do trabalho e dando ênfase aos resultados e evidências coletadas a fim de validarmos efetivamente a análise realizada.

Por último, podem ser consultados os apêndices A que nos trazem as redações cedidas pela Vunesp, e os apêndices B que nos revelam os dados quantitativos dessas redações.

2 FUNDAMENTAÇÃO TEÓRICO-METODOLÓGICA

O aporte teórico de nossa investigação recorre a estudos da Lexicologia e da Linguística de Córpus.

Com relação à Lexicologia, iniciamos a presente seção apresentando uma conceituação/contextualização. Na sequência, discutimos o ensino do léxico e o uso do vocabulário e, por fim, conceitos de Lexicultura e de Carga Cultural Compartilhada.

No que se refere à Linguística de Córpus, apesar do pronunciado viés metodológico, optamos por inserir as discussões na Fundamentação Teórica, principalmente, por se tratar de uma nova forma de entender a língua, direcionando questionamentos e motivando reflexões.

2.1 Lexicologia

Barbosa (1992, p. 122) assinala que “todo sistema linguístico contém unidades lexicais, inventário à disposição dos falantes, [...] com regras que permitem aos usuários a criação de novas palavras mais adequadas às suas necessidades de comunicação”. A mesma ideia é complementada por Biderman (2001) quando nos leva a compreender que os processos de nomeação e categorização dos seres, objetos, etc., é que formam o léxico de uma língua e que originam os signos linguísticos. A autora nos esclarece que todo o conhecimento do universo difunde-se no léxico de uma língua, porque cristaliza-se, ou materializa-se nos signos linguísticos, proporcionando, um armazenamento sob a forma de memória.

A língua é um sistema formado por signos, que expressam o conhecimento de tudo que está à nossa volta e que, se pensados em conjunto, dão origem a uma parte importantíssima da língua: o léxico. Nesse sentido, Barbosa (1992) compreende que o léxico funciona como uma forma de estruturação de conhecimentos de mundo, tanto de modo linguístico quanto extralinguístico. Pode ser visto, então, como um espaço no qual ocorre a designação das coisas ou do estado das coisas existentes no mundo.

Para Biderman (2001, p. 169), lexema é unidade léxica abstrata da língua e lexia é a forma que aparece no discurso. Desse modo, podemos entender que um lexema é um paradigma que engloba diferentes flexões: por exemplo, o lexema verbal TER abarca formas concretas como ter, tendo, tido, tenho, tinhas, teve, teremos, tiveram, etc. Todavia, quando a oposição entre as formas ocorre em termos derivacionais, formam-se lexemas distintos, como em análise, analisar, analisador, analisável, etc. Os lexemas costumam ser divididos em duas grandes categorias: de conteúdo (nacionais) e instrumentais (gramaticais) (BIDERMAN, 2001, p. 333).

Os primeiros se caracterizam por ter significação externa representando o mundo extralinguístico; os segundos desempenham a função de estabelecer relações entre os de conteúdo. Teoricamente, a divisão se sustenta, mas, na prática, é um problema complexo, principalmente na classe das preposições, já que algumas podem ser referenciais ou relacionais. Biderman (2001), enfatiza a necessidade de se combinarem critérios de natureza fonológica, morfosintática e semântica para a identificação de unidades lexicais na fala. A mesma autora reconhece a imposição operacional de se analisar corpus extensos a partir de critérios gráficos, assim o fez em sua pesquisa para um dicionário de frequência do português brasileiro contemporâneo (BIDERMAN, 1998).

O léxico é certamente o de maior interesse para quantificação de informação, pensando em todos os níveis de organização da estrutura linguística, dado principalmente o valor referencial dos lexemas nocionais, o que não quer dizer, todavia, que outros níveis de organização da linguagem não veiculem também informação. Por conta da facilidade de coletarem automaticamente suas unidades, o léxico ocupa um lugar privilegiado no que diz respeito à estratégia de quantificação de informação por meio de recursos da informática.

O objeto de estudo principal da Lexicologia é a palavra, o alicerce da língua e se dá por meio da linguagem. Estuda ainda a categorização lexical e as estruturas do léxico. A noção de palavra é uma preocupação da Lexicologia e nos direciona a pensar em estratégias para ensinar o léxico. Por mais que esse conceito tenha sido renegado, com o passar dos anos percebeu-se que a noção de palavra não poderia ser abandonada haja vista que constitui também elemento fundante para as demais noções.

Com relação à palavra entendida, portanto, como possuidora de padrões de regularidade variáveis, Basílio (1987) enfatiza que é possível que tenhamos em uma mesma palavra diferentes níveis estruturais, isto é, a sequência dos elementos que a constitui não é inalterada. A palavra pode ser reconhecida por seus falantes com facilidade, no entanto, sua definição é complexa. Ressaltemos aqui que na língua formas são combinadas, estas podem ser aceitas ou não e isto ao estudo do léxico constitui um desafio.

O estudo das relações e das estruturas do léxico é um domínio privilegiado para descobrir as interações entre o sistema formal da língua e a atividade humana que a torna possível, a linguagem. Esta atividade é exercida no tempo, no espaço e na sociedade pelo fenômeno geral que é a comunicação. (REY, 1970, p. 149, tradução nossa).²

² *l'étude des relations et des structures du lexique est un domaine privilégié pour découvrir les interactions entre le système formel de la langue et l'activité humaine qui la rend possible, le langage. Cette activité s'exerce dans le temps, l'espace et la société, par le phénomène général qu'est la communication.* (REY, 1970, p. 149).

De acordo com Biderman (1999, p. 90), a definição de palavra, bem como sua identificação no nível do discurso, tem que ser feita língua a língua. Assim sendo, do ponto de vista linguístico, é natural que cada língua possua suas particularidades, que a diferenciam das demais.

Em vista disso, o léxico constitui todo o estoque vocabular de um grupo sociolinguístico e cultural; é nele que as informações sobre o mundo são condensadas e depositadas, a partir de nomenclaturas e interpretações distintas da realidade.

É por meio da palavra que determinadas entidades podem ser identificadas. A designação e a nomeação dessas realidades estabelecem um significativo universo em meio à linguagem, dentro da qual Biderman (2006) chama de “processo de categorização”. Nesse contexto, o léxico exerce um papel imprescindível na propagação e transmissão do significado, objeto da comunicação.

O léxico é a janela da língua para o mundo e reflete a história cultural e social do povo que dela se serve como instrumento de comunicação e expressão. Tais assertivas constituem um campo de pesquisa profícuo concernente aos estudos linguísticos, às análises estatísticas e as suas aplicações. (BECHARA, 2011, p. 9).

O léxico pode ser subentendido como um universo aberto de palavras em constante expansão. Por se tratar de um organismo vivo, cabe ao falante, usá-lo em momento oportuno, o que constitui o seu léxico individual.

No leque das Ciências do Léxico, podemos encontrar a Lexicologia, uma ciência que estuda o vocabulário de uma língua. Para Krieger e Finatto (2004), pode ser definida como o estudo científico do léxico, isso é, das palavras disponíveis aos usuários de uma determinada língua. A Lexicologia, como área de conhecimento direcionada especificamente ao uso da língua, pode fornecer subsídios significativos ao ensino de língua portuguesa, pelo fato de encontrar-se no entrecruzamento dos campos fonético-fonológico, semântico, morfológico, sintático e pragmático.

A Lexicologia é o ramo da linguística que se ocupa do léxico de uma língua. Esse léxico pode ser entendido como um conjunto de palavras de uma determinada língua. O termo palavra é utilizado com várias acepções, considerando isso, de modo mais preciso, podemos utilizar os termos *lexema* e *lexia*.

A Lexicologia de modo tradicional pode se ocupar da formação das palavras. Os lexicólogos têm se preocupado também com as questões do neologismo e da criação lexical. Muitas são as pesquisas léxico-estatísticas cujos resultados são aplicados no ensino do vocabulário, bem como na busca de famílias linguísticas e suas origens. Assim sendo, o objeto

primordial dos estudos lexicológicos é o exame do léxico, o que nos permite conhecer a história e os costumes de um povo.

Como disciplina, a Lexicologia ocupa-se do estudo do vocabulário das línguas sob aspectos morfológicos, fonéticos, fonológicos, sintáticos, etc., abrangendo a observação e análise do repertório lexical.

2.1.1 O ensino do léxico e o uso do vocabulário

Sabemos que a inter-relação entre os seres humanos se dá de modo contínuo e permanente. Para Perini (1995, p. 51), “aprender uma língua não é apenas aprender suas regras, mas ainda memorizar uma grande parte de seu léxico”. As palavras ao serem materializadas em um texto, oral ou escrito, constituem um vocabulário próprio, escolhido pelo indivíduo. Desse modo, consideramos que o ensino do léxico seja uma tarefa de grande eficácia em sala de aula, no sentido de orientar os alunos na seleção e variabilidade dos itens lexicais que usarão em suas produções.

Nesse sentido, apresentamos de modo breve alguns trabalhos realizados por Rocco (1981) e Isquerdo (2003) que trataram da dificuldade na escrita de textos por parte de alunos de Ensino Médio. Essa dificuldade se remete, além de outros fatores, à não contextualização do ensino de léxico nas séries básicas, o que mantém uma crise na linguagem escrita manifestada por meio da qualidade das produções textuais e de um repertório lexical reduzido. Tal crise perdura até as séries mais avançadas.

Rocco (1981, p. 54) nos diz que numa sociedade onde tudo nos é dado como pronto, são tantos clichês percebidos na escrita, escritos sem lógica e coerência, *nonsense* e cheio de contradições. O ensino descontextualizado, ou aquele que não considera os aspectos semânticos e lexicais, pode remeter ao mecanicismo e não fazer sentido para quem aprende. A autora nos traz um estudo realizado sobre 1.500 redações de vestibulandos da FUVEST. Ela examina alguns textos com base nas novas tendências dos estudos da linguagem, que buscam evidenciar uma gramática do texto, uma teoria do texto. Os objetos de seu estudo são a coesão, o clichê, a frase feita, o “não-texto” e o discurso indefinido. O objetivo principal da pesquisa é caracterizar a linguagem escrita dos vestibulandos e a existência de uma crise na linguagem escrita, particularmente desses indivíduos. As redações de vestibular foram escolhidas pela oportunidade da compilação de um *cópus* homogêneo. Sob a hipótese de uma crise na linguagem, a autora buscou estabelecer relações entre os textos e o nível de estruturação mental de seus produtores.

Rocco (1981) identificou diferentes problemas, dentre eles, ressaltam-se a carência de nexos, de continuidade e quantidade de informações, ausência de originalidade. Um dos critérios utilizados para a análise é a utilização do conceito de coesão, a progressão discursiva, o discurso tautológico, as contradições lógicas evidentes, o nonsense. Com o seu trabalho, chegou à conclusão de que 34,8% dos vestibulandos demonstram incapacidade de domínio dos termos relacionais; 16,9% apresentam problemas de contradições lógicas evidentes. No total, 15,2% dos textos são redundantes. Dos textos analisados, 69,0% apresentam o uso excessivo de clichês e frases feitas. Somente em 40 textos, verificou-se a presença de linguagem criativa. Em seus resultados, a autora ressalta que, muitos textos analisados estruturam-se com frases bombásticas, pretensamente de efeito.

No mesmo escopo de pesquisa, Isquierdo (2003) nos traz uma pesquisa léxico-estatística realizada com um *cópus* de 450 redações de vestibular, dos anos de 1999 a 2000, de duas universidades do interior de São Paulo, uma pública e outra privada. Na pesquisa, ela traz resultados estatísticos de uma amostra de vocabulário de universitários iniciantes. Os resultados obtidos assemelham-se aos de outras pesquisas de mesmo cunho, principalmente, no que diz respeito à frequência concentrada em um número reduzido de unidades lexicais, a saber: 11.151 unidades léxicas diferentes totalizaram 113.638 ocorrências do *cópus*.

O confronto desses dados com os do Dicionário de frequências do léxico do português brasileiro contemporâneo, de Biderman (1998), mostrou um pequeno grupo de palavras (cerca de trezentas) comuns nas duas bases de dados e com índices de frequências muito parecidos, isso indica que há um possível núcleo do vocabulário analisado que reúne palavras possivelmente utilizadas em qualquer tipo de texto. Fenômeno destacado nos vinte verbos mais frequentes no *cópus*. Esses dados foram resultantes de recortes e de cruzamentos sociolinguísticos, levando em consideração variáveis como universidade de ingresso dos estudantes; conclusão do Ensino Médio em escola particular ou pública; sexo e renda familiar. Dois fatores que merecem destaque na análise da pesquisadora, são: a amplitude maior do vocabulário daqueles que ingressam na universidade pública e a suplantação das demais variáveis pela escolaridade.

A partir do exposto, tendo em vista a contextualização de algumas importantes pesquisas realizadas, ao recorrermos à Lexicologia, tornam-se possíveis investigações acerca do ensino do léxico. De acordo com Kleiman (1998), o léxico é o indicador mais seguro de dificuldade do texto e se existem deficiências dos estudantes na compreensão da escrita. A partir dessa ideia, podemos deduzir que a prática de atividades adequadas ao desenvolvimento de um vocabulário ativo não tem sido costumeira no ensino, o que acarreta uma série de problemas,

haja vista que é por meio da linguagem que a herança cultural é transmitida e o veículo por excelência é a própria língua, o léxico.

Ao pensarmos na escrita de textos por parte dos nossos alunos, em especial os de Ensino Médio, algumas preocupações surgem no sentido da qualidade do que escrevem em relação à idade que possuem e à estrutura de pensamento que deveriam ter. É perceptível nos textos dos alunos (queixas de colegas professores) um baixo nível de desempenho linguístico que resulta em uma série de dificuldades em estabelecer articulações lógicas elementares nos textos.

Há que se pensar em um cenário de crise na linguagem, conforme salienta Rocco (1981), especialmente quando se avaliam textos cuja qualidade é altamente contestável. Esse cenário de crises no ensino de língua pode servir como um motor propulsor para buscarmos a superação desse panorama, assim como a melhoria da qualidade do ensino de língua. Pode-se buscar um trabalho prático que amplie a aquisição da linguagem e os seus níveis por meio de um trabalho caracterizado pela originalidade. De acordo com Guerra e Andrade (2012), novas práticas são necessárias para o ensino de línguas. Os saberes teóricos e os da prática têm de ser articulados. Pensar sobre novas táticas para o ensino de línguas implica considerar nosso contexto social e educacional de fracasso escolar.

Partindo de tais pressupostos e considerando que tal crise na linguagem se deva a uma série de fatores que vão desde a gramática do texto às questões do discurso, inclusive a elaboração verbal criativa de modo que permita ao aluno expressar o mundo, limitamos a discussão ao ensino do léxico para a escrita de textos e às possíveis contribuições para a superação dessa crise.

Dominar o vocabulário de uma língua com qualidade traz à tona dificuldades, principalmente na modalidade escrita. O vocabulário pode ser entendido como um complemento da gramática empregado pelo locutor nos atos de fala de acordo com a necessidade. Quanto mais palavras conhecemos, mais enriquecemos o nosso vocabulário e, possivelmente, melhor escreveremos. Todavia, a ampliação dele depende da atualização do léxico individual daquele que fala e escreve.

De acordo com Genouvrier e Peytard (1985, p.35), existe uma noção equivocada de que basta que o aluno domine o maior número de palavras possíveis e ele terá um grande domínio lexical e conseqüentemente uma “boa fala”. Pensando no fato de que a palavra é mais que um sentido, seus significados não se justificam em si mesmos, a palavra não é unívoca, é de estatuto totalmente ambíguo, presente em toda e qualquer situação a partir de seus esquemas relacionais, onde é inserida, ela funciona.

Para Guerra e Andrade (2012), um ensino que não considera os aspectos semânticos, pode não fazer sentido ao aluno, gerando conseqüentemente dificuldades de aprendizagem. Assim, podemos inferir que o léxico é extremamente relevante ao ensino de língua materna e, por conta disso, deve ter significativa relevância nas grades curriculares do ensino de Língua Portuguesa. Quando pensamos no ensino de língua, nos vem à memória um ensino ultrapassado, oriundo de técnicas e ideias positivistas e tradicionais, isolado da realidade dos alunos e dado por metodologias descontextualizadas que acabam remetendo a um mecanicismo que não faz sentido ao aprendiz.

O ensino descontextualizado, ou aquele que não considera os aspectos semânticos, pode remeter ao mecanicismo e não fazer sentido a quem aprende. Acreditamos não ser imprudente considerarmos esses modos de entender o ensino de línguas como pontos relevantes a serem percebidos como parte do sentimento de fracasso escolar referente ao ensino de línguas. (GUERRA; ANDRADE, 2012, p. 9).

Para o professor, de acordo com Guerra e Andrade (2012, p.10), o conhecimento acerca dos documentos que regem o ensino não são obrigatórios, mas são essenciais. Cabe aos docentes o desejo por conhecê-los, considerando que a reflexão na prática é um elemento que constitui o ser docente, o saber ser e o saber-fazer. São documentos importantes para o professor em sua prática pedagógica.

O advento de documentos oficiais, como os Parâmetros Curriculares Nacionais e as Diretrizes Curriculares Nacionais, contribuiu para uma visão transformadora do ensino de língua materna no Brasil, tanto nas concepções teóricas quanto no ideal de uma prática pedagógica. (GUERRA; ANDRADE, 2012, p. 10).

Guerra e Andrade (2012, p. 11) também apontam que, ao pensarmos em um ensino de língua inovador, um tripé deve ser considerado: a concepção de língua, linguagem, palavra. É necessário, portanto, dar mais atenção às áreas do conhecimento que cuidem dessas noções e que tenham como preocupação algum tipo de mudança no quadro de ensino de língua brasileira.

Hoje, novas práticas são necessárias para o ensino de línguas. Os saberes teóricos e os da prática têm de ser articulados. Pensar sobre novas táticas para o ensino de línguas implica considerar nosso contexto social e educacional de fracasso escolar, evidenciado quantitativamente, por meio do SAEB e da prova Brasil. (GUERRA; ANDRADE, 2012, p. 11).

Guerra e Andrade (2012, p.12) acrescentam que o vocabulário visto como um conjunto de palavras usadas pelo locutor no ato de fala preciso é uma atualização de um certo número de palavras pertencentes ao léxico individual. Na prática pedagógica, a proximidade entre léxico e vocabulário deve ser considerada. O vocabulário constitui o léxico individual e este está

inserido no léxico de forma ampla. O vocabulário global, que não deixa de ser léxico, compreende um conjunto de palavras vigentes num dado momento sócio histórico, o meio social tem extrema relevância na escolha de palavras que o aluno faz no uso efetivo da língua. O cotidiano e a realidade sociocultural possibilitam o armazenamento de novas palavras em seu acervo lexical individual, o que perpassa a vida. É uma aprendizagem que jamais cessa, é durante a vida toda.

A definição dos níveis de língua é muito delicada, pois aqui se misturam gramática e léxico, normas sociais e intuição pessoal. Na prática, é mais fácil sentir o nível de língua adotado na comunicação, do que descrevê-lo. (GENOUVRIER; PEYTARD, 1985, p. 36).

Os mesmos autores citados usam o termo *estruturas lexicais* para remeter àquilo que a palavra encerra, por sua vez, nunca está vazia. Não se trata de um só sentido, mas de um jogo mútuo e múltiplo. Enfatizam que há diferença existente entre os termos léxico e vocabulário. Aquele traz em si todas as palavras de um dado idioma, estas são usadas pelo falante em situações específicas e de acordo com sua necessidade, constitui assim o léxico individual. O léxico de uma sociedade em determinado momento histórico torna-se geral ou global e a partir do uso do léxico individual que o falante faz, serve-se de termos específicos, para situações concretas, nesse caso ele faz uso do vocabulário.

O que ocorre é que nenhum professor irá ter acesso à totalidade do léxico de seus alunos, mas em situações específicas, como na escrita, por exemplo, poderá perceber a riqueza desse vocabulário, seja em qualidade ou quantidade. Cabe aí uma pergunta: como enriquecer esse léxico individual? Fornecendo o maior número de palavras possíveis ao aluno de modo que sejam sensatamente usadas em seus sentidos, fazendo com que esse léxico aumente e o vocabulário seja precisamente selecionado e/ou adaptado.

Com relação aos meios socioculturais, é importante lembrar que o controle sobre o vocabulário do aluno não está nas mãos do professor. Os indivíduos vêm de um berço familiar e muitas palavras vão sendo adquiridas no engajamento das relações sociais. Esse momento escolar e familiar concomitantemente é decisivo na vida do aluno e ocorre alternadamente. “O enriquecimento do léxico supõe no sujeito uma atividade de comparação, classificação e agrupamento das palavras, mesmo que essa atividade escape à consciência clara. (GENOUVRIER; PEYTARD, 1985, p. 37). É na família que a língua se produz e que os recursos de linguagem são adquiridos. Na escola, o número de palavras e frases vão aumentando.

Tais teóricos acrescentam que a relação está no fato de que o léxico do aluno será aguçado sobremaneira pela forma como seu ambiente familiar irá direcionar discussões e diálogos, orientar a sua curiosidade para temas diferenciados, diversos, cercando-o do que chamamos de clima cultural. Na medida em que as trocas linguísticas vão sendo multiplicadas com o meio, o aluno aprende o sentido das palavras de acordo com as entregas ampliando assim a área de seu léxico.

O papel da escola é o de completar e compensar essa cultura verbal que é primeiramente recebida no ambiente familiar. É lamentável que o vocabulário infantil pertença primeiramente a gíria do bairro, a variedade dos dialetos de uma aldeia ou da região. Ao professor cabe então levar as crianças a expressarem seus pensamentos e sentimentos, seja oralmente ou de forma escrita. Significa que ao entrar para escola o aluno pode mudar o seu meio linguístico, ainda que em casa ele fale um português correto, sempre vai enfrentar na escola algumas censuras da língua culta, como obrigações de uma fala que tenta se aproximar cada vez mais de uma língua modelo, cheia de beleza, riqueza e precisão.

Um dos papéis do mestre consistirá, contudo, em contribuir para uma boa formação léxica, o que quer dizer que lhes cabe favorecer esse trabalho de comparação, classificação e agrupamento que o aprendizado das palavras estimula e exige. (GENOUVRIER; PEYTARD, 1985, p. 37).

A escola é um lugar de ciência, o aluno é posto frente a diferentes disciplinas que exigem atenção, aplicação, da escrita ao cálculo. E em cada uma dessas disciplinas o aprendiz deve abrir caminho por um vocabulário específico, que por sua vez, apresenta um conjunto de termos que ele não vai reencontrar nas ciências naturais. Podemos inclusive afirmar que as disciplinas escolares se dão inicialmente pela aprendizagem de um vocabulário.

2.1.2 A atenção dispensada ao vocabulário no Enem e na Vunesp

O vocabulário é uma parte de variadas dimensões que de acordo com a solicitação do momento ativa um léxico individual que ao mesmo tempo está inserido num léxico verbal. Problemas surgem daí: o léxico total de um aluno não é percebido e o aluno nunca usa totalmente o seu léxico nos atos de fala. No material da competência 1, módulo 03, alguns pontos referentes a desvios vocabulares e escolhas lexicais imprecisas são mencionados.

Figura 1 - Sumário presente no manual de correção do Enem

SUMÁRIO DO CAPÍTULO

3.4. Desvios de escolha vocabular.....	44
3.4.1. Escolhas lexicais imprecisas.....	44
3.5. Orientações Gerais.....	46
3.5.1. Indicações de desvio.....	46
3.5.2. Mais de um desvio em um mesmo vocábulo.....	47
3.5.3. Mesmo desvio em palavras repetidas ou com mesmo radical.....	47
3.5.4. Algumas considerações sobre o nível 5.....	49
3.5.5. Falha de estrutura sintática ou desvio de pontuação?.....	52
4. DESCRIÇÃO DOS NÍVEIS.....	53
4.1. Nível 0 (nota 0).....	53
4.2. Nível 1 (nota 40).....	55
4.3. Nível 2 (nota 80).....	59
4.4. Nível 3 (nota 120).....	63
4.5. Nível 4 (nota 160).....	70
4.6. Nível 5 (nota 200).....	74
5. CONCLUSÃO.....	77
6. REFERÊNCIA BIBLIOGRÁFICA.....	77

Fonte: Brasil (2020, p. 4).

Esses desvios de escolha vocabular mencionados no sumário dizem respeito às escolhas lexicais equivocadas para a situação que se queira expressar no texto, por conta de semelhanças de palavras ou de sentido fora de contexto. Além disso, usos imprecisos, devido à confusão de palavras. Como imprecisão vocabular, são consideradas aquelas palavras inexistentes na Língua Portuguesa ou que de algum modo foram morfologicamente alteradas, também devem ser consideradas como problemas de imprecisão vocabular.

Ao lermos a cartilha do participante do ENEM (BRASIL, 2019), menciona-se que a redação deverá ser redigida de acordo com a modalidade escrita formal da língua portuguesa. O candidato deve atentar-se à estrutura dos períodos, à acentuação e à ortografia das palavras, ao emprego adequado do hífen e das letras maiúsculas e minúsculas, à separação silábica (translineação), à regência e à concordância, nominais e verbais, à pontuação, ao paralelismo sintático, ao emprego dos pronomes e da crase, à adequação à escrita formal da língua portuguesa, sem informalidades e marcas de oralidade, bem como à adequação vocabular.

Tratando especificamente do manual do candidato ao vestibular, organizado pela Fundação para o Vestibular da Universidade Estadual Paulista para o 1º semestre de 2021. Nele, no § 3º da página 40, os quesitos exigidos para a redação são os seguintes: as propriedades de progressão temática, coerência e coesão e a escrita culta. Na página 60, alguns critérios de avaliação da redação são mencionados, a saber: o tema; a estrutura (gênero/tipo textual e a

coerência); expressão (coesão, modalidade), nesta, mais precisamente assinala-se que serão examinados os aspectos gramaticais, tais como ortografia, acentuação, pontuação, regência, concordância (verbal e nominal) etc., bem como a escolha lexical (precisão vocabular) e o grau de formalidade/informalidade expressa em palavras e expressões

Figura 2 - Critérios de avaliação de redações da Vunesp

CRITÉRIOS DE AVALIAÇÃO DA REDAÇÃO

Na prova de redação, espera-se que o candidato produza um texto dissertativo-argumentativo (em prosa), de acordo com a norma-padrão da língua portuguesa, a partir da leitura e compreensão de textos auxiliares, que servem como um referencial para ampliar os argumentos produzidos pelo próprio candidato. Ele deverá demonstrar domínio dos mecanismos de coesão e coerência textual, considerando a importância de apresentar um texto bem articulado.

A prova de redação será avaliada conforme os critérios a seguir:

A) Tema: avalia-se, neste critério, se o texto do candidato atende ao tema proposto. A fuga completa ao tema proposto é motivo suficiente para que a redação não seja corrigida em qualquer outro de seus aspectos, recebendo nota 0 (zero) total.

B) Estrutura (gênero/tipo de texto e coerência): consideram-se aqui, conjuntamente, os aspectos referentes ao gênero/tipo de texto proposto e à coerência das ideias. A fuga completa ao gênero/tipo de texto é motivo suficiente para que a redação não seja corrigida em qualquer outro de seus aspectos, recebendo nota 0 (zero) total. Na avaliação do gênero/tipo de texto, observa-se como o candidato sustenta a sua tese, em termos argumentativos, e como essa

partes da macroestrutura dissertativa, a falta de um posicionamento (por parte do autor da redação) na defesa de um determinado ponto de vista, a falta de autonomia do texto, a presença de contradição entre as ideias, a falta de desenvolvimento dos argumentos e a presença de conclusões não decorrentes do que foi previamente exposto.

C) Expressão (coesão e modalidade): consideram-se, neste item, os aspectos referentes à coesão textual e ao domínio da norma-padrão da língua portuguesa. Na coesão, avalia-se a utilização dos recursos coesivos da língua (anáforas, catáforas, substituições, conjunções etc.), de modo a tornar a relação entre palavras, frases, períodos e parágrafos do texto mais clara e precisa. Serão considerados aspectos negativos as quebras entre frases ou parágrafos e o emprego inadequado de recursos coesivos. Na modalidade, serão examinados os aspectos gramaticais, tais como ortografia, acentuação, pontuação, regência, concordância (verbal e nominal) etc., bem como a escolha lexical (precisão vocabular) e o grau de formalidade/informalidade expressa em palavras e expressões.

Será atribuída nota zero à redação que:

a) fugir ao tema e/ou gênero propostos;

b) apresentar nome, rubrica, assinatura, sinal, iniciais ou

Fonte: Vunesp (2020, p. 60).

A partir de tais considerações, o enriquecimento do vocabulário e do léxico do aluno torna-se um desafio. O caminho a ser seguido pode ser o de fornecer o maior número possível de palavras diversificadas norteando um emprego sensato das mesmas e assim fazer com que esse léxico aumente e torne-se mais preciso. O que não podemos esquecer é que os níveis socioculturais com os quais os alunos entram em contato sucessivamente determinam a sua função de locutor influenciando em seu domínio lexical.

Tratar no ensino do léxico somente os aspectos morfológicos das palavras, suas significações básicas e relações semânticas primárias cria barreiras para que as competências lexicais mais sofisticadas sejam alcançadas. Com base em tais pressupostos, pensando numa possível superação do panorama de crise na escrita propriamente dita, reafirma-se a urgência da ampliação do ensino do léxico, não como um subsistema autônomo, separado e estático, mas como uma organização sistêmica que atenda às necessidades comunicativas de seus usuários.

Partimos do princípio de que o léxico não pode ser visto como objeto à parte de sua situação de produção e uso em contextos reais, uma vez que são as condições de produção que desvendam as intenções dos interlocutores envolvidos. O uso vai redefinindo as funções e as necessidades da comunicação.

Uma mensagem é veiculada por meio do léxico e das palavras lexicais que constituem os enunciados. Desse modo, há que se dizer que na veiculação de significados, o papel do vocabulário é imprescindível. É pelos signos linguísticos que os discursos humanos se perfazem e que as realidades extralinguísticas são referidas.

2.1.3 Conceitos de Lexicultura e Carga Cultural Compartilhada

Esse trabalho recorre à Lexicultura por considerarmos que a herança linguística é o maior patrimônio de um povo, de acordo com as diferentes mudanças, acontecimentos, e diversidades culturais novos sentidos vão surgindo na língua, o que pode resultar em um desuso e marginalização de possíveis itens lexicais. De modo inverso, alguns termos podem voltar ao uso, porém com conotações diferentes. Guerra e Andrade (2012, p. 12) afirmam que na palavra não há valor absoluto, cada uma remete a categorias linguísticas e mentais exclusivas. Desse modo, cada palavra demonstra uma realidade, um valor específico. Nessa constante, o léxico vai sendo enriquecido por novos vocábulos ou novas significações.

Na produção escrita dos alunos, é possível perceber o léxico como um sistema aberto que pode se expandir e sofrer alterações a partir de mudanças sociais e culturais que são refletidas nos usos vocabulares. Assim, o ensino de língua não pode, em nenhuma instância, ignorar as modificações multiculturais que influenciam diretamente no valor lexical.

A maneira pela qual os falantes se apropriam do léxico, compreendem e interpretam novas palavras e assim constituem e reconstituem os seus vocabulários, são questões as quais devemos nos atentar. Basílio (1987, p. 17) coloca em xeque uma discussão muito pertinente acerca do fato de formarmos novas palavras o tempo todo, tanto em situações comunicativas mais monitoradas quanto no cotidiano informal, isto por meio de adições de palavras, novas composições, combinações, mudanças em suas bases semânticas, de modo que lançamos mão de uma série de palavras que provavelmente não estavam expostas para serem usadas, mas surgiram de acordo com a necessidade de um contexto específico.

Partindo do pressuposto da não separação das dimensões linguísticas e culturais existentes no ensino de línguas, para que o implícito do cotidiano seja aprendido, propõe-se uma associação entre léxico e cultura. Barbosa (2009, p.10) discute conceitos de Lexicultura e

de palavras com carga cultural compartilhada objetivando evidenciar o quanto são pertinentes no contexto de ensino-aprendizagem de língua estrangeira em geral e de português como língua estrangeira.

Portanto, evidencia-se que o estudo da lexicultura tem como foco não o significado da expressão ou da palavra em si – tarefa empreendida pelos dicionários – mas o dado cultural – coletivo – evocado pelo signo, uma vez que ele vai revelar as escolhas feitas por aquela coletividade. (BARBOSA, 2009, p. 10).

Segundo Barbosa (2009, p. 11), significados podem ser compreendidos por meio do léxico, uma vez que possui uma relação direta com os aspectos cognitivos, sociais e culturais de uma língua. Desse modo, podemos considerar que as palavras privilegiam o acesso à cultura trazendo em si concepções e visões de mundo distintas; a identidade cultural de uma sociedade é refletida nas características específicas de cada língua, os fatos culturais são classificados dentro de um sistema linguístico, tanto por meio do léxico como pela gramática.

A palavra é pois um instrumento de compreensão social que simboliza as nossas ideias e é por meio do léxico que podemos compreender e/ou explicar a sociedade da qual fazemos parte, por isso podemos aproximar a Lexicologia da Sociologia pois incidem sobre os fatos sociais; a primeira, parte do estudo do vocabulário para explicar a sociedade, utiliza as palavras como material linguístico constituindo-se uma disciplina sociológica; é por meio do léxico que podemos associar os binômios língua-cultura. A segunda, parte do estudo dos processos que interligam os indivíduos a partir de uma base teórico-metodológica voltada para o estudo dos fenômenos sociais.

Cada visão de mundo, segundo Barbosa (2009, p. 12), é organizada dentro da língua por meio de um recorte lexical; nas palavras dessa língua, reconhecer suas crenças, regras de conduta e sua organização social, é compreender os seus dados culturais a partir da evidência de concepções de mundo que uma dada sociedade compartilha. Por conta disso, ter acesso a um conjunto lexical distinto exige do aluno esforços que lhe permitam integrar conhecimentos linguísticos e culturais para a compreensão de valores, crenças, usos e costumes numa dinâmica de convergências e divergências culturais.

Podemos entender o conceito de Lexicultura como um modo de acesso ao estudo do léxico culturalmente marcado. Percebemos de modo singular e também diverso a cultura em uma língua, já que sabemos que o léxico a descreve associando-se à realidade extralinguística. A partir dessa composição, a lexicultura privilegia o léxico e a cultura de modo consubstancial, designando o valor que as palavras adquirem pelo uso que se faz delas.

Por se voltar mais à prática do que à teoria, pode ser vista como um conceito instrumental. Ao conceber a cultura no e pelo léxico faz de seu objeto de estudo um elemento que age e intervém de modo disponível ao aprendente de uma língua estrangeira.

Segundo Antunes (2012, p. 15), toda mensagem transmite uma informação e isso só é possível por meio do léxico e das palavras lexicais que integram os enunciados. Nossas escolhas lexicais preferidas são “pistas” claras do nosso pertencimento aos grupos sociais onde tecemos nossa identidade. É impossível negar a vinculação do léxico da língua às experiências socioculturais que caracterizam cada um dos grupos humanos.

Duas concepções de cultura são apresentadas por Barbosa e Souza (2016, p. 15), a de Cultura com “C” maiúsculo (aquela ensinada pela arte, pela literatura, pelas instituições sociais, oriunda das Ciências Humanas) e aquela com “c” minúsculo proveniente das Ciências Sociais, que engloba valores de um povo, pensamentos, costumes, modos de agir, festejar, comer e conversar, por exemplo. Pensa-se aqui em uma abordagem comunicativa de ensino de língua que vai além do conhecimento linguístico, mas considera habilidades discursivas e sociolinguísticas englobando o que é importante para se ter acesso às habilidades citadas.

As autoras também falam sobre cultura ensinada e comportamental. A primeira é privilegiada pela escola (erudita) e a segunda advinda do cotidiano, do reconhecimento mútuo e do compartilhamento entre falantes; cultura do dia a dia, das massas, terreno de crenças e de comportamentos.

Para Barbosa e Souza (2016, p. 16), muito de um povo é revelado por meio do léxico, uma vez que exterioriza costumes, crenças e visões de mundo de uma determinada comunidade linguística. Por vezes, o aprendente de língua estrangeira, por exemplo, encontra dificuldades para se apropriar de aspectos culturais brasileiros. Nesse sentido, o léxico deve ser um possibilitador de ação na língua-alvo.

Com relação ao léxico marcado pela carga cultural compartilhada, podemos compreendê-la como um valor agregado ao sentido referencial da palavra, entendido e compartilhado de maneira próxima à cultura da experiência, dos modos de vida e do dia a dia, do reconhecimento mútuo. É quase imperceptível aos falantes nativos, por conta disso não é alvo de tantas descrições. É uma forma de adentrar a uma outra cultura com a finalidade de entender e fazer-se entender além do que se pode ver: nos implícitos culturais existentes na língua.

Barbosa e Souza (2016, p. 15) trazem o conceito de Lexicultura, vista como o estudo de unidades lexicais que trazem em si um valor agregado ao uso da língua, dotadas de polissemia,

opacidade e referenciais culturais implícitos. Em sua dimensão social, está o conhecimento, por conta da relação entre o cotidiano e as práticas sociais; o significante, por sua vez está referido à dimensão linguística. Não se trata de realizações escritas ou discursivas dos falantes, mas do conjunto de conhecimentos implícitos originários das práticas sociais, costumes, crenças e valores compartilhados por uma comunidade. O seu principal objetivo é descrever e estudar palavras com carga cultural compartilhada com fins pedagógicos específicos.

2.2 Linguística de Córpus

O *córpus* é uma coletânea de textos disponíveis em formato eletrônico. Existem algumas possibilidades de contribuição/vantagens para a prática docente do profissional de línguas com o uso da Linguística de Córpus: aprendizagem da frequência das palavras; colocações; frases e fragmentos lexicais; gramática; exemplos autênticos; observação, reflexão e conhecimento do léxico em uso; abordagens descritivas; palavras novas; *córpus* especializado/específico e revolução digital. Para a constituição de um *córpus*, alguns aspectos devem ser levados em conta, dentre eles a sistematização da coleta, a representatividade e a forma como os arquivos serão dispostos.

Com relação aos *córpus*, sabemos que os *córpus* escritos ainda dominam a produção na área, mas a compilação de *córpus* multimodais tem se ampliado de forma rápida. Os *córpus* têm encontrado uma aplicabilidade crescente não somente nos estudos linguísticos canônicos, mas no desenvolvimento de tecnologias da fala, do reconhecimento e da síntese. Questionamentos sobre a importância da coleta de dados dos usos linguísticos para as pesquisas vêm sendo recorrentes desde os inícios dos anos 60, no contexto histórico de dominância linguística gerativista, o que traz um contraste radical com o panorama atual.

A afirmação de que todo *córpus* traz questões novas ou questões que não imagináramos encontrar é amplamente conhecida, ainda que nenhum *córpus* nos responda tudo. Desse modo, tanto as observações como os experimentos e hipóteses inicialmente formuladas no âmbito da investigação nos orienta a uma revisão à luz das comprovações e dos resultados.

Assim sendo, sistematizar dados e observações é um ponto crucial, talvez mais importante do que a simples aplicação e contraste de teorias. Descobrir e identificar padrões a partir da observação são os problemas fundamentais, assim toda a teoria resultante de um trabalho consciente sobre os dados conduz o nosso olhar para uma compreensão relevante dos

processos de observação, o que é indispensável nas pesquisas com *córpus* e nos diferentes níveis de descoberta.

Analisar a língua a partir de *córpus* nos leva a compreender uma abordagem que privilegie o uso de dados, pode-se, assim, estudar o funcionamento de uma língua. Ao invés de se recorrer ao conhecimento intuitivo, a observação é sobre como a língua é utilizada por seus usuários. Sendo assim, entendemos que uma linguagem não é inventada, ela é capturada, útil de ser estudada por base de dados textuais, *córpus* oriundos de falantes reais.

Para Dubois *et al.* (1993, p. 21), o *córpus* é considerado como um conjunto de enunciados a partir do qual se estabelece a gramática descritiva de uma língua, ele não pode ser considerado como constituindo a língua, mas como uma amostra dessa língua. Deve ser representativo e capaz de ilustrar uma gama de características estruturais. Se há um indefinido número de enunciados possíveis, não há exaustividade verdadeira no *córpus*. Além do mais, grandes quantidades de dados inúteis de serem analisados só complicam a pesquisa, tornando-a pesada. O linguista deve procurar obter um corpo significativo desconfiando de tudo o que torna o seu *Córpus* não representativo.

Para Biber (1993, p. 33), o *córpus* ao ser elaborado constitui um processo de dois ciclos: inicialmente escolhe-se os textos baseados em critérios externos e culturalmente aceitos e em segundo lugar procede-se com investigações empíricas da língua ou variedade linguística em análise; finalmente o projeto é todo revisado.

Para McEnery e Wilson (1996, p. 17), o termo *córpus* possui conotações específicas e a sua noção traz características fundamentais, dentre elas a amostragem e a representatividade. Um *córpus* deve ter uma amostragem suficiente da língua ou variedade linguística que se quer analisar para que o máximo de representatividade dessa língua seja alcançado.

Para Biber *et al.* (1998, p. 18) uma amostragem proporcional não é adequada para um *córpus* linguístico, pois sua organização é demográfica. Contudo, esse tipo de *córpus* não representaria os gêneros textuais, pois poderia conter uma grande parte de conversação, cartas e notas de pedidos, entre tipos de textos jornalísticos, acadêmicos, literários e escritas não publicadas, considerando que a quantidade de pessoas que publicam ou falam para públicos é pequena, o que importa para o estudo de língua é um *córpus* com amostras representativas que trazem em si a variação linguística existente.

Com relação à diversidade, Biber *et al.* (1998, p. 19) dizem que não existe língua geral, cada gênero tem padrões de uso específico. Se um *córpus* serve a estudo de variação ou representa uma língua, a preocupação deve ser com a diversidade de gêneros, tipos de textos, dialetos e tópicos.

Já para Ducrot e Todorov (2001, p. 12), *córpus* é um conjunto tão variado quanto possível de enunciados que são emitidos por falantes da língua em dada época. Para Trask (2004), *córpus* é um conjunto de textos escritos ou falados numa língua com dados disponíveis.

Sinclair (2004, p. 21) propõe o *córpus* como sendo uma coleção de peças de linguagem e textos eletrônicos selecionados de acordo com critérios externos de representatividade e que na medida do possível atenda as variedades linguísticas.

De acordo com Aluísio e Almeida (2006, p. 170), o que mudou com o passar dos anos foi a concepção de *córpus* e essa mudança pode estar atrelada à concepção que se adquiriu de *Linguística de Córpus*.

Geralmente, aumentando-se o tamanho do *Córpus* aumenta-se também a sua representatividade, pois sua dimensão se aproxima à população. Por outro lado, aumentar o tamanho apenas pode não significar um ganho de representatividade, caso o *Córpus* não respeite a distribuição da população. Por exemplo, se na população de textos jornalísticos existam 10% de editoriais e 20% de notícias e no *Córpus* constem apenas editoriais (100%), esse *Córpus* não será representativo daquela população, mesmo que cresça exponencialmente. Ele poderá chegar a ser representativo do gênero 'editorial', mas não de 'jornais' como um todo. (BERBER SARDINHA, 2012, p. 326).

Para Berber Sardinha (2004, p. 11), *córpus* pode ser definido como uma coletânea de exemplos naturais de linguagem que vão desde algumas frases até conjuntos de textos escritos ou gravações orais coletados para serem usados como dados para pesquisa linguística. De modo mais recente, a palavra *córpus* tem sido usada para referir-se à coletânea de textos armazenados e que podem ser acessados por meio de computadores. Textos escritos retirados de jornais ou revistas podem ser escaneados, retirados de um CD ou da internet, textos orais como conversas, são gravados e depois transcritos. Isto é, são copiadas palavra por palavra de forma que os textos dessas conversas sejam alimentados por computadores, assim torna-se possível analisar a língua contida no *córpus* através de alguns softwares como por exemplo o *Wordsmith Tools* (Scott, 1996), exemplos naturais de linguagem podem ser entendidos como aqueles que não tenham sido produzidos ou criados para o *córpus* em específico. A ideia de natural inclui o fato da linguagem ser produzida por humanos, excluindo assim programas que geram textos.

É necessário admitir pensando no formato computadorizado do *córpus* que o surgimento do computador atingiu de modo direto não só a concepção que se tem de *córpus* hoje em dia como também a sua forma de armazenamento e manuseio. Já que as ferramentas oferecidas pelo computador abrem caminho para que uma quantidade inimaginável de textos possa ser processado na tela em segundos, fazendo com que inúmeras hipóteses sobre alguns fenômenos linguísticos possam ser atestadas de modo rápido e eficiente. Essa nova forma de armazenar

textos permitiu que fenômenos linguísticos antes imperceptíveis pudessem ser observados e descritos, o que não era possível antes por contarem apenas com recursos manuais

Outro aspecto importante é o formato eletrônico; quando se emprega o termo *córpus* automaticamente se admite que os textos estejam no formato eletrônico, diferente da ideia que se tinha de *córpus* antigamente, que se referia somente a textos impressos. O formato eletrônico de textos possui benefícios consideráveis aos *córpus*, pois podem ser manuseados de forma mais ágil e podem ter agregados em si informações extras. Como referência padrão, acredita-se que o *córpus* deve estar disponível para outros pesquisadores. O que chamamos de reuso do *córpus* por constituir uma referência padrão para variedade de língua que representa.

Para Trask (2004, p. 16), a partir de *córpus* pode-se fazer observações precisas, trazer informações confiáveis, exemplos de opiniões e julgamentos acerca dos fatos de uma língua. Desse modo aspectos morfológicos, sintáticos, semânticos e discursivos podem ser observados; o que é bastante relevante para uma pesquisa linguística; a produtividade e o emprego das palavras, as expressões e formas gramaticais também podem ser explicadas, novos fatos na língua não perceptíveis intuitivamente podem ser descobertos. Em suma, a língua pode ser descrita de forma objetiva.

No entanto, há um conjunto de requisitos que devem ser observados para o projeto de um *córpus* computadorizado. Esses requisitos impactaram de modo direto na validade e confiabilidade da pesquisa baseada no *córpus*, são eles: a autenticidade, representatividade, balanceamento, amostragem e diversidade de tamanho, os quais serão descritos a seguir.

Os textos devem ser autênticos, escritos em linguagem natural e não produzidos com alvo de serem pesquisados linguisticamente, devem ser escritos por falantes nativos a menos que se trate de *córpus* de aprendizes, aqueles *córpus* cujos textos são oriundos de falantes que estão aprendendo uma língua estrangeira.

Aluísio e Almeida (2006, p. 165) enfatizam que existem muitos *córpus* disponíveis de forma livre ou paga, a partir desses, subcorpos de estudo podem ser gerados, dependendo da questão de pesquisa ainda pode ser necessário compilar um *córpus* próprio. Para que seja compilado existem três estágios: o projeto do *córpus* que inclui a seleção dos textos e os cuidados com os requisitos que já foram mencionados; compilação ou captura, manipulação e nomeação dos arquivos textuais, permissão de uso; e anotação.

Com relação ao projeto de *córpus*, inicialmente seleciona-se os textos pertinentes e relevantes para a pesquisa. É preciso definir que tipo de *córpus* será compilado, outras decisões também são importantes com relação à composição e aos gêneros que farão parte. Existem várias tipologias de *córpus* que indicam os parâmetros importantes a serem considerados.

O *córpus*, as autoras acrescentam, precisa ser preparado para o processamento computacional, precisa ser limpo e formatado, isso significa tirar imagens, gráficos, tabelas e demais anotações que não estão no texto propriamente dito. Essa limpeza e formatação possibilitam que o *córpus* seja processado por ferramentas computacionais, como o concordanciador, o contador de frequência e a extração automática de termos.

Na compilação de *córpus* deve-se seguir as regras legais para que o uso do material obedeça às regras de privacidade, os direitos de uso e o seu reconhecimento. É uma etapa da compilação não técnica e demorada, marcada por uma série de negociações que podem durar anos, na maioria das vezes é este o motivo do *córpus* não estar disponível publicamente. Para que a negação desse pedido de permissão seja vencida, uma boa estratégia é coletar um grande número de textos dentro de cada categoria de um *córpus*

Não há uma abordagem amplamente aceita para que o anonimato dos indivíduos em *córpus* da modalidade oral seja preservada. A tendência é pelo completo anonimato, ou seja, o apagamento de nome e sobrenome, de títulos profissionais, nome de animais de estimação, endereço e telefone. Esses dados são removidos ou trocados por códigos. Outro procedimento é a troca por nomes similares que são equivalentes prosodicamente aos originais. Na medida que os aspectos éticos e legais da preservação do anonimato tratam do interesse do informante, os aspectos sociolinguísticos e computacionais tratam do interesse da pesquisa; informações sociolinguísticas relacionadas às características socioeconômicas e étnicas podem ser fornecidas por nomes e apelidos por exemplo, todavia fazer a troca por outro nome demanda bastante tempo e raramente isso é feito

Com relação à anotação, Alúísio e Almeida (2006, p. 178) dizem que os níveis de representação das informações presentes no *córpus* são estruturais. Em Linguística, a anotação estrutural diz respeito à marcação de dados externos e internos dos textos. Os dados externos são a documentação do *córpus* em forma de cabeçalho que inclui os metadados textuais, bibliografia, catalogação, autoria, tipologia textual e a distribuição do *córpus*. Já os dados internos podem ser entendidos como a anotação de segmentação do texto cru que envolve a estrutura geral, os títulos e subtítulos, notas de rodapé, elementos gráficos com tabelas e figuras ou sub parágrafos. A anotação linguística pode se dar em qualquer nível, morfossintático, sintático, semântico e discursivo e se encerra em três formas: manual, automática ou semiautomática.

A noção de um *córpus*, segundo Bonelli e Sinclair (2005, p. 214), tem sido fortemente afetada por movimentos externos à língua em estudo, e para muitos utilizadores um *córpus* é apenas uma conveniente coleção de textos. Esta ambiguidade que envolve a ideia de um *córpus*

pode ser explorada, nem sempre de forma deliberada acerca de fonte de dados que não têm direito a ser considerados *córpus*. Devemos estar conscientes dos perigos para os investigadores que procuram a segurança de algo chamado *córpus* em vez de apenas alguns dados convenientes.

Corpus can be of varying sizes, are compiled for different purposes, and are composed of texts of different types. All *córpus* are homogeneous to a certain extent; they are composed of texts from one language or one variety of a language or one register, etc. They also are all heterogeneous to a certain extent, in that at the very least they are composed of a number of different texts. Most *córpus* contain information in addition to the texts that make them up, such as information about the texts themselves, part-of-speech tags for each word, and parsing information. (FLOWERDEW, 2009, p. 402).

3

Um risco é que os dados não podem ser escolhidos com rigor ou clareza suficientes de critérios para que os resultados das consultas possam ser interpretados como declarações verdadeiras sobre uma língua ou variedade; é necessário que se tenha experiência para a interpretação e avaliação dos resultados obtidos. Cabe à comunidade de linguistas, na utilização de *córpus*, determinar se os méritos particulares de um *córpus* valem a pena para preservar a terminologia, ou se a linguagem de criação eletrônica de um texto é de uma ordem tão diferente que a ideia de um *córpus*, tornar-se-á em breve uma camisa de forças limitando a procura de informações definitivas.

Por um lado, acrescentam Bonelli e Sinclair (2005, p. 216), há segurança e estabilidade em ter 'um *córpus* de língua X'. Alega-se que a compilação de um dicionário de uma língua tem vários efeitos, e uma vez que os principais dicionários serão com base em *córpus* nos próximos anos, um *córpus* ao lado do dicionário nacional serviria bem. Por outro lado, o trabalho de compilação de um *córpus* é ainda considerável, apesar das enormes melhorias em tecnologia nas últimas décadas. Mesmo bastante básico, aspectos da concepção ainda não são informados com seriedade de investigação, tais como as proporções desejáveis de linguagem escrita num *córpus* de amostra que procura representar o estado de uma língua. No mundo do reparo rápido, a atmosfera prevalecente da linguística computacional é particularmente a do processamento de linguagem natural, há pouco interesse em refinar as técnicas de desenho e métodos construtivos que vão ser utilizados para a criação de *córpus*.

³ Os *corpus* podem ter tamanhos variados, são compilados para finalidades diferentes e são compostos de textos de diferentes tipos. Todos os *córpus* são homogêneos até certo ponto; são compostos por textos de uma língua ou de uma variedade de uma língua ou de um registro, etc. Também são todos heterogêneos até certo ponto, na medida em que, no mínimo, são compostos por vários textos diferentes. A maioria dos *córpus* contém informações além dos textos que os compõem, como informações sobre os próprios textos, marcações gramaticais para cada palavra e informações de análise. (FLOWERDEW, 2009, p. 402).

O *córpus*, Bonelli e Sinclair (2005, p. 217) acrescentam, deve também manter uma distância de arquivamento de documentos, que, sob a forma de digitalização é, atualmente, uma grande empresa. Os campos sobrepõem-se e podem cooperar com benefício mútuo, as suas prioridades são diferentes; um construtor de *córpus* requer apenas um esboço da proveniência de um documento para saber onde colocá-lo num *córpus*; enquanto para um arquivista, este tipo de informação é o alvo principal da disciplina.

De uma perspectiva intelectual global, então, um *córpus* parece uma entidade bastante frágil; fustigada por muitas forças mais poderosas de áreas como a 'engenharia da linguagem'. Uma vez que as línguas são centrais para a comunicação internacional, há um interesse contínuo e um apoio financeiro de algumas empresas de base política e de defesa de fontes, que podem atravessar as prioridades da ciência.

Para Bonelli e Sinclair (2005, p. 218), o contexto pode ser organizado em padrões principais por meio do estudo das colocações e do uso de palavras específicas. O funcionamento da língua utilizada no texto pode ser acessado por meio do conhecimento desses padrões, o que facilita a distinção entre o significado de uma palavra isolada ou ainda a determinação da variação das características sintáticas.

Em última análise, um fator que um *córpus* pode usar firmemente para se distinguir é o fato de que foi concebido para refletir o mais fielmente possível o tipo de linguagem a partir do qual é selecionado. Essa característica poderia atrair o seu próprio apoio se a tendência normativa do *córpus* em construção fosse enfatizada. Muitas pessoas veem a linguagem da Internet com considerável preocupação pelo tipo de consistência esperado da palavra impressa.

Bonelli e Sinclair (2005, p. 218) enfatizam que uma característica da Internet é que é global, por isso a ideia de todos usarem o mesmo conjunto de convenções é provavelmente uma batalha perdida. Para a compreensão e interpretação de resultados da pesquisa a partir da Web, alguns aspectos devem ser considerados, como: melhorar a qualidade das informações obtidas e replicar os resultados. Este tipo de progresso pode levar a uma ideia mais nova e moderna acerca de um *córpus*, não o *córpus* de amostra de que quase todos os nossos *córpus* atuais são derivados, mas algo mais próximo da noção de um *córpus* monitor, que deixe o texto fluir sobre um conjunto de filtros afinados, com amostragem contínua, atualização de registros e manutenção, um conjunto estável de ferramentas descritivas para os utilizadores, em vez de uma descrição estável. A linguística de *Córpus* está geralmente associada a uma abordagem fraseológica de análise que tem uma visão sintagmática em oposição a uma visão puramente paradigmática da linguagem.

Para Danielsson (2007, p. 51), o eixo paradigmático que testa cada palavra para ver se a unidade permite alguma alternativas, permite descobrir se o córpis oferece alternativas de aprendizagem. Essas alternativas devem refletir a natureza sintagmática de prosódia semântica (a frequente co-ocorrência de um artigo lexical com artigos que expressam uma avaliação positiva ou negativa) e a preferência semântica (as frequentes co-ocorrências de um artigo com um determinado conjunto semântico).

Charles (2007, p. 38), medeia entre o processamento de cima para baixo e de baixo para cima do processamento acadêmico, redação de textos para introduzir à função retórica de "defender o seu trabalho contra a crítica", um padrão retórico em duas partes, "no qual o escritor primeiro admite a possibilidade de crítica e depois avança para a neutralização do seu efeito potencialmente negativo" Charles (2007, p. 40) consegue isto ao perceber que primeiro os seus alunos envolvem-se em tarefas iniciais baseadas no discurso em nível macro, como atividades de conscientização do padrão discursivo. Nesta fase inicial, os estudantes são apresentados à função retórica, a sua finalidade no texto e as diferentes formas em que pode ser realizado. Os estudantes discutem os conhecimentos que adquiriram no seu grupo de discussão numa sessão de feedback de toda a classe.

Charles move-se então para a análise da base para o topo, fazendo com que os alunos realizem pesquisas em córpis que concentrem-se em estruturas léxico-gramaticais específicas dentro de um quadro baseado no discurso, concordância dos estudantes sobre artigos com o objetivo de formular uma generalização sobre a utilização e o posicionamento, enquanto se constrói uma concessão. Os estudantes são convidados a examinar o lexicograma para a construção de uma concessão.

Alguns pesquisadores veem uma abordagem associada à Linguística de Córpis como um catalisador na redefinição de aspectos linguísticos teóricos. Esta é considerada como a abordagem baseada em córpis, em que os dados são abordados sem quaisquer noções pré-concebidas em relação à forma como deve ser analisada. Outros linguistas de córpis tomam uma posição menos extrema, por exemplo, McEnery *et al.* (2006), ao considerar o córpis linguístico como "uma nova abordagem filosófica à investigação linguística" com o seu próprio estatuto teórico, deixa de ser visto como uma disciplina por direito próprio com a sua própria teoria, mas sim como uma metodologia. A mesma posição pode ser aplicada aos dados de aprendizagem DDL (*data-driven-learning*), geralmente aceita-se que não existe uma teoria subjacente enquanto tal, mas uma metodologia que pode desvendar fatos sobre a língua até agora inexplorada.

Flowerdew (2009, p. 403) afirma que não obstante as vantagens desta abordagem para o DDL, durante os últimos anos alguns relatos na literatura têm adotado uma postura mais crítica, desenhando atenção aos potenciais inconvenientes da utilização de *corp*pus no DDL.

Pode-se ver, dessa forma, que muito recentemente os esforços pedagógicos adotaram uma abordagem muito mais discursiva, um desenvolvimento que foi defendido por Flowerdew (1998, p. 404) há mais de uma década. A pedagogia do *corp*pus progrediu para além de olhar para as linhas de concordância, e agora abrange as "unidades de significado". A questão da contextualização continua a ser problemática. É previsível que, no futuro, seja dada mais atenção à marcação do texto escrito com características contextuais como norma para os *corp*pus falados.

It can be seen that utilizing a more top-down approach to processing corpus data provides more co-text, and hence more contextual information on the *corp*pus under investigation by shedding light on different practices of different academic disciplines, as revealed by differences in lexico-grammatical patterning. However, whether the starting point should be with a bottom-up or top-down approach is not an easy question to answer and very much depends on the nature of the query and composition of the *Corp*pus. (FLOWERDEW, 2009, p. 402).⁴

Os *corp*pus, porém, tem se mostrado não completamente desprovidos de contexto, e que o ambiente co-textual pode fornecer pistas contextuais úteis. Embora existam poucos relatos na literatura sobre a "mediação pedagógica" de dados de *corp*pus, estes são poucos e distantes, o que indica que esta é uma área de discussão e expansão. Deve-se prestar mais atenção aos tipos de *corp*pus mais adequados.

Os *corp*pus influenciaram o ensino de línguas em três formas distintas. Em primeiro lugar, os resultados da investigação do *corp*pus têm sido amplamente utilizados para melhorar as referências em materiais para alunos, tais como dicionários e gramáticas. Em segundo lugar, os alunos estão cada vez mais encorajados a explorar *corp*pus por si próprios. Finalmente, técnicas de *corp*pus foram aplicadas para o estudo de língua dos aprendizes.

Desde a publicação, em meados da década de 1980, do primeiro dicionário de aprendizes baseado na pesquisa de *corp*pus (Sinclair *et al.*, 1987, p. 210), os *corp*pus tornaram-se um indispensável recurso para lexicógrafos e gramáticos. Os dicionários dos estudantes tipicamente dão mais atenção à fraseologia, e em particular a colocação. Da mesma forma, a

⁴ Pode-se ver que a utilização de uma abordagem mais de cima para baixo para processar dados do *corp*pus fornece mais co-texto e, portanto, mais informações contextuais sobre o *corp*pus sob investigação, lançando luz sobre diferentes práticas de diferentes disciplinas acadêmicas, conforme revelado por diferenças no léxico. padronização gramatical. No entanto, se o ponto de partida deve ser uma abordagem de baixo para cima ou de cima para baixo não é uma pergunta fácil de responder e depende muito da natureza da consulta e da composição do *Corp*pus. (FLOWERDEW, 2009, p. 402).

gramática dos livros para alunos prestam mais atenção ao registro da variação, ao uso falado, e ao papel das lexias em gramática (Sinclair *et al.*, 1990, p. 214; Biber *et al.*, 1999, p. 15). Em menor medida, os livros de curso também mudaram, agora colocando mais ênfase na co-instalação e fraseologia do que anteriormente.

Os *córpus* influenciaram o método, bem como o conteúdo, do ensino de línguas. Os alunos avançados são frequentemente convidados a participarem de uma aprendizagem orientada por dados, podendo fazer as suas próprias generalizações sobre o uso da língua. Uma das consequências disto é que os aprendizes são expostos a toda complexidade de uma língua, e a tarefa de ensinar explicitamente todos os aspectos dessa língua parece menos viável do que antes.

Como resultado, a aprendizagem orientada por dados coincide com a vista de aprendizagem de línguas que enfatiza a observação guiada por parte do aprendente em vez de exposição do professor. Finalmente, a língua dos próprios aprendizes é estudada extensivamente através do desenvolvimento de *córpus* aprendiz (Granger, 1998, p. 35), ou seja, *córpus* compostos de coletâneas de textos escritos ou falados produzidos por aprendizes de uma língua. Estes permitem a produção dos alunos a ser comparada com a dos nativos falantes.

Recorrendo a Biber e Reppen (2015, p. 14) para conceituar a Linguística de *Córpus*, entendemos que trata-se de uma abordagem de pesquisa que facilita as investigações empíricas relacionadas à variação e ao uso da língua. A Linguística de *Córpus* se ocupa do *córpus*, sua análise e compilação. Por *córpus*, entendemos como sendo um conjunto de textos que podem ser escritos, falados ou multimodais. Esses textos são coletados de forma criteriosa e mantidos em formato eletrônico, constituindo o alvo da pesquisa linguística.

Na Linguística de *Córpus*, os dados constituem um *córpus* que é sistematizado atendendo a alguns critérios amplos e profundos de tal forma que representem a totalidade do uso linguístico. Esses dados podem ser processados pelo computador a fim de gerar resultados que sejam úteis à descrição e análise.

O termo Linguística de *Córpus*, segundo Sarmiento (2010, p. 88), pode ser compreendido como um estudo da linguagem que baseia-se em exemplos do cotidiano real. A Linguística de *Córpus* não pode ser vista como a sintaxe, a semântica e a pragmática. É um ramo da Linguística que concentra-se na descrição e explicação de algum aspecto linguístico em uso, ela é uma metodologia que pode ser aplicada a uma gama de estudos linguísticos, até mesmo ao ensino de línguas. Ou seja, é uma das várias maneiras de se fazer Linguística. Biber, Conrad e Reppen (1998, p. 4) trazem algumas características cruciais na Linguística baseada em *córpus*: ela é empírica, pois analisa padrões reais de uso em textos naturais, utiliza a coletânea de textos para

análise, faz uso de computadores, usando também técnicas automáticas e interativas, é dependente de técnicas analíticas, quantitativas e qualitativas para gerar conhecimento empírico. Os *córpus* são usados para conhecimentos sobre a língua que podem suplementar ou suplantar informações oriundas de fontes de referência e de introspecção (LEECH, 1991, 1992).

Para Novodvorski e Finatto (2014, p. 7), é importante dizer que a Linguística de *Córpus* se coloca como uma nova perspectiva para a Linguística, mas não como um novo tipo de Linguística. Mostra-se àqueles que se aproximam dela como uma metodologia ou como uma abordagem teórica diferenciada dos estudos da linguagem. Se alguém quiser se aproximar da Linguística de *Córpus* apenas por ter um interesse em seu instrumental ou em seus procedimentos, não será nada cobrado em termo de filiação teórica ou epistemológica, ainda que de modo insistente ressaltamos que a Linguística de *Córpus* é também um modo de compreender a língua, um modo próprio de defini-la como objeto de estudo. Ela é vista como sistema probabilístico de combinatórias no qual as associações mantidas com outras unidades definem uma unidade em específico.

Um *córpus* é uma coleção de textos em idiomas que ocorrem naturalmente, escolhidos para caracterizar um estado ou variedade de um idioma. Na linguística computacional moderna, um *córpus* normalmente contém muitos milhões de palavras; isso ocorre porque é reconhecido que a criatividade da linguagem natural leva a uma variedade imensa de expressão. Essa é a dificuldade de isolar os padrões recorrentes que são as pistas para a estrutura lexical da linguagem. (SINCLAIR, 1991, p. 171, tradução nossa).⁵

Ocupa-se da exploração de grandes extensões de *córpus* textuais digitais, que são reunidos de modo criterioso para representar um dado estudo de uso de língua e “minerados” com o apoio de informática. Dando destaque às explorações estatísticas de elementos lexicais e a observação de frequências e combinações de palavras, temos uma trajetória de estudos realizados no Brasil. São estudos considerados amplos e que podem ser aproveitados em diferentes tipos de pesquisa, servindo hoje minimamente para que gêneros textuais sejam caracterizados. Ao longo do processo investigativo quase todos os gêneros textuais escritos foram objetos de alguns estudos em Linguística de *Córpus*, desde literário ou jornalístico,

⁵ A corpus is a collection of naturally-occurring language text, chosen to characterize a state or variety of a language. In modern computational linguistics, a corpus typically contains many millions of words; this is because it is recognized that the creativity of natural language leads to such immense variety of expression. That is difficulty to isolate the recurrent patterns that are the clues to the lexical structure of the language. (SINCLAIR, 1991, p. 171).⁵

manuais e técnicos, textos de culinária, dentre tantos outros vários, sem nos esquecermos dos *córpus* que são especialmente dedicados aos registros orais.

As técnicas da Linguística de *Córpus* têm sido criticadas por enfatizar uma abordagem mais ascendente em vez de processamento de texto de cima para baixo, em que linhas de concordância truncadas são examinadas de alguma forma sem recurso ao conjunto do discurso (Swales, 2002, 2004). Swales, Kaltenböck & Mehlmauer-Larcher (2005, p. 71) expressaram sentimentos semelhantes: "Existem, contudo, certas partes de um texto que nem mesmo um concordanciador pode alcançar. Estes são aspectos da macroestrutura, tais como movimentos textuais, ou seja, uma unidade de texto que expressa uma função comunicativa". No entanto, nos últimos anos, a Linguística de *Córpus* tem dado muito mais atenção a estes dois diferentes modos de processamento de texto (Flowerdew 2003, 2005) com Biber *et al.* (2007a) a explicar o conceito subjacente a estas duas abordagens diferentes, mas complementares.

Flowerdew (2009, p. 191) diz que a análise de *córpus*, de fato, nos traz tanto uma visão paradigmática quanto sintagmática da linguagem. A saída de concordância pode ser "lida" verticalmente, ou seja, paradigmaticamente, em concordância com a noção de preenchimento por tabelas de substituição; ou horizontalmente (sintagmaticamente) por uma perspectiva fraseológica. Sinclair (1999, 2004) diz que o item léxico tem primazia, com o seu significado central e a sua prosódia semântica, vistas como categorias obrigatórias, colocações, coligações e preferências semânticas consideradas como categorias opcionais. O entrelaçamento de algumas ou todas estas categorias resulta numa "unidade alargada de significado", que mais tarde, será tratado por Sinclair (2004, p. 280) como "abordagem máxima" na qual as dimensões de uma unidade de significado seriam prolongadas até que todas as unidades relevantes sejam incluídas até que a ambiguidade desapareça

Tanto Gavioli (2005, p. 63) como Meunier (2002, p. 61) notaram os inconvenientes de uma abordagem indutiva, na qual os estudantes extrapolam as regras, a partir de exemplos: apesar das suas vantagens, as atividades do DDL têm alguns inconvenientes. As várias estratégias de aprendizagem (dedutiva vs. indutiva) que os estudantes adotam podem conduzir a problemas. Alguns estudantes detestam trabalhar de forma indutiva e os professores devem visar uma abordagem combinada.

Esta área representa uma nova abordagem filosófica para os estudos linguísticos. Svartvik (1996, p. 16) acredita que a Linguística de *Córpus* não é uma definição apenas de metodologia emergente para os estudos linguísticos, porém é um novo modo de pesquisa e realmente uma nova abordagem filosófica para este assunto. Como uma ferramenta tecnológica de poder inestimável, o computador tornou esse novo tipo de Linguística possível, mas depende

dos linguistas com suas intuições acerca da língua instruir tais programas para que evidências linguísticas sejam extraídas. Coleções de textos naturais de uma língua ao serem organizados sistematicamente e digitalmente para representar as áreas de uso linguístico podem ser considerados como *córpus* linguístico de base computacional, já que deles podemos extrair novas informações

Stubbs (2001, p. 10) diz que a Linguística de *Córpus* enxerga a linguagem como um sistema probabilístico. Isto é, ainda que possua várias combinações e características possíveis, nem todas são prováveis de ocorrer. Desse modo, por se tratar de uma técnica que se adequa à análise estatística, os *córpus* fornecem informações sobre a frequência de muitos aspectos da língua. Berber Sardinha (2004, p. 31), enfatiza que na diferença de frequências entre os traços, o mais importante é que não são aleatórios, se essas diferenças fossem aleatórias, a frequência não seria significativa e não acrescentaria informações acerca da língua e suas estruturas. Porém, grupos de características linguísticas apresentam variações sistêmicas em textos específicos, variações provenientes de situações comunicativas determinadas e da avaliação sistêmica que pode ser compreendida como a recorrência de traços linguísticos (colocação, coligação, padrão sintático e outros) indicando a padronização da linguagem motivada por diferentes fatores, além das próprias necessidades comunicativas.

Segundo Oliveira (2009, p. 50) a Linguística de *Córpus* pode ser considerada como a face moderna da Linguística empírica, na medida que enxerga a língua como fenômeno social e a analisa a partir de atos concretos de comunicação, ou seja por meio de textos reais, buscando os significados e as negociações que são feitas no discurso. É uma perspectiva própria sobre a linguagem, fenômeno que estuda, e uma maneira específica de fazer pesquisa, por meio do estudo de textos reais auxiliados por programas de computador cujo objetivo principal é que evidências linguísticas do *Córpus* sejam extraídas, o que nos leva a crer que é um campo de estudos que constitui uma área de conhecimento de bases teóricas próprias e uma maneira de fazer análise linguística específica.

Na Linguística de *Córpus* trabalha-se com dados reais tão exaustivos quanto possível e que portanto possam reproduzir com a máxima fidelidade a realidade linguística. Consequentemente, o volume de dados torna-se tão gigantesco que só uma codificação, ordenação e organização desses dados pode permitir seu uso sem que os pesquisadores fiquem naufragados em um mar de informações que não conseguem manipular. (BIDERMAN, 2001, p. 81).

A Área da Linguística de *Córpus*, acrescenta Oliveira (2009), está em expansão e possui uma história ainda recente se comparada a outras subáreas linguísticas. Todavia, existem fatores

que podem acelerar ou reter o seu desenvolvimento. A seu favor, temos o fato de que é uma área altamente relacionada ao uso computacional. Como a tecnologia vem se desenvolvendo aceleradamente, futuramente poderemos contar com máquinas ainda mais robustas para armazenamento de quantidades ainda maiores, tornando os *córpus* cada vez mais complexos. Mas para analisá-los, programas cada vez mais sofisticados são necessários e para que sejam criados e desenvolvidos dependemos de pesquisa de diferentes áreas que trabalhem em colaboração.

Halliday (1993, p. 33) já havia se surpreendido com a oposição criada por alguns pesquisadores entre a Linguística de *Córpus* e a Linguística Teórica como se fossem divergentes. Naquela época, a Linguística de *Córpus* já era vista como altamente teórica no sentido de modificar os pensamentos sobre o léxico e sobre padrões vocabulares, inclusive sobre a gramática. Com os novos dados surgidos a partir do *córpus*, pode-se trazer problemas para as teorias, a dicotomia acaba sendo utilizada entre teoria e dados quando seria mais adequado considerar a complementariedade entre eles, cada lado alimentando e redefinindo o outro de forma constante

Kaplan (2002, p. 18) afirma que a Linguística de *Córpus* está estreitamente ligada aos desenvolvimentos da Linguística Aplicada, cuja previsão é que possui uma ligação ainda maior com a Linguística Descritiva. Para ele, o desenvolvimento da Linguística de *Córpus* revela fatos sobre usos linguísticos e de variações entre registros, indispensáveis para lidar com questões práticas, muitas vezes incompatíveis com a maioria dos modelos teóricos da linguística.

Por outro lado, Oliveira (2009, 51) acrescenta, que estudar a partir de *córpus* traz caracterizações pela busca de tendências, probabilidades ou padrões de ocorrência dentro de uma grande quantidade de dados. Os números nesses casos servem de base para que os padrões sejam identificados e para que possam ser interpretados pelos pesquisadores. Os resultados quantitativos produzidos baseados em *córpus* indicam numericamente o que deve ser discutido a luz dos distintos posicionamentos teórico-metodológicos para que sejam entendidos. Do mesmo modo que o *córpus* traz somente evidências linguísticas e não informações, os números que são extraídos dos dados linguísticos também não são informações em si mesmos, precisam ser interpretados pelos pesquisadores para que sirvam de base para outras descrições linguísticas ou para novas propostas de perspectivas teóricas

Sendo assim, podemos propor que a Linguística de *Córpus* seja vista como uma teoria exofórica complementar com os pontos de vista de outras teorias com as quais estabelece as interfaces.

Se considerarmos que uma teoria pode ser entendida como uma perspectiva sob a qual um fenômeno é observado, entenderemos facilmente o porquê de existirem múltiplas teorias de linguagem, que correspondem a diferentes maneiras de se olhar esse mesmo objeto de estudo. (OLIVEIRA, 2009, p. 51).

A Linguística de Córpus em suas interfaces está situada na interdisciplinaridade e na complementaridade, relaciona-se com outras áreas do conhecimento, outras teorias e abordagens linguísticas, que ao fazerem junção de conhecimentos contribuem para um melhor conhecimento do seu objeto comum de estudo que é a linguagem. Dessa forma, pontos de contato podem ser observados entre a Linguística de Córpus e a Linguística Sistêmico-Funcional, a Linguística Computacional, Linguística Aplicada e outras áreas.

Ao pensarmos nos aspectos metodológicos, o interesse maior é identificar as probabilidades de colocação de algumas palavras com outras, em alguns contextos de uso da língua, por isso *concordancers* são utilizados. Em programas computacionais específicos, outras ferramentas computacionais também existem para analisar córpus com base na teoria sistêmico-funcional investigando níveis léxico-gramaticais.

No que diz respeito à Linguística Computacional, esta se relaciona à Linguística de Córpus por basear-se no córpus para extração de evidências linguísticas, por estar ligada à tecnologia e por ter como foco o uso de linguagem em seus estudos. Porém, possui objetivos diferentes, já que a linguística computacional explora as relações entre as áreas da linguística e da informática, possibilitando a construção de sistemas capazes de reconhecer e produzir informações. Apresentadas em língua natural, a maioria dos trabalhos nessa área estão voltados ao PLN e à construção de programas capazes de interpretar ou gerar informações em linguagem natural. A linguística computacional utiliza os córpus para acessar os materiais que necessitam estudar. Isto é, grande quantidade de textos que ocorrem naturalmente na língua.

Pensando nessas quatro características apontadas pelos autores, damos ênfase à última que constitui uma diferença marcante entre o que a Linguística defende como córpus e o que a Linguística de Córpus também define. Compreende-se que a disponibilização de córpus compilados para outras pesquisas é uma característica que já faz parte do córpus, de tal modo que todo o empenho gerado para que seja construído não seja útil apenas para uma pesquisa, já que se tem uma referência padrão de língua ou de variedade linguística que pode servir a outros pesquisadores. Vê-se assim que os dois grandes eixos diferenciadores entre Linguística e Linguística de Córpus são o fato de o córpus ser computadorizado e estar disponível para outros pesquisadores.

A partir da década de 90 adquire um papel crucial nas pesquisas linguísticas o uso dos *córpus*. É nessa época que as contribuições da Linguística de *Córpus*, ou melhor, da linguística computacional começam a ser datadas. Podemos destacar de modo principal o desenvolvimento de ferramentas computacionais para o processamento de língua natural (PLN). Essas ferramentas tiveram um enorme feito para o processamento de *Córpus*.

2.2.1 A questão da extensão

Segundo Hoffmann (1998, p. 25), o tamanho mínimo necessário para o *córpus* nas pesquisas linguísticas diverge de modo amplo. Na pesquisa de linguagem especializada, alguns resultados úteis já foram obtidos com amostras de 35 mil palavras, mas sugere-se 200 mil palavras como mínima. O autor defende que esse tamanho vai depender dos objetivos da pesquisa e do tipo de *córpus*, assim, não se trata de uma fórmula matemática aceita que informa a quantidade e distribuição de palavras que um *córpus* deve ter para ser representativo, contudo a maioria das palavras têm frequências de ocorrência muito baixas e para que elas apareçam em um *córpus* precisa-se que ele tenha um grande número de palavras. O mesmo pode ser dito no que diz respeito aos diferentes sentidos ou significados de uma mesma palavra, há os mais e os menos frequentes, os sentidos mais raros terão uma probabilidade maior de aparecer em um *córpus* maior.

Para Hunston (2002, p. 235), um outro detalhe relacionado ao tamanho de um *córpus* é a velocidade e eficiência do software de acesso a esse *córpus*, da mesma forma, se a capacidade do computador em acessá-lo para obter a listagem das formas do presente e passado de um verbo levar mais de alguns minutos, o pesquisador pode optar por um *córpus* menor que dê resultado tão confiável quanto o de um *córpus* maior, mas para o qual o software trabalhe com mais velocidade.

Segundo Sarmiento (2010, p. 87), o tipo de conteúdo que um *córpus* deveria conter não é especificado. Um *córpus* pode conter desde a obra completa de Shakespeare até instruções expressas de uma caixa de sabão em pó, ou um texto jornalístico sobre time de futebol quando foi campeão brasileiro. Com relação à sua dimensão, também não há um consenso quanto ao seu tamanho mínimo ou máximo.

Outro aspecto influenciador do tamanho de um *córpus*, diz Sarmiento (2010, p. 85) é o objetivo da pesquisa. Para estudar gramática da linguagem falada, por exemplo, *córpus*

relativamente pequenos podem ser suficientes, pois as palavras gramaticais tendem a ser muito frequentes. Por outro lado, itens de baixa frequência vão exigir um *cópus* bem maior, geralmente os *cópus* coletados com base em um projeto de pesquisa linguística específica, tal como o fornecimento de informações sobre frequências para dicionários e para verbetes ou para produção de material didático para o ensino de língua. Muitas vezes, porém os *cópus* coletados sem que haja um propósito específico ficam disponíveis como recurso da língua geral para linguistas, professores de línguas, lexicógrafos, entre outros. Existem vários tipos de *cópus*, dependendo do tamanho da proposta e a forma como foram compilados.

2.2.2 A questão da representatividade

Existe o *cópus* geral contendo muitos tipos de texto. Por ser de cunho geral, provavelmente esse *cópus* não será representativo de nenhum todo e ainda um *cópus* geral precisa ser muito maior do que um *cópus* específico. Muitas vezes, ele é utilizado em contraste com relação aos *cópus* mais especializados. Há também o *cópus* monitor que é projetado para que mudanças atuais em uma língua sejam verificadas. Ele é alimentado anualmente, mensalmente, ou até mesmo diariamente; cresce com muita rapidez, mas a proporção do tipo de texto é constante, de modo que cada período de tempo possa ser comparado com o ano anterior. O *Bank of English* atualmente possui 400 milhões de palavras.

Sarmiento (2010, p. 83) acrescenta que com relação à questão da representatividade envolve ainda conhecer a totalidade, no caso da linguagem, não é conhecida. Deve-se tentar dividir esse todo estimado em partes. Por exemplo, um *cópus* de linguagem jornalística deve trazer esse vários tipos de jornais, desde os populares aos mais tradicionais, devem incluir também textos de diferentes seções, como esportes, negócios, editoriais, para que seja considerado como representativo e equilibrado. Deve incluir o número aproximado de palavras em cada categoria, negócios nos populares, negócio nos tradicionais, esporte nos populares e assim sucessivamente.

2.2.3 Tipos de cópus

Sarmiento (2010, p. 85) acrescenta que há ainda o *cópus* comparável: dois ou mais *cópus* em línguas diferentes ou em variedades de uma língua, eles são compilados com as mesmas diretrizes, isto é conterão a mesma diversidade de gêneros. O exemplo mais citado é ICE (International *Cópus* of English) que contém mais de um milhão de palavras. Há também

o *Córpus* paralelo (dois ou mais *córpus* em línguas diferentes), contendo os textos que foram traduzidos de uma língua para outra ou que foram produzidos de forma simultânea de duas ou mais línguas. Além dos tipos de *córpus* citados acima, Hunston (2002, p.14) adiciona ainda os seguintes:

O *Córpus* aprendiz: coletânea de textos e redações produzidos por aprendizes de uma língua cujo propósito é identificar em que aspectos esses aprendizes diferem-se entre si e em relação à falantes nativos. O *córpus* pedagógico que consiste na linguagem exposta ao aprendiz podendo estar em livros didáticos e gravações. O *córpus* histórico diacrônico: de diferentes períodos de tempo, utilizado para ver igualmente o desenvolvimento de diferentes aspectos de uma língua ao longo dos anos e o *córpus* especializado: um *córpus* contém um tipo específico de texto ou gênero, como os resumos (abstract), artigos acadêmicos sobre um assunto específico, conversas telefônicas e etc. É um tipo de *córpus* que tem por objetivo ser representativo de certo tipo textual de linguagem.

2.2.4 *Visão panorâmica de projetos e definições de córpus*

Vejamos, de acordo com Aluísio e Almeida (2006, p. 157), alguns *córpus* disponíveis na web para pesquisa:

Arquivos da folha: está disponível na web o texto integral de todas as edições do jornal desde 1994; é um material extremamente útil para buscar conteúdos ou para testar frequências e o emprego de determinadas palavras no uso da língua e no gênero jornalístico

Lácio-Web: é um projeto organizado pelo núcleo internacional interinstitucional de linguística computacional em parceria com o Instituto de matemática e estatística e a faculdade de Filosofia, Letras e Ciências Humanas, todos pertencem à USP. O *Lácio-Web* disponibiliza na web vários *córpus* do português brasileiro de escrita contemporânea e ferramentas linguísticas computacionais como contadores de frequência, concordanciadores e etiquetadores morfossintáticos.

O projeto COMET: composto por *Córpus* multilíngue para ensino e tradução, elaborado junto ao centro interdepartamental de tradução e terminologia da Faculdade de Filosofia, Letras e Ciências Humanas da USP, disponibiliza o *Córpus* eletrônico que objetiva servir de suporte às pesquisas linguísticas, principalmente nas áreas de tradução, terminologia e ensino de línguas. É composto por três sub*córpus*: técnico-científico, *córpus* multilíngue de aprendizagens e *córpus* de tradução.

Línguateca: é um centro de recursos para o processamento computacional da Língua Portuguesa que objetiva servir a comunidade dedicada ao processamento do português em seu site. Os corpúscos disponíveis são bem públicos. Corpúscos de aproximadamente 180 milhões de palavras em português europeu; corpúscos de cerca de 24 milhões de palavras em português brasileiro com base no jornal Folha de S. Paulo e o Corpúscos paralelo que baseia-se em textos em português e suas traduções para inglês e vice-versa.

Está disponível na web uma série de ferramentas que auxiliam no processamento de corpúscos, dentre elas a *WebCorp* que é um conjunto de ferramentas que permitem acesso à web como recurso linguístico permitindo a extração de fatos sobre várias línguas como se a *Web* fosse um Corpúscos. Pode ser usado por pesquisadores e professores de língua que tenham interesse em analisar o uso de determinadas palavras e expressões, especialmente os neologismos e palavras raras que não aparecem em dicionários e em corpúscos padrões

O Unitex: é um conjunto de programas de processamento de corpúscos linguísticos que possui uma interface gráfica em Java. O que permite que a ferramenta possa ser cortada por uma série de plataformas sem que o seu desempenho seja perdido durante o processamento do *Corpúscos*. Oferece recursos linguísticos como dicionários e tabelas lexicogramaticais, possui dicionários com palavras simples e compostas de um idioma, além de informações gramaticais sobre cada palavra. As palavras são representadas em forma autômata de texto e as tabelas mostram as propriedades de algumas palavras.

No processo de conversão de um arquivo de texto ainda não formatado, o texto original é segmentado em sentenças de unidades lexicais. Além do mais, repetições desnecessárias de caracteres, separação e espaços, linhas e tabulações são removidas e formas não ambíguas do texto são normalizadas para se ter operações de busca que sejam simplificadas. Essas normalizações são definidas pelo próprio usuário e não pode ocorrer para palavras ambíguas.

Concordanciador: está presente na ferramenta e possibilita que padrões de expressões regulares sejam buscados, sequências de símbolos reservadas são utilizadas como forma de denotação a uma expressão regular, é possível representar nas expressões regulares informações codificadas nos dicionários.

O Unitex fornece recursos para tratar dicionário, é possível que um dicionário seja comprimido e verificado quanto a erro de formatação ou ordenado. Caso ainda não esteja em ordem alfabética, o usuário pode definir os próprios critérios de ordenação por um arquivo chamado *Alphabet_txt*, os mesmos símbolos podem ser ordenados de maneiras diferentes em concordância com o idioma em uso. Alguns códigos gramaticais permitem a flexão automática de uma forma canônica, formas flexionadas podem fazer parte de um novo dicionário que pode

ser automaticamente gerado pelo Unitex a partir do dicionário original e de uma gramática de flexão definida previamente.

Vários projetos envolvendo *córpus* têm se desenvolvido nos últimos anos. Vejamos, ainda, de acordo com Aluísio e Almeida (2006, p. 158), o detalhamento de alguns outros projetos que partem dos princípios da linguística de *Córpus*.

O NILC (Núcleo Interinstitucional de Linguística Computacional) possui um *córpus* do português do Brasil que foi compilado a partir de 1993, contém cerca de 35 milhões de palavras com textos em prosa, divididos em subgrupos de textos corrigidos, não corrigidos e semicorrigidos. Contudo, ele apresenta alguns problemas: a classificação textual é problemática pois o *córpus* foi construído sob demanda, as amostras foram integrando as categorias textuais de acordo com que foram adquiridas a partir de parâmetros irregulares de classificação.

Sobre a quantidade de textos, é importante dizer que alguns conjuntos de textos são pouco representativos, isto é, quantitativamente insuficiente com relação ao rótulo que carrega, como no caso dos textos jornalísticos. O *córpus* científico, por exemplo, traz poucas teses e dissertações completas e em sua maioria são relacionadas à área de informática. A quantidade de textos é um empecilho para que o *córpus* seja aproveitado para pesquisas gerais.

Alguns textos foram compilados e regularmente são relacionados ao padrão de amostragem aplicada em quase todos CN (*Córpus* NILC). O procedimento ideal é o de trazer textos integrais, mas essa regra foi quebrada e obras parcialmente compiladas poderão ser encontradas. O acúmulo de textos é também uma característica insatisfatória de determinados conjuntos do CN. Por conta de uma escolha de formatação das amostras do *Córpus*, optaram por anexar no único arquivo muitos textos pequenos, o que ocultou algumas especificidades textuais, como por exemplo as diferenças de assunto e autoria.

De sucesso inestimável por conta da importância de um recurso que se baseia em *córpus* de uma língua para o avanço de estudos variados e para a construção de sistemas computacionais, o processamento de língua natural (PLN) tenta preencher uma lacuna em termos de recursos para pesquisa e suporte à criação de ferramentas para a língua portuguesa no Brasil. Os *córpus* disponíveis são: *Lácio-Ref* (*córpus* aberto composto por textos escritos em português brasileiro respeitando a norma culta, possui 4278 arquivos e totalizam 8291818 ocorrências); *Mac-Morpho* (*córpus* fechado e anotado morfossintaticamente, formado por artigos jornalísticos da Folha de São Paulo e dos cadernos de esporte, agronomia, informática e cotidiano. Por último, o *Par-C*. (*córpus* aberto, paralelo, que possui textos de um ano de edições da Revista *Pesquisa Fapesp*, 646 textos em cada língua e 893.283 ocorrências).

Outro projeto é o TermEx, extração automática de termos e elaboração colaborativas de terminologias para intercâmbio e difusão de conhecimento especializado. Teve início em 2003, constituído por uma parceria entre a UFSCar (Universidade Federal de São Carlos) e a USP (Universidade de São Paulo) de São Carlos com o objetivo de pesquisar e implementar métodos para extração automática de termos e gerar um ambiente computacional que auxiliasse na pesquisa terminológica, elaborando um dicionário de terminologias para a área de revestimento cerâmico; os textos foram agrupados dos anos de 1996 a 2003, totalizando 196 textos de 78 páginas com aproximadamente 4000 palavras.

Foram passados para o formato TXT a partir da ferramenta EXTEX (extração de textos de ficheiros formatados). Essa ferramenta ao realizar a transformação, não deixa como o texto original, traz a junção de algumas palavras para reservas e índices de referência bibliográfica, as notas de rodapé anexadas às palavras e a hifenização dos textos no formato PDF (*Portable Document Format*). Todos os arquivos do *Córpus* para esse projeto foram pré-processados para que as informações de autoria e filiação fossem retiradas, bem como tabelas, quadros e referências bibliográficas, fazendo com que o tamanho médio do *Córpus* fosse diminuído de oito para cinco páginas. Uma grande lição aprendida com o projeto TermEx foi o fato do *Córpus* não ter sido balanceado, incluindo gêneros diferentes. Foi um erro observado posteriormente quando se procuravam os contexto definitórios e explicativos para a elaboração do dicionário

O projeto Nano Term intitulado (*Terminologia em língua portuguesa da nanociência e nanotecnologia e sistematização do Repertório vocabular em elaboração de dicionário piloto*) é financiado pelo CNPQ (Conselho Nacional de Desenvolvimento Científico e Tecnológico) e teve início em 2006. Também é uma parceria entre a UFSCAR e a USP de São Carlos tendo como objetivo a Constituição de um *Córpus* em língua portuguesa da nanociência e nanotecnologia, com a busca de equivalentes em português a partir de nomenclaturas inglesas e uma ontologia em língua portuguesa para a elaboração do primeiro dicionário de N&N em língua materna

O *córpus* foi construído a partir da exploração de textos existentes em língua portuguesa. Os tipos de textos que compuseram o gênero científico foram fundamentalmente dissertações e teses. Todos os textos foram obtidos na *Web* e a compilação foi feita utilizando os itens de busca nanociência e nanotecnologia, genômica. Após as buscas serem realizadas, o prefixo *nano* foi utilizado para marcar termos como nanotubo, nanorede, nanocápsula e etc.

Mencionamos ainda, o projeto intitulado *dicionário histórico do português do Brasil* (séculos XVI, XVII e XVIII) no âmbito do programa “Institutos do Milênio” do CNPQ, foi financiado por este órgão e teve início em 2005. A execução do projeto se deu por uma equipe

composta por 10 universidades, 17 doutores e 17 alunos de graduação e pós-graduação com o objetivo principal de elaborar um dicionário de português correspondente aos séculos mencionados. É importante dizer que não há nenhuma obra lexicográfica sobre seu vocabulário dos primeiros tempos de formação do português brasileiro no Brasil, portanto uma proposta bastante original.

Inicialmente os textos foram selecionados de modo que tivessem que ser escritos originalmente em português por indivíduos nascidos no Brasil ou que estivesse residindo no país há muitos anos. Para que a autenticidade fosse garantida, os documentos foram selecionados abarcando diferentes domínios, a saber: gêneros discursivos e tipologias textuais; os gêneros foram distribuídos nos três séculos que envolvem a pesquisa e foi coletada uma quantidade suficiente de textos para o dicionário da diversidade lexical desses séculos fosse produzida, o que se refere às classes abertas: substantivo, adjetivo, verbo e advérbio.

Todo o material foi digitalizado com os cuidados necessários para que os documentos não tivessem uma imagem suja, gerando caracteres estranhos ou falhas no texto digitalizado que precisassem ser removidos em sua revisão. Todo o material digitalizado foi organizado de forma que cada unidade de texto estivesse constituída em dois arquivos, em forma de imagem e em forma de texto. Cada texto com cabeçalho extenso, organizado em pastas que correspondem às determinadas obras.

De modo geral, procurou-se aqui apresentar uma concepção de *córpus* para a Linguística de *Córpus* a partir de questões essenciais para que o *córpus* computadorizado seja elaborado pensando em suas etapas metodológicas para a compilação e a citação de alguns *córpus* de ferramentas disponíveis na *Web*. Foi traçado aqui um Panorama das práticas da Linguística de *Córpus* com reflexões e relatos que podem nortear futuras pesquisas, levantar outros questionamentos e trazer sedimentação às práticas da linguística de *Córpus* no Brasil.

Para Novodvorski e Finatto (2014), podemos considerar aventuras como mais ou menos adequadas no terreno dos estudos da linguagem no nosso país. O *Córpus Brown* fez 50 anos sendo ele o inevitável ponto de referência quando se propõe uma retrospectiva da linguística de *Córpus* em nível Mundial. A partir de diferentes trajetórias delineadas por uma comunidade de pesquisadores da linguística de *Córpus* do Brasil, podemos pensar que trazer à memória alguns marcos históricos de grau de importância elevado nos fazem pensar em nossos próprios percursos até chegar aqui, o que nos conduz a uma reconstrução do próprio processo na Constituição da área reconhecida como linguística de *Córpus*.

Oliveira (2009, p. 155) explica que a exemplo disso, temos O CORPOBRAS (*Córpus do português do Brasil*) que em 2008, ultrapassou a meta de um milhão de palavras,

equiparando-se a *córpus* considerados como médio e grandes em relação ao seu tamanho. Composto por vinte e sete gêneros discursivos distribuídos em vinte gêneros do discurso escrito, cinco do discurso oral e dois do discurso escrito falado, o *córpus* totaliza 1.361 textos e 1.149.600 palavras, contendo até o momento artigos científicos, cartas ao editor carta de reclamação, de recomendação, pessoais, profissionais, acadêmicas, circulares, contos, crônicas dissertações, editoriais, meios acadêmicos, meios pessoais, notícias de jornal, redações de alunos de ensino médio, redação de alunos universitários, redações de vestibular romances, teses, conversas cariocas ,conversas de crianças, entrevistas acadêmicas, grupos de enfoque, atendimento ao cliente, discursos políticos, roteiros cinematográficos.

O CORPOBRAS, de acordo com Oliveira (2009, p. 156), objetiva uma descrição ampla da língua ou análises específicas, ele é representativo do português do Brasil, considerando que para que um *Córpus* representativo de uma língua seja montado, é necessário que amostras de vários gêneros do discurso sejam armazenadas, sejam gêneros orais e escritos. Para que o *Córpus* representativo do português do Brasil seja criado, é necessário considerar principalmente que os textos devem ser reais, refletindo a língua em uso, produzidos por falantes nativos da língua, ou seja brasileiros; produzidos por falantes, escritores únicos ,de modo que cada texto tenha um autor participante distinto; produzidos em diferentes regiões do país para representar a variedade regional de modo abrangente; selecionados não aleatoriamente, tendo conteúdo variado e principalmente distribuídos em gêneros discursivos também variados para que possam representar a maior variedade possível de ações sociais

De acordo com Sarmiento (2010, p. 153), a LC possui limitações, mas possui também inúmeras vantagens. É a metodologia disponível mais indicada para averiguação de dados reais sobre a língua, pois constitui um método de rigor para que dados atestados na língua viva sejam obtidos, de modo que um conjunto de dados reais possam ser acessados. E se o *córpus* for representativo de uma certa porção de linguagem, de modo claro, as unidades de comunicação mais utilizadas e menos utilizadas aparecerão. Além do mais, pode-se acessar padrões semânticos e associativos que as palavras estabelecem entre si, podem também ser acessados suas colocações, a variação das unidades lexicais e outras características.

Os *córpus*, para Sarmiento (2010, p. 154) simplificaram a vida dos linguistas. Quem deseja verificar os verbos modais, por exemplo, facilmente reúne todos esses verbos em um só lugar para observação. Reunir evidências torna-se uma atividade simplificada, liberando os esforços dos pesquisadores para o ato da interpretação. Com o *córpus* percebemos uma padronização na língua de uma forma mais detalhada do que anteriormente foi sugerido. Regras até então vistas como gerais podem ser aplicadas a contextos específicos e o resultado disso é

o surgimento de novas ideias sobre língua e velhas ideias que precisam ser reavaliadas. Podem ser realizadas as afirmações mais objetivas tendo em vista observações baseadas em *córpus* quando comparadas a observações introspectivas.

The first corpus were deliberate attempts to record the normal usage of members of a language-using community. There was an implication that most if not all of the authors and speakers would be native speakers, and probably fairly accomplished native speakers, of the language concerned. The Brown *Córpus* was called a ‘standard’ *Córpus*, and by being restricted to printed documents it acquired all the features of standardization that are provided in the printing and publishing process. (BONELLI E SINCLAIR, 2005, p. 212)

Uma língua pode ser sabida por falantes nativos de modo perfeito, mas nem sempre sabem o que eles dizem ou como o fazem. Do mesmo modo, o sentido intuitivo priorizado e o mais frequente são discrepantes, os *córpus* provêm a possibilidade da prestação de contas total das características linguísticas e não somente dos traços salientes individuais.

*2.2.5 Linguística de *Córpus* e Ensino*

Berber Sardinha (2010, p. 290) nos traz considerações importantes acerca de como preparar materiais de ensino de línguas com *Córpus* e repensar acerca da LC (Linguística de *Córpus*) educacional no país, que virtudes e mazelas traz em si e de que modo pode tornar-se mais atuante no cenário nacional.

Para que as aulas aconteçam há uma condição essencial, a preparação de materiais de ensino. Contudo, é necessário lembrar que para o ensino com a LC, o professor não pode ser dependente de si mesmo, de sua eloquência ou de um quadro, é necessário ter dados concretos oriundos de um *córpus*. Berber Sardinha (2010, p. 291) acrescenta que existe uma metáfora de que o professor é um artista, além dessa existem outras subjacentes ao trabalho docente. Se cada aula é única e possui muitas horas, estaria essa arte sendo feita com alta qualidade? A prática docente deve estar alicerçada em artefatos prontos, rápidos e reutilizáveis, de preferência baseados em *córpus*.

Além disso, pelo que sabemos, cada aula é única e irreproduzível; então, nesse sentido o trabalho do professor é essencialmente uma arte mesmo. Porém, a realidade também nos mostra que o dia a dia da vida do professor é feito de muitas horas em sala de aula, diante do que é quase impossível ser artista continuamente e produzir arte de boa qualidade. As duas metáforas (do profissional e do artista) estão presentes, se complementam e entram em conflito, muitas vezes. (BERBER SARDINHA, 2010, p. 294).

Apesar de as palavras representarem um ponto inicial de muitos estudos na área da linguística de *Córpus*, não significa que a área lida única e exclusivamente com elas, é uma atividade de pesquisa que tem como foco também, além de outros assuntos, o modo como as palavras se combinam em *córpus* representativos, tanto de produção oral quanto escrita de uma dada população, são combinações lexicais.

Os linguistas de *córpus* não consideram o estudo de uma língua por possibilidades, mas por probabilidades. A compreensão da linguagem e o objetivo das pesquisas não é descrever combinações possíveis de serem realizadas, mas o interesse são as combinações ou padrões mais prováveis de ocorrerem em atividade nessa área. Dá uma atenção especial à seleção de itens. Isto é, o uso de uma palavra irá demonstrar provavelmente certa preferência por um outro item em um determinado ambiente linguístico.

Desde meados de 1980 e início dos anos 1990, a LC vem sendo utilizada no ensino de língua estrangeira. O seu nome inicial era *classroom concordancing*, sua culminância deu-se na Universidade de Birmingham, no Reino Unido. Nesta mesma época operava-se o projeto *COBUILD* de John Sinclair; a peça-chave desse trabalho era a concordância utilizada na lexicografia para descrever por meio de padrões de uso e de frequência o significado das palavras. A lista de concordâncias era feita (listagem de palavras de um *Córpus*) para conferir a frequência de palavras em um texto, textos esses em sua maioria religiosos e de prestígio. Introduzir concordâncias no ensino em sala de aula traz vestígios de sua origem, já que um dos desígnios do *classroom concordancing* é que o aluno torne-se um pesquisador, descobrindo sentidos e expressões da LE como se fosse um lexicógrafo. Os alunos nesse engajamento formulam hipóteses sobre os dados apresentados e tentam confirmá-las ou refutá-las.

Berber Sardinha (2010, p. 296) afirma que os materiais de ensino de línguas ainda hoje são marcados por concordâncias. Esses materiais podem ser chamados na filosofia do *classroom concordancing* de "centrados na concordância", sendo ela a peça central, talvez a única. Todavia, existem outras alternativas que a incorporam: lista de palavras, palavras-chave e pacotes lexicais, não estando restritas à concordância das atividades, enfocando textos escritos, visuais e ou audíveis.

Falando sobre atividades de ensino com *córpus*, baseando-nos em Berber Sardinha (2010, p. 297), podemos diferenciar duas modalidades: a centrada na concordância (a atividade é constituída pela LC); e aquela cujas atividades poderiam existir até mesmo sem os *córpus*. O uso da LC com essas últimas modalidades se justifica pela inclusão dela no ensino de modo geral. Uns dos elementos que mantem vivo o sistema linguístico são a frequência e a co-

ocorrência de itens lexicais. O modo como as construções linguísticas são montadas e as coisas são ditas constituem peças essenciais na lexicogramática.

Ao analisarmos pesquisas com *córpus* percebemos que a linguagem possui um modo esperado, padronizado, típico por conta de seus usuários; ao apresentar correlações diferentes entre uso e contexto, é expressa de forma distinta ajustada ao contexto social, situacional, histórico dos falantes, inserindo-se num sistema probabilístico. Quando ensinamos aos alunos por meio de *córpus*, esse sistema pode ser apresentado de modo mais claro do que com outras metodologias linguísticas.

Com a pesquisa em *córpus* a natureza do conhecimento de uma língua se altera. Dizer que sabemos uma língua implica dominar a forma como ela é dita e escrever de acordo com convenções variadas e específicas de uma língua, seja um gênero ou um registro de um determinado contexto. Para tanto, é necessário conhecer a lexicogramática das escolhas pertinentes e desejadas para uma dada situação. Para que a lexicogramática seja usada eficientemente, as probabilidades das escolhas, a frequência dos elementos e as combinações destes devem ser considerados.

Os aprendizes ao se apropriarem dos conceitos acima tornam-se aptos a lidar com ferramentas de análise de *córpus*, tornam-se pesquisadores e não receptores da língua. A atividade de pesquisa seja totalitária ou específica os confere poder e controle sobre o aprendizado.

O conhecimento acerca da língua é oriundo do professor, do dicionário e da gramática para o *córpus*. A informação esmiuçada da lexicogramática não pode ser acessada intuitivamente. Quando dicionários e gramáticas são baseados em *córpus*, são repassadas informações lexicogramaticais preciosas, porém de modo geral e totalitário e não de contextos específicos que sejam interessantes ao ensino. Além do mais, a maior parte dos dicionários e gramáticas feitos com *córpus* restringem-se à língua inglesa. Existem poucos materiais desse tipo para outras línguas.

A primazia do falante nativo, central no conhecimento da língua, é perdida. Isso acontece porque as probabilidades lexicogramaticais não podem ser preditas pelos falantes, a não ser de modo muito geral. Os *córpus* revelam padrões típicos de uma variedade linguística em particular, esses padrões lexicogramaticais tendem a ser seguidos pela maioria dos falantes nativos. Isto é, podemos acessar de modo confiável a lexicogramática por meio da produção linguística em si, mas não intuitivamente. A lexicogramática de qualquer língua possui uma dimensão colossal, haja vista que registra possibilidades de escolha de todos os itens isolados ou coletivos de uma língua. Não se pode representar a lexicogramática de modo econômico, ela

é extravagante e compõe-se de elementos infinitos modificados continuamente. Poderíamos pensar nas demasiadas opções que não podemos controlar, mas, como sabemos, ao explorar-se um *córpus*, a lexicogramática apresenta-se de modo estável; para a maioria das escolhas são poucas as opções que respondem.

Os aprendizes, munidos dos conceitos acima e capacitados a lidar com as ferramentas e artefatos de análise de *córpus*, tornam-se pesquisadores e não meros 'receptores' da língua. A atividade de pesquisa, seja da 'língua como um todo', seja de variedades específicas que lhes sejam pertinentes, os empodera, visto que eles mesmos podem definir metas, criar hipóteses, coletar e fazer levantamento de dados, observar padrões e tirar suas próprias conclusões tornando-os capazes de assumir o controle de seu próprio aprendizado. (BERBER SARDINHA, 2010, P. 298)

Berber Sardinha (2010, 299) enfatiza que se trata de um processo autônomo, mas não individualista, ao contrário, o ideal é que seja conduzido de forma coletiva, por equipes, os alunos ajudam uns aos outros a superarem dificuldades e a entenderem o processo, sem contar o enriquecimento gerado pela interpretação dos achados. Todavia, contrariando o que possa parecer, o professor não é, em nenhuma instância, dispensável; ele é o mediador e colaborador que auxilia todo o aprendizado. Assim como não se dispensa o professor, também não se renunciam outros elementos de ensino, como dicionários, gramáticas, falantes nativos. São elementos que enriquecem o contexto, contudo com um papel bem menos determinista do que em outras abordagens, não existe mais a palavra final sobre a linguagem.

A própria matéria-prima dos *córpus* (os textos e variedades linguísticas) é o que compõe as modalidades centradas em texto e em gênero. Todo *Córpus* é composto de textos orais, escritos, verbo-visuais de variedades linguísticas. Desse modo, toda análise de *Córpus* revela importantes descobertas sobre textos e variedades linguísticas. A LC disponibiliza em seu arsenal, instrumentos para a análise de textos e gêneros: listas de palavras, palavras-chave, segmentadores, etiquetadores e outros.

Por conta disso, há inúmeros benefícios imediatos para o ensino de línguas com textos ou gêneros. Ainda existem possibilidades para o desenvolvimento de novas ferramentas complementares às que já existe, visando especificamente ao ensino. Com as existentes, excelentes materiais de ensino focados em texto e gênero podem ser criados. A preparação de materiais de ensino pode ser sistematizada da seguinte forma:

- Material: objeto que veicula e dá suporte à atividade;
- Atividade: sequência de trabalhos veiculados pelos alunos em um material de ensino.

- Texto de apoio: parte do material que possui uma porção de linguagem a ser explorada na atividade. Podem ser um ou mais exemplares.
- Instrumento de apoio.: parte do material que contém concordâncias, lista de palavras, palavras-chave, dentre outras ferramentas de análise de Córpus.
- Exercícios: perguntas, quadros etc. a serem trabalhados pelos alunos a partir do texto de apoio.

Considerando a modalidade de atividades centradas na concordância, a ideia sustentada é de que os instrumentos podem ser úteis tanto para aprender inglês e outras línguas quanto para produzir dicionários. O desejo central é o de tornar alunos pesquisadores, de modo que busquem nas concordâncias regularidades na autenticidade da língua presente no uso. A partir daí, encontraria por si mesmo padrões de uso que responderiam questões importantes do aprendizado de uma língua estrangeira, como o significado de palavras e expressões, o uso de classes gramaticais, questões acerca de textos acadêmicos e da cultura de língua estudada.

Em continuidade, da tradição de levar para o aluno de inglês instrumental, concordâncias, deixando que ele mesmo as explore, surgiu o (*data-driven-learning*), o DDL. Sua definição diz respeito ao uso de concordâncias geradas por computadores cuja finalidade é fazer com que os alunos explorem as regularidades de padronização da língua alvo e as atividades relacionadas às concordâncias.

As abordagens que mais influenciaram a criação do DDL foram as que promoviam a conscientização, a necessidade do aprendiz de língua ter consciência de seu aprendizado. Essa conscientização pode ser subentendida como a compreensão por parte do aluno do porquê aprender e aplicar determinadas habilidades que lhe são ensinadas; de que modo esse aluno as coloca em prática e até onde vai a eficiência dessas estratégias.

Outra ideia fundamental é a de que o aluno se torne um pesquisador, um detetive linguístico, buscando nas concordâncias descobertas de padrões e sentidos nos dados. O professor assume o papel de orientador e não de detentor das respostas, sabendo que diferentes e desconhecidas facetas do uso da língua podem ser descobertas pelos alunos.

Na época, os professores podiam levar concordâncias prontas para a sala de aula ou dar aula em um laboratório de computação, deixando os alunos pesquisarem. Os programas mais utilizados eram o *MicroConcord*, *WordCruncher*, *OCP*, *Mini-Concordancer* e alguns outros. Até hoje o DDL continua sendo usado na preparação de materiais de ensino com *córpus*, não só no ensino instrumental ou de leitura. O objetivo básico das atividades é fazer com que os alunos identifiquem padrões lexicogramaticais nas concordâncias. Os padrões existentes nos instrumentos de apoio não precisam ser de antemão conhecidos pelo professor, porém ele deve estar apto a auxiliar os alunos na busca por padrões mais relevantes sem desconsiderar que

padrões desconhecidos possam existir nos dados. Todavia, esses padrões podem ser analisados e identificados antes de serem ensinados.

Algumas variáveis precisam ser consideradas para a preparação de atividades de ensino nessa linha, como por exemplo: a escolha da palavra de busca (sequência de caracteres) digitada no concordanciador; modo de apresentação: decidir se a aula será dada em sala de aula ou em laboratório; tipo de concordância; a versão mais básica de um material de ensino é a própria concordância, que pode ser: integral, selecionada, editada, preenchida ou lacunada.

A atividade pode ser realizada com uma (os alunos trabalham para descobrir os padrões da palavra) ou mais concordâncias (os alunos comparam os padrões de duas ou mais palavras de busca). Essas concordâncias podem vir sozinhas ou acompanhadas de exercícios.

Para a preparação do material, primeiramente, deve-se ter um *cópus*. O professor pode usar os seus próprios recursos ou recorrer a compilações que já existem. Em segundo lugar, deve-se ter um *software (desktop)*, como por exemplo o *WordSmith Tools*, ou (*on-line*) como por exemplo o CEPRIL (Centro de Pesquisa, Recursos e Informação em Linguagem).

Inicialmente, antes de aplicar a atividade aos alunos, algumas etapas precisam ser compreendidas, como por exemplo, o modo como uma concordância deve ser lida (nunca da margem esquerda para a margem direita), como nos textos tradicionais. Essa leitura deve ter como foco a área central da concordância, na região que rodeia a palavra *nódulo*.

Deve-se dar atenção ao modo pelo qual os padrões lexicogramaticais podem ser observados. O professor precisa explicar aos alunos de que se trata um padrão e como encontrá-lo. Usar concordâncias em sala de aula pode gerar algumas dificuldades quanto ao desprendimento da gramática por parte do aluno, e isso não pode ser encarado como uma frustração pelo professor. A gramática não deve ser esquecida, pois é necessária grande parte das vezes para interpretar as concordâncias e chegar à lexicogramática.

O professor também, por sua vez, assume outro papel, o de orientador em vez de detentor das respostas. Ele precisa aceitar que os alunos podem descobrir facetas do uso da língua que ele desconhece. (BERBER SARDINHA, 2010, p. 298).

O professor ao produzir as concordâncias deve ter alguns cuidados. Primeiramente quanto ao número de linhas (não há um número ideal, mas costumam ser uma página impressa). É necessário que o professor selecione quais linhas vão figurar na concordância a ser usada no material. O ideal é que escolha as linhas mais relevantes pensando no tipo de padrão necessário para que os objetivos da atividade sejam atingidos.

Um princípio essencial é que o processo de seleção não distorça a evidência; as linhas de concordância devem representar a totalidade das características linguísticas e comunicativas dos dados. A distorção se dá por duas fontes: se a seleção for feita com base em critérios linguísticos desejáveis e não reais e se for realizada por critérios pedagógicos justificáveis por si mesmos, porém com efeitos inesperados sobre a amostra em favor de formas e sentidos que expressam.

É necessário ainda que o professor ensine aos seus alunos cuidados com o uso do *software* de busca, deixando claro como ocorrem a formulação de buscas, a ordenação de concordâncias, a exclusão das que são indesejadas, a extração de colocados e a comparação de resultados.

Pensando na aplicação de atividades de DDL que sigam o esquema (padrão - sentido - concordância), podemos considerar duas orientações: a indutiva (cujas evidências de uso são encontradas pelos alunos que farão generalizações sobre formas e sentidos da língua) e a dedutiva: (formulação de hipóteses sobre o foco da atividade antes de iniciá-la).

Berber Sardinha (2010, p. 299) lista alguns prós e contras referentes à utilização do Classroom Concordancing. Vejamos alguns deles:

- Prós: tempo de preparação de atividades relativamente pequeno; autonomia conferida ao aluno; a linguagem passa a ser vista como sistema probabilístico; novos elementos de linguagem são descobertos pelo aluno, muitos deles ignorados em outros materiais de ensino; compreensão de que muitos elementos da linguagem são quantificáveis; o aluno aprende a correlação entre os contextos de uso e os padrões mais frequentes nesses contextos; aprende que ao mudarmos um padrão lexicogramatical alteramos o sentido expresso por ele.

- Contras: se houver análise prévia de concordâncias, o tempo para a preparação da atividade pode ser longo; exposição contínua de materiais baseados somente em concordâncias pode gerar a desmotivação por parte dos alunos; eles podem sentir falta de uma situação concreta de comunicação que demonstrem os padrões existentes;

Algumas concordâncias podem ser buscadas em regras gramaticais, como de costume e não lexicogramaticais; pode ocorrer a tendência por parte de professores de transformar padrões em regras; alunos de cursos não instrumentais possuem menos motivação para lidar com textos escritos do que os alunos de cursos instrumentais; alunos mais jovens podem apresentar dificuldades na disciplina e na concentração ao analisarem as concordâncias, trabalho este que deve ser cuidadoso.

O autor acrescenta que pensando no ensino com *cópus*, a concordância não é a atividade central, pode-se ter atividades tendo como foco o texto de apoio. A concordância assume um

papel secundário, mas não deixa de ser importante, já que o objetivo principal das atividades com ensino de *córpus* é aprender a lexicogramática e a concordância constitui um instrumento indispensável para que o aluno seja engajado nesse aprendizado.

Como resposta a algumas desvantagens do *classroom concordancing* surgem algumas atividades centradas no texto. Os textos utilizados contextualizam uma situação concreta de uso da língua, na qual padrões estejam representados. Um texto contextualiza situações sociais, históricas, de produção e de reprodução, em um gênero específico que tornem a linguagem discurso.

Embora não seja a única opção em termos de atividades centradas no texto, acredito que ilustre muito bem o que seria essa vertente de materiais baseados em *córpus*, com a vantagem de ser feita automaticamente. E preciso que fique claro, contudo, que o professor pode criar esse tipo de atividade sem utilizar esse *software* específico, usando outros programas de análise de *Córpus* e editores de texto para chegar a resultados parecidos. (BERBER SARDINHA, 2010, p. 309)

Por se sentirem mais familiarizados com o texto, os alunos se sentem mais confortáveis do que na lida com concordâncias, assim, barreiras psicológicas que dificultavam o uso efetivo de *córpus* são quebradas em determinados contextos. É bom lembrar que mesmo tendo como centro um texto, a essência da atividade não muda, pois os *córpus* e padrões ainda existem. No entanto, sobre atividades centradas no texto, Berber Sardinha (2010. P. 300) também aponta prós e contras. Vejamos:

- Prós: o texto como foco contextualiza os padrões; economia de tempo na preparação de materiais inéditos. - Contras: como o foco é o texto, o ensino de alguns padrões lexicogramaticais podem ficar em segundo plano; muitos padrões exibidos nas concordâncias podem não estar no texto e a contextualização pode não ocorrer de fato.

Vejamos agora o que Berber Sardinha (2010, p. 301) pontua como prós e contras nas atividades multimídia/multigênero. Nas vertentes anteriores percebemos que os textos escritos eram a fundamentação essencial. Ainda que a LC seja limitada a arquivos-texto, os materiais de ensino nela inspirados não precisam limitar-se a isso.

Prós: considerando os meios atuais, podemos citar o texto escrito para ser lido, ouvido, visto. Entre as mídias podemos citar as tradicionais (jornal, revista, música, telefone, rádio, televisão) e as digitais (Youtube, podcast, sites, e-mail e tantas outras)

Existem incontáveis gêneros e subgêneros. Do ponto de vista teórico existem divergências em alguns pontos e convergências em outros. Tendo em vista a LC, Berber Sardinha (2009, p. 114) elencou algumas características centrais: são relativamente estáveis e com diferentes dimensões; nenhum gênero é absolutamente independente, todos se relacionam

em maior ou menor grau; são socialmente estabelecidos, oriundos da necessidade humana de comunicação em sociedade para que determinados fins sejam atingidos.; revelam as atividades sociais e intelectuais de seus usuários; são definidos culturalmente (tanto no nível macro quanto micro).

Ao longo da história o conteúdo, organização interna e lexicogramática dos gêneros historicamente marcados vão sendo alterados. São internamente sequenciados por meio de estágios e segmentos cujo conteúdo representa os temas e subtemas enfocados de modo coerente.

Contras: Por serem numerosos, nos impede de chegar a um inventário definitivo de todos os gêneros existentes. Temos uma intuição linguística falha, embora acreditemos conhecer os gêneros plenamente. Inúmeras características fogem à nossa percepção e só podem ser levantadas de modo confiável por meio de ferramentas de análise de *córpus* eletrônicos. A lexicogramática que os compõe é diversificada e se revela pela probabilidade diferenciada de emprego de palavras, padrões e estruturas gramáticas típicas. Com a mudança de gêneros de uma mídia para outra e com o surgimento de novas tecnologias, a quantidade de mídias de gêneros tende a crescer ainda mais e tornar-se mais complexa.

A proposta de materiais de ensino baseados em *córpus* surge num contexto de complexidade cada vez maior de gêneros e mídias. Isto por conta da modificação da sociedade, bem como dos gêneros que nela circulam. Assim, professores e alunos terão cada vez mais contato com esses gêneros (produzindo ou recebendo), ao mesmo tempo em que tais gêneros se tornam parte do cotidiano. Para que não fique ultrapassada ou se torne um lugar anacrônico, a sala de aula deve capacitar os alunos a terem bom desempenho no contato com esses gêneros, para não correr o risco também de ser questionada em relação a constituir um local formador de cidadãos aptos a enfrentar os desafios de uma sociedade globalizada, digital e em constante transformação.

Existe uma infinidade de arquivos, salienta Berber Sardinha (2009, p. 116) em formato digital em rede, não é mais necessário (como antes) esperar que os *córpus* sejam disponibilizados por terceiros, os praticantes de LC podem criar seus próprios *córpus* atendendo a seus propósitos. Este é um ponto positivo, haja vista que o cenário dos gêneros digitais sofre rápida modificação já que está atrelado ao desenvolvimento de novas tecnologias e à melhoria da infraestrutura de transmissão de dados. Outro aspecto importante que designa os materiais multigênero/multimídia é que eles dialogam entre si. O usuário de um gênero passa a interagir com outro, geralmente em torno de uma mesma atividade social ou de um mesmo assunto.

O essencial é que nas atividades com gêneros, elementos linguísticos sejam compreendidos para o desempenho da *atividade social* enfocada. Isto depende do encontro com muitos gêneros. É uma proposta que visa capacitar os alunos a lidar com a linguagem envolvida em determinada atividade social.

Assim como as demais atividades, a modalidade multimídia/multigênero, para Berber Sardinha (2010, p. 302) possui alguns prós e contras. Prós: são atividades que reproduzem o mundo real e os objetivos só podem ser atingidos com o engajamento de diversos gêneros, isso requer uma compreensão lexicogramática de vários gêneros. Gêneros de diversos níveis de generalidade são trabalhados com os alunos desde um contexto cultural macro até uma lexicografia específica. Os alunos entram em contato com muitos padrões, além de frequências de palavras isoladas e de classes gramaticais. O trabalho adquire maior coerência por estar centrado numa atividade social pivô que orienta todos os exercícios de ensino. O material pode ser produzido com base em um *template*, o que diminui o tempo de preparação. Muitas perguntas são padrões e podem se aplicar a outro gênero escolhido.

Contras: por se tratar de um trabalho com uma quantidade de gêneros diferentes, a preparação do material é complexa. Pode ser confundida com uma abordagem para o ensino de gêneros, mas não visa habilitar os alunos nisto. Pode ser vista erroneamente como uma atividade comunicativa centrada em um tópico ou como uma sequência de vários estudos isolados. Exige muitos recursos e infraestrutura por lidar com diferentes mídias e requer professores especializados em LC e em teoria de gênero.

Conforme Hunston (2002, p. 236), a falha mais grave do uso de um *córpus* é que ele apresenta a língua fora de seu contexto natural. Quando os textos estudados possuem ilustrações, elas devem por exemplo ser descartadas, em outras palavras, transcrições de dados orais não representam de modo fiel as informações sobre entonação, linguagem corporal e outras características paralinguísticas. É um fato que aponta a necessidade de um *córpus* ser somente uma das ferramentas entre outras em um estudo linguístico.

O trabalho nas listas de frequências continua até hoje derivado de *córpus* eletrônico cada vez maior, como o British National *Córpus* (BNC: Oxford, 1995) e o *Córpus* of Contemporary American (COCA: Davies, 2009), que espalhou-se para outras línguas, como visto em séries recentes de listas de Routledge baseadas em *córpus* de Espanhol, Alemão, Português, Chinês, Checo, Árabe, Francês e Japonês. Obviamente, a frequência aplica-se não só às palavras, mas também às unidades maiores como frases e fragmentos. Embora não seja este o único critério, a ideia básica é que a frequência de formas e significados podem ser facilmente ensinadas em

diferentes pontos no processo de aprendizagem, como argumentado por Cobb (2007, p. 450) para as fases iniciais, ou Schmitt e Schmitt (2014, p. 210) para mais tarde, estágios.

Este tipo de trabalho pode, assim, informar o design dos testes como a escolha e sequência de formas e significados para ensinar e testar, tornam-se mais baseada empiricamente, por exemplo, na concepção de testes TOEFL, *Test of English as a Foreign Language*, (Biber et al. 2004, p. 11) e testes de vocabulário baseados em frequência (Nation e Beglar, 2007, p. 15). A análise de frequência de córpus também pode ajudar a determinar o que os alunos normalmente podem e não podem fazer em diferentes níveis, por análises baseadas em diferenças estruturais qualitativas, como argumentado por Granger (por exemplo, 2009).

O projeto *English Profile* da Universidade de Cambridge é um grande exemplo deste tipo de trabalho informado por ambos os nativos-falantes e córpus aprendiz. A investigação do córpus não só informou o silabário e os testes, como também tem sido a força motriz por trás de muitas outras ferramentas na descrição da linguagem. Todas as grandes editoras têm seguido essa máxima de que hoje é inconcebível produzir um dicionário em uma língua principal sem uma substancial entrada de córpus. A influência não para no léxico, mas também pode ser explorada na produção de manuais de utilização e livros de gramática, tais como o Gramática Longman de Inglês Falado e Escrito.

Para Cobb e Boulton (2015, p. 482), as tarefas e atividades de aprendizagem baseadas em Córpus podem ser projetadas ao longo de um amplo espectro de muito complexo a simples. O professor pode decidir a pergunta, consultar um córpus relevante, e escolher a informação apropriada, que é então modelada em uma atividade com instruções e respostas focadas, levando a resultados pré-determinados.

Isso nos leva ao segundo maior uso do córpus na sala de aula de idiomas, quando os professores consultam os dados do córpus diretamente, em vez de depender dos tomadores de decisão a montante. Primeiro, as ferramentas córpus podem ser aplicadas a textos individuais, ajudando a decidir se um texto é apropriado e em quais elementos focar. (COBB; BOULTON, 2015, p. 480, tradução nossa).

Com o tempo, qualquer ou todas estas decisões e etapas podem, no entanto, ser assumidas pelos próprios alunos. A consulta do aprendiz de córpus envolve técnicas que são essencialmente semelhantes às atividades do linguista de Córpus: "Como um investigador, o aluno tem que formar hipóteses com base na intuição ou na evidência; as hipóteses em seguida, devem ser testadas e rejeitadas ou alimentadas de acordo com outras provas dentro de um modelo global" (JOHNS, 1988, p. 14).

A consulta a um *cópus* pode desta forma focar-se na aprendizagem em si, ou ser utilizada como uma ferramenta de referência ao lado de dicionários e outros recursos tanto para a compreensão quanto para a produção, especialmente da linguagem escrita. Em leitura, os alunos podem verificar rapidamente padrões específicos que podem não ser frequentes o suficiente para justificar uma menção em dicionários.

Outros benefícios propostos, assim como afirma Cobb e Boulton (2015, p. 484), incluem a motivação inerente à utilização das TIC para fins individualizados e relevantes, sempre que os alunos constroem os seus conhecimentos com base nas suas próprias necessidades e interesses; ⁶é, portanto, um trabalho construtivo à aprendizagem de línguas, à descoberta e à resolução de problemas e uso de procedimentos que favorecem o desenvolvimento cognitivo e metacognitivo, o pensamento crítico e o desenvolvimento de habilidades, oportunizando a aquisição de consciência e sensibilidade linguística ao lidar com textos autênticos, bem como autonomia e aprendizagem significativa.

Resta saber se o *cópus* cumpre as expectativas, com benefícios suficientes para justificar o investimento. Conseguir que os alunos explorem a linguagem não é novidade: frequentemente são orientados a comparar frases de exemplo no quadro, ou identificar características de textos escritos ou falados (Boulton e Tyne, 2014). A utilização de *cópus* apenas os moverá para um outro nível, aumentando a quantidade de dados disponíveis para exame, sistematização dos procedimentos de consulta e linguagem de saída, permitindo aos alunos em potencial um maior papel no processo. De acordo com McEnery e Wilson (1997, p. 12), os primeiros usos do *cópus* se deram no final da década de 1960 na Universidade de Aston em Birmingham; outros inícios podem ser encontrados em cursos na Universidade de Nottingham, no início da década de 1970.

O primeiro artigo publicado é por McKay (1980) na Universidade de São Francisco, descrevendo o uso do aluno em materiais impressos à base de *Cópus*; a primeira descrição de concordância prática pode ser encontrada em Ahmad *et al.* (1985) na Universidade de Surrey. Mas a abordagem é em grande parte associada a Tim Johns na Universidade de Birmingham, onde ele e outros colegas permitiram aos seus alunos o acesso ao COBUILD e a outros *cópus* e software na década de 1980 para propósitos pedagógicos.

⁶ This brings us to the second major use of corpus in the language classroom, when teachers consult corpus data directly rather than relying on decision-makers upstream. First, corpus tools can be applied to individual texts, in helping decide whether a text is appropriate and what elements to focus on. (COBB; BOULTON, 2015, p. 480).

Cobb e Boulton (2015, p. 485) dizem que desde então, avanços consideráveis têm ocorrido. Muitos *córpus* grandes estão disponíveis gratuitamente na *Web*, assim como *software* que ajudam na compilação rápida, a partir de fontes de internet em apenas alguns minutos. Há ainda ferramentas simples, estáveis, rápidas e gratuitas, fáceis de usar, frequentemente acompanhadas por vídeos tutoriais e ajuda *on-line*.

Com base nos estudos até a presente data sugere-se, portanto, que o *córpus* possa ser eficaz, e eficientemente comparável a outros tratamentos. O uso de *córpus* demonstra total efeito para uma série de finalidades: tanto para a aprendizagem e uso da linguagem em qualquer lugar no *continuum* lexico-gramatical (incluindo a colocação e o idioma) tanto para fins receptivos como produtivos, bem como em tarefas de leitura e escrita mais extensas ou em tradução.

O trabalho com *córpus*, acrescentam Cobb e Boulton (2015, p. 486), está pronto para expandir-se para além da universidade, onde tem sido amplamente usado. Os seus efeitos podem continuar a ser investigados e as condições do seu sucesso elaboradas. No entanto, inevitavelmente, alguns cuidados devem ser tomados. Anexar uma única figura a uma meta-análise ajuda a fazer sentido a um corpo de pesquisa com limitado risco de prejulgamento ou subjetividade, fornece um critério conveniente para medir estudos individuais do passado e do futuro, e pode ser politicamente oportuno para atrair interesse para a área. Isto, pode levar alguns a suporem que esta é a palavra final, e que nenhuma investigação futura é necessária. No entanto, muito pelo contrário, é essencial notar a variação dentro dos estudos, de modo algum todos produzem os mesmos resultados: os detalhes são tão importantes quanto as principais descobertas.

As exibições gráficas devem ajudar ainda a visualizar a variação em tamanhos de efeito entre indivíduos e estudos, e talvez sugerir pistas sobre o que os maiores tamanhos de efeito têm em comum e, inversamente, o que os subvariáveis são mais dignos de acompanhamento. As formas como os *córpus* são usados e integrados também precisam de um estudo mais aprofundado: como são controlados, e como é que estes se relacionam com os perfis do aprendiz (motivações, estilos ou níveis de proficiência).

Cobb e Boulton (2015, p. 486) questionam se existem alguns alunos para quem o trabalho de *Córpus* é mais ou menos adequado? Talvez o mais impressionante em necessidade de estudo são os efeitos a longo prazo da concordância regular, trabalhar na consciência e sensibilidade linguística, autonomia, motivação, além de notar outras habilidades cognitivas e metacognitivas, e assim por diante.

Os *córpus* encontraram muitos usos no campo do ensino e aprendizagem de línguas nas mãos de tomadores de decisão, professores e alunos. A pesquisa publicada cobre aplicativos de sala de aula para uma ampla variedade de perfis de aluno e para usos extremamente diferentes, de trabalho altamente controlado a totalmente autônomo, de materiais em papel a concordância prática, de recurso de referência a ferramenta de aprendizagem. Essa variedade destaca o papel altamente flexível dos *córpus* - não há uma maneira "certa" de usá-los. (COBB; BOULTON, 2015, p. 496, tradução nossa).

O uso de *córpus* tem crescido no campo do ensino de língua, para professores e alunos. Existem aplicações em sala de aula para uma grande variedade de perfis de aprendizes e para utilizações extremamente diferentes, a partir de um trabalho inteiramente autônomo, desde⁷ materiais impressos até as concordâncias. O papel do uso do *córpus* é altamente flexível, não há um único "direito" ou forma de usá-los.

Cobb e Boulton (2015, p. 487) acrescentam que do ponto de vista da investigação, isto pode levar a uma percepção/fragmentação do campo, que uma meta-análise minuciosa pode resolver. A meta-análise como forma de investigação é, por definição, exploratória, parte de perguntas em vez de hipótese. A consulta de *córpus* tradicional é, de certa forma, um aspecto relativamente marginal, para ser encontrado em poucas salas de aula em todo o mundo. No entanto, é de muitas formas análoga com pesquisas na internet e com o uso de outras tecnologias para consulta dos vastos dados disponíveis, que tem indiscutivelmente tornando-se o modo de aprendizagem dominante na nossa cultura.

De acordo com Cobb e Boulton (2015, p. 486) os investigadores recorrem cada vez mais ao uso de *córpus* para a percepção de longa data de que a maioria das palavras comuns devem ser ensinadas primeiro, realmente existe um bom corpo de investigação para apoiar essa noção (por exemplo Laufer e Ravenhorst-Kalovski ,2010; Nation, 2006; Staehr, 2009). Já o senso comum tem tradicionalmente ditado que se deve focar no vocabulário mais comum primeiro. O que a linguística de *Córpus* tem contribuído na sua essência, é com a identificação dos itens léxicos mais comuns e mais fáceis, em conjunto com os dados informados. Com grandes quantidades de textos informatizados - agora mais facilmente obtidos do que antes. Pode-se gerar uma lista de frequências numa fração do tempo que teria sido necessário para realizar a mesma tarefa à mão. O trabalho inicial nas listas de frequência foi realizado manualmente; os

⁷ *Córpus* have found many uses in the field of language teaching and learning in the hands of decision-makers, teachers, and learners. Published research covers classroom applications for a wide variety of learner profiles and for extremely different uses, from highly controlled to entirely autonomous work, from paper-based materials to hands-on concordancing, from reference resource to learning tool. This variety underlines the highly flexible role of *córpus* – there is no single “right” way to use them. (COBB; BOULTON, 2015, p. 496).

córpus informatizados trouxeram uma contribuição inestimável para a nossa compreensão da relação frequência-utilidade.

Embora a utilização de córpus informatizados para melhorar a análise lexical soe como recente, Cobb e Boulton (2015, p. 487) afirmam que a percepção da necessidade de utilizar um Córpus para ordenar a vastidão lexical para fins pedagógicos já existe há algum tempo. O (GSL), *New General Service List*, uma lista de 2.000 artigos léxicos considerados como especiais, útil para fins de pedagogia linguística, foi publicada pela primeira vez por Michael West em 1953, mas o trabalho de base para o GSL remonta ao início da década de 1910, onde foi percebido que mesmo os professores mais peritos têm noções muito inadequadas e inexatas da frequência relativa e importância das palavras" (Thorndike, 1921, p. 360).

Desde o apuramento manual de Thorndike de 5 milhões de palavras de livros escolares para crianças, muita coisa mudou. No entanto, a um nível qualitativo, o córpus de investigação e as questões concomitantes assumidas por ele e colegas de mais de um século atrás permanecem em grande parte os mesmos. Muito parecido com os de hoje, os investigadores tinham de tomar decisões sobre o tamanho do Córpus necessário (ou adequado), e como e porquê escolheriam os textos que incluíam o Córpus. Uma questão colocada, no entanto, foi: seria o suficiente para simplesmente fornecer uma lista de palavras? Ou, como diz Michael West no seu relatório sobre o início do trabalho de Thorndike, "Devemos contar com o emprego de uma palavra ou duas? Reconhecendo a importância de tais perguntas, Thorndike juntou-se a Irving Lorge para melhorar a lista existente. Os investigadores voltaram-se aos dados originais de Thorndike (1932) e cruzaram os dados de cada palavra com sentidos individuais no *Oxford English Dictionary*.

Além disso, a lista tornou-se ainda mais acessível pela adição de percentagens, facilitando a priorização do sentido ensinar/alvo primeiro. É esta versão da lista, publicada sob o título *General Service List of English Words* (West, 1953), que provou ser uma das listas mais duradouras e influentes do vocabulário inglês, informadas pelo Córpus. Poder-se-ia adivinhar que foi a capacidade do Ocidente de fazer a lista mais "compreensível" que ajudou a tornar a lista tão bem sucedida; no entanto, pode ser também este mesmo atributo que eventualmente se torne uma das principais fraquezas da lista.

Em menos de uma década após a publicação do GSL, a natureza do Córpus em linguística começou a mudar. Outra lista que trouxe benefícios com a utilização de um computador foi a Avril Coxhead's Academic Word List (Coxhead 2000). Não muito diferente de Thorndike e as motivações dos colegas para as suas listas, Coxhead quis identificar palavras que os estudantes precisavam saber na escola; no caso da *Academic Word List* (AWL), o

interesse estava mais centrado no sucesso no ensino superior. Antes da AWL, existiam outras listas acadêmicas que tinham sido publicadas e que não utilizavam *córpus* informatizados, tais como a Lista de Palavras da Universidade, que por sua vez foi informada pelo Champion e Elley (1971). No entanto, o *Córpus Coxhead* foi recolhido em formato eletrônico, e era muito maior do que a maioria com 3,5 milhões de palavras, constituídas por quatro sub*córpus* equilibradas (artes, comércio, direito e ciência). As palavras eram incluídas no *Córpus* se não constassem do GSL (West, 1953), e se ocorressem pelo menos 10 vezes em cada uma dos sub*córpus* (e em pelo menos 15 das 28 disciplinas e áreas).

Tal como Thorndike e as que o seguiram, a unidade de contagem não eram formulários de palavras, mas formulários "base" que incluíam qualquer formas derivadas, o que Bauer e Nation (1993, p. 11) chamavam de uma "palavra família". A análise de Coxhead acabou por perfazer um total de 570 famílias de palavras que preenchiam os critérios para a inclusão na AWL. No entanto, podem ser apresentados argumentos contra a suposição de que, se alguém entender uma palavra lematizada, é razoável supor que as mudanças na morfologia não devem apresentar demasiadas dificuldades.

Embora o GSL do Ocidente (1953) tenha servido bem os professores e investigadores durante vários anos, os problemas com a lista vêm surgindo ao longo do tempo. A lista acadêmica de palavras representou um importante avanço na pesquisa de vocabulário informado em *Córpus* na área acadêmica; no entanto, ao contrário da GSL, alguns problemas potenciais com a lista foram apontados por investigadores desde a sua publicação inicial. Gardner e Davies conceberam uma nova lista, que incluiria lemas e não famílias de palavras, e não excluíam necessariamente palavras gerais de alta frequência se fossem encontradas em *córpus* inglês acadêmico. Os *córpus*, por sua vez, seriam muito maiores do que os 3,5 milhões da AWL.

Assim, reunindo textos de uma ampla faixa de disciplinas acadêmicas, Gardner e Davies extraíram uma lista acadêmica de "núcleo" de um *Córpus* de mais de 120 milhões de palavras. Tendo em conta a importância dos dados relacionados com a gama de cada lema (a sua frequência nas disciplinas acadêmicas) e dispersão, os investigadores chegaram a uma Nova Lista de Vocabulário Acadêmico (AVL) de pouco mais de 3.000 palavras (a lista completa pode ser explorada em www.wordandphrase.info/academic). Os autores portanto, parecem justificados na sua afirmação de que a AVL pode agora ser considerada "a lista mais atual, exata e abrangente do núcleo de vocabulário acadêmico existente hoje em dia.

De fato, já houve uma série de tentativas de desenvolvimento de listas de artigos com várias palavras. Biber *et al.* (1999), por exemplo, usando o termo "pacote lexical" para

descrever "sequências de formas de palavras ocorram certo número de vezes num cópús. Embora os pacotes lexicais nem sempre sejam estruturas gramaticais completas ou idiomáticas, de acordo com os proponentes da investigação do pacote lexical, podem ser considerados "blocos básicos de construção do discurso" (Biber, Conrad, e Cortes 2004, p. 371).

Biber *et al.* (2004), utilizando um Cópús de mais de 2 milhões de palavras de universidade, ensino e manuais escolares, procurou identificar os feixes lexicais mais comuns no gênero acadêmico. O fato das expressões se repetirem no cópús é sem dúvida importante, e potencialmente de grande valor pedagógico. Embora estes pacotes em si mesmos possam não parecer ter muito significado, não é essa a questão; como salientam Biber *et al.*, os pacotes lexicais devem ser pensados como "fatos descritivos que requerem explicação" (2004, p. 400).

Por outro lado, poderia também dizer que tal lista provavelmente não é de grande utilidade óbvia para parte dos professores ou estudantes de línguas, e por isso talvez possua um limitado valor pedagógico em termos práticos. Vlach e Ellis (2010, p. 111) procuraram compilar uma lista da fórmula mais útil com sequências utilizadas tanto em inglês acadêmico escrito como falado. Para tal, os autores utilizaram uma combinação de cópús, incluindo o *Michigan Cópús of Academic Spoken English* (MICASE) e os ficheiros BNC que continham inglês falado academicamente, depois optou por um método de identificação para as frases da sua lista.

Os investigadores decidiram evitar as frases puras, uma vez que os feixes lexicais apresentam sequências comuns ao mesmo tempo, apresentando a mesma saliência psicolinguística, ainda que instintivamente não o façam. Por outro lado, os autores continuam, os métodos de identificação de frases que dão prioridade à intuição pura podem ser também abertos à subjetividades.

Os pesquisadores puderam então correlacionar os dados de julgamento qualitativo com as estatísticas quantitativas e, através da regressão múltipla, chegar a uma métrica que poderia ser aplicada a todas as fórmulas quantitativamente derivadas e prever quais valeriam a pena ensinar determinado pelo investigador. MI é uma espécie de medida do quanto uma palavra nos diz sobre outra. Em termos práticos, uma pontuação MI elevada indica que quando uma palavra aparece, é provável que apareça também com a outra.

Como evidenciado pela influência duradoura da lista geral de palavras de Michael West (1953), as listas de vocabulário impulsionadas por dados puramente quantitativos são mais suscetíveis de encontrar o seu caminho para aplicações práticas quando são complementadas por julgamentos qualitativos numa apresentação de fácil utilização. Ao mesmo tempo, a nova GSL e AVL ensinam-nos que devemos ter cuidado com julgamentos subjetivos, uma vez que

essas listas mostram como a amostragem cuidadosa pode levar a listas de vocabulário mais confiáveis e importantes.

No entanto, a nova GSL e AVL também nos recordam a importância de considerar dados semânticos e de frequência para evitar a potencial apresentação de uma imagem enviesada do léxico. Temos também de reconhecer que ao desenvolver e divulgar uma lista própria, passamos a fazer parte dela. É indesejável que alguma vez deva haver uma lista prescritiva, isso tenderia a prejudicar a liberdade dos professores e escritores. O que é necessário é uma forma padrão a partir da qual o infinito de divergências possa ser feito, bem como um conjunto de critérios com intenção fundamentada.

2.3 Lexicologia e Linguística de Córpus

Costa e Silva (2020, p. 36) nos trazem uma pesquisa acerca da riqueza lexical de redações escritas em língua portuguesa. Um trabalho baseado em cálculos estatístico-lexicais para possibilitar uma análise quantiquantitativa. O principal intuito foi o de avaliar o conteúdo lexical, entendendo que o léxico mereça mais destaque nas aulas de língua portuguesa, não apenas nos anos iniciais de escolarização ou no Ensino Médio, mas, sim, ao longo da formação como uma prática sistemática.

Tal pesquisa, com o propósito de descrever, analisar e problematizar o repertório lexical em redações escritas em língua portuguesa, trouxe contribuições aos estudos vinculados à Lexicologia e à Linguística de Córpus não apenas para os professores da licenciatura em Letras, mas, também, para os professores já atuantes e para professores em formação inicial. O corpus da pesquisa em pauta foi constituído por 199 redações escritas em um processo seletivo para acesso ao Ensino Superior. Para a análise do *córpus* foi utilizado um programa de computador muito útil na operacionalização da descrição em Linguística de corpus, o *WordSmith Tools*.

Costa e Silva (2020, p. 37) utilizaram o método type-token ratio: a razão forma/item (ou vocábulo/ocorrência) expressa em porcentagem. O total de formas foi dividido pelo total de itens dividido por cem. Como por exemplo, na frase, “o gato viu o rato”, o valor da razão fornecido pelo programa seria 8,0. Valor obtido assim: $4 \div (5 \div 100)$.

Em qualquer *córpus* linguístico são perceptíveis algumas constantes na distribuição das palavras. Dessa forma, observando-se os valores de riqueza lexical, analisou-se na pesquisa, o conteúdo presente nas redações que constituíram o *córpus*.

De acordo com os resultados obtidos por Costa e Silva (2020, p. 38), o aumento de insumo lexical que o aprendiz tem pode impactar em seu desempenho em uma redação, por

exemplo – a depender da exposição e emprego consciente do léxico. Os autores afirmam, nesse sentido, que “é oportuno que esses indivíduos aumentem seu repertório lexical ao longo da formação no que se refere às habilidades de compreensão e produção porquanto sua produção linguística tende a ser avaliada, especialmente no contexto escrito.” (COSTA; SILVA, 2020, p. 39)

Os resultados obtidos com a pesquisa mostraram que as escolhas lexicais das redações foram feitas para atender a um determinado propósito comunicativo, nesse caso alguns termos escolhidos embora não de forma tão ampla induzem-nos a pensar acerca do título escolhido. Costa e Sila (2020, p. 40) salienta que, ainda que o conhecimento lexical seja vasto, é necessário que o pensamento seja organizado de tal modo que ocorra um desencadeamento coerente das estruturas linguísticas na modalidade escrita do texto.

Sabendo que a coesão lexical ocorre quando a intenção é evitar repetições; substituindo termos por pronominalizações, sinônimos ou expressões equivalentes fica claro na análise desvios nesse aspecto, haja vista que determinados termos foram por vezes repetidos, como por exemplo, a palavra país, desenvolvimento, desenvolver, desenvolvido.

Outro aspecto analisado são os desvios de concordância verbal (ocorre quando o verbo se flexiona para concordar com o seu sujeito) e nominal (é a concordância em gênero e número entre os diversos nomes da oração), revelados no texto, como por exemplo nos fragmentos “*algumas empresas são injusta também com o seu empregado*” e “*mas com o tempo os injusto vão sendo tirado do poder*”.

Finalmente, os autores concluíram com a pesquisa, que não se pode negar que os procedimentos estatísticos são muito úteis e devem, sempre que possível, serem utilizados nas pesquisas sociais. Afinal, a Estatística é uma ciência e sua contribuição para o desenvolvimento de pesquisas nos mais diversos campos do conhecimento é inquestionável.

Evers (2013, p. 58) também nos traz um trabalho importante acerca do processamento de língua natural e os níveis de proficiência do português, a partir da detecção de padrões lexicais e coesivos num enfoque computacional. No contexto do exame de proficiência CelpeBras de 2006-1. A hipótese investigada foi a da possibilidade de se classificar automaticamente os textos submetidos ao exame de acordo com os níveis de proficiência preestabelecidos. Foram 177 textos processados, corrigidos de antemão por corretores humanos. O (AM) aprendizado de máquina foi utilizado para cotejar padrões lexicais e coesivos capazes de distinguir os níveis sob estudo. Esse cotejo foi feito pela ferramenta Coh-Matrix-Port que calcula padrões de coesão, coerência e inteligibilidade textual. No corpus foi possível identificar alguns atributos distintivos entre os textos da amostra, dentre eles, a medida da

riqueza lexical, a quantidade de parágrafos, os conectivos negativos, a incidência de adjetivos e de índice Flesch. Foram obtidos resultados lexicais, coesivos e do aprendizado de máquina.

A argumentação exposta nas seções anteriores, bem como as pesquisas citadas, explicitaram as características da Lexicologia e da Linguística de *Cópus*, consideradas individualmente.

Creemos ser válido, a título de reforço, justificar o emprego de ambas linhas de estudo no presente trabalho.

Destacamos que não foi possível usar somente a Lexicologia por conta de que muitas investigações de uso da linguagem se mostram inviáveis ou impossíveis, devido às grandes quantidades de linguagem e de muitos contextos, por conseguinte, a abordagem baseada em *cópus* facilita este trabalho, no qual os padrões de associação representam relações quantitativas. Contudo, a interpretação funcional (qualitativa) é também uma etapa essencial em qualquer análise baseada em *cópus*, e por isso não há como não discutir a relação entre técnicas quantitativas e qualitativas. Ao recorrermos à Linguística de *Cópus*, precisamos entender que uma análise baseada em *cópus* é empírica e analisa os padrões reais de uso em textos naturais; utiliza uma coleção grande de princípios de textos naturais, fazendo uso extensivo de computadores para análise, usando técnicas automáticas e interativas.

Essas características resultam em um escopo de confiabilidade severa. A abordagem baseada em *cópus* vem do uso de computadores e estes possibilitam identificar e analisar padrões complexos no uso da linguagem, permitindo o armazenamento e a análise de um maior banco de dados do que poderia ser tratado manualmente. Além disso, os computadores fornecem análises consistentes e confiáveis cuja manutenção de registros é cuidadosa.

É importante observar que análises baseadas em *cópus* devem ir além de simples contagens de estruturas linguísticas, sendo essencial incluir interpretações qualitativas e funcionais de dados quantitativos. O objetivo das investigações baseadas em *cópus* não reside simplesmente no relato de resultados, mas na exploração da importância desses resultados para aprender sobre os padrões de uso da linguagem.

Do mesmo modo, não foi possível usar somente a Linguística de *Cópus*, a Lexicologia fez-se extremamente necessária para embasar essa pesquisa. Segundo Biderman (1987, p. 164) pelo léxico de uma língua natural que se registra o conhecimento do Universo, a partir de nomeações e classificações de objetos e seres. Sendo assim, nomear a realidade constitui o primeiro passo para a aquisição do conhecimento científico. Nos agrupamentos e categorizações feitas, por traços semelhantes ou distintos, alguns objetos foram sendo individualizados e o homem estruturou o mundo no qual vive.

De acordo com a mesma autora (2001, p. 179), qualquer sistema léxico é a somatória de toda a experiência acumulada de uma sociedade e do acervo da sua cultura através das idades. As mudanças sociais acarretam alterações nos usos vocabulares e o universo semântico se estrutura em torno de dois polos opostos: o indivíduo e a sociedade. O léxico das línguas naturais originou-se, portanto, desse processo de nomeação pelo qual o indivíduo apropria-se do real. Ato sucessivos de cognição desse real, das categorizações e das experiências são cristalizados em signos linguísticos, que podem ser chamados de palavras.

3 MATERIAL E MÉTODO

Gil (2002) define pesquisa com um procedimento sistemático e racional que tem como objetivo principal a propiciação de respostas a determinados problemas. Requer-se uma pesquisa quando as informações dispostas ainda não são suficientes para a resposta a um determinado problema ou quando essas informações estão desordenadas não podendo ser adequadamente aplicadas.

Para que uma pesquisa seja desenvolvida servimo-nos dos conhecimentos disponíveis e da utilização criteriosa de métodos, técnicas, bem como de outros procedimentos científicos. Na pesquisa, é desenvolvido um processo que vai desde a formulação dos problemas à apresentação satisfatória dos resultados. Existem muitas razões para a realização de uma pesquisa; intelectualmente, pelo próprio desejo de conhecer e no aspecto prático pelo desejo de fazer algo mais eficiente e eficaz.

Essa pesquisa é de abordagem quantitativa cujo objetivo maior é a descrição. Ressaltemos que a pesquisa quantitativa está enraizada no pensamento lógico positivista cuja ênfase se dá nos atributos mensuráveis da experiência humana. Parte da localização quantitativa de determinados conceitos e busca compreender a totalidade do fenômeno a partir de ideias acerca da forma pela qual os conceitos estão relacionados. Utiliza procedimentos estruturados e instrumentos formais para a coleta de dados mediante condições de controle que captem o contexto na totalidade. Em nenhuma hipótese dispensa a objetividade na coleta e análise dos dados que por sua vez se dão por meio de procedimentos estatísticos.

Com relação ao seu caráter descritivo, tem como objetivo primordial descrever características de um determinado fenômeno estabelecendo relações entre variáveis a partir de técnicas padronizadas de coleta de dados. A pesquisa descritiva exige do investigador uma série de informações sobre o que se deseja pesquisar para descrever fatos e fenômenos de uma determinada realidade. Nela, são obtidos dados que podem ser analisados estatisticamente. As variáveis podem ser quantificadas e é possível realizar correlações entre procedimentos estatísticos tornando-se possível conhecer a margem de erro dos resultados obtidos na amostragem.

Na seção intitulada Material, descrevemos o corpus de pesquisa, os formulários de avaliação e correção de redações enviado aos professores participantes, o software utilizado na correção dos textos, a saber, *Wordsmith Tools*, a ampliação da lista de palavras CorPop por uma lista deslematizada, o uso do dicionário Houaiss, a impressora virtual Doro Writer, a calculadora estatística online e o cálculo amostral.

Na seção intitulada Método, descrevemos as etapas da pesquisa desde a coleta dos textos ao processamento e tabulação dos dados. O planejamento da análise das estruturas gramaticais dos textos após a correção realizada pelos professores e a tabulação dos dados finais são mencionados.

3.1 Material

Descreveremos na presente seção os materiais utilizados na pesquisa, a saber: *cópus* de redações, formulário de avaliação de redações, software Wordsmith Tools, lista CORPOP, Dicionário Houaiss, Aplicativo Doro PDF Writer 2. 15, calculadora estatística online e o cálculo amostral e coeficiente de correlação de Pearson.

3.1.1 O *cópus* de redações

Para Berber Sardinha (2004, p. 18), “um *cópus* pode ser entendido por um conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios”. Desse modo, entendemos que o *cópus* deve ser suficientemente extenso em amplitude e profundidade, de maneira que seja representativo da totalidade do uso linguístico ou de algum de seus âmbitos, disposto de tal modo que possa ser processado por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise.

O *cópus* dessa pesquisa é constituído por 500 redações escritas por alunos de Ensino Médio. No entanto, foi feito um recorte de 50 redações, um sub*cópus*, para ser corrigido por professores de redação. Ressaltamos que as demais redações estão sendo utilizadas para outra investigação que já está em andamento. Essas redações variam entre 10 (dez) e 30 (trinta) linhas escritas. São dissertações sobre o tema: “o carro será o novo cigarro?”.

Esse *cópus* foi disponibilizado pela Vunesp (Fundação responsável pelo vestibular da Unesp). Com a autorização formalizada, as redações foram concedidas em formato digital.

O total de *tokens* contido nelas é de 134.708 (cento e trinta e quatro mil setecentos e oito) e o total de *types* é de 80.771 (oitenta mil, setecentos e setenta e um). Vejamos um exemplo de redação que constitui o *cópus* dessa pesquisa:

O filme Wall-E retrata uma ideia de futuro para a humanidade e tem como proposta a tecnologia versus a natureza que em análise, choca ao demonstrar seres humanos impossibilitados de caminhar e por isso detentores de carros individuais. Observando a sociedade em sua atualidade, é grande a tendência para tal caminho devido aos

baixos investimentos em transportes coletivos e o fetichismo sobre automóveis. No entanto, a constante mudança de atitudes e ideais humanos já demonstram caminhos alternativos e repletos de conscientização. Seguindo essa ideia, os carros serão os cigarros do futuro e toda a humanidade precisa estar ciente e discutir sobre o assunto a fim de ressalvas e melhorias. Momentaneamente, a sociedade vem circulada pelos resquícios da Guerra Fria e o imperialismo dos Estados Unidos na economia, o que acaba gerando adesão a princípios favoráveis a esse país, como a ideia de fetichismo ou seja, supervalorização de produtos como automóveis, privilegiando empresas estadunidenses e a economia desse país. Ademais, há o baixo investimento para transportes públicos ou coletivos, gerando a desconfiança do usuário ao serviço, tendo em vista a falta de compromisso com horários, a alta lotação e o mau planejamento de linhas, causando transtornos e o caos em massa. Dessa forma, em busca de conforto e segurança, aliado a pensamentos consumistas, a sociedade super utiliza automóveis individuais. Ilustra essa ideia o fato de algumas cidades brasileiras possuírem mais carros que pessoas e ainda, a exemplo do transporte público no Canadá cuja população opta por esses meios devido aos investimentos e a boa funcionalidade do serviço. Em contrapartida, o futuro vem em moldes alternativos e não como o purista em Wall -e, visto que a crescente adesão da educação ambiental iniciada fortemente pelos anos 80 e a ampliação dos problemas ambientais, demonstra resultados e maior apoio da massa, que hoje é mais preocupada e engajada com a causa, exibido em consumidores conscientes e exigentes e demonstrado pela abominação ao uso de couro animal e produtos que utilizam animais para testes. Nesse sentido, o futuro contará com o engajamento da população somado à diminuição da ideia de status com o consumo e a melhoria dos transportes públicos, resultando no baixo uso de carros e veículos individuais. Exemplifica essa ideia e a adesão populacional aos ideais ambientalistas movimentos como os feitos pela ativista Greta Thunberg de 16 anos; incentivando a preservação e o olhar político para a causa e demonstrando como as futuras gerações agirão. Em síntese, atualmente a população é moldada por ideais consumistas e tenta aos poucos se desvencilhar desses, demonstrado pelas novas operações que buscam um modo de vida sustentável e no futuro, abdicarão de maquinários e produtos supérfluos como os transportes individuais.

Pensando nos critérios e no modo de seleção do *córpus* dessa pesquisa, a atenção foi dada a alguns quesitos importantes, apontados por Berber Sardinha (2004, p. 18), dentre eles, a composição, a formatação, a representatividade e a extensão.

Resumindo os quatro pré-requisitos supracitados (BERBER SARDINHA, 2004, p. 19), os textos que compõem o *córpus* desse estudo foram escritos por falantes nativos, são autênticos, isto é, não foram produzidos como alvo de pesquisa linguística, tampouco, são frutos de linguagem artificial ou de programação de computadores. O conteúdo do *córpus* foi criteriosamente escolhido atendendo às condições de naturalidade e autenticidade. No que diz respeito à representatividade, tendeu-se a ver se o *córpus* constituía de fato um conjunto representativo do português brasileiro.

Com relação à tipologia do *córpus* dessa pesquisa (BERBER SARDINHA, 2004, p. 20), pode ser definido como de modo escrito, de tempo contemporâneo, cuja seleção é de amostragem e de conteúdo especializado. No que condiz a sua autoria é de língua nativa para a finalidade de estudo.

3.1.2 O formulário de avaliação de redações

Desenvolvemos um formulário no ambiente *Google Forms* para a avaliação das redações. Inserimos algumas questões sobre o perfil dos participantes e, na sequência, as redações com as respectivas possibilidades de nota, variando entre 0 e 10.

Figura 3 - Seção 1 do formulário de correção das redações

Proposta de correção de redação

Caro professor (corretor) este é um formulário de solicitação de correção de algumas produções escritas por alunos de Ensino Médio. A sua correção será de suma importância para a pesquisa em questão. Antes de corrigi-las, responda as questões abaixo.

Qual é a sua formação acadêmica?

- Graduado(a) em Letras, Pedagogia e afins
- Especialista na área de Linguística
- Mestrado em conclusão ou concluído
- Doutorado em conclusão ou concluído

Você é professor de produção textual ou ministra aula de preparatórios para vestibular?

- Sim
- Não
- Outros...

Em qual etapa de ensino você atua?

- Fundamental II
- Ensino Médio
- Preparatório para vestibular

Após a seção 1 Continuar para a próxima seção

Fonte: Dados da presente pesquisa.

Esclarecemos que essas informações qualitativas acerca dos participantes não foram analisadas na presente pesquisa. Há outra investigação em andamento tendo uma abordagem qualitativa.

Figura 4 - Seção 2 do formulário com as correções digitadas

Seção 2 de 52

Redação 1

Da roda ao automovel milhares de anos que não os separam tanto assim, continuam sendo meios de transporte. A humanidade sempre precisou se locomover, não vivemos parados, mas novos meios de locomoção estão em contante evolução, um dia após o outro sempre estamos mudando nosso jeito de se mover de um ponto A para o ponto B.

Mas todo processo de evolução tem seus males, chegamos ao ponto em que ir de um bairro até o outro virou um desafio, com tantos carros estamos entupindo as veias arteriais de nossas cidades e poluindo seu meio ambiente. A máquina que foi criada para facilitar e deixar nossa viagem mais comoda, está dificultando e se tornando uma experiência desagradavel.

Mas lembre-se, moldamos o mundo a nossa volta e estamos em constante evolução, as dificuldades e os problemas estão ai para serem melhorados e resolvidos. O automovel não é um vicio é uma fase, como todos os outros meios de transporte foram, e logo perdera seu pedestal.

Qual a nota da redação acima em uma escala de 0 a 10, sendo 0 a menor e 10 a maior nota? *

1 2 3 4 5 6 7 8 9 10

Fonte: Dados da presente pesquisa.

3.1.3 Wordsmith Tools

Utilizamos o *software* de processamento linguístico *WordSmith Tools*⁸, versão 8 (SCOTT, 2020), publicado pela Universidade de Oxford. O *software* é composto pelas seguintes ferramentas: *wordlist*, *keywords*, *concord* e possui os seguintes utilitários: *renamer*, *text converter splitter e viewer*. Cada ferramenta com ajustes e funções determinados. Trabalha com os princípios abstratos: ocorrência, recorrência e coocorrência.

Na ferramenta *Wordlist*, ocorre a criação de listas de palavras ordenadas alfabeticamente e pela ordem de frequência, essas listas são apresentadas em duas janelas primárias. É criada também uma terceira janela na qual aparecem as estatísticas relativas aos dados usados para a produção das listas. Dentro dessa ferramenta, temos a lista de dimensões do *córpus* e densidade lexical, uma lista de estatísticas descritivas dos arquivos selecionados. Nela, são apresentadas várias contagens relacionadas ao texto: o tamanho dos itens (*tokens*) e as formas (*types*), a densidade lexical simples em intervalos e o número de parágrafos e sentenças.

⁸ Disponível em <https://www.lexically.net/wordsmith/>

Figura 5 - Ferramenta WordList do software WordSmith Tools

N	Word	#	Freq.	%	Texts	% Lemmas	Set
1	#	5	0,35	1	100,00		
2	A	43	3,00	1	100,00		
3	À	7	0,49	1	100,00		
4	ABERTAMENTE	1	0,07	1	100,00		
5	ACABANDO	1	0,07	1	100,00		
6	ACABAR	2	0,14	1	100,00		
7	ACEITÁ-LO	1	0,07	1	100,00		
8	ACESSO	2	0,14	1	100,00		
9	ACONTECE	1	0,07	1	100,00		
10	ADAPTAÇÃO	1	0,07	1	100,00		
11	ADQUIRIR	1	0,07	1	100,00		
12	AFIRMAÇÃO	1	0,07	1	100,00		
13	AGENTES	1	0,07	1	100,00		
14	AINDA	8	0,56	1	100,00		
15	AJUDAR	1	0,07	1	100,00		
16	AJUDARIA	1	0,07	1	100,00		
17	ALCANÇE	1	0,07	1	100,00		
18	ALÉM	2	0,14	1	100,00		
19	ALGO	1	0,07	1	100,00		
20	ALGUNS	1	0,07	1	100,00		
21	ALTERAR	1	0,07	1	100,00		
22	ALUNOS	2	0,14	1	100,00		
23	AMBIENTES	1	0,07	1	100,00		
24	ÂMBITO	1	0,07	1	100,00		
25	AMBOS	2	0,14	1	100,00		
26	AMPLIADA	1	0,07	1	100,00		
27	AMPLIANDO	1	0,07	1	100,00		
28	ANALISAR	1	0,07	1	100,00		

Fonte: Dados da presente pesquisa.

Já na ferramenta *Keywords*, há uma seleção de itens listas de palavras comparando a ordem das frequências com a lista das referências. Esse contraste gera uma lista de palavras-chave estatisticamente diferentes no *córpus* de estudo e no *córpus* de referência.

A forma como o *córpus* de referência é composto influencia as palavras que podem se tornar chave. Por exemplo, um *Córpus* de características genéricas semelhantes ao *córpus* de estudo tende a filtrar eliminando os aspectos genéricos. A filtragem é removida nas palavras-chave na medida em que as características coincidirem. As características que diferirem entre os dois *córpus* tendem a se manter entre as palavras-chave no modo como aparecem no *córpus* de estudo.

É possível identificar matematicamente a quantidade de palavras-chave num *córpus* de estudo. Elas podem ser estimadas estatisticamente quando o tamanho dos *córpus* de estudo e referencial empregados na análise é conhecido. Por meio do *Key Words* torna-se possível quantificar as vezes que algumas palavras foram chave em várias listas. São chamadas palavras-chave-chave as palavras que foram chave em um determinado número de listas.

Já a ferramenta *Concord* gera concordâncias ou listagens das ocorrências de um item específico. Esse item é chamado de nóculo, que pode ser formado por uma ou mais palavras acompanhado do texto ao seu redor (o *cotexto*). No *WordSmith Tools*, o *concord* pode ser usado de forma separada para analisar concordâncias ou também pode ser utilizado em conjunto com

as ferramentas aqui já citadas. O *Concord* estabelece a concordância do item seccionado a partir dos textos das listas de palavras ou palavras-chave.

De modo geral, o programa *WordSmith Tools* aqui apresentado é essencial para a análise, facilitando a identificação e listagem de palavras no seu contexto original.

3.1.4 A lista CORPOP

Utilizamos o *córpus de referência do português popular escrito do Brasil*, denominado *Corpop* (PASQUALINI, 2018), produto oriundo de uma tese orientada pela professora Dr.a Maria José Bocorny Finatto na Universidade Federal do Rio Grande do Sul. No endereço eletrônico (<http://www.ufrgs.br/textecc/porlexbras/corpop/lista.php>), há a disposição dos pesquisadores uma lista de palavras lematizada. Os verbos aparecem em sua forma infinitiva, adjetivos e substantivos em sua forma singular masculina e sem flexão de grau, por exemplo.

No entanto, para os fins dessa pesquisa, a necessidade era de uma lista deslematizada, por conta da variabilidade de palavras necessárias para traçarmos o perfil lexical dos textos. Por serem diversos os textos processados e corrigidos, quanto mais possibilidades de palavras existentes na lista ao ser deslematizada, mais famílias de palavras tornou-se possível de se dividir para o processamento e análise. Esse processo de deslematização será melhor detalhado no método.

Figura 6 - Lista bruta da frequência de palavras na CorPop

```
<Lista bruta das frequências de palavras no Corpop - wordlist completa do corpus. Ferramenta usada: AntConc./>
<PARA REFERIR: PASQUALINI (2018)>
Corpop: um corpus de referência do português popular escrito do Brasil
Autor: Pasqualini, Bianca Franco
Orientador: Finatto, Maria José Bocorny
Data: 2018
Nível: Doutorado
Instituto: Universidade Federal do Rio Grande do Sul. Instituto de Letras. Programa de Pós-Graduação em Letras.
LINK para acesso da tese sobre o CorPop: <http://hdl.handle.net/10183/177566/>

Wordlist CorPop - 2018

#Word Types: 32311
#Word Tokens: 688788
#Search Hits: 0
1      23564  de
2      22724  o
3      22381  que
4      21629  a
5      19212  e
6      10666  não
7      9082   para
8      8650   se
9      8343   do
10     7720   um
11     7132   da
12     7104   com
13     6943   é
14     6336   os
15     6248   em
16     5895   uma
17     4935   no
18     4416   ele
19     4396   na
20     4329   por
21     4258   as
22     3976   mas
23     3489   mais
24     3348   como
25     3186   eu
26     3002   ela
27     2890   você
28     2650   me
29     2601   dos
```

Fonte: Pasqualini (2018).

Essa lista intitulada como Corpop foi inicialmente formatada pela pesquisadora mencionada, a partir do princípio da simplificação de textos, sobretudo institucionais, referindo-se à simplificação como essencial para a ampliação ou democratização do acesso e entendimento de textos por massas de leitores com baixo letramento. Ele foi compilado a partir de textos baseados no nível de letramento médio dos leitores do país.

Suas bases teórico-metodológicas estão interdisciplinarmente inseridas nos Estudos de Linguagem, a saber os Estudos do Léxico e Linguística de Corpus, Linguística Textual e Psicolinguística, dialoga ainda com os estudos de Processamento de Língua Natural (PLN). O desenvolvimento do CorPop originou-se da análise de dados sobre o nível de letramento dos leitores brasileiros e dos traços que poderiam fazer parte do padrão de simplicidade textual em um *córpus* de textos adequadamente nivelados a esses leitores.

Para o CorPop, os textos selecionados foram do jornalismo popular do Projeto PorPopular (jornal Diário Gaúcho), muito lido pelas classes C e D, em que está o leitor médio brasileiro; textos e autores mais lidos pelos respondentes das últimas edições da pesquisa Retratos da Leitura no Brasil; coleção "É Só o Começo" (adaptação de clássicos da literatura brasileira para leitores com baixo letramento, adaptação esta realizada por linguistas); textos do jornal Boca de Rua, produzido por pessoas em situação de rua, com baixa escolaridade e baixo letramento; e textos do Diário da Causa Operária, imprensa operária brasileira produzida também por pessoas dentro da faixa média de letramento do país.

O objetivo maior desse *córpus* para essa pesquisa, foi o de servir como material de referência para pesquisas linguísticas conectadas com a realidade dos falantes, leitores e redatores.

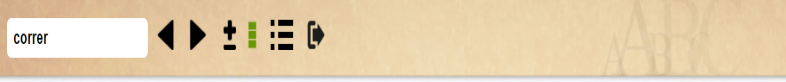
O CorPop se diferencia de outros *córpus* atuais do português, não somente por sua extensão, que é pequena, mas pelo modo como foi planejado e composto, texto a texto, segmento a segmento. Ele foi materializado sobretudo em uma lista lematizada de aproximadamente cinco mil palavras. Esse *córpus* é uma referência, mas salientamos que os dados de frequência precisam ser relativizados.

3.1.5 Dicionário Houaiss

Para a ampliação da lista original da CorPop, realizamos um trabalho de deslematização. Para tal, as flexões verbais foram todas consultadas no dicionário eletrônico *Houaiss* (2009). O Dicionário *Houaiss* possui cerca de 228 500 verbetes, 376 500 acepções, 415 500 sinônimos, 26 400 antônimos e 57 000 palavras arcaicas. Além da quantidade de verbetes, há também as

etimologias de cada palavra e o seu primeiro registro no idioma português. O vocabulário do *Houaiss* abrange as terminologias técnicas e científicas, assim como expressões populares e termos específicos do português europeu, brasileiro, africano e asiático.

Figura 7 - Dicionário eletrônico Houaiss – Conjugações



correr

Resultados

correr (SKIII cf. IVPM)

princ. conj. loc. etim.

Regular

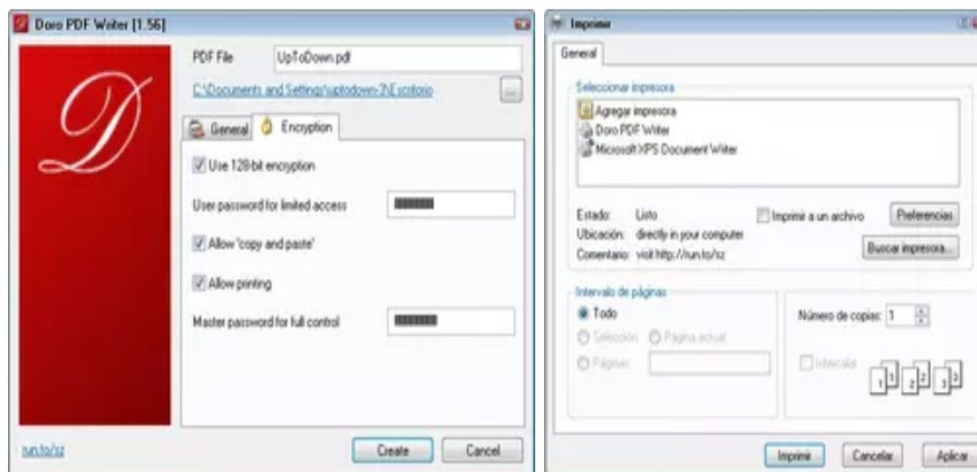
Presente do indicativo	Preterito imperfeito do indicativo	Futuro do presente do indicativo	Futuro do preterito do indicativo	Preterito perfeito do indicativo	Preterito mais-que-perfeito do indicativo	Presente do subjuntivo	Preterito imperfeito do subjuntivo
eu corro	eu corria	eu correrêi	eu correria	eu corri	eu corra	que eu corra	se eu corresse
tu corres	tu corrias	tu correrás	tu correrias	tu correste	tu correras	que tu corras	se tu corresses
ele corre	ele corria	ele correrá	ele correria	ele correu	ele corra	que ele corra	se ele corresse
nós corremos	nós corríamos	nós correremos	nós correríamos	nós corremos	nós corrámos	que nós corramos	se nós corréssamos
vós correis	vós corríeis	vós correríeis	vós correríeis	vós correstes	vós corréis	que vós corrais	se vós corrésses
eles correm	eles corriam	eles correrão	eles correriam	eles correram	eles correram	que eles corram	se eles corressem
Futuro do subjuntivo	Imperativo afirmativo	Infinitivo flexionado	Gerúndio	Participio			
quando eu correr	-	correr eu	correndo				
quando tu correres	corre !	correres tu					
quando ele correr	corra !	correr ele					
quando nós correremos	corramos !	correremos nós	corrido				
quando vós correrdes	correi !	correrdes vós					
quando eles correrem	corram !	correrem eles					

Fonte: Dados da presente pesquisa.

3.1.6 Aplicativo Doro PDF Writer 2.15

Após serem consultadas as flexões verbais no dicionário *Houaiss*, utilizamos um aplicativo específico chamado Doro PDF Writer 2.15. Ele foi necessário para conseguirmos copiar as formas verbais sem os pronomes. Este aplicativo instala uma impressora virtual que nos permite criar e salvar PDFs a partir de qualquer aplicação na lista de impressoras disponíveis no próprio computador.

Figura 8 - Aplicativo Doro – Impressora Virtual



Fonte: Dados da presente pesquisa.

3.1.7 Calculadora estatística online

Após serem corrigidas as redações, foi estabelecida uma correlação entre o perfil lexical e as notas dos corretores, entre a nota do corretor e os types da redação. Isto foi feito num site de cálculos específico (<https://www.socscistatistics.com>). Trata-se de um site que oferece recursos gratuitos para estudantes e pesquisadores que trabalham com estatística nas ciências. Ele possui três seções principais: calculadoras estatísticas, que trata principalmente de testes de hipóteses; calculadoras de valor p, que permitem derivar valores p de Z, t, qui-quadrado e Pearson (r); e estatísticas descritivas - médias, variância, uma calculadora de desvio padrão e criadores fáceis de histogramas e gráficos de barras.

Figura 9 - Calculadora estatística online

The screenshot shows the 'Sign Test Calculator' interface on the 'Social Science Statistics' website. At the top, there are two mathematical formulas for the test statistic t . The main content area includes a navigation menu (Home, Calculators, Descriptive Statistics, Merchandise, Tutorials, Quizzes, Which Statistics Test?, Contact), a Google advertisement, and the calculator's instructions. The instructions state that the calculator is for a repeated-measures study comparing two treatment conditions. Below the text, there are two input fields labeled 'Treatment 1' and 'Treatment 2'. At the bottom left, there is a 'Significance Level' input field.

Fonte: Dados da presente pesquisa.

Por meio deste site, foi possível estabelecer estatisticamente um resultado de influência entre a nota dos corretores e a quantidade de palavras mais formais do texto, palavras fora da lista de palavras utilizada. Essa correlação nos gerou um número que, ao ser comparado a uma escala pré-determinada, também estabelecida estatisticamente, mostrou se a influência do vocabulário na avaliação das redações era fraca, moderada, forte ou muito forte.

A partir dos resultados obtidos, foi possível responder a uma das problemáticas que diz respeito ao fato da qualidade do vocabulário influenciar na avaliação. Com essas respostas torna-se relevante pensarmos em propostas de ensino que priorizem a ampliação do repertório lexical dos alunos e da variabilidade do léxico.

3.1.8 Cálculo amostral e coeficiente de correlação de Pearson

O cálculo amostral é um suporte matemático ao conceito de que uma amostra aleatória de uma população grande tende a estar próximo da média da população completa. Um de seus principais objetivos é determinar a quantidade de elementos necessários para que uma amostra seja composta de modo que os resultados obtidos sejam válidos, sem ultrapassar o que é suficiente. Para o cálculo amostral, são necessárias algumas variáveis: a população (eventos ou indivíduos que compõem o objeto de estudo) e a amostra (que é a parcela aleatória da população).

Para a presente pesquisa, para calcular um valor mínimo de redações que pudesse representar a população, recorreremos à calculadora disponível na página *online* do *Australian Bureau of Statistics*.

Figura 10 - Cálculo amostral

The image shows a web-based calculator titled "Determine Sample Size". It has a green background with white text and input fields. The fields are as follows:

Field	Value
Confidence Level:	95%
Population Size:	43806
Proportion:	0,5
Confidence Interval:	0,05
Upper:	0,55000
Lower:	0,45000
Standard Error:	0,02551
Relative Standard Error:	5,10
Sample Size:	381

At the bottom, there are two buttons: "Calculate" (highlighted with a red box) and "Clear".

Fonte: *Australian Bureau of Statistics* (2020).

Dessa forma, seriam necessárias 381 redações para que suas características pudessem ser extrapoladas para a população. Vale ressaltar que utilizamos um quantitativo de 500 redações, de modo a garantir uma amostra representativa.

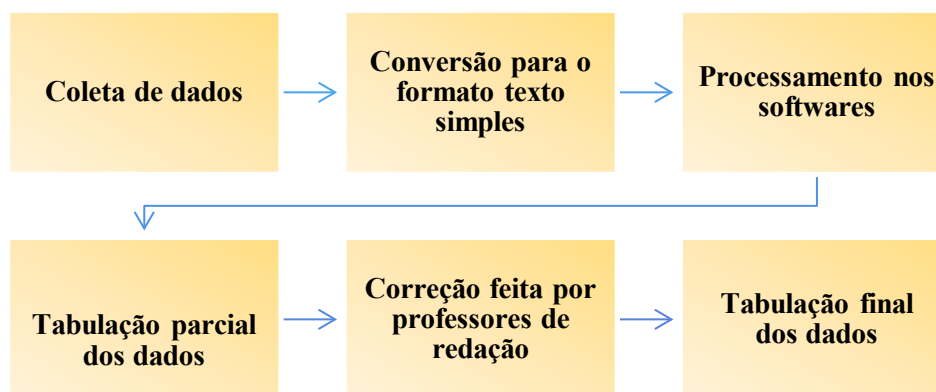
No que tange ao coeficiente de correlação de Pearson, entendemos que para que duas distribuições sejam variavelmente comparadas deve-se usar medidas de variabilidade relativa, a correlação de Pearson é uma delas e é dada por: $CV = S X$ o qual independe da natureza e magnitude da variável X. Esse resultado é multiplicado por 100, para que o coeficiente de variação seja dado em porcentagem.

3.2 Método

A pesquisa foi dividida nas seguintes etapas: a coleta dos textos, a conversão do texto para o formato txt (texto simples), o processamento dos textos nos softwares, a tabulação parcial

dos dados, a correção feita pelos professores de redação, a correlação do perfil lexical e a tabulação final dos dados.

Figura 11 - Fluxograma com as etapas da pesquisa



Fonte: Dados da presente pesquisa.

O primeiro passo foi a coleta das redações realizadas por alunos do Ensino Médio pela Unesp. Essa fundação foi criada em 1979, com personalidade jurídica de direito privado, sem fins lucrativos, cujas atividades são: planejar, organizar, executar e supervisionar o concurso Vestibular da Unesp; realizar vestibulares e concursos diversos para outras instituições e promover as atividades de pesquisa e extensão de serviços à comunidade, na área educacional.

A Unesp atualmente coleta, organiza, analisa e encaminha à Reitoria da Unesp informações técnicas e dados estatísticos relativos aos seus dois vestibulares anuais, assessorando os órgãos colegiados superiores nas decisões relativas às definições das normas no processo e fornecendo o Relatório Vestibular Unesp. Para a aquisição do *córpus*, solicitamos via documento formal um quantitativo de 500 escritas por candidatos ao vestibular do ano de 2019. A solicitação foi atendida e as redações puderam ser processadas para fins da pesquisa atual.

Os textos foram digitados e salvos para o formato de txt (texto simples), sem formatação, exigência do software para que o *córpus* fosse processado. No software WST (*WordSmith Tools*), a quantidade de *tokens* e *types* foi pontuada, permitindo que a variação do vocabulário contido no texto fosse percebida.

Após serem selecionadas as redações, foram digitalizadas e digitadas para a tabulação e análise dos dados. É importante ressaltar que as 500 redações foram digitadas num período de vinte e cinco dias.

O formulário de avaliação enviado aos professores e já mencionado no tópico anterior, foi criado no *Google Forms*, dividido em duas seções. A primeira seção continha perguntas relacionadas à formação acadêmica e à área de atuação dos professores convidados. Já a segunda seção, trazia as cinquenta redações a serem corrigidas, o *cópus* da pesquisa.

A resposta dos professores, bem como a correção dos textos ocorreram num prazo de 20 dias, no período de outubro a novembro. A correção realizada deu-se de forma livre. Cada professor partiu de seus próprios critérios para atribuírem notas aos textos, essas notas variaram de 0 a 10, sendo dez a nota máxima.

Para essa pesquisa, o principal recurso utilizado no *software* WST foi *Wordlist* que fez a contagem das palavras dos textos selecionados gerando alguns cálculos. Os resultados foram apresentados em três telas: as palavras pela ordem de recorrência (da mais para a menos frequente); alfabética, e as estatísticas como *types*, *tokens*, *type/token ratio*. A partir das estatísticas geradas foi possível quantificar a variabilidade lexical contida em cada redação.

Para fins dessa pesquisa, a necessidade era de uma lista de palavras deslematizada; desse modo, partindo da lista de palavras existente, a saber, a *CorPop*, já detalhada, realizamos sua ampliação, no *programa Microsoft Excel*. É importante ressaltar que o *CorPop* também oferece uma lista bruta não lematizada, no entanto, para esta pesquisa eram necessárias ainda mais variações. Essa (ampliação) deslematização foi realizada por Jeane Cardoso Costa (mestranda) e Jéssica Munique Marques Borges (mestranda) que a intitularam como **CorPopex**.

Nessa lista, acrescentamos as variações de cada lema (a classe gramatical pertencente, o plural, as variações de gênero e grau, os superlativos e as flexões dos verbos). Por exemplo, o substantivo abraço recebeu as seguintes variações: abraços, abracinho, abração, abracinhos e abrações; o adjetivo bonito ficou da seguinte forma: bonita, bonitas, bonitos, bonitinho, bonito, bonitona, bonitões, bonitonas, bonitinha, bonitinhas, bonitíssima, bonitíssimo, bonitíssimas, bonitíssimos; para o verbo correr, essas foram as flexões: corro, corres, corre, corremos, correis, correm, corrias, corria, corríamos, corréis, corriam, correrei, correrás, correrá, correremos, correreis, correrão, correrias, correria, correríamos, correríeis, correriam, corri, correste, correu, corremos, correstes, correras, correrá, corrêramos, corrêreis, correram, corra, corramos, corrais, corram, corresses, corresse, corrêssemos, corrêsseis, corressem, correres, correremos, correrdes. É importante ressaltar que todas essas flexões verbais foram retiradas do dicionário Houaiss.

Figura 11 - Lista de palavras deslematizada e ampliada

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
61	ago	substantivo	agos											
62	acolher	verbo												
63	acolhimento	substantivo	acolhimentos											
64	acometer	verbo												
65	acomodar	verbo												
66	acompanhamento	substantivo	acompanhamentos	acompanhamentos	acompanhamentos	acompanhamentos	acompanhamentos	acompanhamentos	acompanhamentos	acompanhamentos	acompanhamentos	acompanhamentos	acompanhamentos	acompanhamentos
67	acompanhante	substantivo	acompanhantes	acompanhantes	acompanhantes	acompanhantes	acompanhantes	acompanhantes	acompanhantes	acompanhantes	acompanhantes	acompanhantes	acompanhantes	acompanhantes
68	acompanhar	verbo												
69	aconselhamento	substantivo	aconselhamentos											
70	aconselhar	verbo												
71	acontecer	verbo												
72	acontecimento	substantivo	acontecimentos	acontecimentos	acontecimentos	acontecimentos	acontecimentos	acontecimentos	acontecimentos	acontecimentos	acontecimentos	acontecimentos	acontecimentos	acontecimentos
73	acordar	verbo												
74	acordo	substantivo	acordos	acordos	acordos	acordos	acordos	acordos	acordos	acordos	acordos	acordos	acordos	acordos
75	acostumado	adjetivo	acostumados	acostumados	acostumados	acostumados	acostumados	acostumados	acostumados	acostumados	acostumados	acostumados	acostumados	acostumados
76	acostumar	verbo												
77	acreditar	verbo												
78	acrescentar	verbo												
79	açúcar	substantivo	açúcares	açúcares	açúcares	açúcares	açúcares	açúcares	açúcares	açúcares	açúcares	açúcares	açúcares	açúcares
80	acumular	verbo												
81	acusação	substantivo	acusações											
82	acusado	adjetivo	acusados	acusados	acusados	acusados	acusados	acusados	acusados	acusados	acusados	acusados	acusados	acusados
83	acusar	verbo												
84	adaptação	substantivo	adaptações											
85	adaptar	verbo												
86	adequadamente	advérbio												
87	adequado	adjetivo	adequados	adequados	adequados	adequados	adequados	adequados	adequados	adequados	adequados	adequados	adequados	adequados
88	adeção	substantivo	adeções											
89	adiantar	verbo												

Fonte: Dados da presente pesquisa.

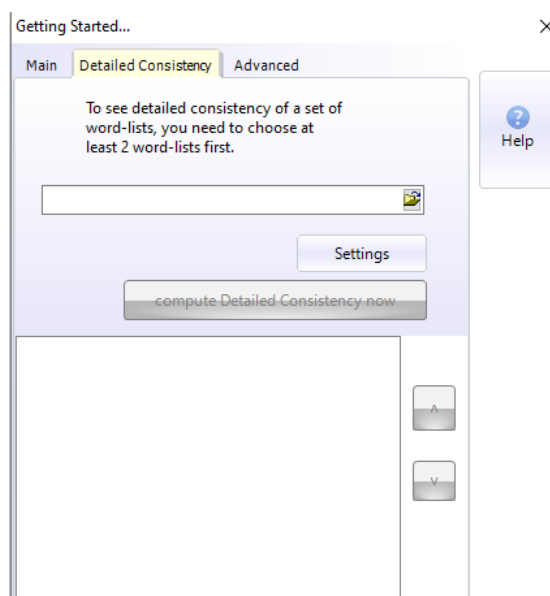
Para que essa lista pudesse ser ampliada, fizemos o *download* do aplicativo Doro, explicado na seção de material. Esse aplicativo foi extremamente útil para as flexões verbais. Seleccionamos cada verbo e colocamos na seção de busca do dicionário Houaiss na versão eletrônica; após a busca do verbo e de todas as suas flexões, a impressora Doro era selecionada, gerando automaticamente um arquivo em PDF; este arquivo ao ser copiado, nos permitia seleccionar somente as formas conjugadas, sem os pronomes; em seguida, as palavras eram coladas em uma segunda aba do *Excel* e por fim copiadas mais uma vez e transportadas para a aba 1 na posição do verbo. Esse trabalho de deslematização foi minucioso, realizado num período de três meses. É importante ressaltar que tanto a planilha quanto o arquivo TXT ficarão disponíveis para futuras pesquisas.

Para que o perfil lexical das redações fosse traçado, a lista de palavras deslematizada foi utilizada no programa de processamento *Wordsmith Tools*. Ao traçarmos o perfil lexical das redações, entendemos de modo mais claro como tem sido a utilização do léxico pelos alunos em suas produções escritas, o que pode nos servir como indicador da necessidade do treino de diferentes faixas de palavras. A partir do perfil lexical, é possível verificar se o vocabulário utilizado pelos alunos nos textos está de acordo com o esperado para o nível de Ensino Médio. Partindo de tais princípios, algumas instruções lexicais podem ser direcionadas pelo professor.

Assim como já mencionado, quando os corretores fizeram a correção das redações, foram instruídos a levar em consideração os seus critérios pessoais, não foram direcionados para o vocabulário. O esperado era verificar se o perfil lexical que tende à formalidade teria impacto nas notas dadas pelos corretores.

Geramos uma *wordlist* para cada redação e uma outra para a lista deslematizada. Todas foram processadas pelo recurso apresentado a seguir.


Figura 12 - Detailed Consistency List do WST



Fonte: Dados da presente pesquisa.

Com o recurso Detailed Consistency List do WST, foi possível quantificar quantas palavras as duas listas tinham em comum (na coluna intitulada *Joint*). No caso do arquivo com as redações, a quantidade de palavras formais foi obtida pela diferença entre a quantidade total de tokens menos as palavras em comum.

Figura 13 - Tela de resultado da Detailed Consistency List

 Detailed consistency relations list (unsaved)

File Edit View Compute Settings Windows Help

N	File 1	Count	File 2	Count	Joint	Relation	Set
1	338.lst	106	lista deslematizada.lst	82,941	92	0.002	
2	309.lst	110	lista deslematizada.lst	82,941	88	0.002	
3	47.lst	247	lista deslematizada.lst	82,941	192	0.005	
4	487.lst	141	lista deslematizada.lst	82,941	125	0.003	
5	486.lst	195	lista deslematizada.lst	82,941	171	0.004	
6	352.lst	102	lista deslematizada.lst	82,941	92	0.002	
7	283.lst	122	lista deslematizada.lst	82,941	106	0.003	
8	80.lst	169	lista deslematizada.lst	82,941	142	0.003	
9	302.lst	115	lista deslematizada.lst	82,941	100	0.002	
10	340.lst	141	lista deslematizada.lst	82,941	119	0.003	
11	367.lst	130	lista deslematizada.lst	82,941	116	0.003	
12	420.lst	203	lista deslematizada.lst	82,941	160	0.004	
13	41.lst	216	lista deslematizada.lst	82,941	183	0.004	
14	488.lst	222	lista deslematizada.lst	82,941	186	0.004	
15	229.lst	144	lista deslematizada.lst	82,941	132	0.003	
16	318.lst	157	lista deslematizada.lst	82,941	124	0.003	
17	188.lst	184	lista deslematizada.lst	82,941	153	0.004	
18	259.lst	150	lista deslematizada.lst	82,941	120	0.003	
19	200.lst	152	lista deslematizada.lst	82,941	128	0.003	
20	70.lst	126	lista deslematizada.lst	82,941	114	0.003	
21	8.lst	184	lista deslematizada.lst	82,941	156	0.004	
22	12.lst	109	lista deslematizada.lst	82,941	102	0.002	
23	374.lst	146	lista deslematizada.lst	82,941	118	0.003	
24	484.lst	194	lista deslematizada.lst	82,941	157	0.004	
25	52.lst	201	lista deslematizada.lst	82,941	166	0.004	
26	325.lst	132	lista deslematizada.lst	82,941	102	0.002	
27	106.lst	158	lista deslematizada.lst	82,941	140	0.003	
28	329.lst	169	lista deslematizada.lst	82,941	143	0.003	
29	169.lst	200	lista deslematizada.lst	82,941	168	0.004	
30	292.lst	126	lista deslematizada.lst	82,941	117	0.003	
31	133.lst	128	lista deslematizada.lst	82,941	116	0.003	
32	396.lst	185	lista deslematizada.lst	82,941	148	0.004	
33	388.lst	146	lista deslematizada.lst	82,941	122	0.003	
34	248.lst	161	lista deslematizada.lst	82,941	135	0.003	
35	132.lst	95	lista deslematizada.lst	82,941	79	0.002	
36	140.lst	171	lista deslematizada.lst	82,941	147	0.004	

statistics consistency consistency relations filenames notes

125,250 entries Row 1 T S < > Help

Fonte: Dados da presente pesquisa.

Ao serem processados, foi possível traçar o perfil lexical pensando na qualidade do vocabulário contido nas redações.

O passo seguinte consistiu nos cálculos. Segundo Fonseca e Martins (1993) as técnicas de testes não paramétricos, adaptam-se aos dados da ciência do comportamento, portanto, a

aplicação dessas técnicas não exige suposições quanto à distribuição da população da qual se tenha retirado amostras para análises. Os testes não paramétricos exigem poucos cálculos e são aplicáveis para análises de pequenas amostras.

Para que fosse gerado o número de correlação que nos possibilitou verificar se a influência do vocabulário dos textos era de (fraca a muito forte) nas notas atribuídas pelos corretores humanos, utilizamos uma calculadora *online* que nos permitiu avaliar as medidas estabelecidas em duas condições de tratamento. Assim foi feito por conta de nosso teste ser não paramétrico no que diz respeito às medidas repetidas; nesse caso, nos baseamos pelas posições e não pelos valores estabelecidos, pressupondo que os nossos dados não seguem uma distribuição específica baseada na análise de medianas.

Usamos o coeficiente de correlação de Pearson para analisarmos a relação linear entre as variáveis postas. Considerando que o coeficiente de correlação pode variar em termos de valor de -1 a +1, quanto maior for o valor absoluto do coeficiente, mais forte é a relação entre as variáveis. Para a correlação de Pearson, um valor absoluto de 1 indica uma relação linear perfeita, o que significa que um valor perto de 0 indica que não há relação linear entre as variáveis. Vejam a tabela de correlação abaixo:

Tabela 1 - Escala de correlação	
<i>Valor de p</i>	<i>Interpretação</i>
0.0 a 0.19	Uma correlação bem fraca
0.20 a 0.39	Uma correlação fraca
0.40 a 0.69	Uma correlação moderada
0.70 a 0.89	Uma correlação forte
0.90 a 1.00	Uma correlação muito forte

Fonte: Larson e Farber (2015).

Após realizarmos os cálculos na calculadora estatística online, obtivemos valores que variaram dentro da escala de *Pearson* e a partir daí pudemos verificar a influência da variabilidade do vocabulário nas notas dos corretores.

Com a obtenção das notas dadas pelos corretores para as cinquenta redações, do perfil lexical dos textos e dos cálculos estatísticos realizados no site já citado para a verificação da influência do vocabulário nas notas, foi possível tabular os dados finais da pesquisa. Vejamos no próximo capítulo a análise dos resultados obtidos.

4 RESULTADOS E ANÁLISE DOS DADOS

No que se refere à nota atribuída pelos corretores, evidenciamos uma grande variação. Apresentamos as notas para as 50 redações e, para fins de análise, indicamos a mediana.

Tabela 2 - Notas atribuídas às redações e suas respectivas medianas

CORRETORES	1	2	3	4	5	6	7	8	9	10	MEDIANA
REDAÇÃO 1	5	5	7	7	8	8	7	6	5	7	7
REDAÇÃO 2	4	6	6	6	8	7	6	6	5	6	6
REDAÇÃO 3	6	7	5	7	10	7	7	7	7	5	7
REDAÇÃO 4	6	7	5	8	10	9	6	7	6	5	7
REDAÇÃO 5	6	6	5	5	10	7	7	7	6	5	6
REDAÇÃO 6	6	7	5	8	10	8	6	6	5	5	6,5
REDAÇÃO 7	4	5	1	4	7	4	5	5	4	1	4,5
REDAÇÃO 8	5	6	8	4	7	6	7	6	5	8	6
REDAÇÃO 9	6	7	6	4	7	5	6	7	6	6	6
REDAÇÃO 10	6	5	5	4	6	6	7	7	4	5	6
REDAÇÃO 11	7	7	9	4	7	8	6	7	7	9	7
REDAÇÃO 12	5	5	6	4	9	6	7	6	6	6	6
REDAÇÃO 13	7	6	6	4	8	7	7	7	6	6	7
REDAÇÃO 14	7	8	5	5	8	7	8	8	7	5	7,5
REDAÇÃO 15	5	5	7	7	10	7	7	6	5	7	7
REDAÇÃO 16	6	6	5	5	7	8	6	8	6	5	6
REDAÇÃO 17	7	8	10	6	9	9	8	8	8	10	8
REDAÇÃO 18	7	8	9	6	8	9	8	8	8	9	8
REDAÇÃO 19	8	8	10	7	9	9	7	9	7	10	8,5
REDAÇÃO 20	7	7	6	5	9	10	7	7	7	6	7
REDAÇÃO 21	7	8	9	6	9	10	8	9	8	9	8
REDAÇÃO 22	6	8	6	5	9	7	7	5	7	6	7
REDAÇÃO 23	7	8	9	6	9	10	8	7	7	9	8
REDAÇÃO 24	7	7	8	7	9	7	6	8	7	8	7
REDAÇÃO 25	6	7	8	6	8	6	6	7	6	8	6,5
REDAÇÃO 26	7	8	6	6	9	10	7	6	6	6	7
REDAÇÃO 27	7	8	10	5	9	8	7	7	7	10	7,5
REDAÇÃO 28	7	8	6	6	10	6	6	8	8	6	7,5
REDAÇÃO 29	7	7	8	4	8	9	7	8	6	8	7,5
REDAÇÃO 30	7	7	7	5	9	9	8	8	7	7	7,5
REDAÇÃO 31	7	7	7	5	8	5	7	7	7	7	7
REDAÇÃO 32	7	8	8	5	6	2	2	7	6	8	6,5
REDAÇÃO 33	7	8	7	5	7	10	8	8	8	7	8
REDAÇÃO 34	8	8	8	5	8	6	5	8	7	8	8
REDAÇÃO 35	8	8	8	7	8	8	8	8	8	8	8
REDAÇÃO 36	7	7	7	6	7	6	2	9	7	7	7
REDAÇÃO 37	7	9	6	8	6	10	8	9	8	6	8
REDAÇÃO 38	8	10	9	6	10	9	6	7	7	9	8
REDAÇÃO 39	9	9	7	6	8	9	7	9	7	7	8,5
REDAÇÃO 40	8	8	8	4	7	8	7	9	8	8	8
REDAÇÃO 41	8	8	7	6	6	9	6	9	8	7	8
REDAÇÃO 42	7	8	8	7	8	10	7	9	7	8	8
REDAÇÃO 43	7	9	6	7	9	9	8	9	8	6	8
REDAÇÃO 44	7	10	7	8	8	9	8	9	7	7	8
REDAÇÃO 45	7	7	7	9	8	8	9	9	8	7	8
REDAÇÃO 46	7	8	6	8	8	10	8	8	8	6	8
REDAÇÃO 47	8	9	7	8	8	10	8	8	8	7	8
REDAÇÃO 48	7	8	8	6	8	9	8	8	8	8	8
REDAÇÃO 49	7	8	7	6	7	6	7	9	8	7	7
REDAÇÃO 50	7	9	7	7	8	9	6	9	8	7	8

Fonte: Dados da presente pesquisa.

Como pode ser visto, o menor valor de mediana foi para a redação 7. E o maior valor foi para a redação 39.

A título de exemplificação, reproduzimos o conteúdo da redação com a menor mediana.⁹

Figura 12 - Redação com menor mediana

06 40 04 11 VUNESP 1901 001162002

REDAÇÃO
Texto definitivo

IBGE

A instância brasileira adverte que o número de acidentes de trânsito andou aumentando decorrente aos anos, a prudência em que dirige e o descuido do pedestre, as mas sinalizações andaram aumentando o número de mortes.

A ponte do Jaguaré uma má sinalização feita pela fiscalização de trânsito, acabou ocorrendo duas mortes de dois, o motoqueiro e o que estava na garupa, tiveram suas cabeças despidas.

A vantagem e oportunismo em que o cidadão encontra para adquirir a carteira de habilitação de trânsito à tal quebra um dinheiro por baixo aos panos, as escolas de formação sendo comprados pelos novos alunos condutores.

As próprias escolas de auto, estão proporcionando um aumento maior de condutores com mal formação, e o condutor comprando uma passagem para a morte.

VUNESP 1901 | Folha 01/02/2020 NÃO ASSINE ESTA PÁGINA 38776.02

Fonte: Vunesp (2020).

Ao analisarmos a redação acima, é possível perceber diversos desvios ortográficos, de concordância e a repetição de palavras. Em um nível lexical, percebe-se o uso de vocábulos que designam um campo lexical específico: trânsito, dirige, pedestre, condutores, sinalização, motoqueiro, habilitação, etc. Vê-se que as escolhas lexicais na redação acima foram feitas para atender a um determinado propósito comunicativo. A depender da exposição e emprego consciente do léxico, o aumento do repertório lexical pode impactar no desempenho em uma redação. É útil que os estudantes ampliem seu repertório lexical no decorrer da formação acadêmica desenvolvendo habilidades que os permitam produzir linguisticamente de forma

⁹ A instância brasileira adverte que o número de acidentes de trânsito andou aumentando decorrente aos anos, à prudência em que dirige o descuido do pedestre, as mas sinalizações andaram aumentando o número de mortes. A ponte do Jaguaré uma má sinalização feita pela fiscalização de trânsito, acabou ocorrendo duas mortes de dois, o motoqueiro e o que estava na garupa, tiveram suas cabeças despidas. A vantagem e oportunismo em que o cidadão encontra para adquirir a carteira de habilitação de trânsito à tal quebra um dinheiro por baixo aos panos, as escolas de formação sendo comprados pelos novos alunos condutores. As próprias escolas de auto, estão proporcionando um aumento maior de condutores com mal formação, e o condutor comprando uma passagem para a morte.

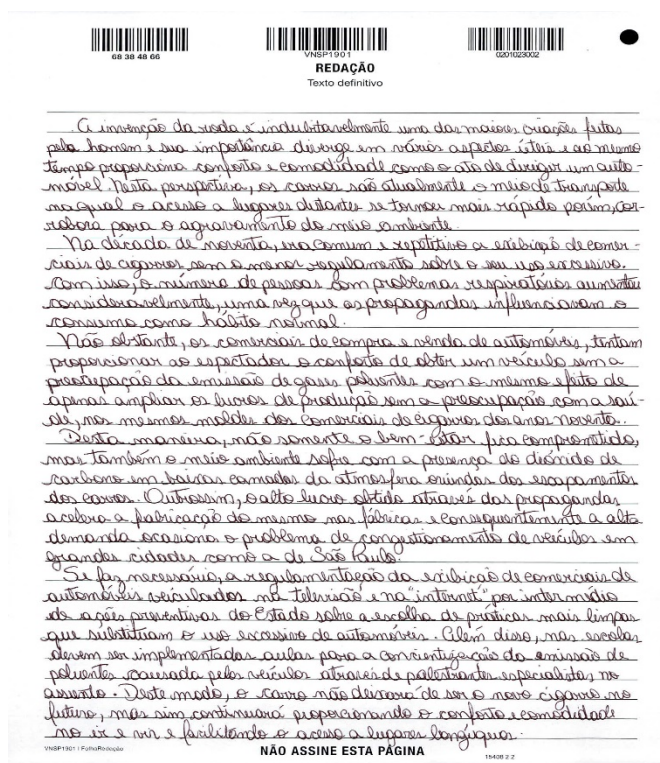
efetiva, haja vista os textos produzidos, principalmente na modalidade escrita tendem a ser avaliados.

Existem outros fatores, afirma Berber Sardinha (2004), que influenciam a seleção de palavras, dentre eles a proficiência linguística do autor, as colocações do tópico, tipo de texto, as normas de redação para os manuais e etc. Existem regularidades e variações nos padrões, essas variações podem ser sistemáticas, com variedades textuais dialetais e etc. Essas regularidades podem ser verificadas não pela intuição de um falante nativo, somente a observação empírica de dados reais em contexto de uso diferentes é que podem nos trazer as suas informações. Assim sendo, cabe dizer que a ocorrência de traços linguísticos e as suas frequências não constituem algo trivial como havia afirmado anteriormente.

Vejamos a redação com a maior mediana estabelecida pelas correções docentes.¹⁰

¹⁰ A invenção da roda é indubitavelmente uma das maiores criações feitas pelo homem e sua importância diverge em vários aspectos úteis e ao mesmo tempo proporciona conforto e comodidade como o ato de dirigir um automóvel. Nesta perspectiva, os carros são atualmente o meio de transporte na qual o acesso a lugares distantes se tornou mais rápido porém, corrobora para o agravamento do meio ambiente. Na década de noventa, era comum e repetitivo a exibição de comerciais de cigarros sem o menor regulamento sobre o seu uso excessivo. Com isso, o número de pessoas com problemas respiratórios aumentou consideravelmente, uma vez que as propagandas influenciavam o consumo como hábito normal. Não obstante, os comerciais de compra e venda de automóveis, tentam proporcionar ao espectador o conforto de obter um veículo sem a preocupação da emissão de gases poluentes com o mesmo efeito de apenas ampliar os lucros de produção sem a preocupação com a saúde, nos mesmos moldes dos comerciais de cigarros dos anos noventa. Desta maneira, não somente o bem-estar fica comprometido, mas também o meio ambiente sofre com a presença do dióxido de carbono em baixas camadas da atmosfera oriundas dos escapamentos dos carros. Outrossim, o alto lucro obtido através das propagandas acelera a fabricação do mesmo nas fábricas e conseqüentemente a alta demanda ocasiona o problema de congestionamento de veículos em grandes cidades como a de São Paulo. Se faz necessário, a regulamentação da exibição de comerciais de automóveis veiculados na televisão e na "internet" por intermédio de ações preventivas do Estado sobre a escolha de práticas mais limpas que substituam o uso excessivo de automóveis. Além disso, mas escolas devem ser implementadas aulas para a conscientização da emissão de poluentes causada pelos veículos através de palestrantes especialistas no assunto. Deste modo, o carro não deixará de ser o novo cigarro no futuro, mas sim continuará proporcionando o conforto e comodidade no ir e vir e facilitando o acesso a lugares longínquos.

Figura 13 - Redação com maior mediana



Fonte: Vunesp (2020).

Analisando lexicalmente a redação acima, percebe-se a variabilidade do vocabulário empregado. As repetições foram poucas e um alto nível de insumo lexical foi demonstrado. As escolhas lexicais também atenderam à finalidade comunicativa estabelecida, os vocábulos se deram dentro da proposta do texto, de modo diversificado, sem inadequações vocabulares. Percebe-se um nível considerável de precisão vocabular, própria ao texto, o que, sem dúvida, interfere na inteligibilidade da produção de modo crucial.

Análises realizadas com auxílio de programas de computador, de acordo com Oliveira (2009) podem orientar novas descobertas acerca de aspectos linguísticos que outrora eram considerados irrelevantes pelos pesquisadores. Isso ocorre por conta da possibilidade de que evidências emergem dos dados para que sejam percebidas, os linguistas envolvidos nos estudos de cópulas precisam confiar no texto, para exercer sobre ele uma observação cada vez mais isenta, deixando que a base para outras descrições e análises sejam os dados que podem levar a novas descobertas teóricas.

Ao observarmos os resultados obtidos podemos refletir acerca das contribuições/vantagens para a prática docente do profissional de línguas do uso da Linguística de Córpulas. Da mesma forma, refletimos o entendimento de Biber e Reppen (2015), uma vez

que nossa investigação empírica permite não apenas o estudo de variação, mas, também padrões reais de uso em textos naturais (BIBER; CONRAD; REPPEN, 1998). Adotamos a chamada nova perspectiva para a Linguística (NOVODVORSKI; FINATTO, 2014). Na investigação da linguagem, o emprego de computadores proporcionou resultados consistentes, permitindo que novos fatos fossem descobertos. Acreditamos que a LC educacional (BERBER SARDINHA, 2004) possa ser enriquecida com os dados apresentados aqui, em pequena escala.

O uso apropriado da língua bem como o seu ensino, conforme Viana (2011, p. 27), se dá a partir do conhecimento de questões semânticas e pragmáticas em diferentes situações comunicativas, nas quais sua produção precisa se aproximar dos falantes nativos de modo alternativo. O conhecimento de tais aspectos possibilita ao aprendiz de língua uma quebra consciente de padrões para que objetivos específicos sejam alcançados, como por exemplo, o uso inovador numa produção literária. Diferentemente da relação colocado-base em que os elementos podem ocorrer, as expressões fixas nas quais uma palavra ocorre são reveladas por agrupamentos lexicais por expressões que estão presentes no conhecimento da frase, lógica indispensável aos usuários de uma língua e na produção discursiva desses falantes.

No que diz respeito à apreciação dos corretores e a respectiva correlação entre as notas e a quantidade de vocabulário mais formal nas redações (nesse caso, contabilizamos apenas types), apresentamos os resultados na Tabela 2:

Tabela 3 - Correlação entre as notas dos corretores e os types da redação

	Corretores									
	1	2	3	4	5	6	7	8	9	10
Valor de p	0.68	0.68	0.58	0.26	0.29	0.18	0.40	0.40	0.62	0.69

Fonte: Dados da presente pesquisa.

Para o entendimento dos números acima, o valor da correlação varia de 0 (muito fraca) até uma influência muito forte (.80 – 1.0). Caso o valor seja 1.0, existe uma correlação perfeita. Isso indicaria que quanto mais palavras mais formais, maior seria a nota. É possível observar que para o corretor 6, as palavras mais formais não tiveram tanta influência. No entanto, para os corretores 3, 7 e 8, a influência foi moderada, o que valida a ideia de que quanto maior a quantidade de palavras mais formais, maior a nota atribuída. Para os corretores 1, 2, 9 e 10, a correlação é forte, indicando que a nota alcançada foi mais alta quando a redação apresentava

mais palavras formais. Em suma, para sete corretores, dos dez que participaram, houve uma correlação de moderada a forte.

Os resultados obtidos aqui indicam que, dentre os professores participantes, existe um efeito positivo entre a presença de vocabulário mais formal e a nota atribuída, o que pode impactar a atividade de professores de Língua Portuguesa e de redação para que dediquem atenção ao ensino de vocabulário – tendo em vista o direcionamento das escolhas vocabulares e o aumento de repertório lexical.

Concordamos com Viana (2011) quando o autor afirma que como um campo de conhecimento dedicado à investigação da linguagem, a Linguística não se distancia da prática de professores de línguas. Para compreender como os alunos expressam as suas ideias na língua é preciso entender o funcionamento dela.

Em outras palavras, para compreender a questão linguística é necessário basear o estudo em um alicerce empírico no qual os resultados advém da observação de exemplos reais. Por ser calcada em dados, a necessidade de se recorrer a um ‘nativo’ fica a margem do processo, podendo a investigação ser realizada por ‘falantes não nativos’ da língua estrangeira. (VIANA, 2011, p. 23).

Segundo Viana (2011, p. 25), podemos entender a linguística de *Córpus* como uma forma de investigação empírica da linguagem por meio da sistemática exploração de um *Córpus*. Não há uma seleção de dados para exclusão ou para inclusão no escopo da análise, o pesquisador tem a função de analisar todas as ocorrências de um dado traço linguístico presente no *Córpus* e somente essas instâncias, de modo prático. Isso corresponde à prática de não fazer o descarte de dados por não parecerem adequados a uma teoria que se desejar provar, assim como exemplos não devem ser inventados para comprovar uma outra teoria. O *córpus* é uma amostra de prática social e a pesquisa deve ser baseada no que ele registra efetivamente

Pensando nas questões do ensino, Viana (2011, p. 31) acrescenta que é preciso definir o que é a palavra na linguística de *Córpus*, principalmente por conta da hifenização; o emprego de hifens corresponde a somente um dos casos, nos quais é preciso determinar o que se entende por palavra. Existem ainda questões de uso do apóstrofo, números e algarismos arábicos por extenso.

Sarmiento (2010, p. 100) afirma que o interesse na aplicação de pesquisas baseadas em *Córpus* no ensino de línguas tem crescido de modo considerável. Essa aplicação pode ser de duas formas, o uso direto de *córpus* com os aprendizes e o uso indireto. No uso direto, os alunos agem como detetives linguísticos, descobrem fatos sobre a língua que estão estudando por meio de exemplos autênticos, além desses usos diretos, os *córpus* vêm sendo cada vez mais usados

na elaboração de materiais didáticos. A linguística de *Córpus* oferece informações relacionadas a vocabulário e à gramática, à formalidade e à informalidade, diferenças entre a linguagem escrita e falada, como as pessoas começam e terminam uma conversa, entre outros aspectos.

Para a utilização de um *Córpus* na elaboração de material didático (livros, polígrafos ou exercícios) é necessário, primeiramente, decidir (no caso da língua inglesa) quanto ao tipo e variedade de inglês que servirá como base para a elaboração do material, uma vez que *córpus* diferentes apresentarão palavras diferentes e, frequentemente, diferentes usos e funções das palavras a serem ensinadas. (SARMENTO, 2010, p. 103).

De acordo com Cobb e Boulton (2015, p. 481), a Linguística de *Córpus* pode ser quase definida como uma Linguística Aplicada. Além da própria disciplina não menos importante no ensino e aprendizagem de línguas, podemos considerar que sua influência tem sido de três tipos principais. A primeira reside em descrições melhoradas das variedades linguísticas e características que podem informar aspectos da linguagem a ensinar; a segunda serve-se de *córpus* e ferramentas para analisá-los, sendo estes disponíveis ao professor; o terceiro coloca-os diretamente nas mãos do aprendiz.

Os estudos que se baseiam em *córpus* sugerem os itens linguísticos e processos mais encontrados por usuários de uma língua e que portanto, devem ser tidos em termos de tempo na elaboração de material didático. Para a utilização de um *córpus*, primeiramente é preciso decidir o tipo de variedade de língua que servirá como base para a elaboração de material, já que o *córpus* apresenta palavras distintas e frequentemente diferentes usos e funções das palavras que serão ensinadas. Sendo assim, a escolha de um *Córpus* pode afetar diretamente as palavras que serão incluídas nos materiais didáticos, do mesmo modo seus sentidos e usos.

Na sequência, apresentamos o perfil lexical das redações:

Tabela 4 - Perfil lexical das redações

Redações	Formal (%)	Fundamental (%)
Redação 01.lst	10,38	89,62
Redação 02.lst	0,00	100,00
Redação 03.lst	8,33	91,67
Redação 04.lst	14,95	85,05
Redação 05.lst	12,15	87,85
Redação 06.lst	5,68	94,32
Redação 07.lst	16,09	83,91
Redação 08.lst	9,76	90,24
Redação 09.lst	4,27	95,73

Redação 10.lst	7,69	92,31
Redação 11.lst	11,96	88,04
Redação 12.lst	8,25	91,75
Redação 13.lst	13,00	87,00
Redação 14.lst	6,90	93,10
Redação 15.lst	2,33	97,67
Redação 16.lst	9,46	90,54
Redação 17.lst	15,00	85,00
Redação 18.lst	12,23	87,77
Redação 19.lst	14,88	85,12
Redação 20.lst	14,19	85,81
Redação 21.lst	17,07	82,93
Redação 22.lst	10,56	89,44
Redação 23.lst	13,59	86,41
Redação 24.lst	6,98	93,02
Redação 25.lst	6,42	93,58
Redação 26.lst	14,29	85,71
Redação 27.lst	13,51	86,49
Redação 28.lst	12,43	87,57
Redação 29.lst	4,00	96,00
Redação 30.lst	15,71	84,29
Redação 31.lst	6,51	93,49
Redação 32.lst	14,63	85,37
Redação 33.lst	18,23	81,77
Redação 34.lst	13,53	86,47
Redação 35.lst	10,12	89,88
Redação 36.lst	21,30	78,70
Redação 37.lst	9,90	90,10
Redação 38.lst	16,73	83,27
Redação 39.lst	15,03	84,97
Redação 40.lst	15,38	84,62
Redação 41.lst	20,41	79,59
Redação 42.lst	13,47	86,53
Redação 43.lst	21,52	78,48
Redação 44.lst	13,73	86,27
Redação 45.lst	12,96	87,04
Redação 46.lst	18,09	81,91
Redação 47.lst	21,86	78,14
Redação 48.lst	19,70	80,30
Redação 49.lst	12,75	87,25
Redação 50.lst	17,49	82,51

Fonte: Dados da presente pesquisa.

Sarmiento (2010, p. 101) afirma que um *cópus* não consegue mostrar nada mais além de seu conteúdo. Por mais representativo que ele se propõe a ser, generalizações são feitas a

partir de resultados, o que podem ser chamadas na verdade de extrapolações. Declarar algo sobre um *cópus* é declarar sobre ele e não sobre a linguagem ou sobre os registros que representa. Dessa forma, concluir a respeito da linguagem a partir de um *cópus* é o mesmo que deduzir, o que não pode ser tratado como fato. Um *cópus* oferece evidências, mas não informações. Fornece uma abundância de exemplos ao pesquisador, mas precisam ser interpretados.

5 CONSIDERAÇÕES FINAIS

Iniciamos o primeiro capítulo realizando uma apresentação das linhas teóricas que nos serviram como aporte: a Lexicologia e a Linguística de Córpus. A pretensão não foi a de esgotar o assunto, mas, de articular ideias de autores conceituados nas áreas, de forma a construir um quadro representativo envolvendo o objeto de pesquisa em pauta. Para o desenvolvimento da nossa argumentação foi de fundamental importância as contribuições de Antunes (2012); Barbosa; Souza (2016), Biderman, (1999, 2006), Guerra; Andrade (2012), Genouvrier; Peytard (1985), Rocco, (1981), Berber Sardinha (2004, 2010, 2012), Biber; Reppen, (2015); Halliday (1993), Sinclair (2004), Barbosa (2009), Barbosa; Souza (2016).

Buscamos comprovar tal influência não somente a nível teórico, mas também quantitativo e metodológico, atentando-nos ao fato de que é de crucial importância a formação lexical dos alunos.

Demonstramos que a Lexicologia pode contribuir teoricamente com o ensino do léxico em uma perspectiva de usos e sentidos. Ao falarmos sobre uma didática de ensino de línguas, não temos que pensar em uma receita pronta, mas no entrelaçamento de saberes teóricos e a prática para fins de um ensino de léxico inter e transdisciplinar, isto é, entre os saberes das diferentes áreas e para além delas.

É importante ressaltar que o trabalho do professor, assim como o arcabouço lexical dos indivíduos sociais, está em constante transformação.

Nessa pesquisa, a Linguística de Córpus viabiliza o processamento das redações de maneira confiável e célere.

Considerando que a língua é um sistema formado por signos, que expressam o conhecimento de tudo que está à nossa volta e que o léxico é uma forma de estruturar os conhecimentos do mundo, tanto de modo linguístico quanto extralinguístico, exploramos conceitos oriundos da Lexicologia vista como uma disciplina, cujo objeto de estudo é o léxico da língua.

Exploramos o vocabulário fundamental, que por sinal, desempenha um papel importante na escrita de textos, um conjunto já consolidado e atestado de adjetivos, advérbios, conjunções, substantivos, preposições e verbos típicos da comunicação de uma língua, imprescindível para uma comunicação básica.

Por se tratar de uma pesquisa léxico-estatística, as reflexões de ordem teórico-metodológicas foram importantes para embasar nossos procedimentos. À medida que a pesquisa avançava, era perceptível o quão importante foi a aproximação com a Linguística de

Córpus e com as contribuições estatísticas. A Linguística de Córpus permitiu que o processamento do córpus com milhares de palavras fosse realizado sem complicações: os dados linguísticos foram manipulados computacionalmente de modo rápido e preciso.

Estamos cientes de que mesmo um córpus colossal não poderá ter exatamente todas as características da língua, tendo em vista a sua infinitude. Continuamos crendo que o córpus não captura todas as realizações da língua nem reflete de forma precisa todas as possibilidades de seus padrões. Todavia, por benefício da investigação científica, sentimo-nos capacitados para, a partir das informações obtidas do córpus, fornecer respostas e indicar perspectivas ainda pouco exploradas pelos linguistas de forma geral.

A ferramenta Wordsmith Tools contribuiu para impulsionar a popularização de estudos baseados em córpus, uma vez que possibilitou-nos traçar o perfil lexical esperado para o seu estudo e análise, ao fornecer informações no que diz respeito aos *types* e *tokens*, permitindo-nos identificar faixas de frequência de palavras superiores, para além do vocabulário fundamental, ponto este crucial para o desenrolar das respostas esperadas. Após a constituição do córpus, foram necessários os seguintes passos: digitar os textos escritos, detalhar a variabilidade do vocabulário e traçar o perfil lexical das redações que constituíram o córpus. A análise linguístico-estatística revela que existe uma influência considerável da qualidade do vocabulário sobre os textos ao serem humanamente avaliados.

São notórios os benefícios trazidos pela junção da tecnologia e da Linguística, aqui em específico, nos estudos lexicológicos. Destacamos a possibilidade de compilar em meio digital uma considerável quantidade de textos (representativos da escrita de textos para vestibular) a fim de estudar o uso real e o comportamento das palavras, conseqüentemente levando à visualização de fatos importantes. Com um córpus controlado por computador, tem-se mais segurança e objetividade no processamento dos dados.

No capítulo 4, os resultados obtidos foram discutidos, onde o embasamento da fundamentação teórica e da metodologia convergiram para a elaboração de uma análise quantiqualitativa. A depender da exposição e emprego consciente do léxico, o aumento do repertório lexical pode impactar no desempenho em uma redação. Nesse sentido, é útil que os estudantes ampliem seu repertório lexical no decorrer da formação acadêmica desenvolvendo habilidades que os permitam produzir linguisticamente de forma efetiva, haja vista que os textos produzidos, principalmente na modalidade escrita, tendem a ser avaliados.

Acreditamos ter atingido o objetivo geral delineado na introdução da dissertação. Procedemos conjuntamente a descrição e a análise do córpus em específico explorando a qualidade do vocabulário contido em redações, partindo de dados lexicais identificados. No que

tange aos objetivos específicos, constituímos um *córpus* de redações de ensino médio, propondo uma metodologia de identificação da qualidade de vocabulário dos textos a partir do perfil lexical obtido com o processamento do *córpus* e das comparações entre a correção dos textos pelos professores de redação e a linguístico-computacional, o que nos permitiu verificar se há, de fato, influência do vocabulário utilizado nas avaliações das redações.

É importante ressaltar que temos condições nessa etapa de fornecer respostas aos problemas da pesquisa, propostos no capítulo introdutório. Apesar de já termos apresentado argumentos que configurariam respostas suficientes, retomamos os dois problemas.

Para o primeiro problema (Qual o perfil lexical das redações?) descobrimos que de modo geral, nos restringindo ao *córpus* analisado, grande parte das redações apresenta considerável repetição de palavras, pouca exploração de sinônimos e baixo insumo lexical. Temos condições de afirmar que ao traçarmos o perfil lexical das redações, entendemos de modo mais claro como tem sido a utilização do léxico pelos alunos em suas produções escritas, o que pode nos servir como indicador da necessidade do treino de diferentes faixas de palavras.

O segundo problema (Qual o vocabulário mais formal empregado?) revelou que o uso do vocabulário indica o quão bem, ou não, os alunos do Ensino Médio conseguem se expressar na língua escrita. Destacamos que apoiar-se no uso de sinônimos e fazer escolhas vocabulares conscientes culminando na variabilidade do repertório de palavras são fundamentais para a qualidade dos textos que serão avaliados.

Percebemos que existe uma correlação significativa entre as correções citadas. Os cálculos estatísticos indicam que existe uma influência do vocabulário formal de moderada a forte na avaliação da maioria dos corretores que participaram da pesquisa. Dessa forma, em termos gerais, professores de Língua Portuguesa e de redação podem recorrer ao ensino de vocabulário tendo em vista o direcionamento das escolhas vocabulares e o aumento de repertório lexical, haja vista que os textos são avaliados.

Conforme o exposto, a análise gera resultados consistentes e respostas frente aos problemas introdutórios. Nessa situação, os resultados obtidos revelam que podemos refletir acerca das contribuições/vantagens para a prática docente do profissional de línguas do uso da Lexicologia e da Linguística de *Córpus*.

Basicamente, nosso vocabulário é composto por um grande conjunto de palavras que conhecemos ao longo da vida, mas existem meios de enriquecimento lexical, o que demanda tempo, estudo, observação e prática. Dessa forma, os alunos devem ser conscientizados de que os vocábulos servem para facilitar o entendimento do leitor e enriquecer o conteúdo do texto, o que sem dúvida, garante a inteligibilidade do que está sendo escrito. Os procedimentos descritos

aqui mostram que identificar o nível de vocabulário utilizado por determinados alunos implica levar em consideração detalhes essenciais, como palavras consideradas como vocabulário fundamental e as faixas superiores a partir de uma lista-base. A lista de palavras ampliada e utilizada nessa pesquisa, pode servir como material de referência, especialmente voltada para o ensino do léxico, uma vez que reflete o uso real da língua na comunicação escrita.

Sob a ótica específica dos estudos e as interfaces aqui exploradas – Lexicologia e Linguística de Córpus – procuramos trazer perspectivas positivas para os estudos linguísticos. Esperamos que a pesquisa aqui apresentada corrobore com pesquisas vindouras.

REFERÊNCIAS

- AHMAD, K., CORBETT, G., and ROGERS, M. **Using computers with advanced language learners: An example.** *Language Teacher*, Tokyo, p. 4-7, 1985.
- ALUÍSIO, S. M.; ALMEIDA, G. M. O que é e como se contrói um Córpus?: Lições aprendidas na compilação de vários córpus para pesquisa lingüística. *Calidoscópico*, São Leopoldo, v. 4, n. 3, p. 155-177, 2006.
- ANTHONY, L. **AntWordProfiler** (Version 1.4.1) [Computer Software]. Tokyo, Japan: Waseda University, 2014. Available from 4
- ANTUNES, I. **Território das palavras.** São Paulo: Parábola, 2012.
- ASTON, G. Córpus in language pedagogy: Matching theory and practice. In G. Cook & B. Seidlhofer (Eds.), **Principle and Practice in Applied Linguistics.** Oxford: Oxford University. Press, 257–270, 1995.
- Assunção, C., & Araújo, C. Linguística de corpus: teoria, perspectivas metodológicas e ensino das línguas. **Filologia E Linguística Portuguesa**, 21(2), 271-288, 2019.
- BAKHTIN, M. A interação verbal. In: _____. **Marxismo e filosofia da linguagem.** Trad. Michel Lahud e Yara Frateschi Vieira. 12. ed. São Paulo, Hucitec, 2006.
- BARBOSA, M A. Lexicologia, Lexicografia, Terminologia e terminografia: identidade científica, objeto, métodos, campos de atuação. In: II SIMPÓSIO LATINO-AMERICANO DE TERMINOLOGIA. I ENCONTRO BRASILEIRO DE TERMINOLOGIA TECNOCIENTÍFICA, 2., 1., 1992. Curitiba. **Anais [...].** Curitiba: IBICT, 1992, s/p.
- BARBOSA, L. O conceito de lexicultura e suas implicações para o ensino- aprendizagem de português língua estrangeira. **Filologia e Linguística Portuguesa**, São Paulo, v. 10-11, p. 31-41, 2009.
- BARBOSA, L. M. A; SOUZA, D. C. J. Linguagens, identidades e letramentos. Lexicultura e hipertextos em letras de canções brasileiras no contexto de português para estrangeiros. **Revista Ecos**, São Paulo, v. 21, n. 2, p. 140-156, 2016.
- BASÍLIO, M. **Teoria lexical.** 8. ed. São Paulo: Ática, 2007.
- BAUER, L. ; NATION, I. S. P. Word families. **International journal of Lexicography**, v. 6, p. 253-279, 1993.
- BÍBLIA, N. T. Provérbios. In BÍBLIA. Português. **Bíblia Sagrada de Promessas: Antigo e Novo Testamentos.** Tradução de João Ferreira de Almeida: Junta de educação religiosa e publicações, 2006. p. 189.
- BECHARA, E. **Dicionário escolar da Academia Brasileira de Letras.** 3. ed. São Paulo: Companhia Editora Nacional, 2011.

BENNETT, G. **Using corpus in the language learning classroom: Corpus Linguistics for teachers.** Michigan: University of Michigan Press, 2010.

BERBER SARDINHA, T. **Linguística de Córpus.** Barueri: Manole, 2004.

BERBER SARDINHA, T. Como usar a Linguística de Córpus no ensino de língua estrangeira. Ou: Por uma linguística de Córpus educacional brasileira. *In: TAGNIN, S.; VIANA, V. (org.). Córpus no ensino de línguas estrangeiras.* São Paulo: Hub Editorial. p. 301-356, 2011.

BERBER SARDINHA, T. *Linguística de Córpus.* São Paulo, Manole, 2004.

BIBER, D. ; REPPEN, R. Introduction. *In: BIBER, D. ; REPPEN, R. (ed.). The Cambridge handbook of English Corpus Linguistics.* Cambridge: Cambridge University Press. 2015.

BIBER, D. Representativeness in corpus design. *Literary Linguist Computing*, Arizona, v. 8 n. 4. p. 244-257, 1993.

BIBER, D. Historical patterns for the grammatical marking of stance: A cross register comparison. **Journal of Historical Pragmatics**, Arizona, v. 5 n. 1. p.107-135, 2004.

BIBER, D. ; CONRAD, S. ; REPPEN, R. **Corpus Linguistics: investigating language structure and use.** Cambridge: Cambridge University Press, 2009.

BIBER, D. Conrad, S., and Cortes, V. **If you look at Lexical bundles in university teaching and textbooks.** *Applied Linguistics*, p. 371–405, 2004.

BIBER, D.; CONRAD, S. e REPPEN, R. **Córpus linguistics: Investigating language structure and use.** Cambridge, 1998.

BIDERMAN, M. T. C. Conceito linguístico de palavra. **Palavra**, Rio de Janeiro, v. 1, n. 5, 1999.

BIDERMAN, M. T. C. **Dicionário didático de português.** 2. Ed. São Paulo: Ática, 1998.

BIDERMAN, M. T. C. Léxico e vocabulário fundamental. **Alfa**, São Paulo, p. 27-46, 1996.

BIDERMAN, M. T. C. A estruturação do léxico e a organização do conhecimento. **Letras de Hoje**, Porto Alegre, v. 22, n. 4, p. 81-96, dez. 1987.

BIDERMAN, M. T. C. **Teoria linguística: teoria lexical e linguística computacional.** 2. ed. São Paulo: Martins Fontes, 2001.

BOULTON, A. and TYNE, H. **Me'thodologie de la de'couverte en didactique des langues: Des documents authentiques aux Córpus.** Paris: Didier, 2014.

BORBA, F. S. **Organização de dicionários: uma introdução à lexicografia.** São Paulo: Unesp, 2003

BRAUN, S. **Integrating corpus work into secondary education: From data-driven learning to needs-driven corpus**”. *ReCALL*, 19 (3), 307–328, 2007.

CAMPION, M. and ELLEY, W. **An academic vocabulary list**. Wellington, New Zealand: Council for Educational Research, 1971.

Cartilha do participante – Inep. **Inep, 2020**. Disponível em http://download.inep.gov.br/educacao_basica/enem/downloads/2020/. Acesso em 24 de julho de 2020.

CHARLES, M. Reconciling top-down and bottom-up approaches to graduate writing: Using a corpus to teach rhetorical functions”. *Journal of English for Academic Purposes*, v. 6, 289–302, 2007.

COBB, T.; BOULTON, A. **Classroom applications of corpus analysis**. In D. Biber & R. Reppen (eds), *Cambridge Handbook of English Corpus Linguistics*. Cambridge: Cambridge University Press, p. 478-497, 2015.

COBB, T. Computing the vocabulary demands of L2 reading. **Language Learning & Technology**, p. 38–63., 2007.

COSTA, J. C.; SILVA, E.B. Uma análise quantiquantitativa de redações em língua portuguesa: algumas reflexões para o ensino. *In: GONÇALVES, R. B.; SILVA, E. B. Recortes linguísticos sob uma perspectiva intercultural*. Maringá: Uniedusul, 2020. cap. 3. P, 36-40.

COXHEAD, A. A new academic word list. *TESOL Quarterly*, p. 213–238, 2000.

DAVIES. **The Cambridge handbook of English Corpus linguistics**. Biber, D.; Reppen, R. Cambridge University Press, 2015.

DUBOIS, J.; GIACOMO, M.; GUESPIN, L.; MARCELLESI, C.; MARCELLESI, J.B.; MEVEL, J-P. **Dicionário de Linguística**. 10. ed. São Paulo: Editora Cultrix, 1998.

DUCROT, O. e TODOROV, T. **Dicionário enciclopédico das ciências da linguagem**. 3ª ed., São Paulo, Perspectiva, 2001.

EVERS, A. **Processamento de língua natural e níveis de proficiência do português: um estudo de produções textuais do exame Celpe-Bras**. Dissertação (Mestrado) – Programa de Pós-graduação em Letras, Universidade Federal do Rio Grande do Sul. Porto Alegre, 2013.

FERREIRA, H.M; VIEIRA, M. S. P. O trabalho com o léxico em sala de aula: desafios para o ensino de uma língua materna. **Revista Letras Raras**, [S.l.], v. 2, n. 2, p. 19 - 33, abr. 2014. ISSN 2317-2347. Disponível em: <<http://revistas.ufcg.edu.br/ch/index.php/RLR/article/view/185/151>>. Acesso em: 29 dez. 2020.

FINATTO, M.B.F. ;CREMONESE, L.E.; AZEREDO, S. O vocabulário na redação de vestibular: do enfoque estatístico às especificidades da enunciação. In: ABREU, S. (org.) **A redação no vestibular: do leitor ao produtor de texto**. COPERSE/UFRGS. Porto Alegre: Editora da UFRGS, 2008, p.95-108. ISBN978-85-386-0026-8

FINATTO, M. J. B. **Redação de vestibular: também cabe pensar sobre o vocabulário.** In: ABREU, Sabrina (Org.). Reflexões linguísticas e redação no vestibular. 1ed. Porto Alegre: Editora da UFRGS, 2010, v. , p. 209-220.

FINATTO, M. J. B. et al. Vocabulário controlado e redação de definições em dicionários de português para estrangeiros: ensaios para uma léxico-estatística textual. In: **Trama.** Marechal Cândido Rondon, PR Vol. 10, n. 20 (2014), f. 53-68, il.

FINATTO, Maria José B. Vocabulário em Redações de Vestibulandos: Contribuições da Léxico-Estatística Textual, Palavras-Signo e Palavras-em-uso. In: REBELLO, Lúcia S.; FLORES, Valdir do N. **O texto de vestibular em perspectiva.** Porto Alegre: COPERSE/UFRGS, 2014, 180p. p. **171-184** ISBN 978-85-386-0232-3

FINATTO, M.J.B; SPAGNOLO, É. BERGMANN, G.L. (2018). **Busca de um vocabulário básico do português do brasil escrito: listas de frequência de palavras – jornais populares, linguagem geral e tradução.** Material de pesquisa para consulta on-line. Inédito. Porto Alegre. Janeiro de 2018, 06p. Disponível em: www.ufrgs.br/textecc/porlexbras/porpopular/index.php

FONSECA, J.S.; MARTINS, G.A. (1993). **Curso de estatística.** 4a ed. São Paulo: Atlas

GARSIDE R, LEECH G & MCENERY A (eds.). **Córpus annotation:** linguistic information from computer text córpus. London: Longman, 1997.

GAVIOLI, L. **Exploring córpus for ESP learning.** Amsterdam: John Benjamins, 2005.

GENOUVRIER, E.; PEYTARD, J. **Linguística e ensino do português.** Coimbra: Livraria Almedina, 1985.

GRANGER S. **The computer learner Córpus:** a versatile new source of data for SLA research. In Granger (ed.), p. 3–18, 1998.

GRANGER, S. **The contribution of learner córpus to second language acquisition and foreign language teaching:** A critical evaluation. In K. Aijmer (ed.), *Córpus and language teaching*, p. 13–32. Amsterdam: John Benjamins, 2009.

GUERRA, M. M. e ANDRADE, K. de S. **O léxico sob perspectiva:** contribuições da Lexicologia para o ensino de línguas. Domínios da Linguagem. v. 6, n° 1, 1° Semestre, 2012.

HALLIDAY, M. A. K. **Quantitative studies and probabilities in grammar.** In: M. HOEY (ed.) *Data, description, discourse - papers on the English language in honour of John McH Sinclair on his sixtieth birthday.* London: HarperCollins, 1993.

HALLIDAY, M; YALLOP, C. **Lexicology: a short introduction.** London e New York: Continuum, 2007.

HARMER, J. **The practice of English Language Teaching.** 4. ed. Harlow, Longman, 2007.

HOFFMANN, L. Possibilidades de aplicação e a aplicação atual de métodos estatísticos

na pesquisa de linguagens especializadas (Título Original: Anwendungsmöglichkeiten und bisherige Anwendung von Statistischen Methoden in der Fachsprachenforschung, Disponível em: **Cadernos de Tradução**, Porto Alegre, nº 20, janeiro-junho, p. 61-76, 1998.

HOUAISS. **Dicionário eletrônico Houaiss da língua portuguesa**. Rio de Janeiro: Objetiva, 2009.

HUNSTON, S. **Córpus in Applied Linguistics**. London: Cambridge University Press, Johns, T. Whence and whither classroom concordancing? In T. Bongaerts, P. de Haan, S. Lobbe, and H. Wekker (eds.), *Computer applications in language learning*, 9–27, 1988.

JAEGER, A. C. O léxico em perspectiva – uma agenda de trajetórias a percorrer. In: BARROS, L. A.; ISQUERDO, A. N. (Org.). **O léxico em foco: múltiplos olhares**. São Paulo: Cultura Acadêmica, 2010. p. 65-78.

KRIEGER, M. G.; FINATTO, M. J. B. **Introdução à terminologia : teoria e prática**. São Paulo: Contexto, 2004.

LARSON, R.; FARBER, B. **Estatística Aplicada**. 4. ed. São Paulo: Editora Pearson, 2010.

LARSON, R.; FARBER, B. **Estatística Aplicada**. 6. ed. São Paulo: Editora Pearson, 2015.

LAUFER, B. and RAVENHORST-KALOVSKI, G. C. **Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension**. *Reading in a Foreign Language*, p.15–30, 2010.

LAUFER, B. Why are Some Words More Difficult than Others? Some Intralexical Factors that Affect the Learning of Words. *Iral-international Review of Applied Linguistics in Language Teaching - IRAL-INT REV APPL LINGUIST*. p. 293-308, 1990.

LEECH, G. N. **Córpus and Theories of Linguistic Performance**. Berlin: Mouton de Gruyter, 1992.

LEECH, G. N. **The State of Art in Córpus Linguistics**. London: Longman, 1991.

M. (orgs). **Using córpus for language research**. London and New York: Longman, McEnery, T. and Wilson, A.. *Teaching and language córpus*, p. 5–14., 1997.

Manual do candidato – Vunesp. **Vunesp, 2021**. Disponível em <https://documento.vunesp.com.br/documento/stream>. Acesso em 18 de novembro de 2020.

Mariano de SOUZA, K., & Regina Duarte XAVIER, V. (2020). Lexicologia e linguística aplicada: algumas aproximações e implicações no ensino de língua materna. **Linguagem: Estudos E Pesquisas**, 23(1), 133-144.

McENERY, T. e WILSON, A. **Córpus linguistics**. Edinburgh, Edinburgh University Press, 1996.

MCKAY, S. **Teaching the syntactic, semantic and pragmatic dimensions of verbs.** TESOL Quarterly, p. 17–26, 1980.

MEUNIER, F. The pedagogic value of native and learner corpus in EFL grammar teaching. In S. Granger, J. Hung & S. Petch-Tyson (Eds.), **Computer Learner Corpus, Second Language Acquisition and Foreign Language Teaching.** Amsterdam/Philadelphia: John Benjamins, p. 119–141, 2002.

NATION, P. and BEGLAR, D. **A vocabulary size test.** The Language Teacher, p. 9–13, 2007.

NOVODVORSKI, A.; BOCORNY FINATTO, M. J. Linguística de Corpus no Brasil: uma aventura mais do que adequada. **Letras & Letras**, v. 30, n. 2, p. 7-16, 18 dez. 2014.

OLIVEIRA, L.P. Linguística de Corpus: teoria, interfaces e aplicações. **Revista Matraga**, v. 16, n. 24, RJ: Programa de Pós-graduação em Letras da UERJ, p. 48-76, 2009.

PASQUALINI, B. F. **CorPop: um corpus de referência do português popular escrito do Brasil.** 250 p. Orientadora: Maria José Bocorny Finatto. Tese (Doutorado) - Universidade Federal do Rio Grande do Sul, Instituto de Letras, Programa de Pós-Graduação em Letras, Porto Alegre, BR-RS, 2018.

PERINI, M. A. **Gramática descritiva do português.** São Paulo: Ática, 1995

REPPEN, R. Building a Corpus: **What are the key considerations?** In A. O’Keeffe and M. McCarthy (eds.), *The Routledge handbook of Corpus linguistics*, p. 31–37. London: Routledge, 2010.

REY, A. **La lexicologie.** Paris: Klincksieck, 1970.

ROCCO, Maria Thereza Fraga. *Crise na linguagem: a redação no vestibular.* [S.l: s.n.], 1981.

SANCHEZ, A. **Definicion e historia de los corpus.** In: A. SANCHEZ et al (org.). *CUMBRE – Corpus Linguístico de Espanol Contemporaneo.* Madrid: SGEL, 1995.

SARMENTO, S. Linguística de Corpus: histórico, metodologia, campos de aplicação. **Revista Trama**, v. 6, n. 2, p. 87-107, 2010.

SCHMITT, N. and SCHMITT, D. **A reassessment of frequency and vocabulary size in L2 vocabulary teaching.** Language Teaching, p. 484–503, 2014.

SCOTT, M. **WordSmith Tools.** (1996) Oxford: Oxford University Press. Versão 5, 2008. Sinclair J *et al.* (1987). *Collins Cobuild English language dictionary.* London: HarperCollins.

SILVA, C. B. M; SILVA, E.B. Efeitos da leitura de excertos de contos no reconhecimento de vocabulário em língua portuguesa: um estudo de caso. In: SERPA, T. ; SILVA, E. B. ; PINTO, P. T. **Corpora, Tecnologias e Web 3.0: leituras e práticas para o ensino de línguas e tradução.** Campinas: Pontes, 2021.

- SINCLAIR J *et al.* **Collins Cobuild English grammar**. London: HarperCollins, 1990.
- SINCLAIR, J. **Córpus, Concordance, Collocation**. Oxford: OUP, 1991.
- STAEHR, L. S. Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal* , p,139–152, 2009.
- STANGROOM, J. Socscistatistics, 2018. Site de cálculos estatísticos. Disponível em: <https://www.socscistatistics.com/>. Acesso em: 25 de novembro de 2020.
- STUBBS, M. **Words and Phrases**. Oxford: Blackwell, 2001.
- SVARTVIK, J. **Córpus are becoming mainstream**. *In: THOMAS, J. and SHORT, Thorndike, E. L.* 1921. Word knowledge in elementary school. *Teachers College Record*, p. 334–370, 1921.
- TRASK, R.L. **Dicionário de Linguagem e Linguística**. São Paulo, Contexto, 2004.
- VIANA V. **Linguística de Córpus: Conceitos, técnicas & análises** [Córpus Linguistics: Concepts, techniques & analyses]. *In: Viana V & Tagnin S (eds.)* Córpus no ensino de línguas estrangeiras [Córpus in foreign language teaching]. São Paulo: HUB Editorial, p. 25-95, 2011.
- WEST, M. **A general service list of English words**. London: Longman, 1953.
- WIDDOWSON, H. G. **Context, community and authentic language**. *TESOL Quarterly*, 32 (4), 705–716, 1998.
- WIDDOWSON, H. G. **Córpus and language teaching tomorrow**. Keynote lecture delivered at the Fifth Teaching and Language Córpus Conference, Bertinoro, *Italy*, 2002.
- WIDDOWSON, H. G. **Text, Context, Pretext**. London: Blackwell, 2004.
- WILLIS, D. **Rules, patterns and words: grammar and lexis in English language teaching**. Cambridge: Cambridge University Press, 2003.

APÊNDICE A – Redações da Vunesp 2020

Redação 1

Da roda ao automovel milhares de anos que não os separam tanto assim, continuam sendo meios de transporte. A humanidade sempre precisou se locomover, não vivemos parados, mas novos meios de locomoção estão em contante evolução, um dia após o outro sempre estamos mudando nosso jeito de se mover de um ponto A para o ponto B. Mas todo processo de evolução tem seus males, chegamos ao ponto em que ir de um bairro até o outro virou um desafio, com tantos carros estamos entupindo as veias arteriais de nossas cidades e poluindo seu meio ambiente. A máquina que foi criada para facilitar e deixar nossa viagem mais comoda, está dificultando e se tornando uma experiência desagradavel. Mas lembre-se, moldamos o mundo a nossa volta e estamos em constante evolução, as dificuldades e os problemas estão ai para serem melhorados e resolvidos. O automovel não é um vicio é uma fase, como todos os outros meios de transporte foram, e logo perdera seu pedestal.

Redação 2

Hoje em dia temos varios meios de transpote publico para se locomover para onde quiser. O governo colocou isto para tentar ajudar o meio ambiente e as pessoas de baixa renda. Mas com o aumento da tecnologia e a rapidez dos carros, este meio de transporte publico está cada dia sendo menos utilizados pela população e isto esta prejudicando a natureza, pois com tantos carros nas ruas, mais poluição esstá soltando. Para ajudarmos nesta causa, poderia deixar os carros, e começar a utilizar este meio, deveria tambem mais insentivo de todos, ter propaganda falando de como isto vai ajudar a natureza, explicar e manter o transporte, se ajudarmos uns aos outros todos irão sair ganhando com isso. Vamos colaborar com isto e usar mais transportes publicos e menos carros. Redação 3 Atualmente nesse mundo globalizado onde viemos, a maior parte da população possui um carro. Hoje em dia esse meio de transporte deixou de ser um luxo para se tornar necessidade na vida das pessoas. Com o aumento da população, o enorme fluxo de carros nas cidades grandes vem causando grandes problemas na atmosfera terrestre, apesar das cidades oferecem transporte público muitos preferem ir com seus próprios carros por ser mais rápido, prático e confortável. No entanto é importante ressaltar que estudos fora do Brasil vem sendo realizados para construir carros menos poluentes movidos por energia carregável ou raios solares. Portanto, acredita-se que no futuro os carros sejam menos poluentes mas até isso acontecer podemos considerar sim, o carro com o novo cigarro.

Redação 3

Atualmente nesse mundo globalizado onde viemos, a maior parte da população possui um carro. Hoje em dia esse meio de transporte deixou de ser um luxo para se tornar necessidade na vida das pessoas. Com o aumento da população, o enorme fluxo de carros nas cidades grandes vem causando grandes problemas na atmosfera terrestre, apesar das cidades oferecem transporte público muitos preferem ir com seus próprios carros por ser mais rápido, prático e confortável. No entanto é importante ressaltar que estudos fora do Brasil vem sendo realizados para construir carros menos poluentes movidos por energia carregável ou raios solares. Portanto, acredita-se que no futuro os carros sejam menos poluentes mas até isso acontecer podemos considerar sim, o carro como o novo cigarro.

Redação 4

No Brasil o automóvel é um grande desejo das pessoas, desde criança já possuem como meta chegar aos 18 anos para tirarem habilitação e comprarem seu primeiro carro. Além das necessidades pessoais, a quantidade de automóveis chega a passar o número de habitantes em algumas cidades, como em São Paulo que foi preciso aderir a prática de rodízio para tentar diminuir o número de veículos. Na Alemanha, por volta dos anos 2000, foi feito um acordo com a população que fizeram a troca de veículos por bicicletas e deu super certo. Se todos países aderissem a ideia da Alemanha diminuiríamos até os efeitos causados ao meio ambiente, como o efeito estufa. Portanto, criar um incentivo desde o ensinamento nas escolas para utilização dos meios de transportes coletivos como ônibus, metros e até mesmo transportes de aplicativos, para cidades grandes e bicicletas para cidades pequenas o número de veículos seriam reduzidos e práticas saudáveis desenvolvidas.

Redação 5

Atualmente há milhares de carros espalhados Brasil à fora. A produção não para, a cada ano novos veículos são lançados no mercado. Por um lado os veículos são ótimos, existem diversas marcas e modelos, sendo essas dispostos ao consumidor, com diversas formas de pagamento que se adequam ao bolso do freguês, fazendo com que o mercado automobilístico aumente numa ampla escala. Por outro lado vemos quão caótico está o trânsito, pois a fabricação desenfreada de carros e a oferta e procura gera enormes filas de congestionamento, acidentes e sem contar a poluição do planeta a emissão de CO2 prejudica o meio ambiente, a falta de infraestrutura nas cidades inviabiliza o fácil acesso a locais que à anos atrás íamos a pé. Hoje para conseguir se locomover, um pedestre tem que passar por diversas dificuldades, como, calçadas mal acabadas, faixas de pedestres sem pintura, falta de sinalização no trânsito, falta de educação de motoristas etc...

Redação 6

Os carros fazem mal como os cigarros. A necessidade do carro como principal veículo de locomoção é realmente questionável, pois além de poluir o meio ambiente também existem gastos específicos para o automóvel. É muito comum, em grandes cidades, ficar preso por horas no trânsito por conta da quantidade absurda de veículos. Isso resulta em estresse e irritação desnecessária que não fazem bem para a saúde. Existem diversas opções de transporte fora o carro, como bicicletas, trem, metro, ônibus e até a caminhada. São formas mais sustentáveis, menos poluentes e mais eficazes. Se no futuro o carro for considerado tão maléfico à saúde como o cigarro, nós teremos um avanço na qualidade de vida.

Redação 7

IBGE A instância brasileira adverte que o número de acidentes de trânsito andou aumentando decorrente aos anos, à prudência em que dirige o descuido do pedestre, as mas

sinalizações andaram. aumentando o número de mortes. A ponte do Jaguaré uma má sinalização feita pela fiscalização de trânsito, acabou ocorrendo duas mortes de dois, o motoqueiro e o que estava na garupa, tiveram suas cabeças desepadas. A vantagem e oportunismo em que o cidadão encontra para adquirir a carteira de habilitação de trânsito à tal quebra um dinheiro por baixo aos panos, as escolas de formação sendo comprados pelos novos alunos condutores. As próprias escolas de auto, estão proporcionando um aumento maior de condutores com má formação, e o condutor comprando uma passagem para a morte.

Redação 8

Certamente os carros serão o novo cigarro do futuro, o índice de pessoas que não se importam com a sua saúde, que não fazem exercícios físicos vem aumentando absurdamente de uns tempos pra cá, a população não quer saber sobre caminhar e sim sobre andar de carro. Os carros hoje em dia são mais usados como um meio de status do que como locomoção. As pessoas se baseiam muito sobre a outra apenas pelo carro. O governo não trás melhorias para os ciclistas o que poderia ajudar muito a incentivar as pessoas a saírem mais de bicicletas do que de carro. Existem várias ciclovias porém algumas que não dão nem para andar de tão prejudicada que se encontram e é um meio de transporte que não causa nenhuma poluição.

Redação 9

Hoje em dia o meio de transporte mais utilizado são os carros, ajudando a chegar nos destinos desejado com mais rápidos, economia e conforto com isso à casa dia que passa fabricantes estão melhorando ou criando outro modelo novo, fazendo assim que as cidade evoluam também. Porém como tudo tem seu lado bom, tem seu lado ruim também e o dele é a poluição do ar com as fumaças emitidas por eles acaba formando neblina cinza sob as cidades, deixando-as sem brilho, cor, pureza, segurando o calor proximo do solo. Se comparar essa poluição do carro com a do cigarro em lugar fechado não fazem o mesmo efeito, mudando o ar bom para um que causa problemas respiratórios até mesmo a morte. Com todo esses problemas o produtores de combustível encontrar algo mais puro e limpo, pois com tanta tecnologia esqueçam do mais importante que era o meio ambiente, mas esses combustivel deveriam que acessivo a todas as classes sociais e com facilidade.

Redação 10

Há quem não goste de carros, seja por ele poluir o meio ambiente ou então pelo fato de seu uso não ser uma prática muito saudável, no entanto, é de se concordar que ele é um instrumento que facilita, e muito, nossa locomoção. Mesmo que seu uso seja grande, é errado acusá-lo de "o novo cigarro", pelo menos em partes. Carro, a ferramenta mais utilizado atualmente, porém, um grande mal para a sociedade. Ele tem sim, suas desvantagens, como o uso excessivo causa, poluição, não só ao meio ambiente, mas também sonora e visual. Seus grandes percursos de engarrafamento incomodam e prejudicam a população, não só pedestres, como transporte coletivo. Aquilo que deveria ajudar, por fim, nos atrapalhou e muito nossa vida. Então, mais transportes coletivos seria a solução para uma melhor locomoção e menor

poluição em geral. No fim das contas, chega ser algo cômico que comparando o carro ao cigarro, este último seja até “melhor”.

Redação 11

Nós esperamos que não, que hajam outros meios de transportes como, já esta sendo implantado em alguns lugares, como a bicicleta o patinete, o Skate , ou até mesmo indo a pé sendo por longas distâncias. Nos grandes centros, sendo terríveis os congestionamentos, por debaixo da terra tem os metros. Mas como vemos em outros países como Holanda, Japão, que as pessoas saem muito de bicicleta, são países onde o governo e a população estão preocupados com o clima, com o trânsito, com a Economia e com o próprio planeta livre e seus habitantes que além de um clima mais agradável, terá tempo de chegar aos seus merecidos lugares. Tomara que aqui no Brasil, não só em Curitiba, ou São Paulo que os outros Estados tenham políticas públicas para que esse desenvolvimento cheguem até eles também .

Redação 12

A essência de ter um automóvel não encanta mais os jovens do hoje, eles tem se tornado cada vez mais consciêntes sobre seu meio e vem lutando por uma melhor qualidade de vida e de transporte público. Atualmente nossos jovens vem se preocupando mais com o futuro de nosso planeta e de nossa sociedade. Estando cientes dos males causados pela poluição dos automóveis, os jovens tem lutado para melhorar a qualidade e acessibilidade dos transportes públicos e das ciclovias. Através de uma maior acessibilidade e qualidade, os transportes públicos podem atingir e facilitar a vida de pessoas de toda e qualquer classe social, e ainda diminuir a emissão de gases e resíduos poluentes. Com ciclovias melhores, aqueles que tem de percorrer distâncias não tão longas podem aderir a um estilo de vida mais saudavel sem ficarem desconfortaveis com carros fumaceando em um engarrafamento ou com pistas em um estado ruim.

Redação 13

O carro é um meio de locomoção muito utilizado em diversos lugares pela sua praticidade, porém tem contribuido para a poluição ambiental, com a emissão de gases poluentes presente no combustível. Muitas pessoas não utilizam o carro apenas como meio de locomoção, mas também como objeto de desejo e aspecto de vida, não se atentando as suas conseqüências. O combustível Etanol seria uma das formas de reduzir o número de gases poluentes, pois é produzido através da cana-de-açúcar e seu bagaço pode ser reutilizado na produção de energia. Assim como a questão do cigarro apresentada à um certo tempo, o carro abriu essa questão, como possíveis formas de melhoria, conclui-se a redução no uso desse transporte, sendo substituidos por mêtros, ônibus, bicicletas, combustiveis menos poluentes, o investimento de cidades em sustentabilidade, através de campanhas e prapagandas.

Redação 14

Hoje em dia é notável a Necessidade do ser humano de algum Meio de transporte podem ser carros, motos, transportes publicos entre outros, e isso torna a vida de muita gente mais facil. Mas existem sim problemas serio sobre eles, um deles e a poluições causadas pelos gases emitidos pelos automoveis, e isso é um dos fatores que contribuem pro aquecimento global. As vezes precisamos deles sim não podemos negar, Mas Não é sempre assim, muitas Vezes poderíamos ir andando ou recorrer a uma opção mais sustentavel. Se pegarmos pelo Menos dois dias da semana se trocarmos o carro por um meio sustentavel, seria um grande começo para diminuirmos a poluição causadas por eles.

Redação 15

O carro sera o novo cigarro ? O carro sera o novo vicio do ser humano? Nos dias atuais vemos mas carros do que pessoas, andando a pé nas cidades. Isso quer dizer que os carros vem cada vez mais tomando espaço e virando um vicio entre as pessoas. Aonde poderia haver mais pessoas andando de bicicleta vemos muitos carros, e muita poluição. O ser humano esta cada vez mais carros novos, luxuosos com mas conforto e comodidade, anos potentes e com muita velocidade, isso para o ser humano é cada vez mas satisfatório e por isso acaba não olhando para outros meio de transporte. Isso é prejudicial para nossa própria saúde pois cada vez mas vem aumentando o sedentarismo, e também a poluição. Ao invés de ir de carro, para serviço escola, e outros lugares deveriam ir de bicicleta, ou até mesmo caminhando, seria benefico para saúde e também para o meio ambiente.

Redação 16

Um automóvel da atualidade tem suas rasões para ser chamado de "o novo cigarro" com sua péssima situação ambiental, porém a tecnologia fará a diferença. É de se indignar como é aceitável poluit tanto por um lado os automóveis é a riqueza da beleza para alguns, apesar do seu conflito. Por outro lado é apenas um meio de locomoção mais fácil. Como diz Jaime Lerner "você poderá continuar a usar, mas as pessoas se irritarão por isso". Com a tecnologia avançada, fará com que os carros futuros tendem a atender as prioridades da sociedade, sendo mais economicos e ecológicos fará com que ajude para o meio ambiente.

Redação 17

A Constituição Federal Brasileira, promulgada em 1988, prevê a todo cidadão o direito ao meio ambiente ecologicamente equilibrado, bem essencial para uma vida saudável, cabendo ao Poder Público e a comunidade preservá-lo para a atual e as futuras gerações. Entretanto, o alto número de carros em todo mundo causa diversos problemas no meio ambiente e na saúde humana. Primeiramente, o uso de veículos para locomoção aumenta a liberação de gás carbônico (CO₂) na atmosfera, o que gera poluição do ar e mudanças climáticas como a ilha de calor. Além disso, o enorme número de carros nas cidades também causa a poluição sonora e o trânsito, problema enfrentado principalmente em grandes cidades que prejudica o bem-estar de toda a população. Segundo Platão, filósofo grego, não se deve esperar uma crise para lembrar do que realmente importa, assim é necessário a adoção de

meios de transportes mais sustentáveis como, por exemplo o metro, popular principalmente nas metrópoles, antes que tanta poluição gere uma crise climática. Infere-se, portanto, a importância de priorizar o transporte público e outros meios mais sustentáveis, medida que alguns países já tomaram, e encerrar que a produção de gasolina e o uso dos carros sai caro, tanto economicamente quanto ecologicamente. Dessa forma usufruiremos dos nossos direitos sem esquecer o que realmente importa.

Redação 18

Com o avanço de novas tecnologias, o carro, que antes era objeto de desejo, vem se tornando cada vez mais popular. Entretanto, através do desenvolvimento tecnológico, problemas como a poluição e a emissão de gases prejudiciais à saúde da população tornaram-se problemas frequentes e que precisam ser combatidos, fazendo com que haja um aumento nas alternativas sustentáveis. Em primeiro lugar, considera-se que através dos benefícios obtidos pela revolução industrial, que teve início no século XIX, desenvolvimento de carros se facilitou. Contudo, ao mesmo tempo que houve um aumento na procura por automóveis, obteve-se uma concentração maior de poluentes, que contribuem para a aceleração em problemas como o aquecimento global e o efeito estufa. Além disso, assim como o cigarro foi e ainda é prejudicial à saúde, os automóveis tomam cada vez mais este espaço, podendo transformar-se no "novo cigarro". Em países como os Estados Unidos, através do desenvolvimento de novas formas de locomoção, somente 60% dos jovens de 18 anos possuem habilitação, demonstrando o aumento na procura por alternativas mais sustentáveis, como a "Yellow", uma bicicleta que é alugada através de um aplicativo para celular, sendo uma forma menos agressiva ao meio ambiente. Dessa forma, entende-se que o carro será o novo cigarro na sociedade com o desenvolvimento de novas alternativas de locomoção, espera-se que haja a diminuição no uso de carros, motos e ônibus para o aumento do uso de meios sustentáveis, pois somente assim será possível reverter este preocupante quadro em sociedade.

Redação 19

Na modernidade os carros tomam conta de todas as ruas. Em cidades como São paulo, precisou ser adotado o sistema de rodízio, para que não ajam tantos automóveis na rua de uma vez só. Enquanto em algumas regiões são feitos planejamentos urbanos para adotar diferentes maneiras de locomoção, outras tão pouco refletem sobre o assunto. Nesse contexto, o uso de carros e motocicletas desencadeiam uma série de problemas, como a emissão de CO₂, prejudicial ao meio ambiente quando liberado em grande quantidade. Além dos acidentes de trânsito. Estudos mostram que os meios de transporte mais perigosos são carro e motos. Dessa maneira, o uso de transporte público facilitaria a vida da população, pois reduziria os acidentes e a emissão de gases de efeito estufa. A poluição causada pela queima de combustível é tão prejudicial quanto a fumaça de um cigarro. Nessa análise, é necessário que as prefeituras em conjunto com os governadores dos estados repensem sobre a questão de locomoção urbana, para que sejam reduzidos os danos causados pela aglomeração de carros. Dessa maneira, aumentará a qualidade de vida dos cidadãos.

Redação 20

O automóvel, criado no início do século XX, marcou esse período como o início de uma era modernizada e tecnológica. Nesse sentido, o mundo evoluiu, assim como os milhares de novos modelos de carros que são lançados quase diariamente. Hoje, nota-se também um consumo desenfreado desses automóveis que prejudica o meio ambiente e a saúde da população mundial. Nesse contexto, observa-se um crescente consumo de diversos materiais produzidos pelas indústrias, como os carros. Desse modo, um maior consumo leva a uma maior extração de matéria-prima e conseqüentemente a degradação do meio ambiente. Além disso, com o aumento no número de automóveis, maior será a emissão de gases do efeito estufa, aumentando a temperatura terrestre, quando o aquecimento global. Eventualmente, o consumismo extremo de carros também pode prejudicar a saúde da população. Além de ampliar o sedentarismo e o estresse, os gases emitidos pelos carros afeta os pulmões e suas funções orgânicas. Mas, nota-se também que essa prática de consumo se tornou um vício da sociedade, em busca de status social, assim como o cigarro há décadas. Com base nos fatos apresentados, pode-se observar que a criação do século XX produziu muitos problemas, como a destruição da natureza e a redução da saúde das pessoas. Desse modo, é necessário repensar sobre a produção e consumo dos automóveis, analisando que esse padrão é somente um vício que nos foi imposto. Assim, utilizar os meios de transporte públicos e sustentáveis, para que haja uma redução dos impactos na qualidade de vida e no meio ambiente.

Redação 21

O futuro na terra Os carros serão os novos cigarros do Futuro? Vai depender do caminho que as ações e opinião popular vão tomar. Num mundo onde egoísmo e despreocupação com o ambiente são comuns, não dá para dizer com certeza o que vai acontecer. Cada vez mais são feitas descobertas na medicina, ciência, avanços na tecnologia, tudo para fornecer uma vida melhor, confortável e mais prática as pessoas. Mas acompanhando tudo isso, a perda de biodiversidade e a degradação do planeta vem tomando o rumo preocupante. O aquecimento global, as mudanças climáticas severas perda de fauna e flora são só algumas das diversas conseqüências da poluição, inclusive da emissão dos gases de carros. É mais do que claro que o ser humano, tomado pela ganância, está afundando o meio ambiente tornando quase irreversível a situação e desperdiçando recursos finitos que deveriam ser valorizados, sendo o planeta, atacado de todos os lados por poluição e degradação ambiental. Sendo assim, a emissão de gases gerados pelos carros é apenas um dos assuntos a serem pautados para chegar num acordo de preservação do planeta. O desenho “Wall-e” exemplifica bem como a situação ambiental pode ficar em situação irreversível. E isso, no desenho, forçou as pessoas a buscarem um novo lar. Ainda há Tempo de tornar o carro o novo cigarro, de cultivar valores sustentáveis e incentivar pessoas a buscar o melhor para o meio ambiente. Uma mobilização social, onde cada um faz a sua parte já seria um avanço, faria aumentar a vida útil de nossos recursos finitos como água e ar. Ainda há Tempo de deixar uma boa herança pro futuro. Ainda há Tempo de tornar o mundo, um lugar melhor!

Redação 22

Um novo vício : carros

Hoje em dia, as ruas são ocupadas por muito por muitos carros, novos, muitas pessoas adquirem esse automóvel por ele ser mais prático, mas não pensam no quanto o carro pode poluir. Várias pessoas trocam de carro todo ano, ele é considerado a nova "beleza" do mundo. Vários prefeitos priorizam transportes públicos, fazem campanhas, mas poucas pessoas aderem a essa ideia de deixar o trânsito mais ecológico. O carro pode ser considerado um novo vício, muitos já dizem que é o "cigarro do futuro". No Brasil, assim que o jovem completa 18 anos, o seu sonho é poder ter a habilitação nas mãos e assim poder ter um carro, ao contrário dos EUA, onde muitos jovens não optam em tirar carta de motorista para poder dar mais sustentabilidade ao ambiente e assim utilizam o transporte público. Segundo estudos, o número de jovens sem habilitação dobrou em 30 anos. Contudo, se continuar assim, não vai ter mais espaço para outros meios de locomoção, os carros dominarão as ruas e os pedestres só irão andar debaixo da terra. As pessoas devem se conscientizar e não deixar os carros virarem a nova droga do mundo, devem optar por transportes públicos, bicicletas, deixando o trânsito muito melhor, sem engarrafamentos, fazendo com que o ar fique mais puro, sem muita poluição, pois só daqui muitos anos, elas vão perceber o quanto um carro faz mal para o ambiente.

Redação 23

Desde o fim da guerra fria, a cultura que hoje é vivenciada é influenciada pelo capitalismo e os carros vem sendo um objeto de desejo da população. Do mesmo modo em que os cigarros na década de oitenta eram tidos com artigos de beleza e pompa porém extremamente prejudiciais, os carros estão indo ao mesmo rumo: se tornarem artigos de luxo e alta periculosidade. Dessa forma devemos lançar um olhar crítico sobre tal desejo e tais máquinas e os impactos negativos causados pelo novo e motorizado cigarro. Primeiramente se faz necessário que a população pare e reflita sobre o mal que os carros trazem ao meio ambiente e às infraestruturas das cidades. O uso de automóveis vem gerando altíssimos níveis de poluição no mundo, e também prejudicando a infraestrutura das cidades com problemas como trânsito em excesso e congestionamento. Dessa forma o cigarro dos anos oitenta volta agora com uma carapuça metálica, criando diversos problemas incômodos que são marcados pelo desejo capitalista das pessoas de se obter algo inulçável, sendo que, existem muitas outras alternativas de transporte, não há mais necessidade de optar por um causador de poluição e condicionamento. Efésto disse "Nada é permanente, exceto a mudança", levando em conta o pensamento de tal filósofo a sociedade deve ser aberta para a mudança antes que um grande câncer se instale no pulmão do mundo. Em virtude dos fatos supracitados, podemos notar que o planeta caminha para colapsos ambientais e sociais (como o efeito estufa e as milhares de mortes no trânsito) causados e agravados pelo uso exagerado de carros, fato que, é estimulado pela necessidade capitalista de posse. Portanto cabe a população criar consciência dos danos causados pelos carros e adotar novas práticas de transporte como transporte públicos e bicicleta, afim de remediar os danos já causados e impedir o surgimento de novos.

Redação 24

É aceitável dizer que no Brasil as pessoas viraram reféns dos automóveis. Nos dias de hoje o vício de ter um carro é como quem fuma cigarro, ou seja, muitas pessoas não consegue ficar sem. O carro faz parte do dia-a-dia, porém o exageiro da população trás grandes consequências. Em primeiro lugar, as pessoas de baixa renda que moram em São Paulo necessitam de um automóvel ou meio de transporte para se locomover, seja ele um Táxi, Uber, Ônibus ou Metro. Porém, a população de classe média e alta transitão de carro todos os dias e atrapalha a vida dessas pessoas que não tem automóveis próprios, trazendo como consequência as dificuldades da mobilidade social no trânsito. Isto é, os indivíduos que moram em bairros periféricos com pouco acesso ao transporte deve sair de manhã da sua casa retornando somente à noite, sendo que, normalmente ficam presas no trânsito por duas horas até chegar no seu local destinado. Além disso, o carro tornou-se um vício pelas pessoas para mostrar os luxos e riquezas, que cada vez mais querem expor suas grandes. Sendo assim, gera outra consequência grave como o aumento de furtos de veículos, ou seja, alguns indivíduos que não tem condições de comprarem um carro, acabam roubando para ter aquilo que deseja mesmo que seja de maneira ilegal. Por isso, vale afirmar que o carro é o novo cigarro, comparando ao vício de uma pessoa que fuma e não consegue parar em relação a outra que não consegue ficar sem utilizar o carro. Portanto, somos dependentes dos automóveis, ele já está incluído na rotina dos indivíduos sem que possamos perceber. Pois, se o automóvel para é como se a vida tivesse parado Descreve em seu poema Carlos Drummond de Andrade.

Redação 25

O carro, o meio de transporte criado para facilitar a vida de quem o utiliza, um objeto que ficou bastante popular entre as pessoas e que hoje é indispensável no cotidiano de grande parte da população. Contudo, com o aumento populacional, principalmente os centros urbanos, os carros também aumentaram. Estima-se que cada família brasileira tenha pelo menos um carro em sua garagem. Com esse aumento, veio também, a elevação de gases poluentes emitidos na atmosfera, o que acaba por prejudicar o meio ambiente. Para incentivar o cuidado com a natureza, e como forma de contornar a situação do aumento de veículos o meio de transporte coletivo está sendo preferido e adotado por muitas pessoas. Assim como na época do cigarro, campanhas e incentivos por incentivado por parte do governo para uma maior inclusão desses meios de transporte. Portanto, com o uso dos metrô, trens e afins, além de diminuir a poluição e o número de carros congestionando as vias, garante-se também um meio ambiente mais saudável e um futuro mais sustentável.

Redação 26

A dificuldade de abandonar o cigarro Apesar dos diversos meios de locomoção encontrados nas grandes cidades, o carro ainda parece ser o preferido da população, mesmo com os gastos de combustível e seguro, falta de vagas para estacionar, tráfego intenso e eventuais manutenções. O motivo dessa escolhas se dá pelas falhas do transporte público, como atrasos, ineficiência, lotação e veículos precários. Isso faz com que as pessoas invistam em automóveis pensando no próprio conforto, e não nas outras questões, por exemplo a

agressão ao meio ambiente. Carros e ônibus são grandes poluidores do ar, mas já existem alguns ônibus que causam menos impacto ecológico. Os metrô e trens são grandes alternativas também para a poluição visual, e há as bicicletas, que além de não causarem impacto ambiental, são saudáveis. No que diz respeito aos carros, há os modelos elétricos, porém são inacessíveis para a maioria da população, além de não solucionar o problema dos altos números de veículos nas ruas. É necessário que haja investimento na mobilidade urbana, criando mais ciclovias e corredores de ônibus, promover maior conforto e eficiência para aqueles que utilizam trens e metrô, caso contrário, essa alternativa jamais serão a primeira opção de transporte para a população. Portanto, se as melhorias e investimentos não forem realizadas, o carro certamente será considerado um cigarro, emitindo poluentes, causando tráfego. Porque mesmo não sendo de propósito, a rotina de perder tempo dentro de um carro e se estressar por isso, tornou-se um vício difícil de abandonar.

Redação 27

A humanidade nunca mais foi a mesma desde que a Revolução Industrial e trouxe o carro, inicialmente um artigo de luxo utilizado pela alta classe para se locomover, atualmente é um meio de locomoção particular muito utilizado por grande parte da população mundial, considerado um bem necessário de grande conforto, praticidade, segurança e infra insubstituível. No entanto já observamos os malefícios causados por esse meio de transporte, as emissões massivas, principalmente em regiões metropolitanas, de CO₂ na atmosfera, o estresse causado pela extensa congestão no tráfego. Pesquisas afirmam que metade das mortes atualmente são causadas por acidentes no trânsito. Uma comparação muito efetiva seria relacionar o uso de cigarros de nicotina, produto também responsável por intensa emissão de gases nocivos ao ser humano, suscetibilidade à instabilidade emocional e principal responsável pela morte por câncer de pulmão. Contudo muitos movimentos têm ganhado destaque, como a locomoção por bicicletas e vias que facilitem essa locomoção, investimento no transporte público e corredores de ônibus, entre outros. Movimentos como esse garantem a desmistificação do transporte privado como bem insubstituível e promovem qualidade de vida a uma maior gama de classes. Concluindo-se que o investimento em infraestrutura no transporte público e a promoção ao uso de locomoções sustentáveis serão os grandes responsáveis pela diminuição, se não forem pelo fim, da dependência ao transporte privado.

Redação 28

Menos poluição e mais sustentabilidade. Tendo em vista que, o automóvel nos dias de hoje é o transporte mais usado pelas pessoas, facilitando a locomoção e proporcionando maior rapidez no transporte de mercadorias e de passageiros. Entretanto, o uso deste já é considerado o novo cigarro (dependência), pois o aumento de carros a cada ano só cresce e esse número irá se expandir cada vez mais com o auxílio da tecnologia. Acredita-se, que 80% dos jovens americanos possuem habilitação, e que 60% dos jovens com idade de 18 anos já possuem seu documento para dirigir. No entanto, esses dados acabam sendo preocupante, pois o carro por ter uma demanda cada vez maior de consumo, acaba tendo uma influência direta no meio ambiente, por meio de resíduos tóxicos e gases liberados durante a sua fabricação, até a chegada dos veículos nas ruas. Diante disso, podemos perceber que esse

meio de transporte de rápido acesso, se tornou uma dependência da população, deixando principalmente a vida de quem depende desse transporte mais prático. Embora, nas décadas anteriores essa procura também era grande, mas o transporte desses passageiros eram feitos de trem, um mecanismo que não agredia tanto o meio ambiente, como os carros fabricados nos dias de hoje. Portanto, a tecnologia tem um papel fundamental na criação dos novos métodos e na elaboração de projetos sobre transporte da população. A conscientização dos consumidores deste, devem cumprir com seu papel também, todos tem a consciência que o automóvel polui, mas poucos praticam optar por escolhas sustentáveis, que irão contribuir para a preservação do meio ambiente e com isso possibilitará uma melhor qualidade de vida ao praticante deste.

Redação 29

Por conta do excesso de veículos nas cidades e pelo o que eles produzem ao funcionar, já são considerados os cigarros para o meio ambiente. Assim como, o próprio cigarro é prejudicial para a saúde do ser humano o excesso de veículos é prejudicial ao meio ambiente. O que muda é a consciência sobre esse fato. É preciso que o ser humano crie essa reflexão de que o veículo torna-se prejudicial ao meio ambiente por produzir substâncias que afetam o ambiente de forma negativa. E que essa consequência nos atingirá de alguma forma. Além de toda a poluição que ocorre, o espaço físico que todos esses veículos ocupam é surreal, os engarrafamentos nas grandes cidades das capitais não são mais exclusivas delas, a quantidade de veículos também crescem nas pequenas cidades, tornando tudo muito insustentável. Esquece o espaço físico dos centros das cidades, as ruas são exclusivas das enormes filas de veículos que se fazem presentes nelas. Causando cansaço físico, psicológico e muito estresse. A relação que o carro tem com o cigarro dentro dos textos da proposta é até mesmo sobre o tema indagado de uma pergunta, é se as pessoas terão a consciência de que é preciso agir como agiu diante do cigarro. Se a forma de ver o cigarro hoje e ter a consciência que se tem sobre ele será a mesma com os carros. Mudar os hábitos ruins para hábitos saudáveis é fundamental e não seria diferente sobre a consciência do uso dos carros. Se o carro será o novo cigarro, não sabemos, mas se for, que seja a respeito da consciência sobre o funcionamento dele ao meio ambiente. Assim como o texto 4 nos traz, "um dia as pessoas [...] em que cigarro era aceito em restaurantes, aviões e lugares fechados".

Redação 30

Mobilidade urbana e sustentabilidade: Um Desafio possível. Atualmente a sociedade está caminhando rumo a Revolução 4.0 e as novas relações de trabalho e transporte. Com a ascensão de carros modernos e o poder social e aquisitivo que o indivíduo demonstra ter quando se torna proprietário dos mesmos acaba deixando a seguinte problemática: será realmente necessário ter um automóvel particular? Qual o benefício social que o mesmo trás? Os carros são produtos de alto custo comercial e valor social, são movidos com combustíveis fósseis e altamente poluentes. A maioria das famílias de classe média possuem no mínimo um carro particular, utilizam raramente os meios de transportes coletivos e públicos e são responsáveis por uma considerável parcela de poluentes soltos na atmosfera. Um problema latente dos grandes centros urbanos e metrópoles é a falta de mobilidade urbana, o trânsito e

engarrafamentos inviabiliza a agilidade na locomoção dos indivíduos, sendo o grande número de carros o principal agente causador desta problemática. A fim de reduzir a poluição de gases liberados por carros e proporcionar de fato a mobilidade urbana o Poder Público deve investir em políticas públicas de transporte e combustíveis de fontes renováveis, aumentar a malha ferroviária dos municípios e linhas de ônibus. Apenas a coletividade dos meios de transporte e a redução de carros podem desenvolver a mobilidade do município e de democratizar acesso à cidade.

Redação 31

Atualmente possuímos a lei que é proibido fumar em locais públicos fechados. Essa Lei foi feita por conta de que muitas pessoas que fumam acabam incomodando as que não fumavam. Em algumas décadas atrás surgiu o automóvel que no começo não poluía muito, porém com um tempo foi evoluindo onde houve a criação do combustível, que era utilizado carvão mineral, onde também acabava poluindo até chegar onde estamos, com veículos bons e ruins, mas que poluem da mesma forma só que com diferentes combustíveis. Fosseis que acaba resultando na poluição do ar e da camada de ozônio. Mesmo poluindo bastante nós acabamos adquirindo algum automóvel por questões de estética, beleza e necessidades. Cada dia que se passa cresce muito a quantidade de automóveis, ocasionando o aumento da poluição. Algumas cidades para realizar uma diminuição investem em ciclofaixas, corredores de ônibus, veículos sustentáveis e combustíveis que diminui o efeito da poluição. Apesar de investimentos a maioria das pessoas não conseguem viver sem um automóvel, porém elas poderiam optar por outras formas de locomoção evitando assim o aumento e até a poluição. Porém mesmo com ajudas governamentais o carro ou qualquer outro veículo automotor acaba sendo o chamado cigarro do futuro porque as pessoas vão utilizando e poluindo do mesmo jeito que se fosse um cigarro qualquer, para isso as pessoas para não acender esse novo cigarro muitas delas acabam utilizando outros meios de locomoção, porém alguns deles acabam tendo o mesmo efeito de poluição e pode prejudicar até quem não utiliza o carro com a fumaça onde acaba poluindo da mesma maneira também, dando como vilão os combustíveis utilizados nos carros, para que aconteça uma diminuição poderiam ter outros meios e formas de combustíveis mais ecológicos e menos poluentes.

Redação 32

Progressismo Automobilístico

É fato que os veículos automobilísticos, de certa maneira auxilia o homem a chegar com uma rapidez maior a seu destino. Porém, os mesmos são movidos por compostos que liberam substâncias como o carbono para a atmosfera, e o número de carros no mundo é exorbitante, resultando em uma poluição atmosférica. “No meio do caminho tinha uma pedra, tinha uma pedra no meio do caminho”. De acordo com o Poeta Modernista Carlos Drummond de Andrade, há um obstáculo que dificulta a passagem do desenvolvimento, precisando ser uma vida. Nesse contexto, infere-se portanto, que é imprescindível que o governo juntamente com o Ministério da Ciência e Tecnologia, reduza ou extingue programas que visam a pesquisa científica acadêmica, porquê através desses programas como CNIQ ou Fapesp, a humanidade tem a Chance de explorar fontes sustentáveis, como o Etanol, para movimentar

veículos sem a degradação e poluição do ar. Convoca-se, as autoridades competentes para a solução da problemática, desta forma, a concentração de poluentes no mundo será menor e estaremos expostos a uma segurança maior, caso o contrário, sem pesquisa e inovação o carro será o novo "cigarro" e o mundo será auto destruído.

Redação 33

Século XXI, a era da reinvenção do carro e de como usá-lo. Com o processo de globalização, além do aumento na velocidade com que a informação é transmitida, houve também o aumento na velocidade em que as pessoas transitam. Um dos responsáveis por isso foi o americano Ford, que popularizou o automóvel, com a fabricação em massa do modelo Ford T. Em seus primórdios, o veículo cumpria com seu objetivo de facilitar a locomoção, contudo na contemporaneidade, tornou-se um grande problema no trânsito das grandes cidades e para o meio ambiente. A Metrópole de São Paulo é um dos exemplos de como o automóvel tornou-se prejudicial no cotidiano do trabalhador. Uma das maiores cidades do mundo e mais populosa, tem sérios problemas para com o tráfego de veículos, com filas de carros e mesas nos horários de “pico”. Para tanto, o transporte público, principalmente o metrô, virou uma das melhores opções no trajeto de casa ao trabalho, além da existência de aplicativos de carona, que visam a redução de automóveis no trânsito e retiram a obrigação do indivíduo em possuir um carro, visto o bom custo-benefício desses aplicativos. Outro cofator prejudicial na utilização do carro, é a alta emissão de gases poluentes para a atmosfera, tal como o gás carbônico (CO₂), gás de efeito estufa. Mas como diria Steve Jobs, uma das mentes por trás da grande empresa de tecnologia Apple: "a tecnologia move o mundo", e graças a ela que já existem carros movidos a energia elétrica, e portanto não poluidores, com são os carros da montadora Tesla. Todavia, esses automóveis sustentáveis ainda não são acessíveis para todas as classes sociais, sendo assim a bicicleta (como é visto em Amsterdã) ou o patinete (como acontece em São Paulo), continuam sendo a melhor opção para o transporte rápido e sustentável. Dificilmente os carros serão extintos no futuro, trata-se de um mercado multibilionário. Contudo é certo que, a forma com que a humanidade o utilizará será diferente, tal como já acontece no presente com os aplicativos de carona, além de que o combustível fóssil será substituído pela energia elétrica, e esse tipo de carro se tornará mais acessível. Portanto no futuro existirá menos carros e a gasolina perderá o seu papel como principal combustível para os veículos.

Redação 34

Ao longo da história observa-se a importância das rodas, ao qual privilegiou a invenção do transporte locomotivos, o carro. Por um lado a criação do automóvel facilitou a locomoção de um lugar à outro, diminuindo consideravelmente o tempo percorrido. Por outro, transformou as grandes cidades, principalmente as que estão no centro, em verdadeiras pistas estáticas. Além da enorme poluição que os combustíveis; muitas vezes não renováveis, acarretam a vida dos moradores das cidades. Contudo, o que era para facilitar o dia a dia do homem acabou se contradizendo. De fato, a 1ª e 2ª Revolução Industrial ampliou o desenvolvimento tecnológico, e de acordo com Tommaso Marinetti em "Manifesto Futurismo", enriqueceu o mundo com a beleza da velocidade, mas sem dúvidas trouxe juntos suas

consequências. A medida em que o sistema econômico se estabeleceu pelo capitalismo houve grande influência pelo consumo e possibilitando que grande parcela da sociedade obtivesse seu próprio automóvel. Com isso, ao longo do tempo, as ruas e estradas foram preenchidas pela grande imobilidade ao que se diz do trânsito, dificultando ainda mais a locomoção das pessoas e aumentando, consideravelmente o tempo gasto nos trazesos. Esta onda tomada pelas máquinas foi refletida em um período literário como o Arcadismo, que trouxe em seus temas a fuga das cidades para o campo, pois a primeira estava repleta da agitação causada pela própria tecnologia e a cidade agora tornava-se um local de refúgio, sendo um local calmo e ameno que distraía a angústia vivida nas cidades. Uma das consequências também está na poluição provocada pela queima de combustíveis em níveis anormais e que, ao longo do tempo prejudicarão os habitantes, principalmente por problemas respiratórios. Assim o uso do transporte público visa uma das soluções para amenizar esse risco, trazendo mais mobilidade ao trânsito e diminuindo os gases eliminados na atmosfera. Conforme tudo isso, uma invenção que ajudaria à todos acaba por causar ainda mais danos tanto em sua mobilidade, quanto à sua saúde. O incentivo ao uso de transporte público é fundamental para amenizar estes problemas e evitar futuros desastres ambientais causados pela poluição dos automóveis.

Redação 35

Com a rápida urbanização, o carro tem sido uma escolha cada vez menos utilizada pela população, sendo considerado por muitos um vício. Porém fatores com o alto custo para se manter, inúmeros riscos no trânsito e a conscientização da sustentabilidade são exemplos de que o carro não será o novo cigarro da população. Os altos custos para se manter um carro atualmente é um dos principais motivos para a população optar pelo transporte público. Embora inúmeras empresas automobilísticas tentem inovar em suas tecnologias, os elevados preços de manutenção fazem com que as pessoas não tenham carros e utilizem os meios públicos de transporte, visto que, são mais econômicos econômicos- entretanto não possuem o conforto de um carro e assim mais utilizados. Logo, há mais pessoas nos transportes urbanos do que em carros. Com os elevados índices de roubos de carros, o medo tomou conta da população. Com o passar do tempo esse medo fez com que as pessoas utilizassem menos seus carros com o receio de sofrerem algum assalto. Além de assalto o risco de acidentes aumentou gradativamente nos últimos anos, posto que, muitas pessoas que utilizam automóveis não possuem habilitação e com isso colocando em risco a vida de inúmeras pessoas. Atualmente movimentos em prol da sustentabilidade são mais frequentes, já que é um assunto cada vez mais debatido. Grupos como o Green Peace lutam pela diminuição do tráfego de carros, posto que, o gás liberado pela fumaça dos automóveis são um dos principais causadores do efeito estufa e do aquecimento global. A população ao se solidarizarem com esses grupos optam em utilizar bicicletas ao invés de utilizarem seus veículos. Com esses fatores e muitos outros e que são os motivos da utilização de carros está diminuindo. Portanto entende-se que as pessoas estão perdendo interesse nos carros e com isso colocando a afirmação de que os carros serão os novos cigarros da população para baixo, pois as pessoas não possuem um vício em seus automóveis

Redação 36

Durante o governo do Presidente Rodrigues Alves, a cidade do Rio de Janeiro recebeu uma reforma inspirada na capital francesa. Um dos objetivos foi acomodar a chegada da mais nova tecnologia: os carros. Eles passariam a significar poder aquisitivo, luxo e desenvolvimento e se tornariam o centro da vida urbana. Porém, o tempo passou e as vontades mudaram e atualmente existe o desejo por novas alternativas, que considerem o coletivo, a acessibilidade e a sustentabilidade, enquanto o carro é abandonado no sinal vermelho. De acordo com o filósofo polonês Zygmunt Bauman, na sociedade de consumo, o indivíduo e suas relações são definidas por aquilo que ele consome. Por décadas, os automóveis exerceram este papel definidor: revelavam a classe social, o estilo de vida e os gostos. Desse modo, não tê-los não era uma opção e conseguir uma carteira de motorista e o seu primeiro carro era como um rito de passagem. Esse significado é exposto em séries norte-americanas como “That 70’s show”, na qual o personagem principal, Eric, sonha com esse grande momento. Entretanto, essa importância não ficou restrita à vida pessoal, mas também começou a determinar todo o espaço urbano: o automóvel passou a reger o desenvolvimento das cidades e a construção de vias tornou-se uma prioridade - como disse Washington Luiz, presidente do país "governar é abrir estradas". Assim, criou-se uma dicotomia entre o individual e o coletivo, no qual o segundo preterido em favor do primeiro, pois este era o mais confortável. Mas esse conforto teve como consequência problemas ambientais como as ilhas de calor e a poluição do ar, a qual também agrava a inversão térmica. Portanto, inicia-se uma nova era em que essa preferência é invertida e novas necessidades emergem: os transportes individuais perderam o posto de grande sonho e novas configurações que priorizam a mobilidade coletiva ganham espaço. Como disse o poeta lusitano Camões, na abertura de um de seus sonetos: "Mudam-se os tempos mudam-se as vontades". O novo tempo pede por modelos sustentáveis e inclusivos e espera-se que os futuros governos sejam capazes de empreender essas reformas.

Redação 37

Tabaco de quatro rodas

No início do século passado, era muito comum o transporte de pessoas em charrétes de aluguel. Entretanto, esse serviço fez com que as cidades ficassem cheias de motoristas particulares. Diante desse contexto, ao olhar o Brasil nos dias de hoje, é possível concluir que muitos deixam seus empregos para trabalhar nesse ramo, fazendo com que o número de automóveis aumente e as ruas fiquem congestionadas, causando mais poluições, mortes por acidentes e irritando os pedestres e motoristas ainda mais, fazendo do carro o novo cigarro. Nesse sentido, é preciso entender suas verdadeiras causas para solucionar esse problema. A princípio, é possível perceber que essa circunstância deve-se a questões político-estruturais. Isso está relacionado ao fato de alguns Governos oferecerem uma baixa qualidade no transporte público, com falta de segurança e super-lotações. Prova disso é infelizmente os metros de São Paulo, fazendo com que o indivíduo peça um transporte pelo aplicativo que é mais seguro e confortável. Dessa forma, cada dia mais o aplicativo de locomoção particular ganha mais carros cadastrados. Outrossim, vale ressaltar que essa situação é corroborada por fatores socioculturais. Durante a Formação do Estado Brasileiro, os carroceiros realizavam

a locomoção dos Brasões e suas famílias, sendo assim, as classes mais baixas não tinham esse privilégio. Bem exemplo disso, são os índices, que indicam um aumento de motoristas de aplicativos. Dentro dessa lógica, logo, nota-se que os automóveis estão ganhando cada vez mais espaço nas ruas e com isso a poluição só tende a crescer, gerando complicações nos pedestres e irritando os mesmos. Torna ver evidente, portanto, que essa situação pode ser nomeada de Tabaco de quatro Rodas, já que, os automóveis veem ceifando vidas, poluindo o mundo e causando desconforto com isso. Entretanto, para diminuir todas as estatísticas é necessário a melhoria em todos os aspectos do transporte público. Com essas medidas, talvez, irá ocorrer uma diminuição na Utilização de Aplicativos de locomoção.

Redação 38

O Uso do carro: Necessidade ou malefício? No século XX, o cigarro era visto como artigo de luxo. Milhares de pessoas utilizavam-no depois de conhecê-lo por meio de propagandas utópicas apresentadas na televisão. Com o passar dos anos, a conscientização humana evoluiu e os malefícios do cigarro passaram a ser reconhecidos e reprimidos pela população. Porém, atualmente, outro objetivo tem se tornado um artefato de luxo: o carro; e, mesmo após tantas notícias alarmantes sobre o atual estado do meio ambiente, as pessoas não param de comprá-los e exibí- los pelas ruas. Em algum momento ocorrerá tamanha mudança como ocorreu com o cigarro? Isso dependerá da evolução de conscientização e da transformação na rotina da população. Em primeiro lugar, a maior dificuldade não se encontra na mudança de hábitos, mas sim na mudança do pensamento social, e a mídia, assim como as empresas automobilísticas, são as grandes vilãs dessas transformação. Com suas mais diversas propagandas, muitas vezes ilusórias, os meios de comunicação apresentam os carros como máquinas cada vez mais potentes e luxuosas, que devem ser consumidas como um item de necessidades básica, gerando à cada venda mais lucro para a indústria automobilística. Essa, por sua vez, segue um modelo de economia capitalista, e utiliza de recursos naturais em sua produção, o que degrada aos poucos o meio ambiente. Por conseguinte, é necessária uma transformação nos hábitos da população. Atualmente, no Brasil, grande parte dos indivíduos possuem ao menos um automóvel. Esses, são uns dos principais poluidores ambientais, já que a queima da gasolina gera emissão de CO₂, um dos principais gases agravantes dos efeitos estufa. A solução para àqueles que se preocupam com o futuro do meio em que vivemos é o uso transporte público (que minimaliza essa emissão ao transportar diversas pessoas em uma só viagem) ou ainda o uso de transportes que não geram gases poluentes, como a bicicleta. Tendo esses aspectos em vista, fica claro que é complexo obter uma mudança drástica no pensamento da população. No entanto, é possível ver uma transformação futura se a educação ambiental e a participação do automóvel como fonte de poluição forem discutidos nas escolas, gerando nas futuras gerações um pensamentos sustentável. Entretanto, como diria Paulo Freire, a educação é a base da mudança social mas, sozinha, não consegue transforma a sociedade. Sendo assim, à medida e às empresas cabe o papel da apresentação e produção dos automóveis de maneira Sustentável, reforçando, também seus malefícios. Da mesma forma, cabe ao governo a organização do transporte público, Tudo isso, contribuirá para a formação e conscientização ambiental da população trazendo, com o decorrer dos anos, o "esquecimento" do uso do carro, assim como, atualmente, ocorre com o cigarro.

Redação 39

A invenção da roda é indubitavelmente uma das maiores criações feitas pelo homem e sua importância diverge em vários aspectos úteis e ao mesmo tempo proporciona conforto e comodidade como o ato de dirigir um automóvel. Nesta perspectiva, os carros são atualmente o meio de transporte na qual o acesso a lugares distantes se tornou mais rápido porém, corrobora para o agravamento do meio ambiente. Na década de noventa, era comum e repetitivo a exibição de comerciais de cigarros sem o menor regulamento sobre o seu uso excessivo. Com isso, o número de pessoas com problemas respiratórios aumentou consideravelmente, uma vez que as propagandas influenciavam o consumo como hábito normal. Não obstante, os comerciais de compra e venda de automóveis, tentam proporcionar ao espectador o conforto de obter um veículo sem a preocupação da emissão de gases poluentes com o mesmo efeito de apenas ampliar os lucros de produção sem a preocupação com a saúde, nos mesmos moldes dos comerciais de cigarros dos anos noventa. Desta maneira, não somente o bem-estar fica comprometido, mas também o meio ambiente sofre com a presença do dióxido de carbono em baixas camadas da atmosfera oriundas dos escapamentos dos carros. Outrossim, o alto lucro obtido através das propagandas acelera a fabricação do mesmo nas fábricas e conseqüentemente a alta demanda ocasiona o problema de congestionamento de veículos em grandes cidades como a de São Paulo. Se faz necessário, a regulamentação da exibição de comerciais de automóveis veiculados na televisão e na "internet" por intermédio de ações preventivas do Estado sobre a escolha de práticas mais limpas que substituam o uso excessivo de automóveis. Além disso, nas escolas devem ser implementadas aulas para a conscientização da emissão de poluentes causada pelos veículos através de palestrantes especialistas no assunto. Deste modo, o carro não deixará de ser o novo cigarro no futuro, mas sim continuará proporcionando o conforto e comodidade no ir e vir e facilitando o acesso a lugares longínquos.

Redação 40

Obsoleto como carro

A individualização do consumo revela um estilo de vida influenciada por fatores midiáticos e culturais, em uma sociedade Capitalista, que provém de uma noção caracterizada pelo consumo, vem sendo transformado nas últimas décadas, como por exemplo, o carro e o cigarro. Portanto, em meio a uma efervescência cultural dos anos entre 1930 e 1960, o cigarro era um dos objetos de grande consumo, repercutindo em diversos meios, como o cinema, comerciais de televisão, novelas, etc. Com o tempo, a aceitação do cigarro foi sendo repudiado, pelo fato de que pesquisas foram feitas sobre os componentes que nele há, foi descoberto que suas substâncias eram prejudicial ao corpo de quem consome e de quem recebe a fumaça do cigarro, chamado este fumante passivo. Com preocupações fisiológicas do seu bem-estar, as pessoas começaram a diminuir a aceitação tanto suas como de outrem o uso. Por conseguinte, o carro foi o modelo das inovações tecnológicas da Revolução Industrial, Ford foi a mente responsável por trazer a possibilidade de um cidadão comum de conseguir comprar um carro, preços próximos a realidade popular. Com o desenvolvimento tecnológico no passar das décadas, carros mais novos e sofisticados surgiram, se tornando o consumo destes como forma de ascensão social. Descobre-se que a fumaça emitida pelo carro é prejudicial ao meio ambiente e ao ser humano, da mesma forma que a fumaça do cigarro faz mal ao corpo, a fumaça do

carro também, seu prestígio vem decaindo ao passar dos anos, preferência por transportes públicos vem aumentando e novos mecanismos para a redução do uso do carro vem demandando mais. Por fim, estes objetos de consumo, carro e cigarro, se tornam cada vez mais obsoletos, e culturalmente revistos suas importâncias para o modelo de vida contemporâneo, o consumo é diminuído. A ciência, como auxílio de novas descobertas influenciam o novo homem e sua vida que faz parte deste planeta, pensa e revê, seus conceitos que se entranham para que se construa um mundo melhor.

Redação 41

O bem veloz sendo repensado Durante o século XX, a vanguarda futurista exaltava o automóvel e a velocidade, mas com o passar dos anos, e a percepção dos efeitos dos produtos como esse geram ao meio ambiente e ao cotidiano, essa visão foi alterada. Com a popularização desse bem, a conta a ser paga pelo planeta tem sido cara, algo que mobilizou grandes transformações no modo de vida contemporâneo e a visão para o futuro. Os grandes problemas envolvendo o carro são a questão ambiental e a qualidade de vida, com uso e ocupação do espaço urbano, por exemplo. Os combustíveis fósseis, que ainda são amplamente utilizados, são um dos responsáveis por parte das emissões de gases poluentes na atmosfera, devido ao processo de combustão. Esses gases acabam por agravar processos naturais como o efeito estufa, tendo consequências negativas, a exemplo do aumento da temperatura terrestre e derretimento de calotas polares. Além da questão ambiental, com a popularização do automóvel, as cidades começaram a ser projetadas pensando nisso e não no pedestre- no Brasil, medidas com a ampliação de rodovias no Governo JK e a construção de Brasília e suas superquadras demonstram a importância dada ao carro, e a negligência à questão ambiental e reais necessidades e otimização da vida urbana. Esse protagonismo do carro vem sendo alterado, devido a consciência ambiental, de mobilidade urbana e econômica para se pensar a cidade e vida urbana. Dentre os novos modelos adotados para a melhoria das questões supracitadas, empresas e aplicativos como a Uber e bancos, começaram a disponibilizar patinetes e bicicletas para aluguel. Esse fenômeno vem sendo recorrente em grandes cidades como São Paulo. Medidas como essa, associada a melhoria na infraestrutura de transportes coletivos, mostram como o carro, aos poucos, vem perdendo espaço, ou sendo visto com outros olhos. Portanto, o senso de urgência quanto as diferentes formas e consequências do uso excessivo do carro motivaram alterações no estilo de vida contemporâneo. A tendência está voltada para alternativas mais sustentáveis de como melhorar o atual cenário, já bem afetado e congestionado.

Redação 42

O termo "cigarro do Futuro", empregado por Jaime Lerner, arquiteto e ex-prefeito de Curitiba, pode ser visto como um termo válido, já que pessoas fumantes em diversos locais do mundo, e em diversos momentos históricos, eram bem vistas na sociedade, pois, além de seu valor estético, possuía também uma grande valorização social. Portanto, essa ideia também pode se refletir nos carros, já que quando ocorreu seu surgimento, nem todos podiam ter acesso a ele. O carro pode ser considerado uma das maiores criações do ser humano, pois além de ganhar muitos adeptos desde o seu aparecimento, a Vanguarda Europeia de- nominada

Futurismo costumava representar a beleza da velocidade desses automóveis em suas obras de arte, o que instigava muito a sociedade. Ainda que tenha tido uma abrupta repercussão, cada vez mais esse veículo pode ser visto caindo em desuso nas cidades grandes, ao ser feito o uso de transportes coletivos para a diminuição de poluentes transmitidos por conta do grande número de habitantes. Nos Estados Unidos, também é possível observar a queda de usuários desse automóvel, visto que, jovens a partir de seus 18 anos não tiram carta de motorista com a mesma frequência de antes, pois país prefere garantir seus investimentos em sustentabilidade para atrair turistas. E mesmo que essa ideia vise a economia, pode ser vista como uma boa forma de manter a saúde do ambiente e do cidadãos, além de servir como "inspiração" para que outros países ajudem a cuidar de suas moradias. Por conseguinte, é aceitável dizer que o carro pode se tornar o "cigarro do futuro", já que, assim como o cigarro, pode cair em desuso e ter maior substituição por transportes coletivos. É importante salientar que tal substituição contribui para a queda da poluição ambiental e melhoria da saúde urbana, facilitando também a locomoção de indivíduos. Assim, muito além de contribuir para que não haja crescimento de doenças devido aos gases poluentes, melhora a condição do meio ambiente, que é o maior bem da sociedade.

Redação 43

A vida parou ou foi o automóvel? A tela entre "São Paulo [Gozo]", da pintura modernista Tarcila do Amaral, evidencia a dinâmica do automóvel no cenário de crescimento paulistano. De modo paralelo com a temática da obra, o carro desempenha papel de destaque no cenário da mobilidade cotidiana. Sob tal pano de fundo, a expansão dos grandes e médios centros urbanos e os impactos ambientais atuam como gargalos da terra. Assim, a crônica sensação da vida parada se confundindo com um automóvel é premente. De fato, os processos de crescimento das cidades brasileiras constitui um desafio para os governos no que tange à mobilidade. O 2º Seminário da Folha de São Paulo "Desafios da Mobilidade Urbana", expõe a eficácia de programas de locomoção implantadas em algumas capitais do país com o uso de transporte alternativos, como bicicletas, patinetes e veículos leves sobre trilhas (VLT'S). Nesse viés, Tem-se o constante enfrentamento em torno da imobilidade, uma vez que grande parte das cidades não possuem planos de planejamento urbano. Com efeito, investimentos nos processos de locomoção tornam-se escassos ou restritos às principais cidades, o que coíbe, de modo parcial, a problemática. Portanto, deve-se ponderar que a utilização de novas e eficientes modais é necessário. Salienta-se, ainda, os impactos ambientais do uso recorrente dos automóveis no atual panorama. Nesse norte, o engajamento por problemas de poluição e desconforto pessoal no espaço urbano vêm ganhando destaque, de modo análogo ao uso do cigarro em locais públicos no século XX. A partir de tal prisma, o carro ocupa as vias brasileiras de modo intenso, pois é visto como símbolo de consumo e ascensão social. Por consequência, há o confronto entre o uso do automóvel para a locomoção e o uso de transportes alternativos para a redução de problemas ambientais. À luz disso, a mudança na utilização de transportes poluidores e nocivos à saúde humana é urgente. Sendo assim, torna-se visível que o automóvel constitui a mola-motor contemporânea -tal como no quadro de Tarcila. Nesse sentido, a necessidade de recursos para o uso de alternativas meios de locomoção, aliado com a preocupação ambiental vêm demandando o acendimento dos holofotes a causa, visto o constante crescimento da população e do poder de consumo. Sob esse enfoque, quiçá, o carro possa ser o novo cigarro da modernidade, afinal o uso indiscriminado e o estímulo constitui

base do crescimento. Logo, o dualismo entre a vida e o automóvel, tal como no poema de Drummond, caracteriza-se como rápido, senão urgente.

Redação 44

A evolução e regressão a favor da Fumaça a Revolução Industrial foi uma série de mudanças que ocorreram na Europa no século XVIII, tais eventos trouxeram consigo as máquinas a vapor e a agilidade por trás delas. Não demorou muito para o ser humano se acostumar com a novidade e fazer parte dela. Agora uma pessoa pode correr altas velocidades conduzindo uma máquina, não a vapor, mas se a álcool ou gasolina. Todavia, os preços a se pagar por tal corrida vai muito além do monetário, também envolve questões de saúde pessoal e coletiva. Primeiramente, é notório que a quantidade de carros aumentou com o passar do tempo na grande São Paulo, fazendo com que o município adote o sistema de rodízio de placas em horários de pico, porém o problema vai muito além do congestionamento causado por esses veículos, ele também inclui a fumaça e poluição provenientes desses. Algo que pode prejudicar não só o meio ambiente e intensificar o efeito estufa e o aquecimento global, mas também a saúde da população local. Além disso, é preciso lembrar que anos atrás foi aprovada a Lei em que proíbe o uso do cigarro em lugares públicos e fechados, pois a fumaça exalada por esses é prejudicial ao demais que estão presentes no devido local. Entretanto não deve se esquecer que os automóveis também liberam altas quantidades de substâncias tóxicas para a comunidade, o que faz necessário a criação de leis para controlar ou diminuir o descontrole de fumaça liberada pelos carros e motos. O filósofo e sociólogo Bauman defendeu em seu livro "modernidade líquida" que a população atual vive em constante mudança, e que muitas vezes ao evoluir ela acaba cometendo uma regressão, e isso pode ser explicado juntamente com fato da sociedade incomodar-se com a fumaça do cigarro e ter uma postura neutra em relação a poluição ocasionada pelos veículos. Em suma, é preciso notar que o uso dos automóveis vieram com a intenção de agilizar a vida do ser humano, mas seu uso excessivo faz mal para os mesmos. Para resolver tal problema, caberia as prefeituras de grandes cidades criarem projetos com o objetivo de diminuir o fluxo de carros, como por exemplo maior frota de transporte público, rodízios de placas para diminuição de trânsito, até mesmo disponibilizar bicicletas gratuitas para o uso da população, assim também estimulando um hábito de vida mais saudável para o cidadão, para que desse modo, os veículos movidos a combustível não se torne um novo cigarro futuramente.

Redação 45

O carro será o novo cigarro? Faz-se importante destacar inicialmente, que o carro, assim como outros produtos mais nocivos do que benéficos, é um artigo tão desejado graças à cultura do consumo exacerbado vendida pelo capitalismo. Ao longo de séculos a raça humana criou e aperfeiçoou inúmeras ferramentas para de fato melhorar sua vida, o automóvel foi fruto do anseio por agilidade no ato de ser locomover. Mas, o fato é que o sistema capitalista de produção em larga escala, criou e cria necessidades que não são tão necessárias à humanidade, com o único intuito de uma acumulação financeira sem justificativa. No Brasil, diferente do que acontece em outros países, é comum os jovens-ao atingirem a maioria - terem como

principal objetivo obter a Carteira Nacional de Habilitação (CNH) para dirigir e assim financiar um carro ao invés de uma casa própria. A justificativa por essa escolha é a praticidade na locomoção que um carro pode proporcionar, porém, esses jovens - sobretudo os das classes mais populares - entram em financiamentos morosos e com altas taxas de juros... Mas, é exatamente isso que o sistema financeiro quer, quer se sustentar por via da especulação e arrecadação de juros. O mais interessante é ver que o salário mínimo nacional não cresce na mesma velocidade que essas necessidades impostas pelo capital e nem na velocidade dessas altas taxas de juros. Além disso, essa ideia de praticidade, ter muitas vezes, acaba findando no sedentarismo, já que muitas pessoas acabam dependendo de um automóvel para absolutamente tudo, deixando de caminhar - e assim melhorar sua saúde - para realizar até as tarefas mais básicas. Já é comum escutarmos que o automóvel emite no ar e inúmeros gases nocivos para o ambiente e para todos nós, porém, se tornou tão corriqueiro saber disso que a poluição do ar resulta - para muitos - num problema da natureza e não um problema também nosso. Então, para além de repetirmos sempre o mesmo jargão, é de suma importância que nós entendamos que esse consumo desenfreado das indústrias automobilísticas, tem desde seu início (idealização e produção) até o seu fim (consumidor final) um conjunto de inúmeros problemas e que no final os prejudicados sempre seremos todos nós, dessa forma, é possível afirmar, que colocando na balança os benefícios e malefícios de ter um carro, o mesmo não será, mas já é o novo cigarro.

Redação 46

Logo nos primeiros episódios da série norte-americana "Mad Men", ambientada na metade do século XX, é explorada a questão da cultura do cigarro, que vendia a imagem do produto como algo essencial para o ser humano moderno. Na série, os personagens trabalhavam em informes publicitários de cigarro e também faziam intenso uso do mesmo. O carro tornou-se popular em meados dos anos 1970 e rapidamente adquiriu a mesma cultura de modernidade do cigarro. Infere-se, assim, que o carro já é o novo cigarro. A cultura do carro foi vendida com os princípios do futurismo, corrente filosófica e artística que exaltava o movimento, o novo, as tecnologias. No Brasil isso se fez muito presente na chegada do fusca, que também moldou a imagem do brasileiro como ser "moderno". Assim, torna-se difícil desapegar-se dessa cultura, visto que muitos preferem a liberdade, rapidez e conforto do carro do que a superlotação e lentidão do transporte público. Em contrapartida, sabe-se que nem tudo sobre o carro é "mil maravilhas". Se o cigarro traz já conhecidos malefícios para a saúde humana, o carro traz consigo a poluição ao meio ambiente com o intenso uso de gasolina e a emissão de gases poluentes. Além disso, ele tem causado problemas na mobilidade urbana, como engarrafamentos, que incomodam tanto quem está dentro como quem está fora do veículo. Assim como o uso do cigarro caiu devido a conscientização sobre seus malefícios, o uso do carro também pode cair se o mesmo conscientizar-se sobre os problemas crescentes dele. Isso, evidentemente, se o mercado global achar economicamente viável esta mudança, visto que é ele quem orienta o que é consumido pelo mundo. Há também a necessidade de melhorias no setor de transportes públicos para que ele atenda perfeitamente a população. Mediante o exposto, o carro teria de entrar em desvantagem socioeconômica para não ser tão utilizado, e a mídia, como em "Mad Men", faria seu papel de influenciadora vendendo a imagem do carro como ultrapassado e inútil. Mas até essa mudança acontecer, o carro continuará sendo

cigarro: por um lado, símbolo da modernidade e liberdade; por outro, problemático e prejudicial a quem usa, a quem está por perto e ao meio ambiente.

Redação 47

O filme Wall-E retrata uma ideia de futuro para a humanidade e tem como proposta a tecnologia versus a natureza que em análise, choca ao demonstrar seres humanos impossibilitados de caminhar e por isso detentores de carros individuais. Observando a sociedade em sua atualidade, é grande a tendência para tal caminho devido aos baixos investimentos em transportes coletivos e o fetichismo sobre automóveis. No entanto, a constante mudança de atitudes e ideais humanos já demonstram caminhos alternativos e repletos de conscientização. Seguindo essa ideia, os carros serão os cigarros do futuro e toda a humanidade precisa estar ciente e discutir sobre o assunto a fim de ressalvas e melhorias. Momentaneamente, a sociedade vem circulada pelos resquícios da Guerra Fria e o imperialismo dos Estados Unidos na economia, o que acaba gerando adesão a princípios favoráveis a esse país, como a ideia de fetichismo ou seja, supervalorização de produtos como automóveis, privilegiando empresas estadunidenses e a economia desse país. Ademais, há o baixo investimento para transportes públicos ou coletivos, gerando a desconfiança do usuário ao serviço, tendo em vista a falta de compromisso com horários, a alta lotação e o mau planejamento de linhas, causando transtornos e o caos em massa. Dessa forma, em busca de conforto e segurança, aliado a pensamentos consumistas, a sociedade super utiliza automóveis individuais. Ilustra essa ideia o fato de algumas cidades brasileiras possuírem mais carros que pessoas e ainda, a exemplo do transporte público no Canadá cuja população opta por esses meios devido aos investimentos e a boa funcionalidade do serviço. Em contrapartida, o futuro vem em moldes alternativos e não como o purista em Wall -e, visto que a crescente adesão da educação ambiental iniciada fortemente pelos anos 80 e a ampliação dos problemas ambientais, demonstra resultados e maior apoio da massa, que hoje é mais preocupada e engajada com a causa, exibido em consumidores conscientes e exigentes e demonstrado pela abominação ao uso de couro animal e produtos que utilizam animais para testes. Nesse sentido, o futuro contará com o engajamento da população somado à diminuição da ideia de status com o consumo e a melhoria dos transportes públicos, resultando no baixo uso de carros e veículos individuais. Exemplifica essa ideia e a adesão populacional aos ideais ambientalistas movimentos como os feitos pela ativista Greta Thunberg de 16 anos; incentivando a preservação e o olhar político para a causa e demonstrando como as futuras gerações agirão. Em síntese, atualmente a população é moldada por ideais consumistas e tenta aos poucos se desvincular desses, demonstrado pelas novas operações que buscam um modo de vida sustentável e no futuro, abdicarão de maquinários e produtos supérfluos como os transportes individuais.

Redação 48

Durante o período das gerações X e Y o cigarro era visto como símbolo de status na sociedade. Entretanto, com o passar das décadas esse objeto maléfico a saúde humana, passou a ter sua função substituída pelos carros. Com a evolução tecnológica, é visível o crescimento de automóveis em circulação, principalmente em países subdesenvolvidos, como o Brasil. Tal

fato faz necessário mitigar a causa do aumento desse transporte nas cidades e suas possíveis consequências no futuro. Em primeiro plano, convém ressaltar que nos dias hodiernos os veículos de quatro rodas deixaram de ser meros meio de locomoção. Atualmente, este transporte é visto como objeto de desejo social. No Brasil, um dos principais objetivos dos jovens é chegar aos 18 anos para tirar sua CNH (Carteira Nacional de Habilitação), e com isso comprar um veículo, sendo na maioria das vezes o objeto de elevar o status social. Em países desenvolvidos, como os Estados Unidos, as pretensões dos cidadãos são diferentes, a cada 100 jovens americanos, apenas 60 têm habilitação antes dos 18 anos. Outrossim, é visível a consequência ambiental dos carros no ecossistema. Organizações como o Greenpeace lutam pela sustentabilidade, incluindo a diminuição da emissão de gases poluentes daqui uns anos não será mais o cigarro responsável pelas doenças pulmonares, mas os carros passarão a ser protagonistas, emitindo um elevado índice de CO₂ na atmosfera. Com isso, haverá uma redução significativa na qualidade de vida humana gerando: problemas de saúde, aquecimento global, engarrafamento urbano e inúmeras outras consequências. Desse modo, é evidente que o carro tem tudo para se tornar o que o cigarro foi no século XX. Sob essa ótica, temos como causa a busca incessante da população pelo status social e por consequência a poluição e superlotação de veículos nas ruas. Neste contexto, é necessário a mudança de postura das pessoas, em busca de um mundo mais sustentável e habitável, em que andar seja mais importante que dirigir.

Redação 49

Carros não são mais necessários! Na teoria evolucionista do Larkinismo os animais se adaptam ao meio passando as futuras gerações os seus caracteres adquiridos, da mesma forma pode-se pensar na capacidade do ser humano de consumir determinados produtos, através dos anos os homens vão se adaptando a um novo estilo de vida mudando assim o estilo de vida das gerações Futuras. Nessa lógica pode-se pensar nos automóveis que, no início do século XX, era artigo de luxo somente da classe alta conceito que se encontra deturpado nos dias atuais perdendo sua grande importância em detrimento aos transportes de massa, pois o ser humano já começou a se adapta a um modo de vida mais sustentável. Os grandes centros urbanos, muito conhecidos por congestionamentos e acidentes de trânsito, estão investindo cada vez mais em transportes de massa, característicos por transportar um maior número de pessoas de uma só vez, temem vista que, principalmente em países periféricos, o movimento pendular diário somado a grande quantidade de equipamentos urbanos acarretam em congestionamentos, gerando estresse, diminuindo a qualidade de vida dessas pessoas. Faz-se salientar a importância da Sustentabilidade na mudança no modo de encherarmos os carros no mundo atual, tendo em vista que o termo, cunhado em 1972, busca a melhoria da qualidade de vida, possibilitando a utilização de recursos disponíveis de forma consciente e ajudando na preservação ambiental de modo que influência as pessoas a reduzirem suas emissões de gases poluentes e criando formas de melhorar a locomoção sem agredir o meio ambiente, fazendo com que as pessoas criem uma consciência ambiental, aumentando a utilização de meios de transportes mais sustentáveis, deixando seus carros nas garagens. Contudo, apesar de ainda haver uma grande quantidade de veículos em circulação com donos que se recusam a utilizá-los, os carros tem seus dias contados, pois assim como no Larkinismo a evolução leva tempo para se manifestar nas gerações, mas é indubitável que a consciência ecológica está se

manifestando na sociedade de um modo que possivelmente no futuro utilizar qualquer veículo que imita gases poluentes pode ser considerado crime igual a Fumar em locais fechados.

Redação 50

“Raízes Clássicas” A elaboração da "Constituição Federal", há mais de 30 anos, foi baseada no sonho de bem-estar social de todos os indivíduos, incluindo o direito de ir e vir e de viver em um ambiente saudável. No entanto, o poder público não obteve êxito nessa caminhada enquanto agente fornecedor de direitos mínimos, devido não só por incentivar o consumo de veículos individuais, mas também por não propor melhorias para o uso do transporte coletivo. Nesse sentido, evidencia-se que o modelo atual de locomoção mostra-se insustentável e prejudicial a toda sociedade. A priori, é indiscutível que a grande oferta de crédito incentiva o consumo por veículos próprios, fator esse visto no trânsito e a poluição nas grandes e médias cidades atualmente. Consoante ao modelo proposto por Ford, o "fordismo" as indústrias detinham os meios de produção, enquanto os trabalhadores assalariados eram os grandes consumidores e mantenedores do funcionamento da economia. Dessa maneira, o uso do transporte industrializado tem desde os primórdios da "Revolução Industrial" e início do capitalismo, lugar cativo em frente a lei da oferta e da demanda. Igualmente, pode-se destacar as mazelas que o transporte coletivo vem sofrendo ao longo do tempo, visto que, em cidades grandes seu uso vem sendo relegado ao carro de uso comercial devido a precariedade dos serviços prestados pelos locomotores coletivos. A exemplo de comparação, vale citar o modelo da cidade de Portland nos Estados Unidos da América, em que os transportes de uso coletivo foi integrado aos diversos modais existentes, fator esse que contribui na redução da Poluição ambiental e também no incentivamento do deslocamento das pessoas na cidade. Dessa forma, esse modelo norte-americano mostrou-se eficiente, diferente dos demais implantado desde a "Revolução Industrial" e não menos diferente do modo brasileiro e implementado por Juscelino Kubitschek. Destarte, é evidente que o modelo capitalista de produção em massa tem influência nos trânsitos cada vez mais superlotados e estressante atualmente nas grandes cidades. Sendo assim, faz-se mister a necessidade de políticas públicas de mobilidade urbana por órgãos competentes como o ministério das cidades e público. Afinal como sugere no Maslow em sua teoria "A hierarquia das necessidade humanas" o indivíduo é movido a objetivos e ao mesmo tempo necessitam de condições mínimas de vidas era se manterem em harmonia.

APÊNDICE B – Dados quantitativos das redações

Redação	Tokens	Types	Type/Token Ratio
1.txt	216	140	64.81
2.txt	248	139	56.05
3.txt	181	121	66.85
4.txt	258	148	57.36
5.txt	282	164	58.16
6.txt	381	203	53.28
7.txt	236	142	60.17
8.txt	304	184	60.53
9.txt	294	172	58.50
10.txt	346	207	59.83
11.txt	327	168	51.38
12.txt	176	109	61.93
13.txt	249	161	64.66
14.txt	353	217	61.47
15.txt	361	191	52.91
16.txt	221	148	66.97
17.txt	275	169	61.45
18.txt	295	150	50.85
19.txt	323	202	62.54
20.txt	224	140	62.50
21.txt	289	169	58.48
22.txt	195	123	63.08
23.txt	448	245	54.69
24.txt	268	146	54.48
25.txt	241	163	67.63
26.txt	296	172	58.11
27.txt	326	193	59.20
28.txt	246	171	69.51
29.txt	255	153	60.00
30.txt	329	182	55.32
31.txt	342	196	57.31
32.txt	331	193	58.31
33.txt	292	194	66.44
34.txt	210	132	62.86
35.txt	165	106	64.24
36.txt	279	182	65.23
37.txt	417	224	53.72
38.txt	288	172	59.72
39.txt	416	233	56.01
40.txt	228	142	62.28
41.txt	391	216	55.24
42.txt	367	199	54.22
43.txt	221	136	61.54

44.txt	232	134	57.76
45.txt	176	109	61.93
46.txt	278	167	60.07
47.txt	444	247	55.63
48.txt	246	158	64.23
49.txt	292	178	60.96
50.txt	172	113	65.70
51.txt	314	198	63.06
52.txt	319	201	63.01
53.txt	351	204	58.12
54.txt	205	130	63.41
55.txt	331	183	55.29
56.txt	246	129	52.44
57.txt	257	174	67.70
58.txt	381	223	58.53
59.txt	286	176	61.54
60.txt	238	150	63.03
61.txt	243	147	60.49
62.txt	194	126	64.95
63.txt	496	279	56.25
64.txt	303	185	61.06
65.txt	194	110	56.70
66.txt	313	176	56.23
67.txt	299	188	62.88
68.txt	188	102	54.26
69.txt	360	209	58.06
70.txt	197	126	63.96
71.txt	253	162	64.03
72.txt	386	203	52.59
73.txt	265	143	53.96
74.txt	219	143	65.30
75.txt	255	168	65.88
76.txt	345	212	61.45
77.txt	262	166	63.36
78.txt	299	174	58.19
79.txt	206	138	66.99
80.txt	294	169	57.48
81.txt	299	183	61.20
82.txt	250	163	65.20
83.txt	212	143	67.45
84.txt	298	172	57.72
85.txt	213	135	63.38
86.txt	237	136	57.38
87.txt	364	197	54.12
88.txt	282	159	56.38
89.txt	247	159	64.37

90.txt	295	172	58.31
91.txt	237	147	62.03
92.txt	595	292	49.08
93.txt	129	87	67.44
94.txt	295	161	54.58
95.txt	455	279	61.32
96.txt	207	139	67.15
97.txt	284	179	63.03
98.txt	176	109	61.93
99.txt	236	146	61.86
100.txt	123	96	78.05
101.txt	152	107	70.39
102.txt	197	129	65.48
103.txt	300	186	62.00
104.txt	284	163	57.39
105.txt	362	211	58.29
106.txt	283	158	55.83
107.txt	423	197	46.57
108.txt	299	157	52.51
109.txt	229	144	62.88
110.txt	153	107	69.93
111.txt	182	114	62.64
112.txt	248	153	61.69
113.txt	259	149	57.53
114.txt	265	171	64.53
115.txt	299	174	58.19
116.txt	246	147	59.76
117.txt	334	206	61.68
118.txt	261	162	62.07
119.txt	346	200	57.80
120.txt	276	180	65.22
121.txt	212	149	70.28
122.txt	224	146	65.18
123.txt	345	206	59.71
124.txt	333	181	54.35
125.txt	251	163	64.94
126.txt	339	193	56.93
127.txt	247	145	58.70
128.txt	267	151	56.55
129.txt	376	193	51.33
130.txt	229	152	66.38
131.txt	244	156	63.93
132.txt	145	95	65.52
133.txt	191	128	67.02
134.txt	290	177	61.03
135.txt	236	142	60.17

136.txt	409	239	58.44
137.txt	401	225	56.11
138.txt	300	181	60.33
139.txt	394	237	60.15
140.txt	298	171	57.38
141.txt	243	137	56.38
142.txt	308	191	62.01
143.txt	204	126	61.76
144.txt	298	172	57.72
145.txt	278	168	60.43
146.txt	158	97	61.39
147.txt	267	155	58.05
148.txt	357	239	66.95
149.txt	195	132	67.69
150.txt	266	163	61.28
151.txt	272	168	61.76
152.txt	199	140	70.35
153.txt	278	150	53.96
154.txt	239	150	62.76
155.txt	230	138	60.00
156.txt	180	111	61.67
157.txt	256	163	63.67
158.txt	314	184	58.60
159.txt	352	207	58.81
160.txt	412	261	63.35
161.txt	294	186	63.27
162.txt	303	168	55.45
163.txt	167	112	67.07
164.txt	247	144	58.30
165.txt	373	215	57.64
166.txt	238	133	55.88
167.txt	247	147	59.51
168.txt	261	160	61.30
169.txt	348	200	57.47
170.txt	281	165	58.72
171.txt	377	184	48.81
172.txt	229	150	65.50
173.txt	191	121	63.35
174.txt	255	164	64.31
175.txt	221	122	55.20
176.txt	349	201	57.59
177.txt	208	127	61.06
178.txt	334	191	57.19
179.txt	207	130	62.80
180.txt	342	197	57.60
181.txt	253	162	64.03

182.txt	310	195	62.90
183.txt	249	157	63.05
184.txt	291	164	56.36
185.txt	335	177	52.84
186.txt	304	192	63.16
187.txt	330	205	62.12
188.txt	332	184	55.42
189.txt	190	136	71.58
190.txt	430	254	59.07
191.txt	176	111	63.07
192.txt	115	88	76.52
193.txt	240	143	59.58
194.txt	274	164	59.85
195.txt	199	113	56.78
196.txt	199	113	56.78
197.txt	179	113	63.13
198.txt	249	143	57.43
199.txt	214	132	61.68
200.txt	294	152	51.70
201.txt	228	146	64.04
202.txt	221	129	58.37
203.txt	350	186	53.14
204.txt	244	157	64.34
205.txt	327	162	49.54
206.txt	127	87	68.50
207.txt	339	228	67.26
208.txt	238	129	54.20
209.txt	382	212	55.50
210.txt	206	128	62.14
211.txt	277	176	63.54
212.txt	358	187	52.23
213.txt	252	157	62.30
214.txt	201	129	64.18
215.txt	286	181	63.29
216.txt	321	184	57.32
217.txt	340	188	55.29
218.txt	366	215	58.74
219.txt	310	195	62.90
220.txt	324	195	60.19
221.txt	251	177	70.52
222.txt	289	176	60.90
223.txt	210	141	67.14
224.txt	402	228	56.72
225.txt	323	166	51.39
226.txt	246	132	53.66
227.txt	367	202	55.04

228.txt	347	178	51.30
229.txt	251	144	57.37
230.txt	274	155	56.57
231.txt	163	110	67.48
232.txt	378	159	42.06
233.txt	204	126	61.76
234.txt	249	158	63.45
235.txt	231	158	68.40
236.txt	309	176	56.96
237.txt	382	222	58.12
238.txt	264	164	62.12
239.txt	290	167	57.59
240.txt	211	129	61.14
241.txt	288	180	62.50
242.txt	231	136	58.87
243.txt	129	82	63.57
244.txt	334	200	59.88
245.txt	270	165	61.11
246.txt	364	204	56.04
247.txt	265	158	59.62
248.txt	290	161	55.52
249.txt	219	137	62.56
250.txt	316	191	60.44
251.txt	359	209	58.22
252.txt	165	117	70.91
253.txt	409	226	55.26
254.txt	225	132	58.67
255.txt	158	104	65.82
256.txt	199	126	63.32
257.txt	164	120	73.17
258.txt	197	133	67.51
259.txt	210	150	71.43
260.txt	192	125	65.10
261.txt	268	157	58.58
262.txt	220	149	67.73
263.txt	298	173	58.05
264.txt	208	138	66.35
265.txt	259	169	65.25
266.txt	225	133	59.11
267.txt	245	149	60.82
268.txt	420	224	53.33
269.txt	289	188	65.05
270.txt	234	150	64.10
271.txt	135	92	68.15
272.txt	262	152	58.02
273.txt	270	161	59.63

274.txt	277	166	59.93
275.txt	291	180	61.86
276.txt	184	120	65.22
277.txt	352	201	57.10
278.txt	200	123	61.50
279.txt	150	97	64.67
280.txt	140	100	71.43
281.txt	194	142	73.20
282.txt	255	172	67.45
283.txt	180	122	67.78
284.txt	114	87	76.32
285.txt	153	86	56.21
286.txt	205	123	60.00
287.txt	164	111	67.68
288.txt	233	136	58.37
289.txt	107	74	69.16
290.txt	180	125	69.44
291.txt	324	207	63.89
292.txt	200	126	63.00
293.txt	329	185	56.23
294.txt	159	110	69.18
295.txt	135	81	60.00
296.txt	169	110	65.09
297.txt	354	220	62.15
298.txt	160	104	65.00
299.txt	260	171	65.77
300.txt	226	139	61.50
301.txt	222	134	60.36
302.txt	158	115	72.78
303.txt	176	117	66.48
304.txt	274	176	64.23
305.txt	176	113	64.20
306.txt	146	100	68.49
307.txt	256	155	60.55
308.txt	216	137	63.43
309.txt	155	110	70.97
310.txt	320	176	55.00
311.txt	198	124	62.63
312.txt	210	128	60.95
313.txt	388	225	57.99
314.txt	204	134	65.69
315.txt	229	152	66.38
316.txt	262	157	59.92
317.txt	229	132	57.64
318.txt	252	157	62.30
319.txt	274	157	57.30

320.txt	304	184	60.53
321.txt	153	101	66.01
322.txt	173	122	70.52
323.txt	169	114	67.46
324.txt	288	151	52.43
325.txt	197	132	67.01
326.txt	287	156	54.36
327.txt	247	145	58.70
328.txt	357	208	58.26
329.txt	280	169	60.36
330.txt	202	141	69.80
331.txt	207	120	57.97
332.txt	328	190	57.93
333.txt	246	151	61.38
334.txt	206	118	57.28
335.txt	201	124	61.69
336.txt	199	131	65.83
337.txt	436	251	57.57
338.txt	168	106	63.10
339.txt	162	103	63.58
340.txt	226	141	62.39
341.txt	270	174	64.44
342.txt	168	116	69.05
343.txt	338	200	59.17
344.txt	321	183	57.01
345.txt	275	172	62.55
346.txt	225	135	60.00
347.txt	193	133	68.91
348.txt	324	184	56.79
349.txt	336	206	61.31
350.txt	303	166	54.79
351.txt	299	183	61.20
352.txt	161	102	63.35
353.txt	302	179	59.27
354.txt	329	206	62.61
355.txt	389	239	61.44
356.txt	174	101	58.05
357.txt	278	165	59.35
358.txt	382	199	52.09
359.txt	266	156	58.65
360.txt	344	211	61.34
361.txt	343	196	57.14
362.txt	257	149	57.98
363.txt	242	143	59.09
364.txt	228	141	61.84
365.txt	284	174	61.27

366.txt	358	216	60.34
367.txt	201	130	64.68
368.txt	171	112	65.50
369.txt	191	111	58.12
370.txt	178	123	69.10
371.txt	215	107	49.77
372.txt	166	123	74.10
373.txt	255	150	58.82
374.txt	251	146	58.17
375.txt	243	157	64.61
376.txt	349	190	54.44
377.txt	272	168	61.76
378.txt	279	162	58.06
379.txt	277	165	59.57
380.txt	281	168	59.79
381.txt	228	144	63.16
382.txt	293	182	62.12
383.txt	403	229	56.82
384.txt	214	138	64.49
385.txt	252	147	58.33
386.txt	229	142	62.01
387.txt	230	166	72.17
388.txt	224	146	65.18
389.txt	268	178	66.42
390.txt	224	158	70.54
391.txt	244	151	61.89
392.txt	272	157	57.72
393.txt	239	147	61.51
394.txt	261	163	62.45
395.txt	319	187	58.62
396.txt	273	185	67.77
397.txt	339	187	55.16
398.txt	377	221	58.62
399.txt	334	171	51.20
400.txt	344	160	46.51
401.txt	258	154	59.69
402.txt	393	224	57.00
403.txt	259	158	61.00
404.txt	256	166	64.84
405.txt	325	181	55.69
406.txt	336	177	52.68
407.txt	272	149	54.78
408.txt	330	188	56.97
409.txt	353	216	61.19
410.txt	383	239	62.40
411.txt	314	172	54.78

412.txt	217	149	68.66
413.txt	267	193	72.28
414.txt	247	131	53.04
415.txt	317	178	56.15
416.txt	276	151	54.71
417.txt	424	243	57.31
418.txt	245	145	59.18
419.txt	358	192	53.63
420.txt	374	203	54.28
421.txt	350	211	60.29
422.txt	189	120	63.49
423.txt	369	221	59.89
424.txt	192	141	73.44
425.txt	225	135	60.00
426.txt	234	144	61.54
427.txt	348	190	54.60
428.txt	283	160	56.54
429.txt	264	146	55.30
430.txt	358	199	55.59
431.txt	334	197	58.98
432.txt	199	135	67.84
433.txt	281	152	54.09
434.txt	219	135	61.64
435.txt	297	174	58.59
436.txt	275	156	56.73
437.txt	174	117	67.24
438.txt	352	193	54.83
439.txt	346	224	64.74
440.txt	237	140	59.07
441.txt	266	180	67.67
442.txt	288	178	61.81
443.txt	259	167	64.48
444.txt	207	133	64.25
445.txt	260	153	58.85
446.txt	224	150	66.96
447.txt	336	191	56.85
448.txt	394	215	54.57
449.txt	294	169	57.48
450.txt	379	230	60.69
451.txt	268	165	61.57
452.txt	426	269	63.15
453.txt	225	136	60.44
454.txt	316	188	59.49
455.txt	321	177	55.14
456.txt	262	166	63.36
457.txt	340	194	57.06

458.txt	234	146	62.39
459.txt	190	123	64.74
460.txt	349	187	53.58
461.txt	217	140	64.52
462.txt	184	111	60.33
463.txt	295	180	61.02
464.txt	220	129	58.64
465.txt	228	147	64.47
466.txt	236	136	57.63
467.txt	212	130	61.32
468.txt	357	205	57.42
469.txt	326	176	53.99
470.txt	230	128	55.65
471.txt	280	166	59.29
472.txt	368	217	58.97
473.txt	189	115	60.85
474.txt	311	191	61.41
475.txt	258	152	58.91
476.txt	270	177	65.56
477.txt	197	115	58.38
478.txt	259	146	56.37
479.txt	239	146	61.09
480.txt	284	189	66.55
481.txt	268	161	60.07
482.txt	271	174	64.21
483.txt	279	171	61.29
484.txt	379	194	51.19
485.txt	176	112	63.64
486.txt	349	195	55.87
487.txt	248	141	56.85
488.txt	412	222	53.88
489.txt	303	179	59.08
490.txt	206	128	62.14
491.txt	213	140	65.73
492.txt	293	179	61.09
493.txt	225	139	61.78
494.txt	347	194	55.91
495.txt	186	126	67.74
496.txt	376	193	51.33
497.txt	262	151	57.63
498.txt	301	178	59.14
499.txt	366	224	61.20
500.txt	271	170	62.73

APÊNDICE C – Palavras não presentes na lista ampliada

abaixando	acelera	ademais	aérea
abandonamento	aceleração	adentrando	aferindo
abarrota	acelerada	adentrou	afetada
abastada	acelerado	adepta	afetados
abastece	aceleram	adeptos	afim
abastecido	acelerando	adequação	afins
abc	acelerar	adequam	afirmadores
abdicação	aceleraram	adere	afortunado
abdica	acendimento	aderem	áfrica
abdica	acentuada	aderência	afugentando
abdicação	acertivamente	aderida	afundados
aberturas	aces	aderido	afundando
abilitação	acessibilidade	aderidos	agilidade
abitos	acessibilidades	aderindo	agilizar
abolido	acessíveis	aderir	agilizou
abolir	acessível	aderiram	agitada
abominação	acessivo	aderirem	aglomeração
abomine	acessório	aderissem	aglomerações
abotam	ácida	aderiu	aglomerado
abrangência	ácidas	adição	aglomerados
abrangendo	acidamente	adição	aglomeram
abrangente	acinzentada	adquirir	aglomerando
abrangentes	aclamado	adquirir	agradável
abrasileiro	acobertando	administração	agravá
abrupta	acom	admirada	agravadas
absorção	acometidas	admiradores	agravadores
absurdamente	acomodada	ado	agravados
absurdas	acomodidade	adolescen	agravamento
abto	acompanha	adoradas	agravamentos
abundante	acompanhada	adotadas	agravante
abusiva	acompanhado	adotados	agravantes
abusivo	acompanhamos	adptanto	agregado
academica	acompanhando	adptos	agregam
acalentariam	aconchegante	adquirimento	agregando
acaou	acontece	adquirí	agregar
acareta	acontecendo	adquiridos	agregaram
acarreta	acorrentados	adquirio	agregou
acarretado	acreditivo	adrelina	agressivamente
acarretados	acrescendo	aduanheiro	agressor
acarretam	açucar	adultera	agrotóxicos
acarretando	aculmulo	advento	aguas
acarretar	acumulação	adventos	aguele
acarretaria	acúmulo	adverte	ai
acarretou	acusá	advertências	aindo
aceitação	adaptada	advinda	ais
aceitável	adaptados	advindo	alagamento
aceitável	adema	advindos	alarmante

alarmantes	altomotivo	análoga	ápice
alastra	altomoveis	analogamente	apli
alavanca	altomóveis	análogamos	aplicação
alavancando	altomovel	análogo	aplicada
alavancar	altônómia	anceio	aplicadas
alavancou	alugada	and	aplicativas
albert	alugados	andorinha	aplicativo
alcançada	alusão	andrade	aplicativos
alcool	alves	andré	aplle
alcoólatras	alvos	angari	apodrecer
alcoólicos	am	animador	apogeu
álcoolizados	amaldiçoada	anomia	apontados
aleado	amanha	anormais	ápos
alegoria	amarral	anseios	apostado
alem	amazônia	antagônico	app's
álem	ambas	antagonistas	apreciadores
alemanha	ambi	antecedencia	apreço
alérgicos	ambien	antecessor	apresentada
alertam	ambientada	antecipada	apresentadas
alertando	ambientais	antepassados	apresentados
alertarem	ambiental	anteriormentes	aprimorada
aletar	ambientalistas	antiético	aprimoradas
alfa	ambientalmente	antigos	aprimorado
alguem	âmbiente	antiguidade	aprimorados
algumas	âmbito	antipatia	aprimoramento
aliada	âmbitos	antiquado	aprimorando
aliadas	amenizados	antisustentável	aprimorar
aliados	amenizando	antracose	apriori
alicerce	amenizar	antrópicas	apro
alienação	amenizarem	anuais	aprofunda
alienado	amenize	anualmente	aprofundada
alienando	ameno	anunciada	aprofundamos
alimentícios	américa	anuncios	aprovada
allo	american	aonde	aprovadas
almejado	amigos	aparecimento	aproveitados
almejam	amontoado	aparencia	aquaviário
almejando	amontoados	apartir	aque
almejar	ampliados	apática	aquece
almejós	ampliamiento	ape	aqueci
almentando	âmplie	apé	àquele
almentar	amsterdã	apecto	àqueles
almententar	amsterdan	apelidou	aquiescido
almento	analfabétismo	aperfeiçoada	aquilo
alterada	analisá	aperfeiçoado	aquisitio
alterados	análisa	aperfeiçoamento	aquisitivo
alterna	analísada	aperfeiçoando	arcadismo
alternância	analísados	aperfeiçoou	arcáicas
alternatica	análisar	aperfeiçuação	arcando
alternatives	analizado	apetrecho	arcar
altíssimos	análoga	apezar	arcariam

arcensão	assesso	atrativo	automobilimo
árdua	assíduas	atrativos	automobilistica
árduo	assíduo	atraves	automobilística
area	assinam	atraves	automobilística
areas	assinaram	atravez	automobilísticas
aréas	assitiam	atravéz	automobilísticas
arendt	associada	atrazados	automobilístico
argumenta	associadas	atrelada	automobilístico
argumentação	associado	atreladas	automobilísticos
argumentam	assola	atrelado	automobilísticos
aristocráticos	assombroso	atrelados	automobilísticos
aristóteles	assosiavam	atrelando	automobilística
aristotélica	assume	atribui	autoações
armazenamento	assuem	atribuição	automotiva
arquitecto	assumindo	atribuindo	automotivas
arraigar	atacado	atribuir	automotíveis
arraigação	atacados	atropelados	automotivo
arrasadoras	ate	atualidade	automotivos
arrecadação	áte	atualidades	automotor
arrefecer	atemporal	atualização	automotores
arriscada	atendecia	atualizações	automoveis
artefato	atentando	atualizada	automovéis
arteriais	atentar	atualmen	automôveis
artificial	atenua	aturá	autómveis
artificios	atenuação	atutomóvel	autómveis
artificius	atenuado	au	automovel
artística	atenuados	audiovisual	automovél
artísticas	atenuar	auditivos	autómvel
artístico	atenuara	auge	autómvel
arvores	aterrorizador	augusto	automovés
ascensão	atiçam	auliscando	automutilatório
ascender	atmosfera	aumennto	autonativo
ascendeu	atmosféricos	aumentado	autônoma
ascenscio	atinner	ausencias	autónomos
ascenssao	atmosferica	autalidade	autorizados
asfixia	atmosférica	autas	autorresponsabilidad
ásia	atmosféricas	auternativas	e
asiáticas	atmosferico	auternativo	autos
asileiros	atmosférico	auto	autosustentáveis
asma	atmosféricos	autoafirmação	auxiliado
asmáticas	atmosféricos	automação	auxiliando
aspécto	atoa	automáticamente	avaliada
aspiração	atona	automático	avançada
assalariados	atônimos	automaticos	avançadas
assegurando	atra	automatização	avançado
assegurar	atraídos	automatizada	avassaladora
assemelha	atrama	automaveis	avenra
assemelham	atrapalhado	autómel	aversão
assemelhando	atras	automia	aversiva
assemelhou	atrativas	automó	avessa

bacon	bicicletarios	brônquite	caótico
bafômetro	bientais	brou	caóticos
bagaço	biente	brs	capacita
bairo	bike	brt	capita
bajulá	bilidade	brumadinho	capolição
balizado	bing	brusca	car
banais	biocombustíveis	bruscamente	caracteres
banalidade	biocombustível	busao	caracteriza
banalização	biodegradáveis	buscaglia	caracterizada
banalizada	biodísel	businas	caracterizadas
banalizado	biodiversidade	buzinada	caracterizado
banalizar	bioeletricos	buzinas	caracterizam
banalizaram	biogradáveis	c	caracterizando
banidas	biologia	cá	caracterizar
barateado	biológica	ça	caracterizou
barateamento	biológico	cabere	carapuça
baratear	biomas	cabiveis	carater
baratiaram	biosfera	cabivel	caráteres
barragem	biotecnologias	cabível	carbonico
barragens	biscicletas	cabível	carbónico
barreto	bituca	cação	carbônico
baseada	blá	cadastrados	carbono
baseados	black	cado	carbôno
baseam	blico	caíram	carburador
basileira	bloqueia	caido	carburadores
báterias	boff	calçadas	cardíacos
batidas	boicotar	calcula	cardiovasculares
batidos	bojo	calotas	carece
batman	bolsonaro	camaro	carecem
bauman	bombardeados	câmbio	carência
baumaun	bombardeamento	caminhados	carlos
baumman	boom	caminhoes	carpool
belas	bourdier	caminhoneiros	carregadas
belecido	bourdieu	camões	carregados
belezinovas	br	camos	carregamentos
bém	branqueamento	campas	carregável
benção	brasil	campolamento	carrosvidentes
bené	brasileio	campos	carrp
benéfica	brasileras	canadá	carruagens
benéficas	brasil	cancer	cartelas
beneficente	brásilia	cancêr	cartilhas
beneficientes	brasões	cancerígenas	carvão
beneficios	bre	cancerígeno	cas
benefico	brevemente	cancerígeno	ças
benéfico	brigadeiro	cansativo	cassandra
beneficos	brir	ção	cassete
benéficos	britanico	cáos	catalizadora
benjamin	brônquiolos	cãos	catástrofe
bens	brônquios	caótica	catastrofecos
bi	bronquite	caótico	catstrofes

catástrofes	científicos	coferência	comodidade
catastrófica	cientistas	cogita	comodismo
catastroficos	ciêntistas	cogito	cômodo
categórico	ciganno	coibe	compactos
causada	cima	colaborativa	compactuado
causadas	cinco	colapso	companhas
causador	cinematograficas	colapsos	compará
causadora	cinzeiro	colaterais	comparada
causadoras	cio	coleccionam	comparadas
causadores	circuação	coleccionáveis	comparados
causados	circuito	coleccionavel	compartilhados
causos	circulada	coleti	compartilhamento
cautelosamente	circulantes	coletiva	compatíveis
cauzado	circulatório	coletivas	compatível
cavalheiros	circunda	coletividade	competidor
caverna	circunstancia	colhetivo	competitiva
cavernas	circunstâncias	colisões	competitividade
cavidades	citação	colocada	complexibilidade
cds	citada	colocadas	componentes
cegando	citados	colocados	comportamental
ceifando	cities	coloração	composta
cem	cituações	combatida	compostos
cenários	civilização	combatidos	comprá
cendo	clama	combostivel	comprados
censitário	clamados	combustão	comprasones
censo	clásse	combustiveis	compreendê
certeira	clássica	combustiveis	comprovadamente
cessidade	clássicas	combustivéis	comprovado
chamativas	classificado	combustíveis	comprovados
chandler	clean	combústiveis	comprovam
charme	climatica	combustivel	comprovando
charmoso	climática	combustível	comprovaram
charrete	climaticas	combústivel	comprovou
charrétes	climáticas	combústivel	compulsiva
chegariamos	climático	comce	compusivos
chevrolet	climáticos	comecar	compussariamente
chinês	clmática	começem	compussivamente
chique	cnh	comentada	compussoriamente
choca	cnh's	comercializada	comrrrompendo
chocam	cniq	comercializado	contemporâneo
ci	co	comercializam	comumente
cicletas	ço	comercializam	comunais
ciclismo	coagidas	comercializavam	comunicação
ciclista	cobiçado	comercialmente	comunicati
ciclistas	cobiçados	comercio	con
ciclofaixas	coerência	cometidos	concebidas
ciclovia	coerente	cômico	concedida
ciclovias	coerentes	comm	conceituou
cidadões	coexistir	commodites	concentização
cientifica	cofator	comoda	concerne
		cômoda	

concessionárias	conhecimentodo	consumida	contribue
concliderado	conivente	consumidas	contribuente
conciência	conjunção	consumidos	contribuido
conciente	conjuntura	consumismo	contribuidor
concientes	conotação	consumista	contribuídores
concientização	conotou	consumistas	contribuimos
conciêntização	consagradas	contada	contribuinte
concientizadoras	consciencia	contados	contribuintes
concientizando	consciência	contaminadas	contribuiram
concientizar	consciênte	contante	contribuïrem
conciêntizar	consciêntes	contemplaçãõ	controlada
concimento	conscientização	contemporanea	controlado
concluimos	conscientizada	contemporânea	controvérsia
concomitantemente	consciêntizada	contemporaneidade	contúdo
concretas	conscientizados	contemporaneidade	conturbação
concretização	consciêntizar	contemporâneo	conturbado
concretize	conse	contemporâneos	convem
condado	consecionaria	contemporanidade	convém
condenados	consequencias	contemporaniedade	convência
condicionado	consensual	contemporânio	convêncimento
condicionados	consentização	contestação	convencionais
condicionamento	consequencia	contestações	convencional
condiz	consequencias	contestado	conveniente
conectados	consequente	contestar	convertido
conectam	consequêntemente	contéudo	convicênte
conectando	consequinte	contexto	convidativo
conectar	consertadas	contidos	convivio
conectividade	conservassão	contigente	corno
confeccionar	considerada	contingenciamento	corona
conferencia	consideradas	contingente	corporativamente
configura	considerados	continuo	corredos
configuração	consideraveis	continuamente	correlacionada
configurações	consideravel	continuação	correlacionado
confortabilidade	consiende	contradições	correlatas
confortavel	consientes	contradiz	corriqueira
confortavelmente	consientização	contradizendo	corriqueiro
confrontada	consoante	contrafato	corrobora
confrontados	consolidar	contraíndicações	corroborado
conges	consolidaram	contraíndicado	corroboram
congestão	consolidou	contraíndo	corroborando
congestionadas	consonância	contramão	corroborou
congestionado	constância	contrapartida	corrompe
congestionados	constantemente	contrapõe	cortela
congestionam	constatação	contrapõem	cósmica
congestionamento	constrangedoras	contrapostos	costumava
congestionamentos	constru	contras	cotidiana
congestionando	construídas	contraste	cotidianas
congestionar	construídos	contratualista	cowspiracy
congregação	construtivas	contribrindo	crecendo
conhecê	consultorio	contribudo	crescente

crescentemente	debaixo	demasiadamente	desaprovação
crescentes	debatida	demasiado	desaprovar
cresiam	debatidos	demografica	desarcebar do
cresimento	debilitada	democra	desativação
criá	debilitando	democratização	desavenças
criativas	debye	democratizar	desbravar
crima	decada	demografia	descar
criolo	decadas	demografica	descarbonificação
crowell	decaindo	demográfica	descartá
crown	decair	demograficamente	descartada
cuja	decepcionam	demografico	descartado
cujas	decisória	demográfico	descartados
culminando	declineo	demográficos	descartando
culminou	declínio	demons	descartável
culo	declino	demorada	descarte
culpados	declive	demostrar	descartes
cultiveram	decorrente	demostre	descaso
culturalmente	decorrentemente	denominada	descoberta
cumprida	decorrentes	denominado	descobertas
curitiba	decrecentes	denominador	descoberto
curitibano	decrécimo	denominados	descolada
cury	dedicada	denota	descolados
custos	deesel	densidade	desconfortante
custoso	defasada	depara	desconfortáveis
dada	defasagem	deparar	desconfortáveis
dadas	déficit	deparou	desconsiderável
dade	deficiêntes	dependência	desconstruir
dades	deficitário	dependencia	descontrolada
dahmer	definidas	depreciativos	descostume
dámos	definidor	depreende	descrita
dandos	definitivas	depreender	descutidas
danifica	degelo	der	descutido
danificado	degeneração	deriva	desda
danificar	degrada	derivadas	desdobramentos
danifiquem	degradação	derivado	desejadas
danosas	degradações	derivados	desejado
danoso	degradam	dernidade	desejados
danosos	degradando	derretendo	desejável
daquela	degradante	derretimento	desempenha
daqueles	degradantes	des	desempenham
daqui	degradar	desabrigadas	desemprega
daquilo	degradativos	desacelerar	desencadeando
darem	deicha	desafogaro	desencadear
datafolha	deixada	desagrada	desencadeia
dauson	deixados	desagradavel	desencadeiam
dável	deliberado	desagrado	desencadeou
david	delimitações	desapegados	desencadio
dawning	demânda	desapegar	desencorajam
dawson	demandados	desapegarem	desencorajando
dean	demasiada	desapego	desenfreada

desenfreadamente	despensável	determinados	diminuídos
desenfreadas	desperdiçam	determinante	diminuíra
desenfreado	desperdiçando	detiam	diminuíram
desenfreado	despertados	detran	diminuírem
desenvolvida	despeza	detrimento	dinâmica
desenvolvidas	despoluir	deturpado	dinâmica
desenvolvido	despopularização	deuso	dinamicidade
desenvolvidos	desprazer	devasta	dinâmico
desenvolvimentista	desprendendo	devastação	dinamismo
desenvolvimentistas	despreocupação	devastadora	dinamismo
deseçados	despreparado	deveríamos	dinamização
desequilíbrio	despresível	deviaram	dinamizar
desérticos	desprezada	devida	dinossauros
desesperadoras	desproporcional	devidas	dioxido
desestabilizando	desprovida	dez	dióxido
desestimulação	desregradas	dezembro	dispensaria
desestimulam	desregulado	dezoito	diretas
desestimular	desrespeita	diada	diretriz
desfavoravel	desromantização	diagnosticadas	diretrizes
desfavorável	desserviços	diariamente	dirigí
desfavorecido	dessusso	diarias	dirigíveis
desfeito	destarte	dicial	dirigíveis
desfruta	destinguir	dicotomia	discolado
desgasta	destrair	didático	discolândia
desgastando	destruição	dido	discorriam
desgastante	destroe	diesel	discrepância
desgastes	destruída	diferen	discrição
design	destruídas	diferenciados	discução
designs	destruído	diferentemen	discursou
desing	destruidoras	diferentemente	discursões
desintegração	destrutivo	díficeis	discutida
desinteresse	desunião	díficeis	discutidas
desintoxicar	desuso	difícil	discutidos
deslocam	desvairado	difícil	dísel
deslocamos	desvalor	difício	dispensavel
deslocar	desvalorização	difícultosa	dispensável
deslumbrante	desvalorizado	difícultoso	dispersão
desma	desvalorizados	difundida	dispertam
desmatamento	desvalorizando	difundidas	disponibilida
desmatamentos	desvantagem	difundido	disponibilidade
desmatando	desvantagens	difundir	disponibilizado
desmatar	desvencilhar	difusão	disponibilizados
desmedida	detectadas	digital	disponibilizam
desmistificação	detector	dignificação	disponibilizar
desmistificada	detentor	dilema	disponíveis
desnecessaria	detentores	diluído	disseminação
desnecessariamente	deterem	dimensão	disseminadores
desnecessario	deterioração	diminuíra	disseminados
desordenada	determinada	diminuída	dissertar
desordenado	determinadas	diminuído	dissipação

dissipada	dualismo	efóca	emitem
dissolveram	duas	egito	emitida
distancia	dublin	ego	emitidas
distinção	duradoura	egocentrista	emitido
distintas	duraduras	egoismo	emitidos
disto	durant	einstein	emitindo
distorce	durkheim	eixo	emitir
distribuidas	durkhiniana	elaboração	emitirem
distúrbios	dutra	elaborados	emparelhado
ditador	duvidável	elegantes	empasse
diverge	eas	eletrecidade	empatia
divergências	eça	eletricas	empecilho
divergente	ecesso	eletricidade	empecílios
diversificação	eclodiu	eletrico	empenha
diversificando	eclosão	eletricos	empírico
diversificação	eco	elétricos	emplacando
divisor	eço	eletrizante	empportamos
divulgada	ecologica	eletromagnético	empreender
divulgadas	ecológica	elevada	empreitada
divulgados	ecologicamente	elevadas	empresas
dobrou	ecológicas	elevados	emprestimos
doen	ecológico	eliminados	empresas
dois	ecologicos	elite	emquanto
doís	ecológicos	elites	en
dolares	economia	elitista	enaltação
domado	ecônomia	elucidam	enaltecia
dominação	economias	emancipação	enaltecido
dominância	economicamente	embalados	encaixa
donas	econômicamente	embalagem	encaixarem
donde	economico	embalagens	encantadas
doque	económico	embarcada	encentivar
doravante	economicos	embelezá	encentivos
dorax	económicos	embelezamento	enceramento
dosney	ecossistema	embriaguês	enchegam
dotado	ecossistemas	embrólios	enchegavam
dramaturgia	edemas	emerge	encherga
drástica	eduardo	emergem	enchergar
drasticamente	educação	emergencial	enchergarmos
drásticas	educativo	emergente	encima
drastico	efêmero	emergentes	encontrada
drástico	efervescencia	emergindo	encontradas
drásticos	éfeso	emete	encontrados
dreamworks	efésto	emição	encorajadas
drenador	efetividade	émile	encorajado
driblar	efetuado	eminente	encurtamento
drummond	efetuados	emissão	encurtando
drumond	eficácia	emissões	encurtar
dss	eficiência	emissor	endeuzando
dter	eficiente	emissores	energetica
dualidade	eficientes	emite	ênfase

ênfatizam	ênvolvidos	ênpetacularização	ênstoques
ênfatizando	ênxofre	ênspirito	ênstração
ênfatizar	ênólica	ênesportivos	ênstrado
ênfatizou	ênólico	ênexpressão	ênstraga
ênfoque	ênépoca	ênesquecimento	ênstrago
ênfoques	ênépocas	ênesquesam	ênstragos
ênfraquece	ênquestionavel	ênessencialismo	ênestranhamento
ênfraquecendo	ênequilibrio	ênestá	ênestrategia
ênfraquecimento	ênequipara	ênestá	ênestratificação
ênfrentados	ênequiparado	ênestabelecidas	ênestratosfera
ênga	ênequiparam	ênestabilizar	ênestreitas
ênengajada	ênequivocada	ênestacionados	ênestress
ênengajamento	ênequivocadamente	ênestacionar	ênestressadas
ênengarrafadas	ênequivocou	ênestadunidense	ênestressados
ênengarrafamentos	ênéram	ênestadunidenses	ênestressante
ênengarrafamento	êneric	ênestagnação	ênestrita
ênengarrafamentos	ênerrada	ênestagnados	ênestrutural
ênengavetamentos	ênerradicado	ênestagnam	ênestu
ênengenharia	ênerradicar	ênestagnou	ênestudada
ênengengei	ênerrados	ênestam	ênestudadas
ênengenhoca	ênerterna	ênestamortalismo	ênestúdios
ênengerrafamento	ênerving	ênestampado	ênestufa
ênengloba	ênesa	ênestao	ênestufas
ênenglobe	ênesacerbada	ênestara	ênesvaem
ênengrossa	ênesbanjar	ênestarem	ênetanol
ênnovação	ênesbarar	ênestarrece	ênetc
ênenquadra	ênescancarando	ênestática	ênéticos
ênenquato	ênescapamento	ênestáticas	ênua
ênenraizada	ênescapamentos	ênestatísticas	ênueropa
ênenraizada	ênescapatória	ênestatos	ênueopéia
ênenraizado	ênescasso	ênestatus	ênueopeias
ênenriquecendo	ênescassos	ênestavamos	ênueopéias
ênenriqueceu	ênescher	ênestável	êneva
ênenriquecida	ênescoamentos	ênesté	ênevacuação
ênenrraisado	ênescrava	ênesteira	êneventuais
ênental	ênescultura	ênesteiras	êneventual
ênente	ênese	ênestenção	êneventualmente
ênentendida	ênesfeitos	ênestetica	ênevidencia
ênentorno	ênesfera	ênestética	ênvidência
ênentranham	ênesfero	ênestéticas	ênvidenciadas
ênentrelaçando	ênesgotados	ênestético	ênvidenciado
ênentretando	ênesgotamento	ênestimativa	ênvidenciados
ênentupindo	ênesgotaveis	ênestimativas	ênvidenciam
ênenturmar	ênespaçar	ênestimável	ênvidenciando
ênentusiastas	ênespacial	ênestipulado	ênvidenciaram
ênenunciados	ênesparecer	ênestipulando	ênvidências
ênenvadiram	ênespe	ênestiver	ênvidenciava
êninvestir	ênespecialistas	ênestivesem	ênvidenciável
ênenvista	ênespectador	ênestocolmo	ênvidentes
ênenvolvidas	ênespelir	ênestopim	ênevitadas

evitados	exercícios	extingue	fazerem
evitardos	exesivo	extinguidos	fedido
evitável	exessiva	extinto	feiday's
evolucionista	exessivo	extintos	feita
evolua	exesso	extração	feitas
evoluídas	exeto	extrangeiras	feitos
evoluimos	exibi	extrativismo	felicitações
evoluindo	exibí	extrema	felismente
evoluiram	exibição	extressante	fenomeno
evolutivo	exibicionismo	exuberante	fenómeno
evovidos	exibido	exuberantes	fenómenos
ex	exibidos	fa	fermentação
exacerbada	exibir	fabricadores	ferra
exacerbadamente	exigida	fabricados	ferroviária
exacerbadas	exílio	fabricam	ferroviárias
exacerbado	existir	fabricante	ferroviario
exacerbados	exorbitante	fabricantes	ferroviário
exageiro	exorbitantes	fabricar	ferrovias
exagerada	expansão	fabricas	ferrovias
exagerados	expande	façais	festejam
exala	expandidos	facil	fetiche
exalada	expandindo	facilitação	fetichismo
exaltação	expandir	facilitador	fichador
exaltações	expandiu	facilitadores	fictí
exarcebada	expansão	facinaram	fictícia
exarcebado	expelidas	facínio	fictício
exaustivo	expelido	fadismo	filippo
excassez	expelidos	falácia	filosofa
excede	experimentado	falado	filosófica
excence	explicada	falecimento	filosófico
exces	explicado	falescer	filosofo
excessivamente	explícita	falsa	filósofo
excessividade	explícitam	familia	filósolo
excessões	explícitamente	familias	filtro
excludente	explícito	fances	filtros
excluem	explorada	fapesp	financiar
exclui	explosivo	farmaceuticas	findando
excluído	expoente	fascinadas	finitos
excluídos	exponencial	fatais	firmada
exclusiva	exponencialmente	fatal	física
exclusivas	exportações	fatalidades	físicas
exclusividade	expossecialmente	fatigados	fisiologicas
exclusivas	expostos	fatura	fixação
execiva	exprecivamente	fauna	flexibilidade
exelente	exprefeito	favoraveis	flora
exem	expressivamente	favorecidas	florir
exemplifica	extensões	favorecido	flúi
exemplificando	extensa	favorecimento	fluides
exencial	extensivo	fazando	fluidez
exerbado	extenso	faze	fluidez

fluído	freiassem	geleiras	grossa
fluvial	freios	generalizada	grotesca
fobias	frenesi	genético	grotesco
focada	frenetico	geografica	quando
focados	frequentada	geografico	gueima
focoault	frequentadas	geográfico	gulismo
folheto	frequentemente	geógrafo	h
fomemtaram	frequêntes	geomorfológicas	ha
fomentação	friedrich	george	hà
fomentam	friends	gerá	habiente
fomentando	froyd	gerada	habilitação
fomentar	fu	geradas	habilitados
fomento	fuligem	gerado	habilitarem
fomentou	fumaceando	gerador	habilite
for	fumacentos	gerados	habita
ford	fumante	geridas	hábita
fordismo	fumantes	gerir	habitado
fordista	fumassa	gestionamentos	habitam
fordt	funadação	gestores	habitamos
fordy	funcional	gilbes	habitat
formula	funcionalidade	glamour	habitáveis
formulou	fundamenta	glo	habitável
fornece	fundamentado	globalização	habito
fornecendo	futura	globalizada	habitots
fornecer	futurama	globalizado	hailitação
fort	futuramente	globalizando	halito
fortificado	futuras	globo	hálito
fosseis	future	gloriosa	hambiente
fosséis	futurismo	gnv	hannah
fósseis	futurista	goebbeles	hans
fossil	futuristas	goffman	harari
fóssil	futurística	goiânia	harmônia
fossilizados	futurísticamente	governar	harmônico
fotossintese	futurístico	gradativa	harvard
fox	gabriel	gradativamente	harvey
fragéis	gala	gradativo	haven
fragilizado	galinhas	gradualmente	hawking
fragmentos	gambá	gráficos	hegel
frança	ganância	gran	hegemonia
francesas	gananciosas	grandesas	hematose
francis	gandhi	gratos	hemoglobina
frankfurst	gandhin	gravemente	henri
frankfurt	gar	gravíssimas	henry
frankfurtt	garantêm	green	heráclito
franquia	gargalos	greenpeace	herádito
franz	garupa	greg	herby
frases	gase	gregório	héroi
fraturar	gasosas	grepeace	hibridos
fre	gasses	greta	híbridos
free	gees	gritante	hidreletricas

hidrocarbonetos	idealizado	implementadas	incentivadores
hidrogênio	idealizadores	implementado	incentivamento
hidroviario	idealizados	implementados	incentivase
hidrovias	idéia	implica	incêntivo
hierarquia	ídeia	implicam	inçentivo
hiperlotação	ideiais	implicando	inceri
hipotece	idéias	implícita	incerto
historiadores	identitária	impopular	incesante
historico	idh	importados	incessante
históricocultural	idolatrado	importancia	incessão
ho	ignorados	importantissimo	incetivam
hobber	igualam	imposição	inchaço
hobbes	igualando	impossibilitadas	inchanço
hobbez	igualar	impossibilitados	inclinação
hobby	ii	impossível	incluído
hodiernamente	ileso	imposta	inclusivos
hodierno	ilusória	impostas	incô
hodiernos	ilusórias	imprescindível	incoerente
hohbes	ilustra	impressos	incômo
holanda	ilustrando	imprevistos	incômoda
hollywoodiano	ilustrar	imprudencia	incomodadas
hollywoodianos	imedia	imprudência	incomôdas
holofotes	imediatismo	impulsionada	incômodo
homeopáticas	imensas	impulsionou	íncomodo
homo	imensurável	impulsionados	incômodos
homos	imergir	imunológico	incompatibilidade
hon	immanuel	in	incompatível
honda	imobilidade	inaceitável	incompleta
honestamente	imobilismo	inacessíveis	incompreensível
hong	imobilização	inacessível	inconcebíveis
horárias	imobilizam	inacessível	inconcequentes
horario	impace	inacreditavel	inconscientemente
hordiadamente	impacta	inadequação	inconsequentes
hordienos	impactando	inadequado	incontáveis
hormônios	impactantes	inadequados	incontrolavel
hoss	impactará	inala	incontrovertível
hostilizante	impactaram	inalação	inconveniente
houver	impactem	inalada	inconvêniente
humana	impactou	inalado	inconvenientes
humanas	impaquito	inalados	incorporado
humanida	impaquitos	inalamos	incorporando
humanística	impases	inalando	incorporar
humanização	impensevel	inalar	incorporaram
humanizando	imperceptível	inanimado	incorretamente
humorística	império	inatingível	incrédulos
hyundai	ímpeto	inauguradas	incrementações
ibge	implantadas	incabível	incrível
ícone	implantados	incalculáveis	incuravel
idealistas	implementação	incapacita	indecoroso
idealização	implementada	incapacitado	independencia

indepêndencia	ineficiência	inofensivo	intensificadores
independendo	ineficiente	inofensivos	intensivo
indepêndente	ineficientes	inóspito	interação
indesejado	inegável	inovador	interações
indevida	inegavelmente	inovadora	interagem
indevido	inércia	inovadoras	intercepta
indevidos	inercial	inovadores	interdependência
indicador	inerente	inovando	interditados
indicadores	inespido	inovar	interim
índice	inexistencia	inovação	ínterim
índices	inexistência	inquestionável	interligá
índices	infalivelmente	inresponsabilidade	interligadas
indireta	infarto	insentivada	interligando
indiretamente	infelizmente	insentivo	intermináveis
indiretos	infere	insentivos	internamente
indiscriminada	inferindo	inserção	internos
indiscriminado	inferir	insere	interpretá
indiscutível	infernais	inserida	interpretada
indiscutivelmente	infinidade	inseridas	interrogado
indispensável	inflamento	inserido	intertextualidade
individamento	infligir	inseridos	intervenções
individualismo	influciada	inserindo	intuições
individualista	influenciada	insersão	intoxicar
individualistas	influenciado	insolente	intrafegáveis
individualização	influenciadora	inspirada	intrafegável
indivíduo	influenciados	instabilidade	intragável
índivíduo	influências	instância	intransitáveis
indivíduos	influenciar	instantaneamente	intrelasada
indivíduos	influenciou	instantâneamente	intrigante
indríver	influyente	instastica	intrínseca
indubitável	influentes	instaurado	intrinsecamente
indubitavelmente	informacional	instigava	intrisecamente
indubtável	informativas	instiguir	introdução
indústria	informativo	instituído	introduzido
índustria	informativos	instituir	introduzindo
índustriais	infra	insubstituível	introduzir
índustrial	infraestrutra	insubstituível	intuíto
industrialização	infraestrutural	insuficiência	inuljável
índustrialização	infuenciam	insuficiente	inumeras
industrializado	ingere	insustentabilidade	inumeros
industrialmente	ingerir	insustentáveis	inundações
indústrias	ingestão	insustentável	inundando
induz	inglaterra	insustentável	inutilizado
induzidas	inglobados	intalar	invejável
induzido	ingrediente	integrados	invenção
induzir	ingressados	integralmente	invenções
inéditas	inibir	intencificam	invenção
ineficacia	iniciada	intensificação	invenções
ineficácia	ínicio	intensificadas	inventado
ineficaz	inofencivo	intensificado	inventados

inventário	jaguare	kubschick	limeira
inversão	jaimé	kubstcheeck	limeirense
inversões	jaimer	kubstchek	limiar
invertida	james	kubtischek	limitada
invertidos	janeiro	kumbecheque	limitados
inves	japão	kyoto	linear
investam	jargão	la	liquefazendo
investido	jean	lançada	liquida
invez	jetos	lançados	liquidez
invéz	jk	largados	lírico
inviabilidade	jobs	largamente	literária
inviabiliza	jogado	larkinismo	literário
invibializa	jogados	las	lo
invisível	john	laser	localidade
iommapo	jóias	lastimável	localidades
iommaso	jonas	lataria	localização
iorque	jonh	latente	localizadas
ipiranga	jornalísticas	le	locke
ipva	josé	lefebvre	loco
iria	jovial	legalizada	locomacão
iriamos	jovialidade	leigos	locomo
irlanda	juceli	lema	locomobilidade
ironicamente	jucelino	lençóis	locomossão
irônico	julgadas	lener	locomotivo
ironiza	junção	lenner	locomotivos
irracional	juntaposto	lentas	locomotores
irracionalidade	juros	lentidão	lógica
irrefutáveis	jus	leo	logística
irrelevante	juscelino	leonardo	logística
irreparáveis	jusselino	lerme	logísticos
irreparáveis	just	lerner	londres
irreparável	juste	les	longevidade
irreparável	justifica	letais	longínquo
irresponsabilidade	justificáveis	letal	longíquos
irreversíveis	justificável	letalmente	lorax
irreversíveis	k	letivos	lores
irreversível	kafka	levá	los
irreversível	kant	levadas	lôs
irri	kanta	levy	losk
irritabilidade	karl	libe	lotação
irritadas	kilômetros	libeirando	lotações
irritadiços	kioto	liberação	lotadas
irritável	km	liberações	lotados
is	kong	liberada	lotoção
isaac	kubichec	liberador	loves
itália	kubichék	liberados	lu
itamar	kubistchek	lícita	luca
ítem	kubitschek	lidade	lucho
itinerário	kubochek	life	lúcido
ja	kubschequi	lilás	lucrativo

luís	manuel	mecanismos	metropóles
luiz	manuseio	mecânismos	metrópoles
luminosa	manutenção	mecanizado	métrópolis
lumo	maquiável	mecher	metrópolis
lup	maquina	media	metropolitana
lurner	maquinário	mediado	metropolitanas
lusitano	maquinários	médias	mêtros
lutadores	maquinas	medio	metroviária
luxuosidade	máquinas	megacidade	metroviário
luxurass	maquinofatura	megacidades	méxico
luxuri	maravilhados	megalópole	michael
luxuria	marcada	meios	michigan
lyotard	maremoto	meiro	midas
m	maria	meis	midias
ma	mariana	melancolia	miática
macrocefalia	marianas	melefícios	miática
mad	marinetti	melharando	miáticos
made	mário	melhorados	miáticos
mahatma	maritimas	melhoramento	migração
maio	maritma	melodia	migrada
maiore	marketing	men	migrar
majoritariamente	marx	mencionado	migrarem
maléfica	mascarada	mencionados	migrassem
maléficas	mascarados	meno	milênio
maleficio	masculinidade	menores	milênios
maleficio	maslow	mensionados	millenials
malefícios	massante	mentas	milton
malefícios	massificação	mento	minarão
maléficos	massiva	mentos	mineradoras
maléficos	massivamente	mercadológica	mini
malefico	massivas	mercadológicas	minimaliza
maléfico	massivo	mercadoria	mínimas
maléficos	massos	mercadorias	minimiza
males	materialista	mercedes	minimizá
malevícios	materias	merecidos	minimiza
malha	matérias	mero	minimizada
malhas	maturidade	meros	minimizado
malz	max	mesclam	minimizando
mandela	máxima	mesh	minimizar
manipulaçã	maximizar	mesosfera	minimizou
manivela	maximo	metáfora	minimo
manobrar	maxímo	metafóricamente	minúsculas
mante	máximo	metaforiza	mirrow
mantê	mazela	metaforizado	mister
mantenedores	mazelas	metálica	mistros
manterem	mcqueen	metamorfose	mitigação
mantida	mdl	metano	mitigar
mantidas	meados	mêtro	mitologia
mantidos	mecânicos	metrópole	mizando
manu	mecanismo	metropoles	mo

mobilidade	mormente	naves	nômades
mobílidade	morosos	ne	nomeada
mobilística	mortalidade	necessaria	nomia
mobilístico	mos	necessarias	nominada
modais	mostradas	necessárias	normalizando
môdais	mostran	necessario	normatizada
modal	motivadores	necessário	norteiam
modemização	motivando	necessarios	noruega
moderado	motivaram	necessi	notada
moderna	motocicletas	necotina	notadas
modernas	motoqueiro	need	notavéis
modernidade	motorizados	negativista	notavel
modernismo	motorizadas	negligencia	notável
modernista	motorizado	negligência	noticiadas
modernização	motorizados	negligenciados	noticiado
modernizações	motors	negligências	notoria
modernizada	motostaxi's	nela	notoriedade
modernizados	motriz	nelas	notorio
modernizando	movel	nele	notório
modernizar	movida	neles	nova
modernizou	movidos	neoliberalismo	novas
moderno	móvidos	neolítica	nóvas
modernos	movimentações	neoplasias	noventa
modificações	mudada	necessi	nóvos
modificados	mudan	necessidade	nú
modismo	mudanca	necessite	nucleo
mões	multibillionário	nessecidade	nula
mogão	multinacionais	neurótico	nulo
mola	múltiplas	neutra	nume
molda	multiplica	neutralizar	numeração
moldada	multiplicar	newton	numero
moldamos	múltiplos	newtow	numeros
moldando	multirão	nicolau	numeroso
moldes	mun	nicotina	nutrida
moldou	mundança	nietzsche	ö
molécula	mundialmente	ninguem	ob
momentânea	município	nissan	obesidade
momentaneamente	municipios	nisso	óbito
momentâneo	mussolini	niterói	obras
momentâneos	nacionalização	nitido	obrigados
monetário	nacionalmente	nítido	obrigatórias
monotrilhos	nao	nitrogênio	obrigatório
monóxido	naom	nitroso	obrigatório
monoxo	naquelas	nivel	obrigatorios
monstrar	naquele	nô	obsessão
montadora	narguile	nocivas	observação
montadoras	nascidos	nocivo	observações
mordenização	nativos	nocivos	observada
mordomia	naturalizado	nologias	observados
	navegaveis	nômade	observável

obsoleta	onibús	ozonio	patológicos
obsoletas	onibûs	ozônio	patrimonios
obsoleto	ónibus	p	patrocinadora
obsoletos	onu	pa	paul
obstaculo	ônus	padania	paulistano
obtam	opcional	padecerem	paulistanos
obte	opita	padronizada	paulo
obtê	opõe	pagá	paupável
obtenção	optantes	page	pausando
obterem	oral	pagos	pautada
obtérem	órbita	países	pautado
obtidas	orcar	paixa	pautados
obtidos	organica	paiz	peace
ocasiona	orgânica	palestrantes	pease
ocasionada	orgânicas	paletos	pedagogo
ocasionadas	orgânico	paliativos	pedestal
ocasionado	orgânicos	palmente	pejorativa
ocasionados	organismo	paltada	pen
ocasionam	organizada	panhando	pende
ocasionando	orgãos	panorama	pendular
ocasionar	órgaos	papeis	pensada
ocasionou	oriental	parabenizar	pensador
ocassionadas	originados	paração	pensadores
ocassionando	originando	paradigma	pequenino
ocazionando	oriunda	parados	percalços
ocorrente	oriundas	paradoxo	percebida
ocorrida	oriundo	parafraseando	perceptíveis
ocorridos	oriundos	paragrafo	perceptível
oculpe	osonio	paralelamente	perceptível
ocupadas	osônio	paralisações	percevejos
ocupados	ostenta	paranaense	percistência
ocupantes	ostentação	paranaenses	percistir
oduvaldo	ostentam	parecidas	perdidas
oeste	ostentar	particulados	perdidos
of	ostentativo	particularidades	perduram
oferecidos	ótica	partículas	perdurar
oferencendo	otimismo	pas	perdurará
ofertadas	otimização	pasárgada	perfura
ofuscando	otimizar	pasárgadas	periculosidade
oitenta	otimizaram	passacem	periféricas
oito	otimizasse	passagem	periféricos
ok	otimo	passível	periféricos
ole	otmização	passividade	perigos
oli	outono	pastos	perímetro
oliver	outrem	patamar	perímetros
omissão	outrora	patinete	periodicamente
oms	outrossim	patinetes	periódicos
ongs	óxido	patinètes	periodo
ong's	óxidos	patins	permeada
onibus	oxigênio	patologias	permeáveis

permitidos	plântio	poluit	possuimos
perpetuam	platão	pompa	possuitem
perpetuando	plo	ponderar	possuisse
perpetuar	pneus	pontuais	posteriori
perpetuou	po	pontual	postular
perplexos	pocibilita	pontualidade	potanto
perrengue	podera	pontuar	potável
persiste	poderam	pontuou	potencia
persistem	poderem	pool	potência
personalizações	poderíamos	popu	potencialidade
personalizado	poderosa	popula	potencializada
personalizar	pódio	população	potencializador
perta	podução	população	potencializam
pertencente	poibido	populacionais	potencializando
pertencimento	poise	populacional	potencializar
pertinentemente	polares	popularidade	potencializou
pertubádor	pole	popularização	potencialmente
pes	polêmica	popularizada	potências
pesadas	políticas	popularizado	pr
pesquisadas	políticos	popularizados	pra
pesquisados	polo	popularizando	pradões
pessima	polonês	popularizar	pragmatismo
pessíma	polu	popularizou	prapagandas
péssima	polua	popularmente	práti
péssimas	poluam	populista	praticado
péssimo	poluem	populoção	prático
petições	poluente	populosa	práticam
petroleiros	poluentes	populosas	praticante
petroleo	polui	poranto	praticidade
petrolíferas	poluí	porcentagem	praticidades
petroquímicas	poluía	porcento	praticos
ph	poluição	porem	pre
pib	poluição	portador	precaredade
pico	poluições	portadoras	precária
pierre	poluidas	portland	precárias
pilar	poluídas	portuguesa	precaridade
pilotado	poluido	portunismo	precariedade
pilotar	poluído	pos	precário
pioneiro	poluidor	pose	precários
pitágoras	poluídor	positivamente	precarização
pitor	poluidoras	positivo	precarizando
pivo	poluídoras	possabilita	precarizou
pivos	poluidores	possibi	precede
pixar	poluidos	possibilitado	precedentes
plana	poluindo	possíveis	preceito
planaram	poluino	possível	preceitos
planejada	poluintes	possue	preciosissimo
planejadas	poluir	possuia	precisamo
planetário	poluiram	possuiam	precissam
planeto	poluíram	possuidor	precocemente

preconceituosos	primero	proibido	providencias
precursor	primicia	proibidos	provindo
precursores	primordiais	projetadas	proximo
predominante	primordial	projetado	prudência
preenchidas	primordialmente	projetar	psicologa
preferencia	primordios	prol	psicológicas
preferencias	primórdios	pról	psicólogo
preferível	princi	proletariado	psicotécnicos
preferível	principado	proliferação	psíquica
pregando	principio	proliferados	pú
preguiçosa	principios	promissor	pub
preguiçossos	prio	promovida	publicada
preju	priór	promulgada	publicados
prejudi	priori	promulgado	publicidade
prejudicada	priorização	propaga	publicidades
prejudicados	priorizada	propagação	publicitária
prejudicial	priorizados	propagandeado	publicitárias
prejudicial	prisma	propagar	publicitários
prejuísos	privando	propenso	publicos
prejuízo	privilegiada	propícia	públicou
prejuízos	privilegiados	propiciar	pubrico
premente	privilegiamento	propiciou	pul
premissa	privilegios	próprio	pulmonares
preocupante	pro	proporção	punidos
preocupantes	pró	proporcionada	purista
prepulso	probabilidade	proporcionadas	puxada
presenciada	problematica	proporcionados	puxados
preservá	problemática	proporcional	quadrado
preservação	próblematica	proporcionalmente	quadrinista
presidencia	problemáticas	proporções	qualificada
presisando	problematico	proposito	quantidade
pressionados	problemático	propostos	quantificar
pressuposto	problematiza	propria	quarenta
prestados	problematização	proprias	quarta
prestigio	produção	proprio	quarteirão
préstigio	produtividade	proprios	quatro
presume	produtoras	popularizou	quatrossentos
pretensões	produzida	prós	queele
preterido	produzidos	protagonismo	queimadas
preocupará	profetor	protagonista	queimados
prevalece	profícuo	protagonistas	queinando
prevaleceu	progeção	protagonizados	queiroz
previêncial	prograganda	protocolo	queles
preza	progressismo	protocolos	quências
prezar	progressista	prototipos	quentemente
primeira	progressiva	provavel	ques
primeiramente	progressivamente	provém	quesito
primeiras	progressivo	provêm	quesitos
primeiro	proibidas	proveniente	questionada
primeiros	proibido	provenientes	questionadas

questionados	reconsiderado	remuneração	respectivos
questionamente	recordista	renascentista	respiratoria
questionavel	recorrencias	rendimento	respiratória
questionável	recriam	rené	respiratorias
quiçá	recusão	reno	respiratórias
quilométricas	redirecionar	renomado	respiratorio
quilométricos	redução	renovaveis	respiratório
quilometros	redução	renovavel	respiratórios
quimico	reduzida	renovável	responsaveis
quiser	reduzidos	rentabilidade	responsavel
quistas	reedu	rentáveis	responzabilizados
quistos	reeducação	rentável	respos
quos	reeducar	repensá	resquícios
racionalizam	reestruturado	repensâ	ressalvas
radiação	reestruturem	repensada	restos
rafaela	referencial	repensado	restringido
raioativos	referente	repensados	restringidos
ram	referentes	repensar	restringir
rando	refletida	repense	restrita
range	reflorestamento	repensem	restrito
rão	reforçada	repentina	restritos
rapida	reformulação	repercutindo	resultados
rapidamente	refuta	repercutir	resultante
rapides	refutada	repercussões	resultantes
rapidez	refutados	repercussão	retardada
rapido	reger	repercussões	retém
rapper	regiao	repercutindo	retenção
rarem	regidos	repetitivo	retratada
ras	regista	replanejar	retratados
rasões	regressão	reportes	retrógado
raul	regresso	reprensão	reurbanização
raza	regula	represas	reurbanizar
razao	regulação	repressivo	reutilizado
razoável	regularizado	reprimidos	revelados
ré	regularizem	repúdia	reverências
reabilitação	regulatórias	repudiados	revertida
reafirma	reinvenção	repugnante	revertidos
realista	reinventando	repugno	vezamento
realizada	reiventando	repulsão	revidar
realizadas	reivindicar	requeridos	revistos
realizados	rejeitados	requisito	revogada
reavaliar	relatividade	requisitado	revolucionando
rebeldia	relegado	requisitar	revolucionou
recarregada	relevância	resaltar	ridades
recarregado	relevante	resguardado	rígidas
recarregamento	relevantes	residem	rigidez
recarregavel	relevar	residentes	rinite
recipiente	reliquias	residuos	ritimo
reclusa	remediada	resolvê	rítimo
recompensadores	remediar	resolvedor	rito

rização	samento	segmentos	significativa
rizar	samotrácia	segregação	significativamente
ro	sampa	seguidamente	significativas
robô	samsa	segunda	silenciosa
rôbos	samu	segundamente	simbolismo
robóticas	sanar	segundos	simboliza
rodagem	sanidade	seguradora	simbolizar
rodeados	santayana	seguradoras	simbolizavam
rodizio	sao	selvagens	simbolo
rodízio	sapiens	semáforo	simbólo
rodizios	sarte	semáforos	similar
rodízios	sartre	semânticos	similarmente
rodoviária	satayanas	semblantes	simplificação
rodoviárias	satisfatórias	semelhança	simultaneamente
rodoviario	satisfatório	semelhanças	senaliza
rodoviarismo	satisfeitas	semelhansa	senalização
rodrigues	satistação	semelhantemente	senalizações
rolagem	saturação	semi	síndrome
rolorado	saturações	sendentarias	singelamente
romântica	saturadas	sendos	singular
romantizada	saturado	sensacionalista	singularidade
rompida	saturando	sensibilização	sinonimo
rompitura	saturar	sensoriais	sinónimo
rondam	sau	sentada	sinônimo
ros	saudaveis	sentadas	sinônimos
ross	saudavéis	separadamente	síntese
rosseau	saudavel	sequencialmente	sintetizado
rota	saúdavel	sequentemente	sinusite
rotação	saude	sequestrariam	situa
rotas	sáude	sequestrou	situou
rotineiro	saudvéis	sera	skate
rótulo	schopenhauer	seram	skates
rótulos	sé	serao	smart
rousseau	seão	serem	smarthphones
rover	séc	seres	smarts
rrafado	sécu	sergio	smbolo
ruido	seculo	serie	so
rungida	seculos	series	sobrepõe
rusa	secundário	serio	sobrepor
russo	secundarizados	serios	sobrevivencia
rvir	sedentaria	sermos	sobrevivente
saberes	sedentária	serras	socialmente
saga	sedentárias	seru	sociedade
sageiros	sedentario	servar	socieda
saiem	sedentário	severas	sociêdade
saldoso	sedentários	severo	sociedade
salentar	sedentarismo	severos	socio
salienta	sedentarização	sí	socioambientais
salientar	seduzidas	siderúrgicas	socioambiental
sálise	seduziu	significamente	socioculturais

sociocultural	substantivo	suscetibilidade	tê
socioeconômica	substituído	suscetíveis	tec
socioeconômicas	substituidor	sussebilidade	tecnicamente
socioeconômico	substituídos	sustentá	técnicas
sociologia	substituílo	sustentabili	técnicocientífico
sociológica	substituisse	sustentabilidade	tecnologia
sociológico	substituível	sustentação	técno-logia
sociologo	subterrâneo	sustentáveis	técno-logia
sociólogo	sucateada	sustentáveis	tecnológicas
sócrates	sucateado	susténtáveis	técno-logias
sofisticação	sucateados	süstentáveis	técno-logias
sofisticadas	sucatiando	sustentavel	tecnologica
sofisticados	suce	sustentável	tecnológica
solar	sucedida	susténtavel	técno-logia
solares	sucessiva	sustêntavel	tecnologicamente
solidaria	sucessor	sustentavelmente	tecnológicas
solidárias	sucessores	sustentaveu	tecnológicas
solidarizarem	sucumbindo	sutentáveis	tecnologico
solucionados	sueca	sygmund	tecnológico
somados	super	ta	técno-lógico
somen	superdependente	tabaco	tecnologicos
sonetos	superestimado	tabáco	tecnológicos
sonhada	superficiais	tabagismo	tecnológicos
sonora	superficial	tabagistas	telejornais
sonôra	supérfluos	tabela	telejornal
sonoras	supérfulo	tablet	televisiados
sonoro	superinteressante	tabula	televisiva
sonoros	superlotação	tam	televisivas
sovre	superlotações	tambem	televisoes
sp	superlotados	tambêm	tenção
speed	superpõe	támbem	tendão
srur	superpopulação	tambén	tendencia
star	superprodução	tamente	tenebrosos
statu	superquadras	tamento	tenem
statuário	supervalorização	tando	tenologia
status	suposta	tange	tentável
stephen	supraátadas	tanta	teologia
steve	supracitada	tantas	teor
stituir	supracitado	tantos	tera
stop	supracitados	tao	terceira
stress	supressão	tarão	terceiros
stresse	suprida	tarçila	terem
sub	surgidos	tardia	térmica
subdesenvolvido	surpreendente	taxado	térmicas
subdesenvolvidos	surreal	taxi	terminações
subexistir	surte	taxí	termômetros
subjugada	surtiram	taxis	terrestre
sublime	surtiu	taxi's	terrestres
submetida	surto	táxis	territorial
substancias	sus	taylorista	territorio

terríveis	tos	tras	unidas
terrivelmente	totalitário	transporte	uniformizou
tes	tóxica	transporte	urbanistas
tesla	tóxicas	tratá	urbanístico
that	toxicidade	tratores	urbanização
the	tóxico	travam	urbanizadas
thomas	toxicológicas	travas	urbno
thumberg	toxicos	tráz	urge
thunberg	tóxicos	trazera	usá
ti	toxinas	trazerem	usabilidade
tidos	toyorista	trazetos	usada
tiête	toyota	trêm	usadas
tiilizam	toyotismo	trêns	usados
timento	traballadora	três	usalos
tionamento	traduz	trezentos	usinas
tionamentos	trafegam	tributação	usp
tionar	trafegar	trilhagem	usual
tirá	tráfego	trinta	usuario
tiradas	tráfegos	triplo	usúario
tista	tragetos	trivial	usúarios
titulo	tragicamente	tros	usuflua
tivas	trajetória	trouzer	usufluem
tividades	trama	trxtos	usufluir
tivo	tranpor	tumutuo	usufruimento
tizaram	trans	tunel	usufruirem
tnelebús	transformações	tupiniquim	usufruto
to	transformada	turbinado	util
toa	transformers	turbulento	utilidade
tocante	transforte	tv	utilidades
tões	trânsi	uber	utilitários
toiota	trânsiito	úber	utilitarista
tolerância	transitão	ubers	utilizá
tomada	transitável	uber's	utilização
tomadas	trânsito	utiliza	utilizada
tomados	transitos	utilizados	utilizadores
tomão	trânsitos	útilizão	utilizados
tommaso	translocar	utilizar	útilizar
tomóveis	transmitida	ultima	utilizável
tona	transmitidos	últimamente	utilizer
tópico	transmutar	ultimas	utilizado
tópicos	transnacionais	ultimo	utopia
toppazio	transpor	ultimos	útopica
tóquio	transportadoras	ultrapassada	utópicas
tória	transporte	ultrapassadas	uva
torios	transporter	ultrapassados	vagorosamente
tormento	transportes	ume	valar
tornandoo	transpote	umidade	valerizando
tornár	transtorno	umo	válida
tornarvam	transtornos	unica	validar
torndo	traram	unico	válidas

válidez	veiculos	vilãs	wall
válido	veis	vincia	walle
válidos	vel	vínculado	walter
valiosa	velocímetro	vinculados	wars
valioso	veloses	vinda	washington
valorativo	velosos	vindas	washiton
valorização	veloz	vinil	wation
valorizados	velozes	vinte	watson
valorizamento	vendida	virada	way
valoroso	vendidos	vírgulas	waze
vangloriava	venhar	virtuais	wazecar
vanglorização	veras	virtualmente	wazepool
vanguar	verificase	visados	wilde
vanguarda	versáteis	visível	will
vanguardas	versus	vísível	x
vantajosa	vertente	vislumbre	xílio
vantajoso	vertiginosamente	vislumbrem	ximando
vapor	ves	vistoria	ximas
vargas	vés	visualizada	xix
variabilidade	vetores	visualizar	xviii
variações	viabilizam	vital	xviil
variadas	viabilizando	vitimando	xvill
variavel	viabilizar	vitimas	xx
variável	vianna	vivencia	xxi
variedade	viária	vivenciada	xxl
varios	viarismo	vivenciamos	y
vasta	viaveis	vivenciando	yellow
vasto	viáveis	vivida	yenal
vastos	viavel	vívido	york
váveis	viável	vizando	you
ve	viciadas	vlt	youtube
vecal	viciados	vo	yuval
vêm	viciante	vó	z
veganismo	viciantes	voador	zé
vegetarianismo	vicío	voadores	zelam
véi	vicios	voce	zero
veias	vide	voluveis	zigmund
veicolos	viés	vonta	zygmont
veiculados	viéš	voracidade	zygmund
veiculando	vies	vorazes	zygmunt
veicular	vigente	vorecendo	
veiculares	vigentes	vulgarização	
veiculo	vigorosamente	vygostky	