

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

Jéssica Munique Marques Borges

**ITINERÁRIO DE LEITURA DE CONTOS LITERÁRIOS: UM ESTUDO LÉXICO-
SINTÁTICO POR INTERMÉDIO DA LEXICOLOGIA E
LINGUÍSTICA DE CÓRPUS**

GOIÁS
2023

JÉSSICA MUNIQUE MARQUES BORGES

**ITINERÁRIO DE LEITURA DE CONTOS LITERÁRIOS: UM ESTUDO LÉXICO-
SINTÁTICO POR INTERMÉDIO DA LEXICOLOGIA E
LINGUÍSTICA DE CÓRPUS**

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

2023

**TERMO DE AUTORIZAÇÃO PARA PUBLICAÇÃO DE TESES E DISSERTAÇÕES
NA BIBLIOTECA DIGITAL (BDTD)**

Na qualidade de titular dos direitos de autor, autorizo a Universidade Estadual de Goiás a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UEG), regulamentada pela Resolução, CsA nº 1.087/2019 sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9.610/1998, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data¹. Estando ciente que o conteúdo disponibilizado é de inteira responsabilidade do(a) autor(a).

Dados do autor (a)

Nome completo: Jéssica Munique Marques Borges

E-mail: jessicamunique1997@gmail.com

Dados do trabalho

Título: ITINERÁRIO DE LEITURA DE CONTOS LITERÁRIOS: UM ESTUDO LÉXICO-SINTÁTICO POR INTERMÉDIO DA LEXICOLOGIA E LINGUÍSTICA DE CÓRPUS

Tipo:

Tese Dissertação

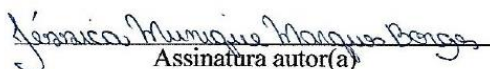
Curso/Programa: PROGRAMA DE PÓS-GRADUAÇÃO STRICTO SENSU EM LÍNGUA, LITERATURA E INTERCULTURALIDADE - POSLLI


Concorda com a liberação documento

SIM NÃO

¹ Período de embargo é de até **um ano** a partir da data de defesa.

Goiás-GO, 30 de junho de 2023.


Assinatura autor(a)


Assinatura do orientador(a)

CATALOGAÇÃO NA FONTE

Biblioteca Frei Simão Dorvi – UEG Câmpus Cora Coralina

- B732I Borges, Jéssica Munique Marques.
O itinerário de leitura de contos literários : um estudo léxico-sintático por intermédio da lexicologia e linguística de córpus [manuscrito] / Jéssica Munique Marques Borges. – Goiás, GO, 2023.
119 f. ; il.
- Orientador: Prof. Dr. Eduardo Batista da Silva.
Dissertação (Mestrado em Língua, Literatura e Interculturalidade) – Câmpus Cora Coralina, Universidade Estadual de Goiás, 2023.
1. Linguística de corpus. 1.1. Métricas léxico-sintáticas. 1.2. Conto literário - prática de leitura.
I. Título. II. Universidade Estadual de Goiás, Câmpus Cora Coralina.

CDU: 81'1:028.1

Bibliotecária responsável: Marília Linhares Dias – CRB 1/2971

UNIVERSIDADE ESTADUAL DE GOIÁS

(Criada pela lei nº 13.456 de Abril de 1999, publicada no DOE-GO de 20 de Abril de 1999)

Pró-Reitoria de Pesquisa e Pós-Graduação

Coordenação de Pós-Graduação Stricto Sensu

UEG CÂMPUS CORA CORALINA

Av. Dr. Deusdeth Ferreira de Moura Centro - GOIÁS CEP: 76600000

Telefones: (62)3936-2161 / 3371-4971 Fax: (62) 3936-2160 CNPJ: 01.112.580/0001-71

ATA DE EXAME DE DEFESA 19/2023

Aos dezanove dias do mês de maio de dois mil e vinte e três às catorze horas, realizou-se o Exame de Defesa da dissertação do(a) mestrando(a) Jéssica Munique Marques Borges, intitulado “**ITINERÁRIO DE LEITURA DE CONTOS LITERÁRIOS: UM ESTUDO LEXICAL POR INTERMÉDIO DA LEXICOLOGIA E LINGUÍSTICA DE CÓRPUS**”. A banca examinadora foi composta pelos seguintes professores: Dr. Eduardo Batista da Silva – Presidente – (POSLLI/UEG), Dra. Talita Serpa (Unesp/Ibilce), Dra. Marília Silva Vieira (POSLLI/UEG). Os membros da banca fizeram suas observações e sugestões, as quais deverão ser consideradas pelo(a) mestrando(a) e seu/sua orientador(a). Em seguida, a banca examinadora reuniu-se para proceder a avaliação do exame de defesa. Reaberta a sessão, o(a) presidente da banca examinadora, proclamou o resultado, segundo o qual a dissertação foi (x) aprovada, () aprovada com ressalvas, () reprovada com as seguintes exigências (se houver):

Cumpridas as formalidades de pauta, às 16 h, a presidência da mesa encerrou esta sessão do Exame de Defesa e lavrou a presente ata que, após lida e aprovada, será assinada pelos membros da banca examinadora.

Goiás-GO, 19 de maio de 2023.



Prof. Dr. Eduardo Batista da Silva (POSLLI/UEG)



Profa. Dra. Talita Serpa (Unesp/Ibilce)



Profa. Dra. Marília Silva Vieira (POSLLI/UEG)

AGRADECIMENTOS

Em primeiro lugar, agradeço a Deus por ter me concedido discernimento e sabedoria ao longo da minha jornada. Reconheço humildemente que, sem a orientação divina, não estaria aqui, alcançando este importante marco em minha vida. Sua presença constante tem me fortalecido para perseverar e me inspirado a buscar o melhor de mim.

Expresso aqui minha profunda gratidão ao meu orientador, Prof. Dr. Eduardo Batista da Silva, que desde a graduação me guiou de maneira excepcional, sendo o grande motivador para que eu ingressasse no mestrado. Obrigado por acreditar em meu potencial, mesmo quando eu mesma duvidava.

Agradeço à minha família, em especial à minha mãe, Valdirene Marques, meu pai, Júnior Cezar, e minha irmã, Esteice Munique. Durante toda essa caminhada, vocês estiveram diariamente ao meu lado, acreditando, ajudando e incentivando a realizar mais esta conquista.

Agradeço ao meu namorado, Matheus Pires, pelo apoio, amor e paciência. Obrigada por ouvir meus desabafos e me amparar nos momentos mais difíceis. Sou grata pela sua presença em minha vida, e por cada instante compartilhado juntos.

Às minhas amigas pessoais, Franciele Souza, Katharina Elias e Pamela Cabral, gostaria de expressar minha profunda gratidão por sempre torcerem por mim. Guardo com carinho cada palavra de incentivo e carinho que recebi de vocês.

Agradeço também à minha colega de turma, Cássia Beatriz, com quem pude compartilhar as diversas emoções que a jornada como mestranda proporcionou, e que se tornou uma amiga especial através do POSLLI.

Por fim, expresso minha gratidão à Universidade Estadual de Goiás, instituição que me proporcionou a oportunidade de cursar a graduação e, agora, a pós-graduação. Sou grata por todo o conhecimento adquirido, pelas experiências

enriquecedoras e pelas oportunidades de crescimento acadêmico e profissional que essa instituição me ofereceu. Agradeço a todos os professores, funcionários e colaboradores que contribuíram para a minha formação ao longo desses anos.

*“Todos os dias quando acordo
Não tenho mais o tempo que passou
Mas tenho muito tempo
Temos todo tempo do mundo.”*

(Legião Urbana)

RESUMO

Considerando que a prática de leitura pode promover o desenvolvimento das competências linguísticas e culturais, desenvolvemos uma análise linguístico-computacional em um corpus composto exclusivamente por contos literários brasileiros. Nosso objetivo geral é propor a criação de um Itinerário de Leitura de Contos Literários em função de cálculos de métricas léxico-sintáticas. Nossos objetivos específicos são: 1) Construir um corpus representativo de contos literários; 2) Identificar o perfil léxico-sintático dos contos literários; 3) Organizar as obras com base no nível de complexidade linguística 4) Aplicar um questionário de percepção com alunos da segunda fase do Ensino Fundamental. Para a fundamentação teórico-metodológica, recorreremos à Lexicologia (ANTUNES, 2018; BIDERMAN, 1998, 2001; COSSON, 2014), à Linguística de Corpus (BERBER SARDINHA, 2004, 2012; BIBER; REPPEN, 2015; HUNSTON (2022) e à Interculturalidade (BARBOSA, 2009, SARMENTO, 2004) Constituímos um corpus com 878 contos literários brasileiros de 78 autores e 1.556.749 de palavras. O processamento foi realizado pela plataforma NILC-Matrix (NILC, 2021). Por fim, um questionário de múltipla escolha foi aplicado a alunos do Ensino Fundamental II. Nossa classificação do nível de complexidade, realizada a partir dos dados do software, foi a seguinte: "muito fácil" (5 contos); "fácil" (211 contos); "moderado" (484 contos); "difícil" (160 contos) e "muito difícil" (18 contos). Além disso, a classificação atribuída aos contos pelos alunos, obtidas por meio de questionário, mostrou-se semelhante, variando em média 4 pontos em relação aos resultados do software. Tal classificação empírica evidencia o potencial do nosso itinerário como uma ferramenta auxiliar para os professores na seleção de contos adequados ao nível dos alunos, tornando-se uma referência pedagógica para as aulas de língua portuguesa. Dessa forma, promove-se uma dupla intervenção: a prática de leitura e a exposição a contos literários variados em termos léxico-sintáticos.

Palavras-chave: contos; NILC-Matrix; métricas léxico-sintáticas; Linguística de Corpus.

ABSTRACT

Considering that reading practice can promote the development of linguistic and cultural competencies, we developed a linguistic-computational analysis on a corpus composed exclusively of Brazilian literary short stories. Our overall objective is to propose the creation of a Literary Short Stories Reading Itinerary based on lexical-syntactic metric calculations. Our specific objectives are: 1) To construct a representative corpus of literary short stories; 2) To identify the lexical-syntactic profile of the literary short stories; 3) To organize the works based on the level of linguistic complexity; 4) To apply a perception questionnaire to students in the second phase of Elementary School. For the theoretical-methodological foundation, we will use Lexicology (ANTUNES, 2018; BIDERMAN, 1998, 2001; COSSON, 2014), Corpus Linguistics (BERBER SARDINHA, 2004, 2012; BIBER; REPPEN, 2015; HUNSTON (2022), and Interculturality (BARBOSA, 2009, SARMENTO, 2004). We constructed a corpus with 878 Brazilian literary short stories by 78 authors and 1,556,749 words. The processing was carried out by the NILC-Metrix platform (NILC, 2021). Finally, a multiple-choice questionnaire was applied to students in the second phase of Elementary School. Our classification of the level of complexity, based on the software data, was as follows: "very easy" (5 short stories); "easy" (211 short stories); "moderate" (484 short stories); "difficult" (160 short stories); and "very difficult" (18 short stories). Furthermore, the classification attributed to the short stories by the students, obtained through the questionnaire, was similar, varying on average 4 points in relation to the software results. This empirical classification highlights the potential of our itinerary as an auxiliary tool for teachers in selecting short stories appropriate to the students' level, becoming a pedagogical reference for Portuguese language classes. In this way, a double intervention is promoted: reading practice and exposure to varied literary short stories in lexical-syntactic terms.

Keywords: short stories; NILC-Metrix; lexical-syntactic metrics; Corpus Linguistics.

LISTA DE ILUSTRAÇÕES

Figura 1 – Principais nomenclaturas da Lexicologia	32
Figura 2 – A BNCC	52
Figura 3 – Armazenamento dos textos	54
Figura 4 – Extrator NILC-Metrix	56
Figura 5 – Formulário de percepção de complexidade	58
Figura 6 – Fluxograma com os ciclos da pesquisa	59
Figura 7 – Resultados das métricas do NILC-Metrix	61
Figura 8 – Planilha com cálculos linguístico-estatístico	63
Figura 9 – Boxplot com a distribuição dos contos	74

LISTA DE TABELAS

Tabela 1 - Características do nosso cópús de estudo	54
Tabela 2 - Distribuição dos contos	75
Tabela 3 - Resultados dos questionários	82

LISTA DE QUADROS

Quadro 1 – Tamanho do córpus	42
Quadro 2 – Classificação de tipos de córpus	42
Quadro 3 – Informações sobre as métricas selecionadas	61
Quadro 4 – Habilidades de formação literária - BNCC	66
Quadro 5 – Habilidades do elemento cultural – BNCC	69
Quadro 6 – Contos categorizados como muito fáceis	77
Quadro 7 – Contos categorizados como fáceis	78
Quadro 8 – Contos categorizados como moderados	79
Quadro 9 – Contos categorizados como difíceis	79
Quadro 10 – Contos categorizados como muito difíceis	81

LISTA DE ABREVIATURAS E SIGLAS

BNCC	Base Nacional Comum Curricular
CNE	Conselho Nacional de Educação
CORPOP	Cópus Popular Escrito do Brasil
CORPOPEX	Cópus Popular Escrito do Brasil Extra
DCL	<i>Detailed Consistency List</i>
LC	Linguística de Cópus
NILC	Núcleo Interinstitucional de Linguística Computacional
PDF	<i>Portable Document Format</i>
RLB	Retratos da Leitura no Brasil
TXT	Formato de Arquivo de Texto

SUMÁRIO

1. INTRODUÇÃO	16
2. FUNDAMENTAÇÃO TEÓRICA	23
2.1. Literatura	23
2.2. Lexicologia	29
2.3. Linguística de Córpus	40
2.3.1. <i>Conceituação</i>	<i>40</i>
2.3.2. <i>Linguística de Córpus e Ensino</i>	<i>45</i>
2.4. Lexicultura	48
3. MATERIAL E MÉTODO	51
3.1. Material	51
3.1.1. <i>Base Nacional Comum Curricular</i>	<i>51</i>
3.1.2. <i>O córpus de contos literários brasileiros</i>	<i>53</i>
3.1.3. <i>NILC-Metrix</i>	<i>55</i>
3.1.3.1. <i>Seleção das métricas</i>	<i>57</i>
3.1.4. <i>Questionário de avaliação</i>	<i>57</i>
3.2. Método	58
4. RESULTADOS E ANÁLISE DOS DADOS	64
4.1. <i>Análise da Base Nacional Comum Curricular</i>	<i>64</i>
4.2. <i>Dados fornecidos pelo NILC-Metrix</i>	<i>74</i>
4.3. <i>Análise do questionário</i>	<i>81</i>
5. CONSIDERAÇÕES FINAIS	89
REFERÊNCIAS	92
APÊNDICE A – Categorizações dos contos literários	96
APÊNDICE B – Dados fornecidos pelo NILC-Metrix	119

INTRODUÇÃO

A literatura pode nos transportar para outros espaços, proporcionando experiências de enriquecimento linguístico e cultural. Exploramos nessa dissertação um corpus composto exclusivamente por contos literários da Literatura Brasileira.

Apesar da importância dessa manifestação artística, fundamental na construção do homem enquanto sujeito, a literatura encontra-se desvalorizada, com baixos índices de leitura.

Ao consultar a pesquisa Retratos da Leitura no Brasil (RLB) (FAILLA, 2020), realizada em âmbito nacional com o objetivo de avaliar o comportamento e hábitos de leitura dos brasileiros, percebemos que o índice de leitores caiu de 56% para 52% ao correlacionarmos os dados dos anos de 2016 e de 2020. Os dados ficam ainda mais preocupantes quando se observa o índice de penetração de livros lidos nos últimos 3 meses apresentados por faixas etárias, 62% das crianças de 11 a 13 anos (correspondente à fase inicial do Ensino Fundamental 2) leram ao menos 2,47 livros. Os índices pioram quando chegamos à faixa etária de 14 a 17 anos (correspondente à fase final do Ensino Fundamental 2 e início do Ensino Médio) com apenas 41%, correspondendo a apenas 1,29 livros – o que indica uma queda com o passar dos anos. Chama a atenção que 59% dos adolescentes não leram sequer um livro nos três meses anteriores.

Segundo Failla (2020), o gênero literário conto é a preferência entre os estudantes. Em se tratando de faixa de escolaridade, o conto é o segundo gênero mais lido, ficando atrás apenas da Bíblia. Os dados por faixa etária também trazem o gênero como um dos preferido do público jovem, compreendido na faixa de 11 a 17 anos.

Frente ao exposto, com o intuito de incentivar a prática de leitura e a valorização da literatura, esse trabalho se propõe a analisar o perfil léxico-sintático de um corpus composto exclusivamente por contos literários da Literatura Brasileira com a finalidade de propor, por meio de cálculos computacionais, o Itinerário de Leitura de Contos Literários, organizado em cinco níveis de complexidade textual.

Nosso trabalho impacta o ensino educacional, uma vez que, a leitura constitui-se uma prática fundamental para o crescimento e desenvolvimento pessoal, intelectual e profissional. Dentro desse contexto, promovemos a literatura no âmbito educacional, uma prática fundamental. A falta de repertório e o pouco/nenhum contato com o

gênero aqui explorado possivelmente corroboram para o surgimento de dificuldades manifestadas ao longo do cotidiano escolar. Frente ao exposto, a presente pesquisa origina-se a partir da necessidade em se refletir acerca do contexto escolar, em especial no que diz respeito às práticas de leitura e aquisição de repertório linguístico.

Considerando que a leitura é essencial na vida de qualquer indivíduo, torná-la um hábito é fundamental para o crescimento pessoal, intelectual e profissional. Creemos que é nosso papel, seja como pesquisadores ou professores, explorar sugestões e métodos que possam contribuir para a educação. Acreditamos que se faz necessário repensar o trabalho com o texto literário e sua seleção em sala de aula, considerando aspectos lexicais e sintáticos para a seleção dos textos a serem trabalhados.

O texto literário pode ser a base para ressignificarmos a leitura e sanarmos dificuldades. Sobre gêneros textuais, Marcuschi (2008) refere-se a todos os textos que rodeiam o nosso cotidiano, uma vez que não há comunicação que não seja manifestada verbalmente através de algum gênero cometido por textos. Devido a isso, o estudo dos gêneros textuais se configuram como “uma fértil área interdisciplinar, com atenção especial para o funcionamento da língua e para as atividades culturais e sociais” (MARCUSCHI, 2008, p. 155-156). Tendo isso em vista, tal afirmação nos leva ao encontro da Literatura, a qual cremos ser a unificação do que o referido autor explica. De acordo com Araújo (2015), o conto é um dos gêneros textuais literários mais populares da Literatura, sendo importante incluí-lo em sala de aula por ser baseado em situações cotidianas, práticas sociais e eventos da história humana, além de retratar experiências e acontecimentos que atribuem a literatura uma função social. No entanto, não basta levar o conto literário para a escola, faz-se necessária uma seleção de textos que estejam compatíveis com o nível de domínio/conhecimento linguístico do aluno.

Vale salientar que trabalhos anteriores já trataram da questão da literatura na escola (COSSON, 2014; COLLIE; SLATER, 1992; FRANÇA, 2022) em uma perspectiva de formação do cidadão.

No que se refere ao uso da literatura na sala de aula, Cosson (2014) diz que a relação entre literatura e educação está distante de ser pacífica. Segundo o autor, para muitos professores e pesquisadores da área de Letras, a literatura só se mantém presente devido à tradição e à estagnação curricular. As multimodalidades, principalmente quando se trata de redes sociais, constituem um desafio para a

capacidade de concentração dos alunos em leituras mais extensas, como contos, por exemplo. Essa é uma das razões frequentemente citadas para explicar a relutância em incluir literatura no currículo escolar nos dias de hoje. Além disso, há alegações de que a linguagem literária, por ser irregular e criativa, pouco contribuindo para o ensino da língua portuguesa culta, e que a complexidade do vocabulário e sintaxe dos cânones literários acarretam desinteresse por parte dos alunos. No entanto, é importante considerar que essa visão representa apenas uma perspectiva e não é compartilhada por todos. Existem várias razões pelas quais a literatura é valorizada e considerada uma parte essencial do ensino da língua.

Essa generalização simplista com noção conteudista minimiza a literatura apenas como uma disciplina marginal, favorecendo para que não seja ensinada em sua forma essencial, de construir e reconstruir a palavra que nos humaniza. Cremos ser urgente a necessidade de romper essa visão, proporcionando ao aluno muito mais que um conhecimento literário, mas experiências de leitura a serem compartilhadas.

Collie e Slater (1992) ressaltam as principais razões em se trabalhar com a literatura, um valioso e autêntico material que transcende o tempo, abordando questões humanas fundamentais. O enriquecimento cultural também é proporcionado aos aprendizes, dando a oportunidade de se aprofundar compreendendo o modo de vida, costumes de outros locais do mundo. O enriquecimento da linguagem também é um benefício vindo da literatura, fornecendo um contexto rico em que elementos léxicos, sintáticos, morfológicos, semânticos e até fonológicos, são contextualizados e aos poucos ganham familiaridade, ampliando vocabulário e obtendo conhecimento. Por fim, o envolvimento pessoal, normalmente o processo de aprendizagem se caracteriza como analítico e fragmentado, já com a Literatura, pode-se tornar mais prazeroso, a partir da mudança de foco e o envolvimento do leitor com o texto, gerando motivação no processo de aprendizagem da língua.

França (2022) defende o ensino da literatura na escola pública, constituindo-se como um direito inalienável dos alunos, que muitas vezes, pertencem às classes socioeconômicas vulneráveis. A necessidade em se debater tal tema reside na discrepância entre as escolas públicas e privadas frente à literatura. Nas instituições mais abastadas, a literatura ainda é vista como uma disciplina, possuindo um espaço de respeito. No entanto, nas escolas públicas, a literatura tem sido cada vez mais ausente. Tal realidade alimenta ainda mais os debates sobre o acesso à literatura juntamente com todos os elementos de exterioridade histórico-sócio-ideológica, que

envolvem a materialização de sujeitos e efeitos de sentidos, seja na composição de textos ou na meditação sobre as práticas de ensino-aprendizagem.

De forma subsidiária, outros trabalhos demonstram a importância em se desenvolver o léxico no contexto educacional (ABBADE, 2011; ASSUNÇÃO, 2015; ANTUNES, 2018; SERRA, 2019; SILVA; SILVA, 2021).

Na seara dos estudos lexicais, Abbade (2011) relata que, apesar da pouca ênfase no estudo lexical, historicamente, já se instituiu como importante e necessário. Compreender o léxico de uma língua é também considerar aspectos históricos e culturais do povo que a utiliza, sendo esses três elementos, língua, história e cultura indissociáveis. Ademais, a ciência do léxico, a Lexicologia, nos instrui a como abordar a palavra, considerando o seu contexto e a polissemia que a mesma carrega consigo dentro dos textos, contribuindo para o domínio lexical e ampliação do vocabulário, incentivos aos quais são fundamentais no ensino escolar.

Assunção (2015) propôs a criação de um glossário criativo, ao perceber a grande dificuldade de compreensão de palavras consideradas difíceis por seus alunos. Esses vocábulos comprometiam o desenvolvimento de práticas de ensino por desconhecerem seus significados e contextos, uma vez que não faziam parte de seu vocabulário ativo. Assim, utilizando um instrumento já conhecido no meio escolar, o dicionário, alunos pesquisavam o significado e transcreviam a sua compreensão acerca da palavra, ampliando seu acervo vocabular. Segundo a autora, tal proposta não se trata de uma receita pronta, mas já auxilia e contribui para um estímulo a prática pedagógica reflexiva sobre o ensino do léxico.

Antunes (2018) ressalta que a hegemonia dada a gramática ocorre devido a crença em acreditar que obter domínio sob a mesma é suficiente para alcançar um bom desempenho comunicativo. No entanto, o conhecimento linguístico vai além de estudá-la, na verdade, é resultado da prática constante, continuada, reflexiva, atenta, crítica e integral, o que considera a complexidade da língua, ou seja, também abarca o léxico.

Estudos como o de Serra (2019) criticam a conduta de escolas da educação básica do Brasil que, em pleno século XXI, continuam se baseando em uma coleção de regras e repetições para construir seus métodos de ensino, priorizando o controle da atenção dos alunos e marginalizando o desenvolvimento de habilidades linguísticas. Segundo o autor, ações como essas contribuem para o cenário de

insucesso que presenciamos atualmente, comprovados através de provas de amplitude nacional.

Silva e Silva (2021) propõem a prática de vocabulário a partir de excertos de contos literários. Essa intervenção foi realizada com alunos de uma escola pública da rede estadual no município goiano de Itumbiara, com o intuito de atestar o uso de obras literárias para a ampliação de vocabulário em sala de aula. Tal ação rendeu bons resultados, com um ganho lexical de 24,1 pontos percentuais entre as meninas e de 15,6 entre os meninos – demonstrando que o estímulo à leitura, apesar de se tratar de trechos de obras, traz ganhos significativos em contexto pedagógico.

Entendemos que o texto configura-se como um espaço privilegiado para desenvolver habilidades linguístico-culturais. Segundo Antunes (2018), o texto não é uma opção, e sim o único espaço capaz de demonstrar a interação verbal. As palavras passam a ganhar sentido somente quando encaixadas dentro de um enredo.

Conforme o exposto, existe uma necessidade de se repensar as práticas perpetuadas em contexto escolar. Acreditamos que ainda não houve atenção dedicada sobre quais textos são ofertados aos alunos, e principalmente, se a léxico-gramática dos textos é compatível com o nível do alunado. Para mudar tal contexto, cremos que o Itinerário será substancial no auxílio da triagem de textos que considerem esses critérios.

Esse trabalho apresenta uma proposta que pode auxiliar profissionais da educação, em especial de língua portuguesa, buscando ressignificar o trabalho com textos literários no cenário escolar, evidenciando a importância em se alinhar o nível léxico-sintático do texto selecionado ao nível de vocabulário já dominado pelos alunos. Argumentamos a favor de um ensino que considera a leitura como uma prática basilar para o desenvolvimento do alunado, pautando-se diretamente na ampliação do repertório linguístico e formação social, profissional, pessoal e acadêmica do ser.

Esse trabalho baseia-se na Lexicologia (ANTUNES, 2018; BIDERMAN, 1998, 2001; COSSON, 2014), na Linguística de Córpus (BERBER SARDINHA, 2004, 2012; BIBER; REPPEN, 2015; HUNSTON (2022) e na Lexicultura (BARBOSA, 2009, SARMENTO, 2004).

A Lexicologia é uma ciência que se ocupa do léxico de uma língua, e possui como objeto de estudo a palavra. A ciência lexicológica auxilia a formular novas estratégias de ensino que considerem o léxico, contribuindo fundamentalmente para o ensino de línguas.

A Linguística de Córpus (LC) é a área de pesquisa que se ocupa da coleta e análise criteriosa de córpus, devendo ser mantidos em formato eletrônico a fim de servirem como objeto de pesquisa linguística. Junto à tecnologia, a LC se apresenta como uma nova abordagem através de dados estatísticos empíricos, auxiliando compreender a língua como um fenômeno social, contribuindo para o ensino de línguas.

A Lexicultura é o estudo do léxico culturalmente marcado. Assim, compreendendo que a língua é a herança de um povo, sendo constituída por um léxico inacabado, uma vez que está em constante transformação. Essa área nos fornecerá subsídios teóricos para identificar as transformações sofridas pelo léxico na sociedade e também literatura.

Tendo as três linhas de pesquisa como orientação, nosso objetivo geral é propor a criação de um Itinerário de Leitura De Contos Literários em função de cálculos de métricas léxico-sintáticas. Nossos objetivos específicos são: 1) Construir um córpus representativo de contos literários brasileiros; 2) Identificar o perfil léxico- sintático dos contos literários; 3) Organizar as obras com base no nível de complexidade linguística; 4) Aplicar um questionário de percepção acerca do nível de dificuldade dos contos com alunos da segunda fase do Ensino Fundamental.

Pensando em perguntas que irão nortear nossa pesquisa, traçamos as seguintes questões:

- 1) Qual a abordagem dedicada a tríade contos, léxico e interculturalidade, no principal documento em vigência atualmente, a Base Nacional Comum Curricular (BNCC)?
- 2) No itinerário literário, qual a distribuição dos contos em função da complexidade léxico-sintática (muito fácil, fácil, moderado, difícil e muito difícil)?
- 3) Existe correlação entre as parciais do NILC-Matrix e os apontamentos dos alunos obtidos através do questionário?

Os materiais a serem utilizados nessa pesquisa são os seguintes: Base Nacional Comum Curricular (BNCC) (BRASIL, 2018); Córpus de Contos Literários Brasileiros (córpus de estudo) constituído por 878 contos de 78 escritores brasileiros; NILC-Matrix (NILC, 2021) plataforma com 200 métricas capazes de extrair informações objetivas em vários níveis da língua, como léxico, morfologia, sintaxe, semântica, além de um questionário de múltipla escolha que será aplicado a fim de correlacionar os dados.

Nossos procedimentos metodológicos iniciaram a partir da análise da seção de Língua Portuguesa na BNCC, a fim de investigar qual a abordagem do documento em relação a textos literários, em especial o conto literário, a abordagem do léxico, e se há contemplação de elementos interculturais. Em seguida, utilizando a plataforma NILC-Matrix, processamos o corpus de estudo para obtermos os cálculos das métricas e quais serviram como objeto de análise. Por fim, aplicamos um questionário de múltipla escolha com alunos da rede de ensino, a fim de correlacionarmos os dados advindos da LC em confluência com a avaliação humana.

Nossa pesquisa se diferencia das anteriores devido a quatro características: construiremos e processaremos um corpus exclusivamente composto por contos literários brasileiros criteriosamente selecionados; recorreremos à Linguística de Corpus e seus instrumentos de análise léxico-sintáticos para obtermos dados empíricos de complexidade textual de contos; criaremos uma referência didático-pedagógica exclusiva para a leitura de contos literários brasileiros, dividido por níveis de complexidade.

Acreditamos que a presente pesquisa trará contribuições, especialmente para a comunidade escolar, especialmente para professores de língua materna que trabalham diretamente com a leitura e práticas de vocabulário e/ou gramática. Salientamos que qualquer indivíduo que deseje entrar ou manter contato com a literatura poderá recorrer ao itinerário aqui proposto.

A dissertação está estruturada em seis grandes seções: a primeira trata das ideias que serviram para orientar as reflexões linguístico-literárias; material e métodos; resultados e análise dos dados, considerações finais; por último, anexos e apêndices.

FUNDAMENTAÇÃO TEÓRICA

Iniciamos a seção 2.1 com uma contextualização da Literatura. Tendo o conteúdo textual literário como pano de fundo, apresentaremos nessa seção os subsídios teórico-metodológicos de nossa pesquisa que recorrem fundamentalmente às linhas da Lexicologia, Linguística de Corpus (LC) e a Lexicultura.

A respeito da Lexicologia, iniciamos com a conceituação de seu objeto de estudo, o léxico, citando conceitos e contribuições dessa ciência no ensino de línguas. Para a Linguística de Corpus, optamos em elencar discussões teóricas por se tratar de uma nova maneira de analisar e compreender o funcionamento da língua. Para a Lexicultura, procedemos sua conceituação na última seção.

2.1 Literatura

Em se tratando do contexto educacional, a literatura é de suma importância, devido a proporcionar um vasto cenário de inúmeras possibilidades de aprendizagem a partir do texto. Mas apesar desse potencial, o cenário para a literatura não é favorável, e a mesma vem travando lutas contra as adversidades impostas a ela, para se manter presente.

Zilberman (2009) afirma que desde a criação da instituição escola, a literatura sempre esteve presente. Considerando desde os povos sumérios, inventores da escrita, até o modelo de escola grega dos séculos V e IV a.C., referência que segue até os dias atuais. No entanto, foi somente entre o final da década de 70 e início da década de 80 que os debates sobre a leitura e o papel da literatura se intensificaram. Em meio a essas discussões, pesquisadores das áreas de Letras e Pedagogia, buscavam maneiras de melhorar a escola brasileira e a qualidade do ensino, visto que a perspectiva de piora era alta. Assim, a literatura surge como uma esperança:

No âmbito dessas discussões, que envolvia a aprendizagem e o uso da língua portuguesa, recebeu a literatura uma valorização específica, pois era nela que se colocavam as esperanças de superação dos problemas experimentados na sala de aula. Com efeito, se os diagnósticos identificávamos dificuldades de leitura e expressão escrita por parte dos estudantes, era à literatura, representada por obras de ficção e de poesia, que se transferiam os créditos e as expectativas de mudança e de sucesso quando do exercício da ação educativa por parte dos docentes. (ZILBERMAN, 2009, p. 13).

Contudo, os anos passaram e em meio a inúmeras reformas nominais e estruturais no ensino, além dos adventos tecnológicos que transformaram a sociedade e cultura do novo milênio, os problemas permanecem os mesmos. O sucateamento da escola pública, a desvalorização do professor e o enfraquecimento da aprendizagem dos alunos, são exemplos graves que presenciamos atualmente.

Tudo o que mudou parece ter mudado para melhor – menos a escola, com suas conseqüências: a aprendizagem dos alunos, a situação do professor, as políticas públicas dirigidas à educação, para não se mencionarem as condições de trabalho, onde predomina a insegurança, e o espaço físico das salas de aula, degradado e degradante. Onde deveria reinar a mesma euforia, predominam a desolação, o desestímulo, os sentimentos de decepção e de fracasso. (ZILBERMAN, 2009, p. 14).

Se antes a literatura era concebida como método de transmissão de padrões linguísticos e de patrimônio cultural, hoje a mesma também é responsável pela formação do leitor. Logo, a literatura se associa a prática de leitura, legitimando ambas. A prática de leitura pautada na literatura trata-se de uma atividade integral, uma vez que proporciona ao sujeito a interação com a linguagem verbal, ampliando seu conhecimento linguístico, e ao mesmo tempo estimula o imaginário, proporcionando através da ficção, a reflexão crítica sobre temas que ocorrem no meio em que está situado.

Sobre textos literários, Candido (2004) destaca que a Literatura é uma necessidade do ser humano e seu acesso é um direito. Estar em contato com textos literários possibilita termos acesso a outras realidades, que apesar de serem ficção, nos fazem refletir a respeito da qual vivemos atribuindo as nossas concepções de mundo, nos fazendo evoluir como seres humanos, sociais e críticos. Além disso, Antunes (2012) destaca que o texto literário é um espaço em que a linguagem não está restrita a nenhum de seus padrões convencionais, permitindo não apenas o contato com funções referenciais, mas também com funções emotivas e poéticas. Nesse sentido, não se pode ignorar a importância do léxico na criação literária.

Entretanto, apesar dos ganhos notáveis que a prática da leitura juntos aos textos literários podem proporcionar, a realidade prática nas escolas é outra. Zilberman (2009) afirma que a prática de leitura é vista como inusitada, em virtude de os alunos não serem adeptos a esse hábito, por sua vez, a literatura é encarada como um alienígena. Esse ponto de vista mencionado pela autora, é complementado com

uma crítica as escolas que não cedem e promovem momentos de leitura que visem o contato integral com o texto literário.

Somado a isto, chegamos aos documentos curriculares vigentes na educação. Tais escrituras são elaboradas contendo orientações e conteúdos para serem referências aos professores na preparação e execução de suas aulas. No entanto, se analisarmos como a literatura é abordada, encontraremos algumas explicações que justificam o contexto em que presenciamos. De acordo com França (2022), a crescente desvalorização da literatura, a fez perder o status de disciplina, desaparecendo até como apêndice dentro das aulas de língua portuguesa. Cosson (2014) também aborda essa distinção da literatura e relata a subordinação imposta a ela. Em uma abordagem conteúdista presente nos programas curriculares, com aulas usualmente uma vez na semana ou no fim do semestre letivo, abandonam a leitura do texto integral e priorizam o ensino de informações de autores, obras e características. Essa abordagem conteúdista subestima a literatura, e desconsidera os benefícios que a leitura pode atribuir, sobre isso os autores Dias e Menezes complementam:

A literatura ajuda o aluno a compreender a si mesmo, a sua comunidade e o seu mundo. Todavia, na escola, o ensino da literatura continua reduzido, em geral, à contextualização histórica, caracterização da obra de acordo com o período literário e seu respectivo autor. O aluno não consegue perceber a plurissignificação do texto literário, os possíveis diálogos entre as obras. A excessiva preocupação com a identificação de características estéticas das escolas literárias, assim como a fragmentação do texto literário nos livros didáticos somado ao despreparo por parte de alguns professores, que veem o livro didático como objeto de salvação, que deve ser seguido e idolatrado, sufoca a leitura por prazer. (DIAS; MENEZES, 2014, p. 129).

Tendo em vista que esse ensino mecanizado está difundido nas escolas, não é surpresa presenciarmos alunos que não são simpatizantes da leitura e muito menos da literatura. Tal realidade é preocupante, considerando o quanto ambas são fundamentais para o desenvolvimento de habilidades, como um bom raciocínio e o ganho de vocabulário, além de ampliar o conhecimento e aprimorar a escrita. Além disso, consideramos a leitura de literatura fundamental para a formação do sujeito crítico, uma vez que, incita a interpretar a realidade e tecer opiniões que contribuam para o meio em que se inserem.

Assim, abandonar metodologias ineficazes é o primeiro passo para que possamos discutir novos meios de se estimular o apreço pela prática da leitura e

literatura. Nesse contexto, o professor será a peça fundamental para essa transformação, uma vez que será ele o condutor deste percurso em ressignificar o papel da leitura e da literatura, através de estratégias que demonstrem aos alunos que ler é muito mais que a tarefa chata de decodificar palavras, e sim uma atividade prazerosa e significativa capaz de atribuir sentidos.

Para isso, é interessante promover a interação entre leitor e texto, principalmente com composições que retratem realidades parecidas com a que o leitor vive gerando o processo de identificação. De acordo com Zafalon (2010), se o leitor se familiariza com o que está sendo narrado, e se reconhece na cultura do texto lido, haverá chances maiores desse sujeito continuar a leitura. Assim, para despertar esse primeiro interesse sobre a literatura, o professor deve ir muito além do domínio de conteúdo, ele precisa conhecer seus alunos, considerar sua realidade e seu conhecimento prévio, para selecionar textos que corresponda o desejo deles, que atraia a atenção e instiguem a lerem.

A partir do momento em que o professor começa a oferecer aos alunos a oportunidade de fazer leituras de textos e obras realmente significativos do ponto de vista de suas aspirações e conhecimentos prévios, pode-se então planejar alçar vôos mais altos, ou seja, o professor, paulatinamente, introduzirá uma literatura que seja mais aprofundada e abrangente, que desperte prazer, sem prescindir de um objetivo prático imediato. (ZAFALON, 2010, p. 7).

Além disso, a autora complementa sobre as características estruturais do texto que influenciam no andamento da leitura. Para que os alunos se sintam estimulados, o texto não pode ser cansativo para que o leitor não se sinta entediado. Textos muito extensos e letras miúdas também cansam e desmotivam o leitor. A respeito do nível de linguagem, a autora se refere diretamente aos alunos do ensino médio, no entanto acreditamos que seja válido para todos os níveis de ensino, onde o vocabulário precisa ser acessível e coerente ao nível do alunado. Dito isso, concordamos com Zafalon, quando aconselha que a princípio sejam trabalhados textos mais contemporâneos, uma vez que a linguagem é mais atual, deixando os cânones para quando alcançarem a maturidade literária.

[...] o texto literário deve ser colocado em evidência nas aulas de literatura, promovendo um espaço no qual sejam realizadas diferentes leituras e construções de sentido, motivando o desenvolvimento das atividades leitoras dos alunos. As estratégias utilizadas devem fugir ao engessamento

das práticas tradicionais de ensino da literatura, valorizando o papel do professor no processo ensino-aprendizagem, selecionando obras que sejam condizentes com a faixa etária dos alunos, abordando temas atuais e que se adequam aos interesses de todos os envolvidos durante o desenvolvimento das aulas. Sobretudo, o ensino da literatura deve primar pela tentativa de compreender melhor e respeitar a literatura em sua historicidade, não desprestigiando as tradições culturais e linguísticas, mas priorizando suas particularidades, e o diálogo constante da literatura com outras artes e linguagens, sem perder de vista o objeto central – o texto literário – e a sua realização em contato com a recepção do leitor. (ZAFALON, 2010, p. 16-17).

Dessarte, pensamos que o gênero conto literário atenda todas essas condições, sendo uma escolha interessante para cativar o interesse dos alunos. A arte de contar é uma prática cultural intrínseca do ser humano, presente desde os primórdios antes do advento da escrita, o conto surgiu como uma tradição oral sendo uma das expressões mais antiga da literatura, que mais tarde se adaptou à modalidade escrita.

De acordo com o Dicionário Houaiss (HOUAISS, 2009), o conto pode ser definido como "narrativa breve e concisa, contendo um só conflito, uma única ação (com espaço geralmente limitado a um ambiente), unidade de tempo, e número restrito de personagens". Ao definir conto, costuma-se fazer comparações com outro gênero literário, o romance. Assim com o objetivo de diferenciá-los, Magalhães Junior (1972, p. 10-11) comenta:

O conto é uma narrativa linear, que não se aprofunda no estudo da psicologia dos personagens nem nas motivações de suas ações. Ao contrário, procura explicar aquela psicologia e essas motivações pela conduta dos próprios personagens. A linha do conto é horizontal: sua brevidade não permitiria que tivesse um sentido menos superficial. Já o romance, em vez de episódico, como o conto, é, ao contrário deste, uma sucessão de episódios, interligados. E exige do autor tratamento diverso, quer na apresentação dos acontecimentos, quer no estudo dos personagens. O romance explora-os em sentido vertical, com uma profundidade a que o conto não pode aspirar. Outra distinção, em que insistem alguns críticos e ensaístas literários, é a de que o conto geralmente narra um acontecimento pretérito, ao passo que o romance história um acontecimento ou série de acontecimentos no tempo presente, à medida que estes se desenrolam.

Desse modo, considerando que o conto se trata em suma, de uma narrativa curta e linear, não requer muito esforço do leitor, sendo um ótimo aliado para a iniciação da leitura através do contato com o texto integral. Ademais, sua extensão

também corrobora para o contexto de sala de aula, sendo facilmente trabalhado e explorado dentro do tempo de duração de uma aula de língua portuguesa, que geralmente possui 50 minutos. Para mais, o gênero conto possui várias categorias como fantástico, terror, policial, misterioso, entre outros, proporcionando ao professor uma gama de temas que podem ser levados para a sala de aula a fim de serem lidos, interpretados e debatidos conforme o interesse e perfil dos alunos.

Posto isso, reconhecendo a relevância em se debater sobre o papel da literatura e rever o seu espaço no contexto escolar e nos documentos curriculares, essa pesquisa objetiva a ressignificar a literatura, em conjunto à leitura e aquisição de vocabulário, que consideramos ser suplementares. A convivência com a literatura de forma integral proporciona ao sujeito o contato com uma pluralidade de textos, propiciando gradativamente o gosto pela leitura, o ganho de conhecimentos e conseqüentemente a ampliação do acervo vocabular, o que inibe algumas dificuldades de leitura, escrita e interpretação manifestadas pelos alunos no cotidiano escolar.

No entanto, para que essa atividade pautada na literatura seja efetiva, concordamos com Zafalon (2010) ao mencionar a importância em selecionar textos que sejam compatíveis com o nível e interesse da turma. Logo, propomos nessa dissertação a criação de um roteiro de leituras, intitulado Itinerário de Leitura de Contos Literários, o qual será composto por um cópús exclusivamente composto por contos literários brasileiros, dividido em cinco níveis de complexidade vocabular. Tal instrumento pedagógico poderá ser consultado por professores que desejam complementar suas aulas, selecionando textos que sejam compatíveis com o nível de conhecimento vocabular dos alunos, visando sempre o aprimoramento dessas habilidades.

Em suma, consideramos de extrema importância que mais trabalhos sejam realizados visando debater sobre o ensino e o espaço cedido a literatura. Pois é dessa maneira, abordando o tema, que seremos capazes de reivindicar mudanças. Não podemos ser cúmplices desse apagamento da arte literária nas escolas, uma vez que sua ausência provoca inúmeros danos a sociedade. Assim, como dizia Antônio Cândido (CANDIDO, 2004), o acesso a literatura é um direito, e nós como pesquisadores e professores precisamos ser a resistência para manter esse direito acessível aos nossos.

2.2 Lexicologia

A língua é um sistema constituído por signos, que transmitem o conhecimento de tudo o que existe ao nosso redor. Antunes (2018) define o léxico em uma visão geral, como o amplo repertório de palavras de uma dada língua ou o conjunto de itens à disposição dos falantes para amparar as suas necessidades comunicativas. O léxico é fundamental, não existe língua sem léxico, uma vez que o mesmo, constitui as "matrizes cognitivas" com as quais operamos a criação e expressão de sentidos. Ou seja, partimos do léxico para atribuímos significados às coisas, ao mundo e a nós mesmos.

Segundo Biderman (2001), o léxico de uma língua natural se constitui como uma maneira de registrar o conhecimento do universo. Esse pensamento é complementado, ao descrever os processos de nomeação e categorização dos seres, objetos, etc., estratégia que foi desenvolvida pelo homem para materializar a sua experiência através da cognição, ou compreensão da realidade que o cerca. Esse modo de associar palavras a conceitos, que representam os referentes, consolida-se em signos linguísticos: as palavras.

É esse processo de nomeação que gerou e gera o léxico das línguas naturais. Por outro lado, podemos afirmar que, ao nomear, o indivíduo se apropria do real como simbolicamente sugere o relato da criação do mundo na bíblia judaico-cristã, anteriormente referido, em que Deus incumbiu ao primeiro homem dar nome a toda a criação e dominá-la. A geração do léxico se processou e se processa através de atos sucessivos de cognição da realidade e de categorização da experiência, cristalizada em signos linguísticos: as palavras. (BIDERMAN, 1998, p. 92).

Esse processo de cognição e de apoderamento do conhecimento desenvolvido pelo homem gerou inúmeros sistemas lexicais de línguas naturais (vivas ou mortas). Apesar de, a princípio, se basearem em uma mesma técnica de categorização, as línguas se organizam em sistemas distintos, constituindo-se de arranjos característico de cada cultura. Sobre isso, Vilela (1994, p. 6) complementa:

O léxico é a parte da língua que primeiramente configura a atividade extralinguística e que arquiva o saber linguístico numa comunidade. Avançose recuos civilizacionais, descobertas e inventos, encontros entre povos e culturas, mitos, crenças, afinal quase tudo, antes de passar para a língua e para a cultura dos povos, tem um nome e esse nome faz parte do léxico. O léxico é o repositório do saber linguístico e é ainda a janela

através da qual um povo vê o mundo. Um saber partilhado que apenas existe na consciências falantes duma comunidade.

Podemos observar que o léxico e a Literatura são reflexos de demandas sociais, capazes de registrar através da língua os principais acontecimentos e descobertas de uma comunidade, perpetuando conhecimento ao longo da história com base nos registros linguísticos.

O léxico de uma língua natural, segundo Biderman (2001), se constitui como um patrimônio vocabular de uma dada comunidade linguística. Concebendo-se como uma espécie de tesouro cultural abstrato, transmitido à geração seguinte. Para Antunes (2018), o léxico é visto como uma "memória", a qual, segundo ela, é dinâmica e em constante movimento, que se reformula conforme a sociedade e a cultura se desenvolvem.

Assim, o léxico, um sistema aberto, inacabado, em expansão (pela criação de novas palavras, por palavras advindas de outras línguas, pela ressignificação de palavras já existentes, entre outros), é concebido como uma fonte de informações, permitindo a compreensão da cultura e do meio social.

(...) o Léxico se expande, se altera e, às vezes, se contrai. As mudanças sociais e culturais acarretam alterações nos usos vocabulares; daí resulta que unidades ou setores completo do Léxico podem ser marginalizados, entrar em desuso e vir a desaparecer. Inversamente, porém, podem ser ressuscitados termos que voltam à circulação, geralmente com novas conotações. Enfim, novos vocábulos ou novas significações de vocábulos já existentes, surgem para enriquecer o Léxico. (BIDERMAN, 2001, p. 179).

Essas transformações sofridas pelo léxico podem ser observadas nos escritos literários. Tal discrepância vocabular, inclusive, serve de pauta para discussões a respeito da linguagem literária ser considerada antiquada para os dias atuais. No entanto, cremos que o processo de transformação e expansão do léxico é resultado cultural da evolução do ser humano com o seu meio social, valorizando ainda mais a importância em se compreender e reconhecer tais termos, mesmo que tenham caído em desuso.

Simões e Assis (2014) afirmam que é humanamente impossível um indivíduo falante de determinado idioma reconhecer, memorizar e usufruir de todo o léxico de uma língua. Os autores consideram que o termo léxico pode caracterizar duas situações: o léxico de uma estabelecida língua, caracterizando-se como aberto e infinito; e o léxico do indivíduo falante, mais conhecido como vocabulário individual,

repertório de palavras adquiridas em meio as interações sociais desse sujeito. Assim, o vocabulário se caracteriza como um recorte do léxico. Logo no léxico de dada língua existem vários desses recortes, por exemplo: vocabulário ativo, vocabulário passivo, vocabulário fundamental, vocabulário comum, vocabulário individual, vocabulário específico, vocabulário básico, etc.

Recorremos ao nosso acervo vocabular para produzir discursos orais e escritos. Torna-se indispensável a busca por métodos que ampliem tal repertório. A prática de leitura a partir de textos literários, é uma maneira de ampliação vocabular, uma vez que a promoção do contato com as palavras através do texto pode facilitar a familiarização.

Assim, consideramos o conceito de Silva (2006) para definirmos vocabulário:

[...] o vocabulário é um subconjunto do léxico de uma língua. É composto detodas as unidades semânticas, graficamente simples e compostas, e locuções indecomponíveis que se atualizam no discurso, chamadas de vocábulos ou palavras. (SILVA, 2006, p. 16-17).

O vocabulário individual de um falante, assim como o próprio termo nos leva a pensar, ocorre de maneira heterogênea com cada sujeito, pois a construção se dará a partir das influências que sofrerá ao longo da vida. Bourdieu enfatiza que a família desempenha um papel central na transmissão do capital cultural e no desenvolvimento cognitivo e lexical das crianças. Famílias que valorizam a educação, proporcionam um ambiente estimulante e promovem a interação verbal tendem a beneficiar o desenvolvimento da criança. No entanto, as oportunidades e recursos disponíveis na família não são distribuídos igualmente na sociedade, o que contribui para a reprodução das desigualdades sociais. Logo, o papel da família é fundamental nessa fase para proporcionar uma base sólida e um meio favorável que estimule e beneficie essa aprendizagem inicial corroborando para um bom desenvolvimento cognitivo e lexical, a partir da construção do vocabulário ativo e passivo desse sujeito.

Para compreendermos a conceituação de vocabulário ativo e passivo, usufruímos das palavras de Castro (2015), no qual o vocabulário ativo é o acervo que o falante utiliza no seu cotidiano, recorrendo-o para a prática de seus discursos orais e escritos; já o vocabulário passivo, seria o acervo de palavras as quais o sujeito reconhece e entende, mas não é capaz de integra-las as suas práticas discursivas.

Assim, com as definições apresentadas, podemos resumi-las com a seguinte figura 1:

Figura 1 – Principais nomenclaturas da Lexicologia



Fonte: Dados dessa pesquisa.

As áreas que constituem as ciências do léxico são Lexicologia, a qual a grosso modo, ocupa-se do estudo, análise e categorização da palavra, a Lexicografia, ciência da produção de dicionários; e Terminologia, subárea que se ocupa em analisar o léxico específico de áreas de conhecimento. Cada ciência aborda o léxico de maneira distinta, porém ambas possuem objetivo em comum, descrever esse mesmo léxico (BIDERMAN, 2001, p. 15).

Para Abbade (2011), a Lexicologia é a ciência que estuda as múltiplas relações com outros sistemas da língua, e acima de tudo sobre o próprio léxico. Jackson e Amvela (2021) define introdutoriamente a Lexicologia como o estudo do léxico, compreendido como o acervo de palavras de um determinado idioma, ou seja, seu vocabulário ou léxico. Orsi (2012) define a ciência lexicológica como a área que "estuda as unidades lexicais de uma ou várias línguas, seja no que tange o significado ou o significante, isto é, o léxico em todos os seus aspectos.". Sobre a essa ciência, Guerra e Andrade (2012, p. 239) complementam:

A Lexicologia apresenta um entrelaçamento de áreas da linguagem que resulta em uma forma complexa de ver a noção de palavra e o ensino do léxico e do vocabulário. Ela tem muito a contribuir com o ensino de línguas, atingindo diretamente tanto os indivíduos em processo de formação

escolar quanto aqueles que buscam a formação continuada, os educadores de línguas.

Desse modo, a palavra é o objeto central de estudo da lexicologia, sejam as palavras simples em todos os seus aspectos, como também as palavras complexas e compostas, as unidades significativas da linguagem. Como as palavras devem ser analisadas compreendendo sua forma e também seu significado, a lexicologia se fundamenta em estudos decorrentes da morfologia, campo de estudo dos morfemas e seus arranjos na formação de palavras e também da semântica, o estudo que busca explicar e descrever os significados. Um terceiro campo também de interesse das pesquisas lexicológicas é a etimologia, estudo das origens das palavras, no entanto não podemos confundir-la com a lexicografia, técnica para a escrita e confecção de dicionários, uma vez que não se trata de um estudo da linguagem.

De acordo com Jackson e Amvela (2021), além da relação entre Lexicologia, Morfologia, Semântica, Etimologia e Lexicografia, é importante ressaltar a conexão intrínseca entre a dialetologia e o léxico. O estudo das variações lexicais presentes nos dialetos desempenha um papel fundamental na compreensão das diferenças de vocabulário entre regiões geográficas ou grupos sociais específicos. Essas variações nos dialetos fornecem insights valiosos sobre a evolução histórica das palavras e as influências socioculturais que moldam o léxico. Por meio dessa análise, é possível obter uma compreensão mais abrangente da diversidade linguística e cultural, revelando como as palavras refletem as características e peculiaridades de diferentes comunidades linguísticas. O estudo da dialetologia em conjunto com o léxico contribui para uma apreciação mais profunda das riquezas e complexidades das línguas, enriquecendo nosso conhecimento sobre as formas como a linguagem é moldada e utilizada pelas diversas sociedades ao longo do tempo.

Além disso, os autores também discorrem sobre a Lexicologia como nível de análise da linguagem. Sobre essa ideia Guerra e Andrade (2012, p. 230) complementam:

O léxico situa-se numa intersecção linguística que absorve informações advindas de caminhos diversos, ou seja, da fonética e da fonologia; da semântica; da morfologia; da sintaxe e das situações comunicativas, ou seja, da pragmática.

No entanto, as práticas realizadas estão distantes do que foi dito pelos autores anteriormente. Existe uma força hegemônica da gramática que ainda perpetua nos

estudos da língua. Tal fenômeno pode ser observado nas aulas de língua portuguesa, nos tópicos dos livros didáticos, nos documentos curriculares. Antunes (2012) comenta que existe uma dificuldade entre pessoas até mais escolarizadas em aceitar a transformação da língua, para elas é muito comum pensar a língua estática, ou como um conjunto fixo e inalterado de palavras, associando as transformações da língua a um processo de deterioração ou decadência da língua.

Essa preocupação também é reforçada ao lermos os documentos educacionais. Guerra e Andrade (2012), ao analisar os escritos dos Parâmetros Curriculares Nacionais - PCNs (BRASIL, 1997; 2006), e os Orientadores Curriculares proposto para o Ensino Médio - OCEM (BRASIL, 2006), notaram que os documentos não conceituam "léxico". As poucas menções são feitas entre parênteses ou em finais de orações, como uma informação extra, minimizando o mesmo. Outro estudo voltado para a análise dos documentos educacionais foi realizado por Borges e Silva (2021), o qual se encarregou em averiguar o Documento Curricular do Estado de Goiás - DCGO (2018), documento regional voltado para a educação exclusivamente nos municípios goianos. A autora relata que as únicas menções ao 'léxico', são vistas no termo léxico-morfologia. As menções a 'vocabulário' são voltadas somente para a atenção a textos técnicos e especializados. Conferindo ao léxico uma abordagem coadjuvante, e a supremacia da gramática no ensino de línguas.

Antunes (2018) cita algumas perspectivas indispensáveis para serem exploradas a partir do léxico, constituindo como uma espécie de programa de estudo do léxico nas escolas. Na concepção da linguagem, o estudo do léxico pode desenvolver competências para a geração e compreensão de discursos, sejam eles orais ou escritos, contemplando variações lexicais e a compreensão de vocabulários especializados. Dentro desse eixo também podem ser explorados de maneira contextualizada, sinônimos, antônimos, hiperônimos, parônimos, homônimos.

Já se tratando do vocabulário como componente de estruturação do texto, pode-se realizar análise quanto a coesão e coerência do texto. Na visão da constante renovação e expansão do léxico, pode-se explorar a variação lexical, as especificidades lexicais do texto oral e escrito, as formalidades e informalidades dos textos, os regionalismos. Dentro desse eixo, ainda pode-se realizar estudos acerca da derivação, composição neologismos, dos empréstimos (estrangeirismos), e da atribuição de novos sentidos a palavras já existentes na língua. Além da perspectiva

da escolha das palavras, podem ser explorados aspectos da polissemia, construções de metáforas e metonímias, estudos de fraseologia, entre outros.

Um estudo que considera o léxico como elemento fundamental no ensino de língua enfatiza a importância do conhecimento do vocabulário e contribui significativamente para o aumento das habilidades linguísticas e comunicativas, beneficiando toda a sociedade ao torná-la mais consciente de suas práticas de discurso. O conhecimento do léxico é fundamental para o desenvolvimento das competências linguísticas de um indivíduo. Quando uma pessoa possui um vocabulário amplo e preciso, ela é capaz de compreender os textos que lê com maior facilidade e produzir textos coerentes e bem estruturados. Além disso, o conhecimento do léxico também é importante para a compreensão das regras gramaticais e da estrutura sintática da língua. Quando uma pessoa conhece o significado das palavras e como elas se relacionam entre si, ela é capaz de seguir o raciocínio e a lógica compreendendo o contexto em que as palavras estão sendo utilizadas. Dessa forma, o conhecimento do léxico contribui diretamente para a expansão das competências linguísticas e comunicativas de uma pessoa, beneficiando tanto a ela quanto a toda a sociedade. Na mesma direção, cremos que a Literatura é uma ótima parceira para potencializar tais benefícios. Considerando que a melhor maneira de se compreender o léxico é dentro de seu contexto, o texto literário é um importante instrumento para desenvolver estudos lexicais.

Estudos anteriores já discutiram a importância em se atentar ao repertório vocabular. De acordo com Simões e Assis (2014), a prática languageira cotidiana é fundamental para garantir a troca de experiências entre diferentes grupos sociais contudo, apesar dessa prática corroborar na ampliação vocabular, não fornece ao falante subsídios suficientes para lidar com a linguagem formal, presente em textos variados que circulam em meio a sociedade. Em consonância, dispomos de falantes possuintes de um repertório restrito, que possuem dificuldades de leitura e escrita, o qual refreia o acesso a bens de cultura e a ascensão socioeconômica.

Bezerra (2018) diz que para ampliar o repertório vocabular, é necessário mobilizar conhecimentos prévios, associando-os aos novos, em uma prática contextualizada. Essa ação de acordo com Koch (2006), é conhecida como processamento textual, que se trata do uso de várias estratégias sociocognitivas para mobilizar diferentes tipos de conhecimentos armazenados em nossa memória e produzir sentidos ao ler um texto. Ainda de acordo com a autora, esses três tipos de

conhecimento podem ser organizados em três grandes sistemas: 1) conhecimento linguístico, o qual inclui informações sobre o léxico e as regras gramaticais de uma língua; 2) conhecimento enciclopédico, que diz respeito ao conhecimento de mundo adquirido pelo indivíduo através de suas experiências; e 3) conhecimento interacional, que é o conhecimento das formas de interação através da linguagem em diferentes contextos sociais e culturais. No entanto para isso ocorrer, é preciso romper com o princípio de que, estudar vocabulário é apenas a prática de memorizar lista de palavras ou realizar ditados.

Basta caminhar por alguns espaços escolares para que o ensino de vocabulários nas escolas brasileiras, quanto tem, existe, quase que exclusivamente, para ensinar os alunos a perceber a morfologia das palavras, a ortografia e para saber qual a classe que elas pertencem. Tais práticas são sempre desenvolvidas sem qualquer finalidade prática, não visam, nem de longe, uma atividade linguística ou, muito menos, a escrita do aluno, desassociando completamente a ideia de aquisição de vocabulário da aquisição de uma habilidade textual-comunicativa. (BEZERRA, 2018, p. 09).

Rodrigues (2014), Simões e Assis (2014) e Borges e Silva (2022) concordam que a aquisição vocabular está estritamente ligada a prática de leitura. Tal hábito é fundamental na vida de qualquer indivíduo, além da obtenção de conhecimento, a leitura auxilia o sujeito a desenvolver suas competências linguísticas, e o contato frequente com as palavras, acarreta a familiarização e conseqüentemente a ampliação do acervo vocabular.

Antunes (2012) por sua vez, diz que o texto é o único lugar onde se pode presenciar toda atividade verbal, compreendendo as funções das palavras (lexicais e gramaticais) e incluindo questões linguísticas (morfossintáticas, semânticas e pragmáticas). Ademais, a autora complementa que "as palavras só ganham sentido pleno quando encaixadas na trama do texto e de seus contextos". Logo, ditados descontextualizados e apenas a memorização de palavras, não possuem eficácia para quem deseja ampliar o vocabulário, uma vez que dessa forma, as palavras estão soltas e sem significado para serem absorvidas.

A nosso ver, esta relação entre leitura e vocabulário é circular, pois se já ficoudocumentado que o domínio do vocabulário potencia o desempenho da leitura, a leitura constitui um meio que possibilita a aprendizagem de palavras novas, o que significa que quanto mais se lê, maior é o desenvolvimento lexical. (RODRIGUES, 2014, p. 63).

Em meio a isso, Simões e Assis (2014) defendem a utilização dos textos literários clássicos como instrumento para a prática de leitura e aquisição vocabular. Segundo as autoras, o texto literário é por si polissêmico, se tornando fundamental para desenvolver leituras proficientes. Concordando com as autoras, Antunes (2018) também diz que apesar de sua especificidade, o texto literário também é feito de gramática e léxico, explorando potencialmente a multissignificação das palavras. Por sua vez, Gil (2022, p. 105) declara que “o uso de textos literários no ensino da língua é importante porque neles há uma exploração ilimitada das virtualidades do sistema da língua para a criação estética.” Assim concordamos com Simões e Assis (2014, p. 25) a respeito de invocar textos literários como base para ampliar o vocabulário, além de obter outros ganhos:

Temos a crença de que trazer de volta os clássicos para a sala de aula (sem abandonar todos os outros materiais de leitura em uso) seja uma estratégia produtiva para a aprendizagem da língua, para o enriquecimento vocabular e para a expansão do conhecimento enciclopédico, a visão de mundo do estudante.

Tendo em conta todos os argumentos que elencamos acima, pesquisar o léxico e elevá-lo como destaque, é fundamental para que sujeitos possam compreender melhor sua língua, ampliar vocabulário e adquirir conhecimento. Consoante a isso, vimos que o texto é o melhor instrumento para se trabalhar a potencialidade lexical, em especial textos literários. No entanto, é preciso romper paradigmas para buscarmos novas maneiras de trabalhar com o texto, além de estimular o gosto pela leitura e a valorização da literatura.

Sabemos que um problema do ensino é o tratamento inadequado, para não dizer desastroso, que o texto vem recebendo, não obstante as muitas alternativas e experimentações que estão sendo hoje tentadas. Com efeito, introduziu-se o texto como motivação para o ensino sem mudar as formas de acesso, as categorias de trabalho e as propostas analíticas. (MARCHUSCHI, 2008, p. 52).

A pesquisa que realizamos evidencia a importância da formalidade vocabular no estímulo à leitura. Reconhecemos que o uso adequado e consciente do vocabulário formal tem um impacto significativo na motivação e engajamento dos leitores. Quando nos preocupamos com a formalidade vocabular, estamos considerando a relevância de utilizar uma linguagem mais precisa, técnica e sofisticada, especialmente em contextos acadêmicos, literários e profissionais. A escolha de palavras apropriadas e o emprego de uma linguagem formal estimulam o

desenvolvimento de habilidades de compreensão, enriquecem o vocabulário ativo e passivo dos leitores e contribuem para a ampliação de seus horizontes linguísticos e culturais. Assim, através de dados léxico-sintáticos apresentaremos um instrumento pedagógico que servirá como um roteiro de leitura de contos literários em diferentes níveis de complexidade, desde o menos formal (mais fácil) ao mais formal (mais difícil).

Acreditamos que para estimular o hábito da leitura, é preciso fornecer textos que possuam o vocabulário condizente em sua maioria, com o vocabulário já dominado pelo sujeito. Se o sujeito está realizando a leitura de um texto, o qual ele consiga compreender o que está sendo lido, ele estará sendo estimulado a continuar essa leitura.

Em se tratando das práticas de ensino, Guerra e Andrade (2012) e Serra (2019) consentem sobre a educação brasileira estar vivendo um momento de crise. Essa situação surge devido à falta de contextualização nas atividades escolares, que ainda se baseiam em regras e na repetição de uma infinidade de exercícios descontextualizados. Isso não quer dizer que o ensino da gramática seja o problema, pois sabemos que a língua inclui a gramática. De acordo com Mascuschi (2008), o problema é a redução da língua a exercícios formais e técnicos, tornando essa abordagem o foco principal dos estudos de língua nas escolas, desconsiderando outros aspectos como o lexical, cognitivo, pragmático, entre outros.

Para Antunes (2018), esse ensino mecânico e tradicionalista ocorre devido a marginalização do léxico e a hegemonia da gramática. Sobre isso a autora discorre:

O lugar do ensino do léxico nos programas de ensino tem sido marginal, pois não tem recebido a devida atenção nem de autores de livros didáticos nem dos professores que criam suas atividades. Basta ver, em grande parte dos livros didáticos, e nas tarefas indicadas pelos professores, o espaço alargado que é concedido às descrições e às classificações das diferentes categorias gramaticais e o espaço diminutivo com que as questões lexicais são exploradas. (ANTUNES, 2018, p. 252).

Ainda segundo a autora, a carência da formação dos professores e o número excedente de alunos em uma única sala solidificam a predominância no contexto educacional. Uma vez que, para o professor é muito mais cômodo ensinar regras e definições (morfologia e sintaxe), do que realizar atividades que desenvolvam leitura e escrita contemplando questões lexicais (semântica e pragmática). Consonante a isso, os documentos curriculares são permissivos, uma vez que, apesar de pregarem

a importância de um ensino contextualizado que vise o desenvolvimento da competência linguística, ainda confere ao léxico um papel coadjuvante, com poucas citações diretas, simplificando-o como segundo plano.

As vezes que o léxico surge nos documentos o papel desempenhado pelo termo é de apêndice, não o de protagonista (ou seja, palavra em posição de relevância). Aparece entre parênteses ou em finais da oração, no entanto, sempre remetendo a uma informação a mais no texto.” (GUERRA; ANDRADE, 2012, p. 238).

Em defesa da abordagem que articula o léxico e a gramática, Gil (2022) apresenta algumas ressalvas.

1) o ensino do léxico não deve se restringir a sua dimensão semântica; 2) a unidade lexical possui dimensão semântica, gramatical e discursiva; 3) as dimensões gramatical e discursiva da unidade lexical contribuem para o estabelecimento de seu sentido; 4) o significado (nível da língua) da unidade lexical é diferente de seu sentido (nível do discurso); 5) o saber lexical, assim como o saber gramatical, devem ser construídos na observação das regularidades dos recursos linguísticos e de sua relação com diferentes normas da língua e gêneros discursivos; 6) o texto deve ser o local privilegiado para que se analisem critérios de escolha das formas do dizer e seus efeitos de sentido.

Por isso, é importante defender um ensino que inclui tanto o léxico quanto a gramática, pois ambos são fundamentais para a compreensão e comunicação em uma língua. O léxico fornece um conjunto de elementos básicos, as palavras, para que as pessoas possam expressar seus pensamentos e ideias. A gramática, por sua vez, ajuda a organizar esses pensamentos de maneira coerente, através do estabelecimento de regras. Dessa forma, entendemos que a combinação dos dois é necessária para que o ensino da língua seja eficiente.

Por sua vez, Simões e Assis (2014) apontam a tendência equivocada das escolas em nivelar a linguagem escolar à linguagem advinda dos alunos. Devido a isso, os aprendizes não desfrutam do contato com outros registros, como a linguagem formal, que possam adiante fazer parte do seu vocabulário, limitando-os a um acervo vocabular restrito, o qual provoca inúmeras dificuldades ao longo do cotidiano escolar, social e profissional, como afirma Borges e Silva (2022). Ademais, os textos que apresentam uma linguagem dessemelhante a dominada pelos discentes, estão sendo substituídos, ou até mesmo esquecidos, como é o caso da literatura e a sua perda de espaço na grade curricular.

O equívoco de nivelar a linguagem escolar pela prática verbal do novo aluno, por conseguinte, resseleccionar os textos de trabalho de modo a facilitar a leitura e a compreensão, resultou na exclusão do texto clássico e na predominância dos quadrinhos, das propagandas, das notícias, enfim dos textos curtos e de linguagem cotidiana. (SIMÕES; ASSIS, 2014, p. 24).

Considerando que o ensino tradicional se mostra pouco eficaz no progresso dos alunos, refletir, debater e propor novas práticas é fundamental para reformulá-lo. Dessa forma, acreditamos que um ensino que leve em consideração todos os níveis de análise da língua é fundamental para potencializar o ensino. Por isso, a ciência lexicológica tem muito a oferecer. A Lexicologia se ocupa de como abordar a palavra, levando em consideração todo o seu contexto e seu potencial polissêmico. Esses fatores são importantes e não devem ser ignorados quando se realizam leituras e atividades que visam além da ampliação do vocabulário e do estudo de estruturas, mas também a expansão de conhecimentos sobre o mundo.

O saber linguístico é fruto da prática, do exercício: constante, continuado, reflexivo, atento, crítico e integral, o que abarca o léxico (que palavras usar) e a gramática (normal de como arrumá-las na sequência do texto) para que consigamos expressar o que queremos. (ANTUNES, 2018, p. 253).

Sendo assim, compreendendo a importância em se ressignificar o ensino de línguas e reascender a motivação pela prática de leitura a partir de textos literários em sala de aula. Acreditamos que a ciência lexicológica tem muito a contribuir, orientando a compreensão acerca da palavra e a sua relação com o texto.

2.3 Linguística de Córpus

Nessa subseção, discorreremos sobre a Linguística de Córpus, apresentando os principais conceitos, seguidos de discussões acerca de sua utilização e contribuição para o ensino.

2.3.1 Conceituação

A fim de tecermos essa seção, apresentamos algumas definições sobre córpus. Egbert, Larsson e Biber (2020) afirmam que o córpus se trata de uma amostra grande e baseada em princípios de textos projetados para representar um domínio-alvo de

uso da língua, possuindo como objetivo final advinda da análise do cópuz, a descrição empírica difundida do uso da linguagem em um discurso alvo.

Para Hunston (2022), os estudos baseados em cópuz auxiliam em descobrir 'como é uma língua'. O'Keeffe e McCarthy (2022) dizem que o cópuz é uma ferramenta essencial para explorar condições de uso da língua, auxiliando na identificação de aspectos do uso da linguagem que antes poderiam passar despercebidos. Uma das definições mais conhecidas da área, é a descrita por Berber Sardinha (2004, p. 18):

Um conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso linguístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise.

Logo, adotando uma definição simplista o cópuz nada mais é que, uma coletânea de textos legítimos, uma vez que não são produzidos para este fim e sim retirados da sociedade garantindo a autenticidade dos dados, reunidos através de critérios pré-estabelecidos, processado em computador para serem utilizados como um objeto de pesquisa.

Como visto é importante que o cópuz reúna informações legítimas que autenticem a sua análise. Para isso, a construção do cópuz se trata de uma tarefa laboriosa, que demanda tempo e atenção a alguns critérios que devem ser seguidos para garantir a sua veracidade.

Berber Sardinha (2004) reúne alguns pontos importantes sobre o processo de construção do cópuz. Segundo o autor, o cópuz deve ter o propósito de servir como um objeto de estudo, logo a origem dos dados que vão compor o cópuz deve ser autênticos, devendo estar em linguagem natural escrita ou falada por falantes nativos, não podendo ter sido criados com o único intuito de compor o cópuz. Além disso, a composição do cópuz deve seguir critérios que garantam além da autenticidade, a representatividade do que se está almejando analisar, logo é importante que se reúna uma variedade de textos que garantam a representação do objeto de estudo, justificando as grandes extensões dos cópuz.

No caso do estudo de contos, a LC é uma aliada, uma vez que, graças às suas ferramentas computacionais, há uma otimização do tempo das análises de grandes acervos literários que antes eram morosos. A LC auxilia a reconhecer a frequência de

palavras, identificar as palavras-chave, detectar o nível de formalidade, localizar itens lexicais, entre outros.

Sobre a extensão, Berber Sardinha (2004) realizou durante quatro anos um processo de monitoramento de eventos sobre LC, o qual resultou em uma lista de classificação de tamanhos do córpus.

Quadro 1 – Tamanho do córpus

Tamanho em palavras	Classificação
Menos de 80 mil	Pequeno
80 mil a 250 mil	Pequeno Médio
250 mil a 1 milhão	Médio
1 milhão a 10 milhões	Médio Grande
10 milhões ou mais	Grande

Fonte: Berber Sardinha (2004, p. 26).

Egbert, Larsson e Biber (2020) destaca que o tamanho do córpus é importante por duas razões, 1. Alguns padrões linguísticos são mais difíceis de serem identificados em coleções de textos menores. 2. Para se obter análises quantitativas válidas e confiáveis, é importante manipular uma grande coleção de textos a fim de levantar estimativas sobre o objeto estudado. Logo, objetivando a construção de córpus genuínos, visando a representatividade, a tendência é se deparar com córpus cada vez maiores.

Berber Sardinha (2004) também observa o surgimento de vários tipos de córpus que foram surgindo ao longo do tempo, de acordo com o objetivo de cada pesquisador, o qual gerou uma lista de classificações de tipos de córpus que podem ser resumidos no quadro abaixo, de acordo com as considerações do autor.

Quadro 2 – Classificação de tipos de córpus

Tipo	Definição
Modo	Falado: composto de porções de fala transcritas. Escrito: composto de textos escritos, impressos ou não.
Tempo	Sincrônico: compreende um período de tempo. Diacrônico: compreende vários períodos de tempo. Contemporâneo: representa o período de tempo corrente. Histórico: representa um período de tempo passado.

Seleção	De amostragem: composto por porções de textos ou de variedades textuais, planejado para ser uma amostra finita da linguagem como um todo. Monitor: a composição é reciclada para refletir o estado atual de uma língua. Opõe-se a corpora de amostragem. Dinâmico ou orgânico: o crescimento e diminuição são permitidos, qualifica o corpus monitor. Estático: oposto de dinâmico, caracteriza o corpus de amostragem. Equilibrado: os componentes (gêneros, textos etc.) são distribuídos em quantidades semelhantes (por exemplo, mesmo número de textos por gênero).
Conteúdo	Especializado: os textos são de tipos específicos (gêneros ou registros definidos). Regional ou dialetal: os textos são provenientes de uma ou mais variedades sociolinguísticas específicas. Multilíngue: inclui idiomas diferentes.
Autoria	De aprendiz: os autores dos textos não são falantes nativos. De língua nativa: os autores são falantes nativos. Paralelo: os textos são comparáveis (por exemplo, original e tradução). Alinhado: as traduções aparecem abaixo de cada linha do original.
Disposição Interna	Paralelo: os textos são comparáveis (por exemplo, original e tradução). Alinhado: as traduções aparecem abaixo de cada linha do original.
Finalidade	De estudo: o corpus que se pretende descrever. De referência: usado para fins de contraste com o corpus de estudo. De treinamento ou teste: construído para permitir o desenvolvimento de aplicações e ferramentas de análise.

Fonte: Berber Sardinha (2004, p. 20-21).

Hunston (2022) destaca que para projetar um *córpus* é necessário ter em mente que o mesmo é um objeto feito propositalmente para investigar os tipos de informações ao qual foi delineado. Segundo a autora, há três critérios pelos quais os *córpus* são tipicamente julgados, sendo eles o tamanho, equilíbrio e representatividade. O primeiro, que diz respeito ao tamanho, é medido pela quantidade de tokens (palavras), ou também pela quantidade de textos que compõe o *córpus*. O equilíbrio, diz respeito ao balanceamento do *córpus*, correspondendo aos números de textos ou fichas de cada componente do *córpus*. Por fim, a representatividade, quesito indispensável que se refere a relação entre a amostra, o *córpus*, e a população de onde foi retirado.

Berber Sardinha (2012) atesta que a Linguística de *Córpus* é área de pesquisa que reúne e estabelece análises de corpora ou conjunto de dados, sejam eles escritos, falados ou multimodais, elegidos através de critérios estabelecidos, e mantidos em formato eletrônico. Para se promover a pesquisa, é necessário estabelecermos alguns critérios para a sistematização da coleta, que selecionados formarão o corpus da pesquisa.

Sarmiento (2010) destaca que desde os anos 60, *córpus* são importantes recursos para o estudo e ensinos de línguas. Percebemos essa contribuição no

primeiro trabalho criado baseado em cópulas de língua inglesa o Brown Corpus, que deu origem a vários outros trabalhos feitos posteriormente. Sua criação se torna ainda mais relevante pelo momento em que surgiu. Os paradigmas de Noam Chomsky eram contra o registro e pesquisas de desempenho linguístico, discordando dos estudos e pesquisas baseadas em probabilidade e estatística.

Sarmiento (2010) diz que o termo Linguística de Córpulas é o estudo da linguagem através de textos retirados da vida real baseada em córpulas. Suas características principais são: ser empírica, pois ela se baseia e analisa textos reais que circulam em nossa sociedade; manuseia de um grande acervo de textos baseado em princípios pré-estabelecidos; utiliza de ferramentas computacionais para análise; depende de técnicas analíticas quantitativas e qualitativas. Compreende-se textos reais, como exemplos de textos que não foram produzidos com o intuito de serem analisados em um córpulas.

Os estudos baseados em córpulas assim como outras áreas também possuem suas limitações. O córpulas não consegue informar se algum fenômeno linguístico é possível ou não. Ele também não consegue apontar nada além de seu conteúdo. Apesar de algumas limitações da Linguística de Córpulas, ela é a metodologia mais indicada para o estudo sobre a língua, tendo um método exigente na reunião de dados retirados de textos reais facilitando a vidas dos estudiosos da área da Linguística. Berber Sardinha (2012) afirma que conforme a Linguística de Corpus se desenvolve, os corpora também se expandem, se tornando cada vez maiores, isso se dá por conta do desenvolvimento das tecnologias e dos meios eletrônicos, onde os textos estão sendo disponíveis através de formatos online presente na rede mundial de computadores.

Shepherd (2009) em seu artigo nos traz indagações a respeito da Linguística de Córpulas, como por exemplo, seria ela apenas uma metodologia de investigação que possui dados coletados e armazenados em formato digital ou seria uma área da Linguística com conceitos próprios? A autora ainda destaca que a inquietação pela busca de respostas iniciou inúmeros estudos ao longo do tempo, gerando obras que buscam esclarecer qual é a área em que a Linguística de Córpulas atua. Entretanto, os estudiosos ainda não nos trouxeram uma resposta unânime.

Para a nossa pesquisa, a Linguística de Córpulas é compreendida como uma linha teórico-metodológica, uma vez que a LC constitui como uma nova forma de se enxergar a língua. Por fim, a LC contribui para a identificação da complexidade léxico-

gramatical em nosso corpus constituído por contos literários, a fim de construirmos nosso Itinerário de Leitura.

2.3.2 Linguística de Córpus e Ensino

Ao longo dos anos, discussões acerca de como aprender e ensinar uma língua acarretaram mudanças no processo de ensino-aprendizagem. Nesse aspecto, segundo Berber Sardinha (2004), a Linguística de Córpus tem contribuído através de evidências empíricas advinda do uso de córpis em diversos campos, como a tradução, a descrição de padrões linguísticos e o estudo de frequências. Outra área importante a qual a LC vem sendo empregue é a Linguística Aplicada, tais descrições linguísticas advindas da LC corroboram para a reflexão sobre o tipo de instrução oferecida nas aulas de línguas.

O interesse em estudos baseados em córpis no ensino de línguas ocorre há muito tempo. Precusores como Edward Thorndike, Harold Palmer, A. S. Hornby, Irving Lorge e Michael West iniciaram seus trabalhos ainda na fase pré-computador, onde todo o processo de compilação e levantamento de frequências era realizado de maneira manual. Tais esforços geraram inúmeros trabalhos pioneiros na produção de materiais para ensinar e testar a leitura nos Estados Unidos e Europa, ficando conhecido como *Reading Movement*.

O que hoje seria trivial para um lingüista do corpus, no início do século passado era tarefa de proporções gigantescas, pois não havia computadores que ajudassem o analista no armazenamento ou na análise do corpus. Mesmo assim, os levantamentos de Thorndike, Lorge e West foram concluídos, publicados e acolhidos com entusiasmo pela comunidade de educadores de língua materna e estrangeira. (BERBER SARDINHA, 2004, p. 253).

O êxito dessa tarefa pelejadora, logo foram substituídos por dúvidas e questionamentos provenientes da Linguística Gerativa, tendo como principal representante Noam Chomsky. Segundo ele, o processamento manual de milhões de palavras não era confiável, uma vez que o ser humano não está preparado para lidar com tarefas dessa magnitude, o que corrobora para o erro e a perda de confiabilidade.

Apesar das duras críticas e do cenário desfavorável pesquisas com córpis continuaram a ser realizadas, e foram otimizadas a partir dos anos 60 com a chegada

dos computadores contribuindo diretamente para o reaparecimento e fortalecimento da pesquisa baseada em *córpus*.

Sobre o uso do *córpus* na área de ensino, Berber Sardinha (2004) atesta que o mesmo pode ser sintetizado em quatro áreas: 1) Descrição da linguagem nativa, área que se encarrega da descrição da linguagem escrita e falada de nativos; 2) Descrição da linguagem do aprendiz, campo também encarregado de descrever a linguagem escrita e falada por sua vez de falantes aprendizes de língua estrangeira; 3) Transposição de metodologias de pesquisa acadêmica para a sala de aula, área que emprega os recursos da LC em sala de aula servindo de suporte para a criação de materiais de ensino, como concordância e listas de palavras; 4) Desenvolvimento de materiais de ensino, currículos e abordagens, por fim refere-se a criação de métodos de ensino a partir da exploração de *córpus* seguindo conceitos da LC, exemplos *Lexical Syllabus* (Currículo Lexical), *Lexical Approach* (Abordagem Lexical) e *Data Driven Learning – DDL* (Ensino Movido a Dados).

Segundo Assunção e Araújo (2019), o *Lexical Syllabus* criado por Dave Willis (1990), baseia-se na utilização de *corpus* composto por textos produzidos de falantes nativos. Sendo assim, a ideia principal desse método seria de que "os sentidos mais usuais da língua são manifestados por meio do vocabulário mais frequente." (p. 281). Se tratando do *Lexical Approach* desenvolvida por Michael Lewis (2000), se caracteriza como o desenvolvimento de atividades que concebem o léxico como o centro através dos estudos de colocações. Por fim, *Data Driven Learning – DDL* criada por Tim Johns (1991), essa abordagem atribui ao aluno um papel autônomo capaz de pesquisar e obter respostas para seus próprios questionamentos através da análise de dados obtidos de linguagem autêntica.

Existem contestações sobre o uso de *corpus* no ensino. Como a concepção de John Sinclair e o grupo Cobuild, o qual defendem a aplicação direta do *corpus*, demandando respeito dos professores a um dos principais princípios da Linguística de *Córpus*, o uso exclusivo de linguagem autêntica, sem intervenções. Outra controvérsia, tem ligação com a anterior, e refere-se a como mediar a relação entre *córpus* e aluno, isto significa dar acesso total aos dados brutos ou simplificá-lo. Nesse sentido, Berber Sardinha (2004) destaca:

Em outras palavras, aquilo que o professor quiser ensinar não precisa ser relacionado ao que o lingüista de *corpus* já descreveu. O professor tem liberdade para fazer suas próprias análises, enfocando o que achar mais relevante e apresentar os resultados como achar mais apropriada,

filtrando e simplificando-os, se necessário. (p. 257)

Szudarski (2018) afirma que a frequência das palavras é uma das utilizações mais básicas utilizadas a partir dos corpora, e também importante, pois a partir das informações que as listas de frequências disponibilizam, se pode traçar objetivos explorando o uso da linguagem em situações específicas. Como fazer comparações entre o vocabulário falado e escrito, palavras de função e conteúdo, analisar a frequência de palavras em textos diferentes, conhecida como a Lei de Zipf, analisar divisões dos vocabulários em categorias, além de fazerem pesquisas baseadas em frequência sobre vocabulário básico e avançado.

Szudarski (2018) diz que essas análises além de servirem para aperfeiçoar o entendimento e descrição do vocabulário, também possui importantes funções para a vida profissional do trabalhador da educação. Essas análises fundamentadas em frequências além de auxiliar no processo de ensino em sala de aula, podem fornecer informações que irão interferir nas escolhas dos professores e autores de materiais didáticos equivalentes do que ensinar e incluir nos materiais para o ensino da língua. Essas informações também são válidas para os lexicógrafos que desenvolvem dicionários. Entretanto o autor ressalta que apesar da frequência seja importante para identificar o vocabulário primordial que deve ser ensinado, existem outros fundamentos que são necessários para melhorar o processo de ensino de idiomas, como por exemplo levar em consideração o aprendizado e a capacidade de ensino dos elementos específicos, pois existem itens que são mais difíceis de aprender do que comparados a outros, idade e estilos de aprendizagens dos alunos, e os objetivos específicos para o quais o vocabulário será estudado.

Chambers (2019), discute e propõe em seu artigo a idealização de uma ponte entre pesquisa e prática em relação ao corpus. Apesar do crescente uso de corpus em sala de aula com a comprovada eficácia dos mesmos, existe uma lacuna entre as pesquisas de corpus e a prática de professores de ensino de idiomas, que apesar de criar, desenvolver e utilizar corpus em sala de aula através de materiais didáticos como gramáticas, dicionários, não são de fato linguistas de corpus, sendo assim não conseguem fazer análises e explorar recursos conscientemente e com a mesma precisão de um especialista em corpus. Dessa forma seria interessante que professores buscassem treinamento com pesquisadores para que aperfeiçoassem suas habilidades, resultando em uma maior eficácia em sala de aula, além de estar contribuindo e impulsionando os estudos sobre corpora no ensino de idiomas, o que

alcançaria maiores resultados em menos espaço de tempo, se caso fossem feitos apenas por especialistas.

A influência da Linguística de Corpus no ensino ocorre de modo indireto, apenas quando os resultados da pesquisa são absorvidos e incorporados, em geral parcialmente, nos materiais de ensino. A entrada no ambiente pedagógico ocorre, primordialmente, pelos livros didáticos e muito pouco por intermédio do professor. Ainda é raro, em sala de aula, o professor usar a descrição da linguagem baseada em corpus, e a razão principal é o distanciamento entre o profissional de ensino e a academia. (BERBER SARDINHA, 2004, p. 255)

Romer (2006) lista alguns desejos para o futuro em relação ao desenvolvimento da Linguística de Corpus e o ensino de línguas. Como a maior disponibilização de corpora falado, por conta de serem mais difíceis de conseguir do que em relação aos corpora escritos, seria interessante busca formas de facilitar a obtenção, pois o currículo ressalta as competências tanto na fala como na escrita, e ter mais fontes confiáveis de corpora falado seria de grande importância para alunos e professores. Sobre as ferramentas de análises, o autor sugere que se possa repensar formas de tornar esses instrumentos mais fáceis de manusear, pois professores e alunos não possuem as mesmas habilidades de linguistas que lidam o tempo todo com esses recursos. Sobre as atividades de pesquisas, podemos concluir que de fato o uso de corpora tem impacto na pedagogia, mas que ainda há muito o que se desenvolver,

Romer (2006) sugere que se faça mais estudos com corpus aprendiz, além de se utilizar métodos contrastivos, integrando esses resultados em materiais de ensino. E por último e não menos importante, como o próprio autor chama de "trabalho missionário", que se trata de divulgar a Linguística de Corpus, para professores, alunos, redatores, entre outros, demonstrando suas vantagens de uso que facilitariam o seu trabalho diário, com o intuito de convertê-los a LC.

2.4 Lexicultura

De acordo com Sarmiento (2004), há duas perspectivas sobre a linguagem. A primeira é que ela é vista como um código lógico ensinado e reproduzido para fins de comunicação, e nesta visão, a linguagem e a cultura não estão interligadas, pois não há interferência da cultura na língua. Por outro lado, há uma perspectiva mais dinâmica que considera a linguagem como um produto derivado de ações sociais, e nesta visão,

a linguagem, a sociedade e a cultura estão interligadas, refletindo as práticas e crenças do meio social.

Por sua vez, Geraldi (2003) enfatiza que a linguagem é uma expressão do pensamento, um instrumento de comunicação e uma forma de interação social. Ele destaca a importância da linguagem como meio de expressar ideias, emoções e conhecimentos, além de promover a interação e construir significados compartilhados. Essa abordagem valoriza a competência comunicativa dos estudantes e enfatiza a importância de práticas autênticas de linguagem.

Considerando as transformações sofridas pela língua ao longo do tempo, é evidente que a língua não é uma entidade estática, mas sim um fenômeno dinâmico, que se adapta e se molda às relações culturais e sociais das comunidades que a utilizam. Assim, nosso trabalho recorre à Lexicultura para compreender a profunda conexão entre léxico, língua e cultura, que se manifesta nas expressões e palavras utilizadas por diferentes comunidades e grupos sociais, bem como nas mudanças e evoluções linguísticas decorrentes das transformações culturais e sociais.

Além de desempenhar seus papéis individuais, ao usarem linguagem, os participantes devem trabalhar juntos em unidades sociais. A linguagem é fundamentalmente usada com propósito social. As línguas, assim como as conhecemos, não existiriam caso não desempenhassem seu papel social. (SARMENTO, 2004, p. 01)

A palavra é um dos principais elementos da linguagem humana e está estreitamente relacionada à cultura. De acordo com a lexicultura, a palavra é uma unidade básica da linguagem que se relaciona com o mundo cultural, social e histórico. Dessa maneira, a palavra tem um significado culturalmente construído e pode ser compreendida como uma expressão simbólica da realidade, sendo que o seu significado não é dado de forma natural, mas sim construído social e culturalmente.

As palavras não são estáticas, mas sim dinâmicas, sua aparição e desaparecimento ao longo do tempo estão relacionadas diretamente às mudanças culturais e sociais. Assim, uma palavra pode surgir para descrever um novo objeto ou conceito, enquanto outra pode desaparecer por não ser mais relevante ou necessária. Logo, as palavras que compõem o léxico de uma língua, nos auxiliam a compreender e explicar melhor a sociedade que a utiliza.

Se levarmos em conta que o léxico não é transparente e nem unívoco, mas opaco e polissêmico, o seu uso mostra, aos poucos, ao aprendente de uma língua estrangeira, as possibilidades de utilização de uma mesma palavra, por exemplo, no sentido denotativo ou conotativo. Nessa mesma perspectiva,

é possível apreender que há palavras mais carregadas de referências culturais que outras, cuja significação nem sempre é acessível a falantes de outras línguas. (BARBOSA, 2009, p. 33)

Portanto, é crucial que integremos aspectos linguísticos e culturais no ensino de línguas, para alcançarmos uma compreensão mais ampla, incluindo não apenas conhecimentos linguísticos, mas também que elementos culturais como valores, crenças e comportamentos sejam aprendidos.

Com relação a isso, o conceito de Lexicultura é proposto. De acordo com Barbosa (2009), a Lexicultura se dedica ao estudo do léxico marcado culturalmente. Essa área é o resultado da junção de dois termos, "léxico" e "cultura", representando, respectivamente, o conjunto de palavras de uma língua e as manifestações culturais de um povo. Assim, de acordo com a autora, dado que o léxico está estreitamente ligado à realidade extralinguística, ou seja, à compreensão do mundo, vivência e experiências pessoais, a definição de Lexicultura destaca a união entre léxico e cultura e enfatiza o valor que as palavras adquirem por meio do uso delas.

MATERIAL E MÉTODO

Lançando um olhar para nossa pesquisa e aos objetivos delineados, podemos classificá-la como descritiva, com abordagem quantitativa de cunho bibliográfico.

Na seção Material (seção 3.1), descreveremos os necessários para o desenvolvimento da pesquisa: documento educacional, o corpus de estudo, os softwares e o questionário.

Na seção Método (cf. 3.2), apresentaremos cada ciclo de nossa pesquisa desde a análise do documento educacional BNCC, a seleção dos contos literários para a composição do corpus de estudo, processamento no software NILC-Metrix o cálculo percentual para definir o perfil lexical, a aplicação do questionário de múltipla escolha, a correlação dos dados e a tabulação final dos dados.

As atividades descritas nas seções a seguir foram realizadas no período compreendido entre os anos de 2021 e 2022.

3.1 Material

Descreveremos nessa seção os materiais utilizados no decorrer dessa pesquisa, a saber: Base Nacional Comum Curricular (BNCC), corpus de contos literários, o software NILC-Metrix, e por fim, o questionário.

3.1.1 Base Nacional Comum Curricular

A BNCC (BRASIL, 2018) é um documento normativo homologado em dezembro de 2018 pelo Conselho Nacional de Educação (CNE), com o objetivo de reunir e garantir as aprendizagens essenciais que devem ser desenvolvidas junto a todos os alunos durante as etapas e modalidades da Educação Básica, como determina o Plano Nacional de Educação (PNE).

Figura 2 – A BNCC



Fonte: Brasil (2018).

A BNCC surgiu para se tornar referência para gestores e professores na elaboração e desenvolvimento de currículos dos sistemas e das redes escolares em amplitude nacional durante as três etapas de ensino: Educação Infantil, Ensino Fundamental e Ensino Médio, visando a contribuir no alinhamento de ações e políticas educacionais em âmbito federal, estadual e municipal, superando fragmentações e fortalecendo a qualidade do ensino ofertado.

O Ensino Fundamental é a etapa mais longa, com 9 anos de duração, atendendo estudantes de 6 a 14 anos. Assim, o documento dividiu o Ensino Fundamental em duas partes a fim de atender as demandas de cada fase. O Ensino Fundamental Anos Iniciais compreende do 1º ao 5º anos e visa a dar continuidade à Educação Infantil, consolidando as aprendizagens anteriores e ampliando as práticas. Por sua vez, o Fundamental Anos Finais abarca do 6º ao 9º anos, aprofundando o aprendizado e estimulando a autonomia.

O Ensino Fundamental, como um todo, está organizado em cinco áreas de conhecimento, sendo elas: 1) Linguagens, correspondendo às disciplinas de Língua

Portuguesa, Arte, Educação Física, e por fim Língua Inglesa, inclusa somente a partir do 6º ano; 2) Matemática; 3) Ciências da Natureza; 4) Ciências Humanas que inclui as disciplinas Geografia e História; 5) Ensino Religioso. Cada área do conhecimento possui competências específicas, onde cada componente possui um conjunto de habilidades, que se relacionam com os objetos de conhecimento, ou seja conteúdos, que são organizados em unidades temáticas.

Assim, considerando a grande importância desse documento para o cenário educacional, nos predispomos a analisar a área de Linguagens, em específico a seção Língua Portuguesa, a fim de averiguarmos qual o tratamento dedicado às práticas de leituras com contos literários. Ademais, averiguaremos qual a atenção dedicada ao léxico, se há menções sobre práticas de léxico-gramática a partir da leitura de contos literários e se há contemplação de elementos culturais nesses momentos.

*3.1.2 O *córpus* de contos literários brasileiros*

Nosso *córpus* de estudo teve a pretensão de ser extenso e representativo, de modo a constituir-se como uma fonte de dados confiáveis no que se refere a presença de contos literários brasileiros. De acordo com Berber Sardinha “o *córpus* de estudo é aquele que se pretende descrever” (2004, p. 97).

Nosso *córpus* é composto por contos literários de origem autêntica, entendemos essa autenticidade como textos que não foram escritos com o fim de servirem de objeto de análise.

A seleção de autores/obras foi feita de maneira cautelosa para que atingíssemos a representatividade necessária (procedimento descrito na seção 3.2).

O conteúdo textual foi retirado da internet e salvo em formato eletrônico (.txt), conforme ilustrado pela Figura 3.

Figura 3 – Armazenamento dos textos

N	Filename	Tokens	Date
1	C:\Users\User\Downloads\contos\Adélia Prado - Sem enfeite nenhum.txt	962	2022.04.28
2	C:\Users\User\Downloads\contos\Afonso Arinos de Melo Franco - Assombramento.txt	6.955	2022.04.28
3	C:\Users\User\Downloads\contos\Alcides Maya - Estaqueado.txt	1.569	2022.04.28
4	C:\Users\User\Downloads\contos\Alcides Maya - Miguelito.txt	1.092	2022.04.28
5	C:\Users\User\Downloads\contos\Alcântara Machado - A apaixonada Elena.txt	973	2022.04.28
6	C:\Users\User\Downloads\contos\Alcântara Machado - A isigne Cornélia.txt	2.103	2022.04.28
7	C:\Users\User\Downloads\contos\Alcântara Machado - A piedosa Tereza.txt	1.142	2022.04.28
8	C:\Users\User\Downloads\contos\Alcântara Machado - A sociedade.txt	1.019	2022.04.28
9	C:\Users\User\Downloads\contos\Alcântara Machado - Amor e sangue.txt	571	2022.04.28
10	C:\Users\User\Downloads\contos\Alcântara Machado - Apólogo brasileiro sem véu de alegoria.txt	1.145	2022.04.28
11	C:\Users\User\Downloads\contos\Alcântara Machado - Armazém Progresso de São Paulo.txt	1.168	2022.04.28
12	C:\Users\User\Downloads\contos\Alcântara Machado - Carmela.txt	1.230	2022.04.28
13	C:\Users\User\Downloads\contos\Alcântara Machado - Corinthians 2 vs Palestra 1.txt	998	2022.04.28
14	C:\Users\User\Downloads\contos\Alcântara Machado - Gaetaninho.txt	781	2022.04.28
15	C:\Users\User\Downloads\contos\Alcântara Machado - Guerra Civil.txt	1.236	2022.04.28
16	C:\Users\User\Downloads\contos\Alcântara Machado - Lisetta.txt	697	2022.04.28
17	C:\Users\User\Downloads\contos\Alcântara Machado - Miss Corisco.txt	2.023	2022.04.28
18	C:\Users\User\Downloads\contos\Alcântara Machado - Nacionalidade.txt	1.074	2022.04.28
19	C:\Users\User\Downloads\contos\Alcântara Machado - Notas Biográficas do Novo Deputado.txt	1.280	2022.04.28
20	C:\Users\User\Downloads\contos\Alcântara Machado - O Aventureiro Ulisses.txt	1.108	2022.04.28
21	C:\Users\User\Downloads\contos\Alcântara Machado - O Filósofo Platão.txt	1.438	2022.04.28
22	C:\Users\User\Downloads\contos\Alcântara Machado - O Ingênuo Dagoberto.txt	2.088	2022.04.28
23	C:\Users\User\Downloads\contos\Alcântara Machado - O Inteligente Cícero.txt	1.857	2022.04.28
24	C:\Users\User\Downloads\contos\Alcântara Machado - O Lírico Lamartine.txt	501	2022.04.28
25	C:\Users\User\Downloads\contos\Alcântara Machado - O Monstro de Rodas.txt	722	2022.04.28
26	C:\Users\User\Downloads\contos\Alcântara Machado - O Mártir Jesus.txt	1.622	2022.04.28
27	C:\Users\User\Downloads\contos\Alcântara Machado - O Patriota Washington.txt	1.548	2022.04.28
28	C:\Users\User\Downloads\contos\Alcântara Machado - O revoltado Robespierre.txt	883	2022.04.28
29	C:\Users\User\Downloads\contos\Alcântara Machado - O tímido José.txt	1.431	2022.04.28

Fonte: A autora.

A respeito da tipologia do *córpus* (Berber Sardinha, 2004), podemos definir nosso *córpus* como sendo de modo escrito, em tempo diacrônico, uma vez que possui contos de vários períodos. Quanto à seleção, trata-se de amostragem composta por porções de textos ou de gêneros textuais, planejado para ser uma amostra finita da linguagem como um todo. O conteúdo é especializado, já que os textos pertencem a um gênero definido. Os autores são falantes nativos.

Tabela 1 - Características do nosso *córpus* de estudo

Característica	Números absolutos
Período	séc. XX
Autores	78
Contos	878
<i>Tokens</i>	1.556.749
<i>Types</i>	69.658

Fonte: A autora

Conforme indicado na Tabela, chegamos a um número de 78 contistas e 878 contos literários brasileiros. Por fim, sobre a extensão, totalizamos os quantitativos de 1.556.749 tokens (um milhão, quinhentos e cinquenta e seis mil, setecentos e quarenta e nove) e 69.658 (sessenta e nove mil, seiscentos e cinquenta e oito) types.

3.1.3 NILC-Matrix

O NILC-Matrix é uma ferramenta de Processamento de Linguagem Natural (PLN) desenvolvida por um grupo de pesquisadores pertencentes ao Núcleo Interinstitucional de Linguística Computacional (NILC). Através dos estudos realizados por mais de uma década (2008-2021) originou-se o programa NILC-Matrix, composto por um conjunto de 200 métricas capazes de extrair informações objetivas de vários níveis da língua, através de uma análise de multiníveis, como o léxico, sintaxe, semântica e discurso, a fim de calcular coesão, coerência e também nível de complexidade textual.

A fim de esclarecimentos, métricas são medidas quantificáveis usadas para mensurar o resultado de um processo ou ação. Assim, cada métrica desenvolvida possui uma fórmula distinta de cálculo para identificar o objeto apresentando-os através de dados numéricos. Segundo Leal et al. (2021), as métricas podem auxiliar pesquisadores a identificar: (i) como as características do texto se correlacionam com a compreensão da leitura; (ii) quais são as características mais desafiadoras de um determinado texto, ou seja, quais características tornam um texto ou corpus mais complexo; (iii) quais textos possuem as características mais adequadas para desenvolver as habilidades dos alunos-alvo; e (iv) quais partes de um texto são desproporcionalmente complexas e devem ser simplificadas para atender a um determinado público.

Figura 4 – Extrator NILC-Metrix

NILC-Metrix

[Switch to English](#)


NILC-Metrix agrupa as métricas desenvolvidas em mais de uma década no NILC, iniciadas com o Coh-Metrix-Port (uma adaptação da ferramenta Coh-Metrix para o Português Brasileiro). O foco principal das métricas é calcular coesão, coerência e nível de complexidade textual.

Essa versão disponibiliza 200 métricas, detalhadas aqui.

Entre com o texto na caixa abaixo (Máximo 2000 palavras por vez).

Entre com o texto aqui

Não sou um robô

 INGLÊS
Português / Inglês

Processar **Limpar**

trator NILC-Metrix

Fonte: NILC (2021)

A versão disponibilizada pelo NILC no si¹te possui uma interface simples, possibilitando ao usuário inserir seu texto na caixa de diálogo que possui um limite de 2 mil palavras por vez. Além disso, caso desejar mais esclarecimentos sobre as métricas poderá acessar o menu encontrando informações sucintas das 200 métricas que estão divididas em 14 categorias, sendo elas: 1) Índice Descritivo; 2) Métricas de Facilidade de Texto; 3) Coesão Referencial; 4) Coesão Semântica LSA; 5) Diversidade Lexical; 6) Conectivos; 7) Léxico Temporal; 8) Complexidade Sintática; 9) Densidade de Padrão Sintático; 10) Informação Semântica de Palavras; 11) Informação Morfossintática de Palavras; 12) Frequência de Palavras 13) Medidas Psicolinguísticas; 14) Fórmulas de Legibilidade.

¹ Disponível em <http://fw.nilc.icmc.usp.br:23380/nilcmetrix>

3.1.3.1 Seleção das métricas

Como foi mencionado acima, a caixa de diálogo possui limite de 2 mil palavras por vez, ou seja, um número insuficiente para processarmos contos literários, visto que grande maioria ultrapassa esse teto.

Entramos em contato via e-mail com um dos organizadores da plataforma, Sidney Leal, com o intuito de solicitarmos ajuda no processamento de nosso cópuz diretamente em seu servidor institucional. Rapidamente, obtivemos resposta e em cinco dias recebemos o lote de processamentos em arquivo Excel.

Para nossa pesquisa, selecionamos 10 métricas com a finalidade de analisar a complexidade do nosso cópuz de estudo, que serão descritas na seção método.

3.1.4 Questionário de avaliação

Elaboramos um questionário na plataforma do Google Forms a fim de coletar informações acerca da percepção qualitativa dos alunos sobre o nível de complexidade dos contos literários. Tal questionário é destinado à segunda fase do Ensino Fundamental, público-alvo dessa pesquisa.

O participante pode selecionar apenas uma opção: 1) Muito fácil; 2) Fácil; 3) Moderado; 4) Difícil; 5) Muito difícil.

Figura 5 – Formulário de percepção de complexidade

Análise de nível de vocabulário em contos literários brasileiros

Caro aluno, esse questionário visa coletar as suas percepções a cerca do nível de dificuldade no vocabulário presente em contos literários brasileiros, apresentados aqui através de trechos. Ressaltamos a importância de serem honestos quanto as suas respostas, que serão primordiais para o andamento de nossa pesquisa.

Conto 1

– Olá! exclamei eu, vendo saltar do bonde de Botafogo o meu querido Ernesto Branco. Bons ares te tragam! Como vais tu? Mas que diabo de cara tens agora? Estás zangado?

– Ora! Não me fales! Não estou zangado; estou aborrecido. Aborrecido com esta vida infernal do Rio de Janeiro; aborrecido com este calor selvagem, este calor inimigo da civilização e do trabalho; e aborrecido principalmente com as nossas patricias, esses monstros de olhos sedutores e sorrisos virginais!

– Ó diabo! a cousa agora é mais grave... Dar-se-á o caso de que o meu espirituoso amigo levasse tábua de alguma moça com quem estivesse para casar?...

– Hein?! Casar?! Eu?! Com quem?!

– Oh! com qualquer moça do teu gosto...

[...]

A partir de seus entendimentos, como você classificaria o nível de vocabulário presente no Conto 1?

Muito fácil

Fácil

Moderado

Difícil

Muito difícil

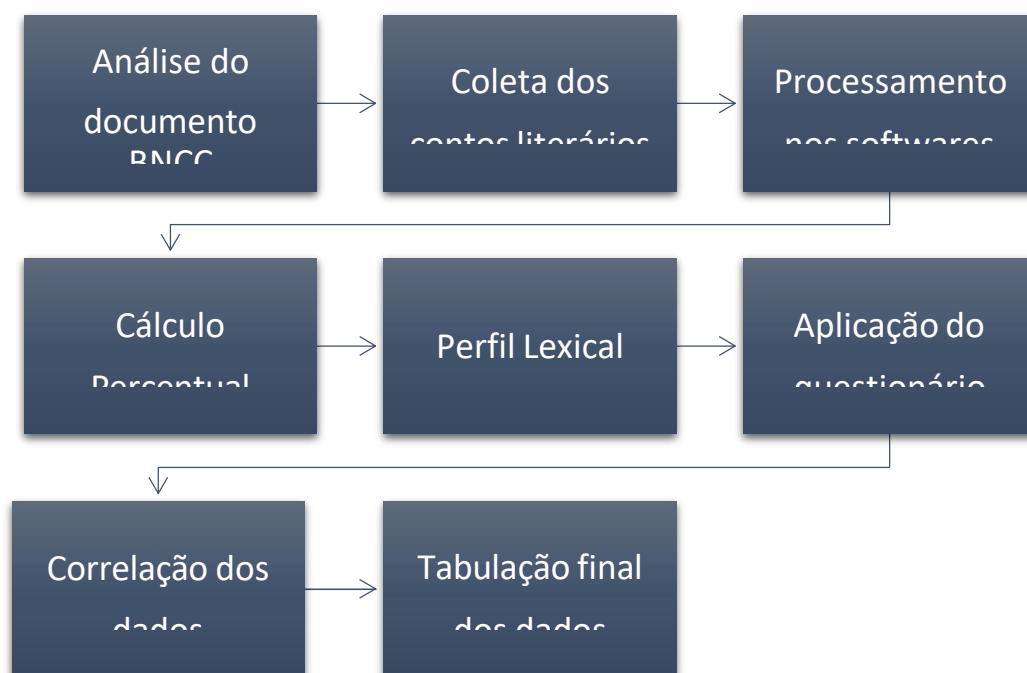
Fonte: A autora.

Esse questionário é o último instrumento que compõe a tríade de processamento do nível de complexidade do corpus de estudo, etapa fundamental para confrontarmos os dados quantitativos com aqueles oriundos da percepção humana.

3.2 Método

Apresentaremos nessa seção o método utilizado. A pesquisa foi dividida nos seguintes ciclos: análise do documento BNCC, a coleta dos contos literários, o processamento dos contos no NILC-Matrix, o cálculo percentual; o perfil lexical, a aplicação do questionário, a correlação dos softwares com o questionário e por fim a tabulação final dos dados.

Figura 6 – Fluxograma com os ciclos da pesquisa



Fonte: A autora.

da

O método de nossa pesquisa inicia com a análise do documento normativo BNCC, mais precisamente, na seção Linguagem, a disciplina Língua Portuguesa na segunda etapa do Ensino Fundamental. Essa análise visa a identificar qual a conduta da BNCC em relação as práticas de leitura, o uso dos contos literários, a aquisição léxico-sintática a partir da leitura de contos literários, e se há contemplação de conhecimentos culturais nesses momentos.

Em seguida, partimos para a coleta de nosso corpus de estudo, que corresponde ao objeto de análise dessa pesquisa. Essa etapa foi realizada de maneira atenta, uma vez que não bastava apenas selecionar o maior número de contos, mas, sim, selecionar os contos mais importantes e representativos da Literatura Brasileira. Fez-se necessária a leitura de críticos com o intuito de identificarmos e selecionarmos nomes de autores contistas brasileiros importantes ao longo da história literária, para, então, selecionar os contos pertencentes a esses autores.

Tivemos a preocupação de selecionar nomes de contistas nativos brasileiros que marcaram a historicidade da Literatura Brasileira. Para tanto, consultamos trabalhos realizados por críticos da área da Literatura Brasileira como Bosi (1982), Massaud Moisés (2000), Candido (2000) e Moriconi (2000). De forma complementar,

consultamos coletâneas de contos como Os Cem Melhores Contos Brasileiros do Século (MORICONI, 2000), Antologia do Conto Brasileiro do Romantismo e Modernismo (TUFANO, 2005), Antologia de Contos: Contos Brasileiros Contemporâneos (LADEIRA, 2005).

Tendo conhecimento dos nomes dos contistas que iriam compor o córpus, a próxima etapa foi a busca e coleta dos contos, etapa realizada via internet. Assim, utilizando uma ferramenta de busca, escrevemos o nome do autor e selecionamos os contos disponíveis na rede. Os contos encontrados estavam disponíveis em arquivo pdf (*Portable Document Format*) ou escritos em páginas online. Sendo assim, copiamos os contos separadamente e colamos no bloco de notas. Em seguida, salvamos com o respectivo nome do autor e obra em formato (.txt) reunindo todos em uma única pasta. Ao fim dessa etapa, obtivemos um córpus de estudo composto por 878 contos literários brasileiros, totalizando 1.556.749 *tokens* e 69.658 *types*.

Como foi dito na seção dos materiais, a versão online do NILC-Matrix suporta apenas o limite de 2 mil palavras por vez, logo se trata de um número insuficiente para contemplar a extensão dos contos literários, visto que muitos ultrapassam esse número de palavras. Assim, através de um e-mail solicitamos aos organizadores do NILC-Matrix o auxílio no processamento de nosso córpus, onde rapidamente fomos atendidos pelo colaborador Sidney Leal.

Os resultados recebidos das 200 métricas foram agrupados em único arquivo em formato Excel (xlsx). As duas primeiras colunas correspondiam a nome do autor e obra respectivamente, as demais equivaliam as parciais de cada métrica, como pode ser melhor observado na figura abaixo.

Figura 7 – Resultados das métricas do NILC-Metrix

contos_metrics_v1.xlsx		Open with					
A	B	C	D	E	F	G	H
group	text	adjective_ratio	adverbs	content_words	flesh	function_words	sentences_per_paragraph
Humberto de Campos	O religio preguiçoso.txt	0,01515	0,04545	0,5303	61,52445	0,4697	1,27778
Humberto de Campos	O ladrão amepido.txt	0,01754	0,03509	0,58772	73,65405	0,41228	1,34783
Clarice Lispector	Atenção ao sábado.txt	0,01778	0,08889	0,61778	72,03364	0,38222	1
Fernando Sabino	A companhia de viagem.txt	0,01892	0,1	0,56216	64,5117	0,43919	12,6
Humberto de Campos	Roubado e contente.txt	0,01905	0,04762	0,55238	82,33419	0,44762	1,48
José J. Veiga	Roupa no coradouro.txt	0,02049	0,08116	0,56173	63,16126	0,43927	3,94444
Coselho Neto	O mentiroso.txt	0,02055	0,08562	0,60274	50,30046	0,40068	1,46667
Alcântara Machado	O Monato de Rodas.txt	0,02067	0,06359	0,64706	69,81113	0,35294	2,21429
Alcântara Machado	A ligue Comélia.txt	0,02068	0,10604	0,6299	77,41197	0,3701	1,94949
José J. Veiga	Uma pedrinha na ponte.txt	0,02074	0,10818	0,60202	77,96723	0,39798	28,8
Humberto de Campos	O purgativo.txt	0,02075	0,09544	0,59187	95,62704	0,44813	2,15385
Alcântara Machado	Amazém Progresso de São Paulo.txt	0,02124	0,07722	0,63996	74,02625	0,36004	1,86301
Femeia Guilar	A estante.txt	0,02247	0,09738	0,62172	76,11735	0,37628	2,29167
José J. Veiga	Tarde de sábado, manhã de domingo.txt	0,02311	0,08955	0,60951	68,85237	0,39076	3,03297
Humberto de Campos	O perigo das profecias.txt	0,02326	0,02907	0,54651	68,00178	0,45349	1,13043
João Guimarães Rosa	Soroco, sua mãe, sua filha.txt	0,02338	0,07768	0,66033	73,7878	0,44042	1,44882
Humberto de Campos	As pesquisas de Moisés.txt	0,024	0,092	0,616	68,51733	0,384	1,5
Alcântara Machado	O Filadelfo Platão.txt	0,02441	0,10822	0,66721	75,70681	0,33279	3,07229
Alcântara Machado	A piedosa Tereza.txt	0,02467	0,05242	0,62795	74,60317	0,37205	1,84821
Arthur Azevedo	O retrato.txt	0,02473	0,08668	0,59341	59,90844	0,40934	1,36
Machado de Assis	Umás férias.txt	0,02482	0,06902	0,57503	70,90838	0,42497	3,95745
Fernando Sabino	O gato sou eu.txt	0,02488	0,08087	0,52722	87,85912	0,47434	16,4
João Simões Lopes Neto	O duelo dos farapos.txt	0,02493	0,0679	0,56074	75,68687	0,43979	1,19872
Fernando Sabino	O canto do gato.txt	0,02498	0,08453	0,56964	73,12136	0,43132	32
José J. Veiga	Onde andam os didangos.txt	0,02522	0,07246	0,58647	68,07563	0,41353	2,42857
Valdomiro Silveira	Eperando.txt	0,02545	0,1018	0,6087	80,00966	0,3913	1,67143
Lygia Fagundes Telles	Que se chama solidão.txt	0,02566	0,0647	0,59565	67,33311	0,40435	4,20588
José J. Veiga	Tia Zi rezando.txt	0,02581	0,08475	0,56317	63,63534	0,43683	3,25
Humberto de Campos	A despedida.txt	0,02583	0,07749	0,56089	68,87253	0,43911	1,5
José J. Veiga	Os cavallinhos de Piauplanto.txt	0,02586	0,09614	0,58157	69,94997	0,41916	2,37662
José J. Veiga	Era só brincadeira.txt	0,02597	0,08472	0,56707	62,60122	0,43318	2,20492
Raul Pompéia	Antes e depois.txt	0,02597	0,0487	0,62652	77,03261	0,37338	1,89286
João Simões Lopes Neto	O anjo da vitória.txt	0,02609	0,05349	0,5486	66,54316	0,45205	1,17323
João Antônio	Leão de chácara.txt	0,02623	0,07335	0,59194	78,27468	0,40855	1,59667

Fonte: A autora.

Nesse ponto, coube a nós realizarmos uma triagem das métricas que seriam selecionadas para nossa análise. Para isso, elaboramos alguns critérios de seleção, métrica não ter apresentado erros nos resultados de processamento e métrica com amplitude (diferença entre o menor valor e o maior valor) maior que 3.

Em um primeiro momento, foram selecionadas 14 métricas. Porém, devido aos critérios estabelecidos, 4 foram eliminadas, a saber: Proporção de palavras de conteúdo em relação à quantidade de palavras funcionais do texto; Proporção de orações que não estão no formato SVO (sujeito-verbo-objeto) em relação a todas orações do texto; Proporção de substantivos em relação à quantidade de palavras do texto; Quantidade de diferentes tempos-modos verbais que ocorrem no texto.

No final, 10 métricas atenderam aos critérios, e precisaram de um peso para que déssemos seqüência. No quadro abaixo, é possível checar quais métricas foram selecionadas, a categoria que pertencem e o peso atribuído.

Quadro 3 – Informações sobre as métricas selecionadas

Categoria	Nome da métrica	Peso atribuído
Informações Morfosintáticas de palavras	Proporção de Adjetivos em relação à quantidade de palavras do texto	1,0
Informações Morfosintáticas de palavras	Proporção de palavras de conteúdo em relação à quantidade de palavras do texto	1,5

Medidas descritivas	Quantidade de sentenças no texto	1,0
Medidas descritivas	Média de palavras por sentença	1,5
Medidas descritivas	Quantidade de palavras no texto	1,5
Diversidade lexical	Proporção de <i>types</i> (despreza repetições de palavras) em relação à quantidade de tokens (computa repetições de palavras) no texto	3,0
Complexidade sintática	Fórmula de complexidade sintática de Frazier	2,5
Complexidade sintática	Quantidade Média de palavras antes dos verbos principais das orações principais das sentenças	2,0
Complexidade sintática	Proporção de conjunções coordenativas em relação a todas as orações do texto	0,5
Complexidade sintática	Proporção de orações subordinadas reduzidas pela quantidade de orações do texto	2,0

Fontes: Dados da presente pesquisa

Finalmente, com as métricas e pesos definidos, passamos para o cálculo para reconhecermos o nível de dificuldade em cada conto.

Para tornar os resultados mais legíveis e objetivos, o cálculo é iniciado a partir do arredondamento dos valores obtidos para uma casa decimal. Em seguida, é identificada a pontuação mais alta em cada métrica, que será utilizada como referência para o cálculo dos valores de cada conto. Esse processo envolve a seleção da pontuação parcial correspondente a cada conto, que é multiplicada por 10 e dividida pela pontuação mais alta encontrada. O resultado dessa operação é então multiplicado pelo peso atribuído à métrica em questão, gerando um valor para cada conto em cada uma das métricas avaliadas, separadas por cores como poderá ser melhor observado na Figura 8.

Figura 8 - Planilha com cálculos linguístico-estatísticos

1	group	text	words_before_main_verb (2,0)			content_words (1,5)			sentences (1,0)			words_per_sentence (1,5)			words (1,5)			ttr (3,0)		
2	Humberto de Campos	O relógio preguiçoso.txt	0,5	5,6	2,8	0,5	5,9	2,9	23	0,4	0,2	8,6	2,5	1,2	198	0,29058	0,3	0,7	9,1	
3	Humberto de Campos	O ladrão arrependido.txt	0,6	6,7	3,3	0,6	6,5	3,3	31	0,5	0,3	7,4	2,1	1,1	228	0,33461	0,3	0,7	9,2	
4	Clarice Lispector	Atenção ao sábado.txt	0,6	6,7	3,3	0,6	6,9	3,4	11	0,2	0,1	20,5	5,9	2,9	225	0,3302	0,3	0,7	8,4	
5	Fernando Sabino	A companheira de viagem.txt	0,6	6,7	3,3	0,6	6,2	3,1	63	1,0	0,5	11,7	3,4	1,7	740	1,086	1,1	0,7	8,9	
6	Humberto de Campos	Roubado e contente.txt	0,6	6,7	3,3	0,6	6,1	3,1	37	0,6	0,3	5,7	1,6	0,8	210	0,30819	0,3	0,7	8,6	
7	José J. Veiga	Roupa no coradouro.txt	0,6	6,7	3,3	0,6	6,2	3,1	142	2,3	1,2	26,5	7,6	3,8	3758	5,51512	5,5	0,7	9,2	
8	Coelho Neto	O mentiroso.txt	0,6	6,7	3,3	0,6	6,7	3,3	44	0,7	0,4	6,6	1,9	1,0	292	0,42853	0,4	0,7	9,4	
9	Alcântara Machado	O Monstro de Rodas.txt	0,6	6,7	3,3	0,6	7,2	3,6	93	1,5	0,8	6,8	1,9	1,0	629	0,9231	0,9	0,8	9,4	
10	Alcântara Machado	A isigne Cornéllia.txt	0,6	6,7	3,3	0,6	7,0	3,5	386	6,3	3,1	4,9	1,4	0,7	1886	2,76783	2,8	0,7	9,3	
11	José J. Veiga	Uma pedrinha na ponte.txt	0,6	6,7	3,3	0,6	6,7	3,3	144	2,3	1,2	12,4	3,6	1,8	1784	2,61814	2,6	0,8	9,4	
12	Humberto de Campos	O purgativo.txt	0,6	6,7	3,3	0,6	6,1	3,1	56	0,9	0,5	4,3	1,2	0,6	241	0,35368	0,4	0,6	6,9	
13	Alcântara Machado	Armazém Progresso de São Paulo.txt	0,6	6,7	3,3	0,6	7,1	3,6	136	2,2	1,1	7,6	2,2	1,1	1036	1,5204	1,5	0,7	9,3	
14	Ferreira Gullar	A estante.txt	0,6	6,7	3,3	0,6	6,9	3,5	55	0,9	0,4	9,7	2,8	1,4	534	0,78368	0,8	0,7	8,8	
15	José J. Veiga	Tarde de sábado, manhã de domingo.txt	0,6	6,7	3,3	0,6	6,8	3,4	276	4,5	2,2	13,8	4,0	2,0	3808	5,58849	5,6	0,8	9,5	
16	Humberto de Campos	O perigo das profecias.txt	0,5	5,6	2,8	0,5	6,1	3,0	26	0,4	0,2	6,6	1,9	1,0	172	0,25242	0,3	0,8	9,5	
17	João Guimarães Rosa	Soroco, sua mãe, sua filha.txt	0,6	6,7	3,3	0,6	6,2	3,1	184	3,0	1,5	7,2	2,1	1,0	1326	1,94599	1,9	0,8	9,5	
18	Humberto de Campos	As pesquisas de Moisés.txt	0,6	6,7	3,3	0,6	6,8	3,4	42	0,7	0,3	6,0	1,7	0,9	250	0,36689	0,4	0,8	9,6	
19	Alcântara Machado	O Filósofo Platão.txt	0,7	7,8	3,9	0,7	7,4	3,7	255	4,2	2,1	4,8	1,4	0,7	1229	1,80364	1,8	0,7	9,2	
20	Alcântara Machado	A piedosa Tereza.txt	0,6	6,7	3,3	0,6	7,0	3,5	207	3,4	1,7	4,7	1,4	0,7	973	1,42794	1,4	0,8	9,7	
21	Arthur Azevedo	O retrato.txt	0,6	6,7	3,3	0,6	6,6	3,3	27	0,4	0,2	13,5	3,9	1,9	364	0,53419	0,5	0,7	9,3	
22	Machado de Assis	Umás férias.txt	0,6	6,7	3,3	0,6	6,4	3,2	186	3,0	1,5	13,9	4,0	2,0	2579	3,78485	3,8	0,8	9,6	
23	Fernando Sabino	O gato sou eu.txt	0,5	5,6	2,8	0,5	5,9	2,9	82	1,3	0,7	7,8	2,3	1,1	643	0,94365	0,9	0,6	7,6	
24	João Simões Lopes Neto	O duelo dos farrapos.txt	0,6	6,7	3,3	0,6	6,2	3,1	187	3,0	1,5	10,1	2,9	1,5	1885	2,76636	2,8	0,8	9,4	
25	Fernando Sabino	O canto do galo.txt	0,6	6,7	3,3	0,6	6,3	3,2	64	1,0	0,5	16,3	4,7	2,3	1041	1,52774	1,5	0,7	9,2	
26	José J. Veiga	Onde andam os didangos.txt	0,6	6,7	3,3	0,6	6,5	3,3	153	2,5	1,2	16,3	4,7	2,4	2498	3,66598	3,7	0,7	9,1	
27	Valdomiro Silveira	Esperando.txt	0,6	6,7	3,3	0,6	6,8	3,4	132	2,1	1,1	7,1	2,1	1,0	943	1,38392	1,4	0,7	9,3	
28	Ligia Fagundes Telles	Que se chama solidão.txt	0,6	6,7	3,3	0,6	6,6	3,3	143	2,3	1,2	12,5	3,6	1,8	1793	2,63135	2,6	0,8	9,5	

Fonte: A autora

Após calcular os valores de cada conto em todas as métricas, é realizada uma etapa final de cálculo. Nesse estágio, todos os resultados finais de cada métrica são somados e o valor resultante é dividido pelo total de peso das métricas (16,5). O resultado dessa operação é o valor léxico-sintático final para cada conto literário do corpus em questão.

Já a aplicação do questionário contou com o apoio e autorização de uma escola municipal localizada em Piracanjuba-Goiás. Com a ajuda da professora regente de Língua Portuguesa, o questionário foi aplicado na turma do 9º ano como uma atividade extracurricular a ser realizada em casa, sem interferência de colegas e professores. No entanto, somente 10 dos 16 alunos responderam todas as cinco partes do questionário, servindo como base de análise a seguir.

RESULTADOS E ANÁLISE DOS DADOS

Neste momento, iremos apresentar os resultados obtidos e a análise dos dados. Para facilitar a visualização e compreensão dos resultados, optamos por dividir em subseções, que são: Análise da Base Nacional Comum Curricular; Análise dos dados fornecidos pelo NILC-Metrix; Análise do questionário.

4.1 Análise da Base Nacional Comum Curricular

Nesta seção, iniciaremos com a análise da Base Nacional Comum Curricular (BNCC), aprovada em 2018, que estabelece os conhecimentos, competências e habilidades que todos os estudantes brasileiros devem desenvolver durante sua trajetória educacional. A BNCC serve como referência para a elaboração de currículos escolares, materiais didáticos e avaliações, contribuindo para uma educação de qualidade e igualitária.

Nesse sentido, espera-se que a BNCC ajude a superar a fragmentação das políticas educacionais, enseje o fortalecimento do regime de colaboração entre as três esferas de governo e seja balizadora da qualidade da educação. Assim, para além da garantia de acesso e permanência na escola, é necessário que sistemas, redes e escolas garantam um patamar comum de aprendizagens a todos os estudantes, tarefa para a qual a BNCC é instrumento fundamental. (BNCC, 2018, p. 8).

Considerando a relevância da BNCC para o campo educacional, analisaremos a sensibilidade do documento em relação ao conto literário, léxico e elementos interculturais presentes na seção de Língua Portuguesa do Ensino Fundamental, buscando compreender como a BNCC aborda esses aspectos ao longo de suas 600 páginas.

Tendo como objetivo a aprendizagem e formação integral dos alunos, a BNCC formulou competências específicas para cada componente curricular, que devem ser contempladas ao longo da educação básica. Para a área de Língua Portuguesa no Ensino Fundamental, a BNCC define dez competências gerais, que podem ser sintetizadas em: entender a língua como meio de comunicação, expressão e interação entre pessoas; conhecer e valorizar a diversidade linguística e cultural; compreender os usos da língua em diferentes contextos; e fazer uso da linguagem para participação social, exercício da cidadania e aprimoramento pessoal.

O documento BNCC atribui ao componente Língua Portuguesa uma perspectiva enunciativa-discursiva, também assumida em outros documentos como os Parâmetros Curriculares Nacionais (PCN). Essa proposta atribui centralidade ao texto como unidade de trabalho, de maneira a relacionar seus contextos ao desenvolvimento e uso significativo da linguagem em diversas atividades, como leitura, escuta e produção de textos. Assim as práticas de linguagem são divididas em quatro eixos: oralidade; leitura/escuta; produção (escrita e multissemiótica) e análise linguística/semiótica.

Essa abordagem tem como objetivo desenvolver a competência comunicativa dos estudantes por meio da reflexão sobre o uso da língua em diversos contextos. Através da leitura e produção de diferentes gêneros textuais, os alunos podem aprimorar suas habilidades linguísticas e discursivas, assim como suas competências sociais e culturais. Essa perspectiva enunciativa-discursiva reconhece a importância do contexto sociocultural e histórico da linguagem na compreensão e produção de textos, incentivando a exploração de diversas perspectivas e pontos de vista em relação ao uso contextualizado da língua. Isso inclui a capacidade de reconhecer a presença da atividade verbal, mobilizando conhecimentos linguísticos e compreendendo plenamente o sentido das palavras contextualizadas na trama de um texto, conforme afirmado por Antunes (2012).

Sobre gêneros textuais, a BNCC destaca que as práticas de linguagem contemporâneas demandam cada vez mais a inserção de gêneros multissemióticos e multimidiáticos, porém isso não quer dizer que gêneros impressos, tais como notícia, reportagem, artigo de opinião, conto, charge, deixem de ser privilegiados no ambiente escolar. Assim, para garantir a contemplação dos cânones e das novas tendências de gêneros textuais a BNCC possui outra categoria organizadora, os campos de atuação, que assim como os eixos temáticos, também se articulam com as práticas de linguagem. No ensino fundamental anos finais, os campos de atuação se dividem em quatro categorias: Campo artístico-literário; Campo das práticas de estudo e pesquisa; Campo jornalístico-midiático e Campo de atuação na vida pública.

A escolha por esses campos, de um conjunto maior, deu-se por se entender que eles contemplam dimensões formativas importantes de uso da linguagem na escola e fora dela e criam condições para uma formação para a atuação em atividades do dia a dia, no espaço familiar e escolar, uma formação que contempla a produção do conhecimento e a pesquisa; o exercício da cidadania, que envolve, por exemplo, a condição de se inteirar dos fatos do mundo e opinar sobre eles, de poder propor pautas de discussão e soluções

de problemas, como forma de vislumbrar formas de atuação na vida pública; uma formação estética, vinculada à experiência de leitura e escrita do texto literário e à compreensão e produção de textos artísticos multissemióticos. (BRASIL, 2018, p. 84).

Logo, cabe ao campo literário da BNCC a inserção e o desenvolvimento do texto literário, como o gênero conto, no currículo escolar. Isso significa que é responsabilidade deste campo abordar os diferentes gêneros e formas literárias, e também as especificidades de cada época e cultura. Além disso, é preciso desenvolver habilidades que permitam ao estudante compreender e apreciar a literatura, como a interpretação de textos, a análise crítica, a reflexão sobre os temas abordados, entre outras.

Abaixo apresentamos exemplos de habilidades desenvolvidas pelo campo literário ao longo Ensino Fundamental anos finais (6º ao 9º ano) as quais propõem tal formação literária defendida pela BNCC (2018):

Quadro 4 – Habilidades de formação literária - BNCC

CAMPO DE ATUAÇÃO	HABILIDADE	SÉRIE
	(EF67LP27) Analisar, entre os textos literários e entre estes e outras manifestações artísticas (como cinema, teatro, música, artes visuais e midiáticas), referências explícitas ou implícitas a outros textos, quanto aos temas, personagens e recursos literários e semióticos.	
	(EF67LP28) Ler, de forma autônoma, e compreender – selecionando procedimentos e estratégias de leitura adequados a diferentes objetivos e levando em conta características dos gêneros e suportes –, romances infantojuvenis, contos populares, contos de terror, lendas brasileiras, indígenas e africanas, narrativas de aventuras, narrativas de enigma, mitos, crônicas, autobiografias, histórias em quadrinhos, mangás, poemas de forma livre e fixa (como sonetos e cordéis), vídeo-poemas, poemas visuais, dentre outros, expressando avaliação sobre o texto lido e estabelecendo preferências por gêneros, temas, autores.	6º e 7º
	(EF67LP29) Identificar, em texto dramático, personagem, ato, cena, fala e indicações cênicas e a organização do texto: enredo, conflitos, ideias principais, pontos de vista, universos de referência.	
	(EF67LP30) Criar narrativas ficcionais, tais como contos populares, contos de suspense, mistério, terror, humor, narrativas de enigma, crônicas, histórias em quadrinhos, dentre outros, que utilizem cenários e personagens realistas ou de fantasia, observando os elementos da estrutura narrativa próprios ao gênero pretendido, tais como enredo, personagens, tempo, espaço e narrador, utilizando tempos verbais adequados	

CAMPO
ARTÍSTICO-
LITERÁRIO

à narração de fatos passados, empregando conhecimentos sobre diferentes modos de se iniciar uma história e de inserir os discursos direto e indireto.

(EF89LP33) Ler, de forma autônoma, e compreender – selecionando procedimentos e estratégias de leitura adequados a diferentes objetivos e levando em conta características dos gêneros e suportes – romances, contos contemporâneos, minicontos, fábulas contemporâneas, romances juvenis, biografias romanceadas, novelas, crônicas visuais, narrativas de ficção científica, narrativas de suspense, poemas de forma livre e fixa (como haicai), poema concreto, ciberpoema, dentre outros, expressando avaliação sobre o texto lido e estabelecendo preferências por gêneros, temas, autores.
(EF69LP44) Inferir a presença de valores sociais, culturais e humanos e de diferentes visões de mundo, em textos literários, reconhecendo nesses textos formas de estabelecer múltiplos olhares sobre as identidades, sociedades e culturas e considerando a autoria e o contexto social e histórico de sua produção.

8º e 9º

(EF69LP46) Participar de práticas de compartilhamento de leitura/recepção de obras literárias/ manifestações artísticas, como rodas de leitura, clubes de leitura, eventos de contação de histórias, de leituras dramáticas, de apresentações teatrais, musicais e de filmes, cineclubes, festivais de vídeo, saraus, slams, canais de booktubers, redes sociais temáticas (de leitores, de cinéfilos, de música etc.), dentre outros, tecendo, quando possível, comentários de ordem estética e afetiva e justificando suas apreciações, escrevendo comentários e resenhas para jornais, blogs e redes sociais e utilizando formas de expressão das culturas juvenis, tais como, vlogs e podcasts culturais (literatura, cinema, teatro, música), playlists comentadas, fanfics, fanzines, e-zines, fanvídeos, fanclipes, posts em fanpages, trailer honesto, vídeo-minuto, dentre outras possibilidades de práticas de apreciação e de manifestação da cultura de fãs.

(EF69LP49) Mostrar-se interessado e envolvido pela leitura de livros de literatura e por outras produções culturais do campo e receptivo a textos que rompam com seu universo de expectativas, que representem um desafio em relação às suas possibilidades atuais e suas experiências anteriores de leitura, apoiando-se nas marcas linguísticas, em seu conhecimento sobre os gêneros e a temática e nas orientações dadas pelo professor.

6º ao 9º

(EF69LP53) Ler em voz alta textos literários diversos – como contos de amor, de humor, de suspense, de terror; crônicas líricas, humorísticas, críticas; bem como leituras orais capituladas (compartilhadas ou não com o professor) de livros de maior extensão, como romances, narrativas de enigma, narrativas de aventura, literatura infantojuvenil, – contar/recontar histórias tanto da tradição oral (causos, contos de esperteza, contos de animais, contos de amor, contos de encantamento, piadas, dentre outros) quanto da tradição literária escrita, expressando a compreensão e interpretação do texto por meio de uma leitura ou fala expressiva e fluente, que respeite o ritmo, as pausas, as hesitações, a entonação indicados tanto pela pontuação quanto por outros recursos gráfico-editoriais, como

negritos, itálicos, caixa-alta, ilustrações etc., gravando essa leitura ou esse conto/reconto, seja para análise posterior, seja para produção de audiobooks de textos literários diversos ou de podcasts de leituras dramáticas com ou sem efeitos especiais e ler e/ou declamar poemas diversos, tanto de forma livre quanto de forma fixa (como quadras, sonetos, liras, haicais etc.), empregando os recursos linguísticos, paralinguísticos e cinésicos necessários aos efeitos de sentido pretendidos, como o ritmo e a entonação, o emprego de pausas e prolongamentos, o tom e o timbre vocais, bem como eventuais recursos de gestualidade e pantomima que convenham ao gênero poético e à situação de compartilhamento em questão

Fonte: BNCC (2018)

No que diz respeito ao léxico, a BNCC estabelece que o desenvolvimento dos conhecimentos lexicais, bem como dos conhecimentos grafofônicos, ortográficos, morfológicos, sintáticos, textuais, discursivos, sociolinguísticos e semióticos que contribuem para a compreensão e produção da linguagem devem ser construídos simultaneamente ao longo do Ensino Fundamental, por meio de práticas de leitura/escuta e produção de textos orais, escritos e multissemióticos. Isso inclui reflexões sobre o uso da língua, bem como sobre conceitos como regras e efeitos de sentidos (BRASIL, 2018, p. 81).

Ao analisarmos o ensino de Língua Portuguesa nos anos finais do Ensino Fundamental, é possível constatar que o estudo do léxico está inserido no eixo Análise Linguística/Semiótica. Esse eixo tem como finalidade promover a compreensão dos mecanismos de produção e interpretação da linguagem, por meio do desenvolvimento de competências relacionadas à identificação, uso e compreensão dos efeitos de sentidos sobre a língua. No entanto, ao considerar essas estratégias (meta)cognitivas, nota-se que os conhecimentos lexicais são vistos de forma limitada, como meros elementos que contribuem para a construção de estilos e significados dentro do texto. Isso pode indicar uma visão simplificada da importância do léxico no estudo da língua como afirma Guerra e Andrade (2012), comprometendo o reconhecimento dos estudos lexicais.

A falta de atenção adequada dada ao léxico é ainda mais evidente quando procuramos a palavra "léxico" na BNCC, encontrando apenas 10 menções em suas 600 páginas. Além disso, ao examinarmos a proposta de ensino da Língua Inglesa, também conhecida como L2, podemos observar que ela possui uma unidade totalmente dedicada ao léxico, denominada "Estudo do Léxico". Por outro lado, na

Língua Portuguesa (L1), o léxico é considerado apenas um elemento complementar, integrado às análises morfológicas, representado pelo objeto de conhecimento "léxico/morfologia".

Sobre o elemento intercultural, a BNCC o aborda de diversas formas, uma vez que a cultura é um dos componentes essenciais para a formação integral dos estudantes. De maneira geral, a BNCC estabelece que a educação brasileira deve contemplar a diversidade intercultural do país, reconhecendo a pluralidade étnico-racial, de gênero, religiosa e outras manifestações culturais presentes no Brasil. O elemento intercultural está presente em diferentes componentes curriculares, como História, Geografia, Artes, Língua Portuguesa, Língua Inglesa e Educação Física, o qual prevê o desenvolvimento de habilidades relacionadas à valorização da cultura e ao respeito à diversidade, que devem ser trabalhadas em todas as disciplinas.

Na seção de Língua Portuguesa, a BNCC aborda o elemento intercultural como um dos eixos centrais para o desenvolvimento de habilidades linguísticas e culturais dos estudantes. O elemento intercultural é compreendido como a capacidade de se relacionar e interagir de forma respeitosa com outras culturas e modos de vida, reconhecendo a diversidade e a pluralidade cultural do país. Para desenvolver essa habilidade, a BNCC propõe ações que promovam a reflexão crítica sobre as diferentes formas de expressão cultural presentes no Brasil, bem como o reconhecimento das múltiplas identidades culturais que compõem a sociedade brasileira. Assim, além de trabalhar as habilidades de leitura, escrita, oralidade e compreensão, a BNCC também prevê o desenvolvimento de competências relacionadas à valorização da diversidade cultural e ao respeito às diferentes manifestações linguísticas presentes no país.

Com base no exposto, elaboramos um quadro contendo habilidades da seção de Língua Portuguesa para os anos finais do Ensino Fundamental que citam ao longo de seu texto o trabalho com o elemento cultural:

Quadro 5 – Habilidade do elemento cultural - BNCC

CAMPO DE ATUAÇÃO	CÓDIGO E HABILIDADE	SÉRIE
CAMPO JORNALÍSTICO-MIDIÁTICO	(EF69LP01) Diferenciar liberdade de expressão de discursos de ódio, posicionando-se contrariamente a esse tipo de discurso e vislumbrando possibilidades de denúncia quando for o caso.	6º ao 9º
	(EF89LP02) Analisar diferentes práticas (curtir, compartilhar, comentar, curar etc.) e textos pertencentes a diferentes gêneros da cultura digital (meme, gif, comentário, charge digital etc.) envolvidos no trato com a informação e opinião, de	8º e 9º

CAMPO DE ATUAÇÃO NA VIDA PÚBLICA	<p>forma a possibilitar uma presença mais crítica e ética nas redes.</p> <p>(EF69LP21) Posicionar-se em relação a conteúdos veiculados em práticas não institucionalizadas de participação social, sobretudo àquelas vinculadas a manifestações artísticas, produções culturais, intervenções urbanas e práticas próprias das culturas juvenis que pretendam denunciar, expor uma problemática ou “convocar” para uma reflexão/ação, relacionando esse texto/produção com seu contexto de produção e relacionando as partes e semioses presentes para a construção de sentidos.</p>	6º ao 9º
CAMPO ARTÍSTICO-LITERÁRIO	<p>(EF69LP44) Inferir a presença de valores sociais, culturais e humanos e de diferentes visões de mundo, em textos literários, reconhecendo nesses textos formas de estabelecer múltiplos olhares sobre as identidades, sociedades e culturas e considerando a autoria e o contexto social e histórico de sua produção.</p> <p>(EF69LP49) Mostrar-se interessado e envolvido pela leitura de livros de literatura e por outras produções culturais do campo e receptivo a textos que rompam com seu universo de expectativas, que representem um desafio em relação às suas possibilidades atuais e suas experiências anteriores de leitura, apoiando-se nas marcas linguísticas, em seu conhecimento sobre os gêneros e a temática e nas orientações dadas pelo professor.</p>	6º ao 9º
TODOS OS CAMPOS DE ATUAÇÃO	(EF69LP55) Reconhecer as variedades da língua falada, o conceito de norma-padrão e o de preconceito linguístico.	6º ao 9º

Fonte: BNCC (2018)

No campo jornalístico-midiático, a habilidade EF69LP01 envolve distinguir entre liberdade de expressão e discurso de ódio, essencial para criar um ambiente respeitoso em espaços sociais. Barbosa (2009) destaca que, uma vez que o léxico está intrinsecamente relacionado com a realidade extralinguística, ou seja, a cultura, é fundamental compreender o valor que as palavras podem adquirir quando utilizadas. Dessa forma, os alunos precisam ser capazes de identificar discursos que promovam discriminação, preconceito e violência, e denunciá-los quando necessário. O desenvolvimento dessa habilidade permite aos estudantes compreender que a liberdade de expressão deve ser exercida com responsabilidade e respeito pelos direitos humanos. Também ajuda a cultivar o diálogo e a empatia, respeitando a diversidade de opiniões e crenças, mas sem tolerar discursos que atentem contra a dignidade e os direitos humanos. O desenvolvimento dessa habilidade pode ajudar a formar cidadãos mais conscientes, críticos e ativos, contribuindo para a construção de uma sociedade mais justa e democrática.

Por sua vez, a habilidade EF89LP02 envolve a análise crítica de práticas e textos da cultura digital presentes nas redes sociais, como curtidas,

compartilhamentos, comentários, memes, gifs e charges digitais, visando uma presença ética e crítica nessas plataformas. É essencial que os alunos compreendam o impacto dessas práticas na disseminação de informações e opiniões na internet, evitando a propagação de discursos de ódio e desinformação. Além disso, essa habilidade pode contribuir para o desenvolvimento de habilidades importantes como o pensamento crítico, a empatia e o diálogo, que são fundamentais para uma convivência harmoniosa e democrática na sociedade.

No campo de atuação vida pública, a habilidade EF69LP21 relaciona-se à capacidade de posicionar-se diante de conteúdos presentes em práticas não institucionalizadas de participação social, incluindo manifestações artísticas, produções culturais, intervenções urbanas e práticas próprias das culturas juvenis. É importante ressaltar que a linguagem desempenha um papel fundamental nesse contexto, pois, além de permitir o desempenho de papéis individuais, é usada com propósitos sociais (Sarmiento, 2004). Através do desenvolvimento dessa habilidade, os alunos podem expandir sua compreensão da diversidade de expressões culturais existentes em nossa sociedade, bem como das questões sociais e políticas que essas produções procuram problematizar e debater. Adicionalmente, eles podem aprimorar sua habilidade de reflexão crítica e diálogo, bem como sua sensibilidade em relação às diferentes formas de expressão cultural e artística.

Já no campo artístico-literário, a habilidade EF69LP44 relaciona-se à capacidade de inferir valores sociais, culturais e humanos em textos literários e entender as diferentes visões de mundo apresentadas. Essa habilidade também envolve compreender como esses textos estabelecem múltiplos olhares sobre identidades, sociedades e culturas, considerando seu contexto social e histórico de produção. Ao desenvolver essa habilidade os estudantes têm a oportunidade de ampliar sua compreensão sobre as relações sociais, culturais e humanas retratadas na literatura, bem como aprimorar sua capacidade de análise crítica, interpretação e sensibilidade em relação à sua realidade e à diversidade cultural que os cerca (Candido, 2004). A consideração da autoria e do contexto de produção dos textos ajuda os alunos a compreender as intenções dos autores, os valores e ideologias presentes em suas obras literárias, e a importância dessas obras como registros históricos e culturais. Por fim, a habilidade EF69LP49 é essencial para a formação de leitores críticos e reflexivos, capazes de apreciar e compreender a complexidade e a diversidade de textos literários e produções culturais do campo. Ao mostrar-se

interessado e envolvido na leitura e estar aberto a textos que representem um desafio, os alunos podem ampliar sua visão de mundo, desenvolver habilidades de análise e interpretação crítica, e aprimorar sua capacidade de apreciação e manifestação da cultura.

A habilidade EF69LP55 é comum a todas as áreas de atuação e tem como objetivo aprimorar a capacidade do aluno em identificar as diferentes variações linguísticas presentes na língua falada, bem como compreender os conceitos de norma-padrão e preconceito linguístico. É fundamental que o estudante compreenda que a língua é um fenômeno em constante evolução, sujeita a expansões, alterações e até mesmo contrações (Biderman, 2001), resultando em diversas formas de utilização que estão intrinsecamente ligadas ao contexto sociocultural e histórico, assim como às características regionais, sociais, culturais e individuais dos falantes. É crucial que o estudante compreenda o conceito de preconceito linguístico, que consiste na discriminação de um tipo de uso da língua em detrimento de outros, com base em critérios sociais e arbitrários, como a região de origem, classe social, gênero, etnia, entre outros. Desenvolver tal habilidade é essencial para que o aluno compreenda que todas as formas de utilização da língua são igualmente válidas e merecem respeito, sem preconceitos ou discriminações.

Para exemplificar e desenvolver habilidades citadas acima, uma atividade interessante seria a leitura e análise de um conto literário que aborda temas relevantes para a sociedade, como preconceito, discriminação, liberdade de expressão, entre outros. "O Negrinho Pastoreiro" de João Simões Lopes Neto é um exemplo de conto literário que pode ser utilizado para essa atividade, já que trata de questões como racismo, exploração infantil e o papel da religião na cultura brasileira. Além de analisar o enredo e a mensagem do conto, os alunos podem discutir a relevância desses temas na sociedade atual e como as questões apresentadas no texto ainda são relevantes nos dias de hoje.

Para essa atividade, sugere-se que os alunos leiam o conto individualmente e em seguida, discutam em grupo suas impressões sobre o texto e os temas abordados. Depois, os alunos podem fazer uma análise mais aprofundada do conto, identificando elementos literários como narrador, personagens, espaço e tempo, além de destacar trechos que chamaram atenção e refletir sobre a mensagem do autor. A partir dessa análise, os alunos podem discutir a relevância dos temas abordados no conto na sociedade atual, fazendo uma conexão com situações vividas no cotidiano. Também

podem realizar uma pesquisa sobre a história e o contexto da época em que o conto foi escrito para entender melhor a mensagem do autor. Por fim, a atividade pode culminar em produções escritas individuais ou em grupo, em que os alunos expressem suas reflexões sobre os temas abordados no conto.

Em síntese, a leitura e análise crítica do conto "O Negrinho Pastoreiro" é uma atividade enriquecedora para os alunos, pois não só os ajuda a desenvolver habilidades críticas e reflexivas, como também os estimula a refletir sobre questões sociais importantes e atuais. É fundamental que os educadores incluam em sua prática pedagógica atividades que desafiem os alunos a pensar criticamente sobre o mundo em que vivemos, e o conto em questão é uma excelente ferramenta para isso. Com base na análise do documento BNCC, é possível concluir que ele é de extrema importância para o contexto educacional brasileiro. Embora seja necessário melhorar a ênfase dada ao léxico na seção de Língua Portuguesa, o espaço dedicado ao trabalho com o texto literário e a formação intercultural é muito valorizado. Esses aspectos são fundamentais para a formação integral dos alunos, já que permitem a reflexão sobre temas sociais relevantes e o desenvolvimento de habilidades críticas e conscientes. Além disso, o trabalho com textos literários e a valorização dos elementos culturais contribuem para a ampliação do repertório cultural dos alunos e para a construção de uma sociedade mais plural e inclusiva. Portanto, é essencial que a BNCC continue sendo um guia para a educação brasileira, buscando sempre aprimorar seus aspectos e promover uma educação mais qualificada e inclusiva para todos.

Quanto ao conto literário, o documento da BNCC valoriza-o como parte integrante da formação literária dos estudantes, recomendando sua abordagem no ensino de Língua Portuguesa em diferentes etapas de ensino. Destaca-se também a importância de proporcionar aos alunos o contato com diversos tipos de textos e autores, incluindo o conto, como uma forma de enriquecer seu repertório linguístico e cultural, bem como desenvolver a compreensão das características desse gênero literário, além de promover habilidades de leitura crítica e interpretação.

Dessa forma, o professor se sente encorajado a desenvolver atividades didático-pedagógicas embasadas em textos literários, visando fomentar suas aulas e cultivar o gosto pela leitura, a fim de estimular o interesse dos estudantes pela literatura. Ao seguir as orientações da BNCC e trabalhar com contos literários de forma criativa e contextualizada, o professor contribui para o desenvolvimento das

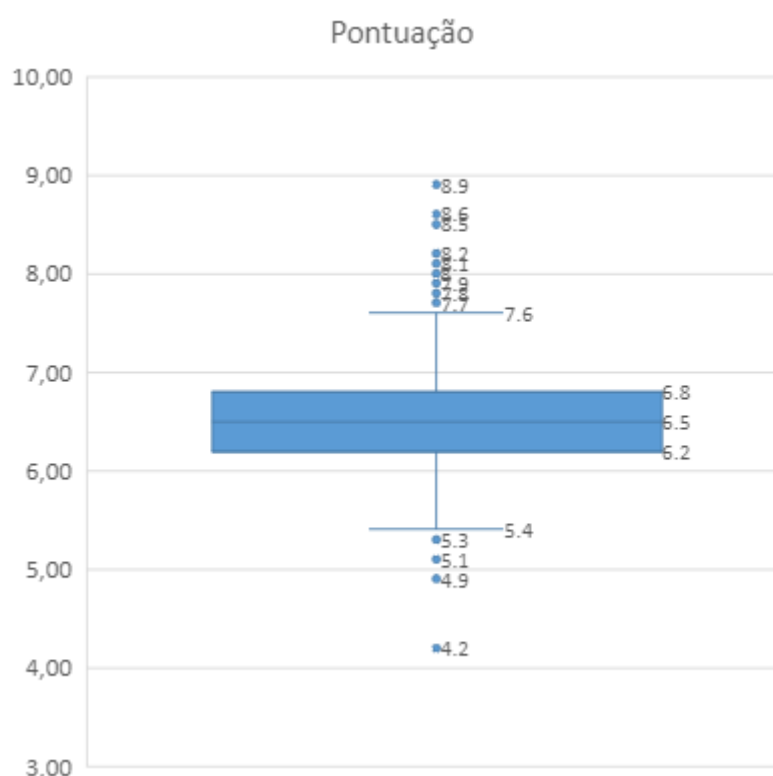
competências leitoras dos estudantes, estimula sua imaginação, criatividade e senso crítico, e desperta neles o interesse pela literatura como forma de expressão artística e de compreensão do mundo. Dessa forma, o trabalho com contos literários pode se tornar uma ferramenta valiosa para a formação integral dos estudantes, auxiliando no desenvolvimento de suas habilidades linguísticas, cognitivas, sociais e culturais, e enriquecendo sua vivência literária e sua relação com a leitura.

4.2 Dados fornecidos pelo NILC-Metrix

A seguir, apresentamos uma seção da planilha que contém os resultados dos cálculos linguístico-estatísticos obtidos através do extrator NILC-Metrix. Esses resultados foram gerados a partir de 10 métricas, as quais foram categorizadas em informações morfossintáticas, medidas descritivas, simplicidade textual, diversidade lexical, complexidade sintática, frequência de palavras e índice de leitura. Para determinar o nível léxico-sintático dos contos literários do corpus, cada métrica recebeu um peso entre 0,5 e 3,0, totalizando um peso total de 16,5. (Quadro 3).

Dessa forma, a distribuição dos contos foi apresentada em um diagrama de caixa, como pode ser visto na Figura 9.

Figura 9 - *Boxplot* com a distribuição dos contos



Fonte: A autora.

Ao analisarmos o gráfico, podemos perceber que os valores obtidos variam de 4.2 a 8.9 em uma escala de 0 a 10, tendo uma mediana de 6.5. Os contos considerados fáceis estão representados do limite inferior até o primeiro quartil. A caixa retangular entre o primeiro quartil e o terceiro quartil representa os contos considerados moderados. Já os contos difíceis estão situados do terceiro quartil ao limite superior. Os pontos que ultrapassam os limites inferior e superior são conhecidos como *outliers*, ou valores discrepantes, representando respectivamente, as categorias muito fáceis e muito difíceis.

Para uma melhor visualização da distribuição dos contos entre as cinco categorias, preparamos a tabela a seguir:

Tabela 2 - Distribuição dos contos

Categoria	Intervalo de pontuação	Quantidade	Porcentagem
Muito fácil	4.2 – 5.3	5	0,6%
Fácil	5.4 – 6.1	211	24,0%
Moderado	6.2 – 6.8	484	55,1%
Difícil	6.9 – 7.6	160	18,2%
Muito difícil	7.7 – 8.9	18	2,1%

Fonte: A autora.

A tabela exibida indica que apenas 0,6% dos contos foram classificados como "muito fáceis", enquanto 24% foram considerados "fáceis". A maioria dos contos da amostra foi classificada como "moderados", correspondendo a 55,1% do total. Em contrapartida, os contos "difíceis" representam 18,2%, enquanto os "muito difíceis" correspondem a 2,1% do total de contos avaliados. O conto mais fácil de nossa pesquisa, com a menor pontuação de 4,2, foi "O purgativo" de Humberto de Campos. Acreditamos que isso se deve ao fato de ser um diálogo breve com apenas 241 tokens e três personagens cujos nomes são repetidos ao longo da história.

Por outro lado, o conto mais difícil classificado nesta amostra é intitulado "Cabeça e coração" e foi escrito pelo Visconde de Taunay, atingindo uma pontuação de 8.9. Acreditamos que a dificuldade do conto se deve à sua extensão e estilo, apresentando um esboço psicológico longo com 5.409 tokens e uma ampla variedade de tipos de palavras, totalizando 2.120 types. Isso requer do leitor um vocabulário mais abrangente para a leitura e compreensão do conto. Além disso, essa exigência implica que o leitor tenha um vocabulário mais amplo para ler e entender os contos de Visconde de Taunay, que não só conquistou o primeiro lugar, mas também ocupa o

segundo lugar com o conto "Rapto original". Isso reforça as características do autor, reconhecido por seu estilo pitoresco e pelo uso de termos e expressões típicas e regionais em suas obras.

As porcentagens obtidas por meio da ferramenta computacional estão de acordo com nossas expectativas iniciais. A categoria moderada correspondeu às nossas previsões, sendo a categoria com a maior percentual, com 55,1% do corpus. No entanto, as categorias fácil e difícil também estiveram dentro de nossas expectativas, embora tenha sido esperado um índice maior na porcentagem de contos considerados difíceis. Ficamos surpresos com a discrepância nas categorias que representam os pontos extremos, muito fácil e muito difícil, o que nos incentiva a investigar em estudos futuros as razões e critérios que podem ter influenciado nesses resultados, que ficaram abaixo de nossa expectativa. Será importante aprofundar a análise para compreender as possíveis variações de dificuldade dos itens avaliados e identificar fatores que possam ter contribuído para essa classificação.

Destacamos o uso de ferramentas computacionais, como o NILC-Matrix, que pertence à área de Linguística de Corpus e se baseia em evidências empíricas obtidas de corpus, conforme mencionado por Berber Sardinha (2004). Essa ferramenta é especialmente útil para pesquisadores e educadores, pois possibilita uma análise mais aprofundada do texto, auxiliando na seleção de materiais educacionais e contribuindo para o ensino em sala de aula, como afirmado por Szudarski (2018).

Com o auxílio do NILC-Matrix, foi possível obter informações detalhadas sobre o uso real das palavras em contextos autênticos, o que pode guiar a escolha de materiais didáticos mais relevantes e adequados às necessidades dos estudantes, em nosso caso contos literários. O uso dessas ferramentas pode aprimorar a qualidade do ensino, proporcionando uma compreensão mais aprofundada da descrição vocabular e sua aplicação prática no contexto educacional. Essa abordagem pode enriquecer a experiência de aprendizagem dos estudantes, permitindo-lhes desenvolver habilidades de leitura e compreensão vocabular de forma mais eficaz e contextualizada, como afirma Bezerra (2018).

Após a análise dos níveis de complexidade dos contos literários selecionados, a próxima etapa será a aplicação de um questionário de avaliação com alunos que cursam o ensino fundamental anos finais. O objetivo principal desse questionário é coletar as percepções dos alunos quanto ao nível de dificuldade léxico-sintático dos contos literários, a fim de estabelecer uma correlação com as classificações obtidas

anteriormente pelo software. Por meio das respostas dos alunos, será possível avaliar se a classificação atribuída pelo software realmente reflete a percepção dos leitores em relação à dificuldade dos contos e se as dificuldades léxico-sintáticas dos textos estão compatíveis com o nível de leitura dos alunos.

Abaixo, o quadro com a classificação dos contos, incluindo os cinco primeiros classificados de cada categoria. O quadro completo pode ser visualizado no Apêndice A.

Quadro 6 – Contos categorizados como muito fáceis

#	Autor	Título	Pontuação
1	Humberto de Campos	O purgativo	4,2
2	Fernando Sabino	O gato sou eu	4,9
3	Olavo Bilac	Como os cães	5,1
4	Lima Barreto	Na janela	5,1
5	Alcântara Machado	O tímido José	5,3

Fonte: Dados da presente pesquisa.

O Quadro 4 exibe os contos categorizados como "muito fáceis" pelo software NILC-Metrix, levando em consideração que quanto menor a pontuação, menor a complexidade. As pontuações variam de 4,2 a 5,3, indicando uma baixa complexidade desses contos, de acordo com a avaliação feita pelo software.

O conto "O purgativo", de autoria de Humberto de Campos, conhecido por seu estilo de escrita de fácil compreensão, recebeu a pontuação mais baixa na análise. Trata-se de uma história cômica composta por apenas 241 palavras, divididas em 56 sentenças com estruturas sintáticas simples, indicadas pelas métricas de avaliação. Com uma pontuação de 4,9, o conto "O gato sou eu", de Fernando Sabino, obteve a segunda posição como o conto mais fácil na análise. Embora seja um pouco mais extenso em comparação ao conto de Campos, ele apresenta uma baixa quantidade de tipos de palavras (types) e de palavras de conteúdo, como substantivos, adjetivos e verbos, que são consideradas palavras de classe aberta. Isso resulta em um índice geral de complexidade baixo para o conto.

Os contos de Olavo Bilac e Lima Barreto intitulados respectivamente como "Como os cães" e "Na janela" receberam a mesma pontuação de 5,1. Embora o conto de Lima Barreto seja um pouco mais longo, o que o torna mais complexo, o conto de Bilac possui uma média de palavras por sentença mais alta, o que também influencia no seu grau de complexidade, resultando na mesma pontuação para ambos.

Finalmente, o conto "O tímido José" de Alcântara Machado obteve uma pontuação de 5,3, colocando-o como o quinto conto mais fácil nessa análise. Contrariando as expectativas de que contos mais curtos seriam mais fáceis, o conto de Alcântara Machado tem um tamanho considerável, com 1.310 palavras. No entanto, as métricas indicaram um baixo nível de complexidade sintática, juntamente com uma baixa média de palavras por sentença, o que o classificou como muito fácil, apesar de sua extensão.

Na sequência, no Quadro 5, apresentamos cinco contos da categoria "fácil", todos avaliados com a pontuação de 6,1.

Quadro 7 – Contos categorizados como fáceis

#	Autor	Título	Pontuação
1	Alcântara Machado	O Lírico Lamartine	6,1
2	Carlos Drummond de Andrade	A melhor opção	6,1
3	Amadeu Amaral	Brinquedo	6,1
4	Olavo Bilac	O defunto	6,1
5	Machado de Assis	O sermão do diabo	6,1

Fonte: Dados da presente pesquisa.

O conto "O lírico Lamartine" de Alcântara Machado, embora seja mais curto que "O tímido José", apresenta uma sintaxe um pouco mais complexa de acordo com a métrica. No entanto, apesar dessa complexidade, a falta de variedade de tipos e palavras de classe aberta torna-o fácil de ser compreendido, o que explica a sua pontuação. O mesmo se aplica ao conto de Carlos Drummond de Andrade, intitulado "A melhor opção", que, apesar de ser breve e ter uma baixa contagem na quantidade de palavras, obteve um número alto na métrica que contabiliza a média de palavras por frase.

Os contos "Brinquedo" de Amadeu Amaral e "O defunto" de Olavo Bilac apresentam uma variedade de tipos de palavras ligeiramente mais alta do que aqueles classificados no quadro 4. Além disso, possuem uma média mais alta de palavras por sentença e a presença de conjunções coordenativas, o que eleva a complexidade sintática, comparados aos anteriores.

Por sua vez, o conto "O sermão do diabo" de Machado de Assis é a obra com a maior pontuação na métrica que analisa a complexidade sintática dos cinco contos, com uma média total de palavras por sentença de 12, o que indica uma maior

complexidade sintática. No entanto, as métricas de variedade de palavras, classe abertas e variedades de adjetivos obtiveram baixos números, o que equilibrou a média final de 6,1 atribuída ao conto.

Embora Machado de Assis seja amplamente reconhecido como um cânone da literatura e seja conhecido por sua linguagem complexa em algumas obras, o autor também tem títulos mais acessíveis que não requerem tanto conhecimento linguístico para serem lidos e interpretados. "O sermão do diabo" é um exemplo disso, pois apesar de ter uma complexidade sintática maior, ainda é um conto relativamente fácil de entender e apreciar.

No quadro 6, temos os contos considerados como moderados apresentando pontuação de 6,8.

Quadro 8 – Contos categorizados como moderados

#	Autor	Título	Pontuação
1	Raul Pompéia	Conto de fadas	6,8
2	José Geraldo Vieira	Uma operação gratuita	6,8
3	Machado de Assis	Verba testamentária	6,8
4	Hugo de Carvalho Ramos	A madre de ouro	6,8
5	Arthur Azevedo	A Marcelina	6,8

Fonte: Dados da presente pesquisa.

Essa categoria já apresenta contos mais extensos superiores a 800 palavras, o que já requer do autor preparação para ler textos mais extensos. Devido ao tamanho dos textos, é esperado que o número de sentenças também aumente o que foi confirmado pela métrica. A complexidade sintática e a variação de types também elevou um ponto nas métricas, o que demonstra um aumento da dificuldade léxico-sintática dos contos dessa categoria comparado as anteriores já apresentadas.

No sétimo quadro, temos os contos difíceis:

Quadro 9 – Contos categorizados como difíceis

#	Autor	Título	Pontuação
1	José J. Veiga	O galo impertinente	7,5
2	Arthur Azevedo	Black	7,6
3	Alcides Maya	O estaqueado	7,6
4	Visconde de Taunay	O estorvo	7,6
5	Gonzaga Duque	Idílio Roxo	7,6

Fonte: Dados da presente pesquisa.

“O galo impertinente” do autor goiano José J. Veiga apresenta 7,5 de pontuação sendo considerado um conto difícil. Embora não seja tão extenso quanto outros contos da mesma categoria, o conto de Veiga apresenta poucas sentenças, com uma média de palavras por sentença bastante elevada, resultando em uma complexidade sintática superior aos contos anteriores. Além disso, a média de palavras antes do verbo principal da oração é alta, o que também contribui para essa elevada complexidade. Por outro lado, a complexidade lexical não é tão alta, indicando que o vocabulário utilizado não é muito diversificado, uma vez que as métricas de variação de tipos de palavras, adjetivos e classes abertas não tiveram diferença significativa comparada a categoria anterior.

Apesar de possuir apenas 697 palavras, o conto "Black" de Arthur Azevedo apresenta uma complexidade sintática considerável e um grande número de conjunções coordenativas, o que pode desafiar os leitores. Por sua vez, o conto de Alcides Maya é ainda mais extenso e foi categorizado como difícil, apresentando uma quantidade maior de sentenças e uma média de palavras por sentença mais elevada, resultando também em um aumento das conjunções coordenativas. Além disso, o conto possui uma riqueza de variação de palavras que eleva a complexidade lexical, exigindo que o leitor possua um vocabulário mais amplo.

Ao final, nos deparamos com os contos de Visconde de Taunay e Gonzaga Duque, cuja extensão demanda dos leitores disposição para uma leitura prolongada, o que pode fatiga-los caso não estejam habituados. O conto de Visconde de Taunay, embora apresente uma sintaxe mais complexa, contém menos variedade de palavras em relação ao conto de Gonzaga Duque. Ambos os escritores são reconhecidos por um estilo rebuscado, que se utiliza de arcaísmos, regionalismos e expressões idiomáticas típicas da época, pouco comuns no vocabulário cotidiano, o que eleva a dificuldade dos contos e exige uma leitura mais atenta e concentrada.

Além disso, é importante ressaltar que os contos de Visconde de Taunay e Gonzaga Duque são obras representativas do movimento literário do Realismo no Brasil. Ambos os autores abordam em seus contos temas sociais, políticos e históricos, retratando a realidade brasileira da época. Desse modo, a leitura dessas obras não só exige uma boa compreensão da linguagem rebuscada utilizada pelos autores, mas também um conhecimento prévio do contexto histórico e social da época em que foram escritos.

Chegamos ao oitavo quadro que abrange os contos mais desafiadores. Conforme mencionado anteriormente, Visconde de Taunay e Gonzaga Duque são reconhecidos por sua complexidade, portanto, era esperado que esses autores também estivessem incluídos na categoria de maior grau de dificuldade nesta pesquisa.

Quadro 10 – Contos categorizados como muito difíceis

#	Autor	Título	Pontuação
1	Visconde de Taunay	Uma vingança	8,2
2	Gonzaga Duque	Agonia por semelhança	8,2
3	Euclides da Cunha	Judas-Ahsverus	8,5
4	Visconde de Taunay	Rapto original	8,5
5	Visconde de Taunay	Cabeça e coração	8,9

Fonte: Dados da presente pesquisa.

Os contos de maior extensão ficaram nessa categoria, conseqüentemente aumentando as métricas de palavras e sentenças.

Os contos "Judas-Ahsverus" de Euclides da Cunha e "Rapto Original" de Visconde de Taunay também receberam uma avaliação de 8,5 cada um. Embora o conto de Taunay seja mais longo em termos de número de palavras e sentenças, o conto de Cunha envolve o uso de estruturas gramaticais mais elaboradas, que vão além das frases simples e diretas, com um número menor de sentenças porém com uma média maior de palavras, o que aumenta o nível de dificuldade e uma variedade de vocabulário ligeiramente mais elevada.

De acordo com a análise, o conto intitulado "Cabeça e Coração" do Visconde de Taunay é o mais complexo, obtendo uma pontuação de 8.9. Este conto em particular possui 5.330 palavras, sendo o maior dentro do corpus analisado. A obra de Taunay apresenta uma média elevada de palavras por sentença e alta complexidade sintática, além de uma grande variedade de tipos de palavras e uma classe de palavras abertas, o que contribui para a complexidade léxico-sintática presente no texto.

4.3 Análise do questionário

O questionário foi elaborado na plataforma *Google Forms*, tendo sido escolhido um conto representativo de cada categoria de complexidade, totalizando cinco contos

distintos. Entretanto, após a seleção dos contos, percebeu-se que alguns deles apresentavam uma extensão considerável, o que poderia tornar a leitura exaustiva para os participantes. Levando em consideração a importância da leitura integral dos contos para a realização da atividade, optamos por dividir o questionário em cinco partes, garantindo assim que os participantes pudessem ler cada conto com a atenção e dedicação necessárias para uma avaliação precisa da complexidade literária. Dessa forma, foi possível obter resultados mais confiáveis e consistentes.

Sobre a organização do questionário, seguiu a sequência de disponibilizar o conto completo para leitura e solicitar a avaliação da complexidade em cinco níveis.

A tabela a seguir apresenta o resultado dos questionários em relação aos contos analisados, classificando-os de acordo com o nível de dificuldade percebido pelos alunos. Os níveis de complexidade variam de 1 a 5 para classificação, sendo 5 (muito difícil), 4 (difícil), 3 (moderado), 2 (fácil) e 1 (muito fácil).

Tabela 3 – Resultado dos questionários

Conto	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
Cabeça e coração Visconde de Taunay	5	4	4	5	5	5	5	4	5	5
O estaqueado Alcides Maya	5	4	4	5	5	4	5	5	5	4
Quem tudo quer, tudo perde Coelho Neto	3	3	2	3	3	3	4	4	3	4
Vista cansada Otto Lara Resende	2	2	1	2	3	1	3	2	3	3
O purgativo Humberto de Campos	1	1	1	1	1	1	2	1	2	1

Fonte: A autora

O conto "Cabeça e coração" de Visconde de Taunay foi classificado por 70% dos alunos como um conto muito difícil, e por 30% como difícil. Os resultados correspondem satisfatoriamente aos resultados do NILC-matrix, o qual o classificou como o conto mais difícil de nossa análise. Fatores como o uso de linguagem formal e sofisticada, a trama complexa e extensa, a estrutura não linear com diferentes linhas do tempo e a presença de metáforas que exigem conhecimento prévio do contexto

histórico e cultural em que o conto foi escrito, são razões que devem ter sido consideradas ao avaliar a dificuldade do conto.

Para contextualizarmos, abaixo temos um pequeno trecho do conto “Cabeça e coração” que corresponde a fala de Bettina, protagonista do conto:

Fora rematada loucura, indigna de sua reflexão, contrária a tudo quanto lhe ensina o conhecimento que tem do coração humano... Pelo menos, assim me afigura... Por mim, quero julgá-lo. São susceptibilidades de exagerado melindre, de exaltado e meticuloso cavalheirismo, que o elevam aos meus olhos... Aliás, para que e como discutir sentimentos? (TAUNAY, 1872).

Podemos notar que o trecho apresenta uma escrita sofisticada, com o uso de palavras de vocabulário refinado e uma estrutura frasal elaborada, o que eleva o nível de compreensão da fala da personagem Bettina. Além disso, a reflexão feita por Bettina aborda questões psicológicas e filosóficas profundas, especialmente em relação à dualidade entre razão e emoção, o que pode se tornar algo difícil para leitores que não possuem uma familiaridade ou uma opinião formada sobre o assunto.

Acreditamos que o conto de Taunay pode não ser a melhor opção para leitores iniciantes ou com dificuldades de leitura e interpretação. Isso se deve ao fato de que a obra apresenta uma profundidade temática e uma complexidade léxico-sintática elevada, o que demanda uma maior maturidade literária, como apontado por Zafalon (2010). Por outro lado, de acordo com as observações de Simões e Assis (2014), a riqueza lexical e o teor reflexivo sobre filosofia e psicologia presentes no conto de Taunay podem ser excelentes para leitores proficientes que buscam leituras desafiadoras, podendo contribuir para o enriquecimento vocabular e a expansão de conhecimentos como destacam as autoras.

Também foi constatado que 60% dos alunos acharam a leitura do conto "O estaqueado" de Alcides Maya muito difícil, enquanto 40% classificaram como difícil. Embora os resultados da pesquisa não tenham sido idênticos aos do NILC-Matrix, que classificou o conto "O estaqueado" como difícil, ainda assim houve uma semelhança significativa. Dentre as razões para isso, está o fato de que a obra apresenta uma linguagem elaborada e refinada, utilizando um vocabulário técnico pouco comum, o que pode dificultar a compreensão de algumas passagens, como pode ser visto nesse trecho que encerra a história do conto: “Estendera, reteso à cólera, o braço; eram-lhe áscuas vivas as pupilas, e, da cima isolada e pedregulhenta do cerro, o seu gesto de

maldição recortou-se enérgico e austero sobre a amplitude melancólica dos campos...”(MAYA, 1926)

Além disso, assim como o conto de Taunay, a narrativa de Maya é não linear e contém saltos temporais que podem confundir o leitor. Adicionalmente, o contexto histórico em que a obra foi escrita no século XIX e a abordagem de temas controversos como violência, crime e justiça podem requerer que o leitor tenha algum conhecimento prévio para apreciar a obra.

Portanto, é importante ter em mente que a escolha de leituras deve levar em consideração o nível de conhecimento dos leitores. No caso de contos complexos como esses, concordamos com Zafalon (2010) ao recomendar o trabalho com textos mais complexos com estudantes já tenham um maior domínio da língua e da literatura. Isso porque, para a apreciação plena das obras, é necessário um vocabulário amplo, uma compreensão avançada das regras sintáticas e um conhecimento prévio do contexto cultural e histórico em que foram escritas, para leitores iniciantes a complexidade da leitura pode gerar desmotivação e afastá-los da literatura, por isso é importante que o processo de leitura seja gradativo, com obras que correspondam ao nível de conhecimento e habilidades dos leitores.

Continuando nossa análise com base na tabela fornecida, podemos ver que o conto "Quem tudo quer, tudo perde" de Coelho Neto recebeu uma variedade de respostas por parte dos alunos. Dos 10 participantes que responderam, 6 classificaram como moderado, 3 como 4 difícil e 1 como fácil. Assim, 60% das respostas se correlacionam com a classificação do NILC-Metrix quanto ao conto, classificando-o como moderado. No entanto, isso também demonstra como a percepção de dificuldade pode variar de acordo com o leitor e suas habilidades e conhecimentos prévios. A seguir, um trecho retirado do conto de Coelho Neto:

Rojaram-se todos, e, esquecidos das palavras prudentes do velho, puseram-se a encher os sacos, sempre achando pouco o que guardavam. O carvoeiro levantou-se, e, com esforço, aos arrancos, arrastou seu saco até o limiar da caverna, sem poder erguê-lo, tão superior às suas forças era a carga. (NETO, 1912).

Por exemplo, os estudantes que classificaram o conto como difícil podem ter enfrentado dificuldades na compreensão do vocabulário, devido a palavras como "prudentes", "carvoeiro" e "limiar". Além disso, as construções sintáticas complexas, como orações subordinadas e pronomes oblíquos, podem ter dificultado a compreensão do enredo. Por outro lado, os estudantes que consideraram o conto

como fácil provavelmente possuem maior familiaridade com textos desse tipo e um vocabulário mais amplo. A percepção de dificuldade de um texto também pode ser influenciada por outros fatores, como o interesse do leitor pelo tema abordado e sua familiaridade com o gênero literário em questão. Isso ressalta ainda mais a importância de incentivar sempre hábitos de leitura e o contato com diferentes tipos de textos literários, uma vez que experiências variadas de leitura podem contribuir para a ampliação do vocabulário como afirmam Rodrigues (2014), Simões e Assis (2014) e Borges e Silva (2022), que por sua vez auxiliam na melhoria da compreensão de textos mais complexos.

O conto "Vista Cansada" de Otto Lara Resende também recebeu diversas classificações por parte dos alunos, resultando em um empate técnico nas respostas. Dos 10 alunos que avaliaram o conto, 4 o classificaram como moderado, enquanto outros 4 o avaliaram como fácil e 2 como muito fácil. A partir da percepção dos alunos, é possível inferir que o conto de Otto Lara Resende pode ser considerado uma leitura de nível fácil-moderado, embora tenha sido classificado como de fácil compreensão pelo NILC-Metrix. Apesar dessa divergência, é possível notar que o conto apresenta uma linguagem mais acessível e direta em comparação a outros contos, como o de Taunay, como pode ser observado no trecho abaixo:

Lá estava sempre, pontualíssimo, o mesmo porteiro. Dava-lhe bom-dia e às vezes lhe passava um recado ou uma correspondência. Um dia o porteiro cometeu a descortesia de falecer. Como era ele? Sua cara? Sua voz? Como se vestia? Não fazia a mínima ideia. Em 32 anos, nunca o viu. (RESENDE, 1963)

O tema comum do envelhecimento é tratado de maneira clara e as personagens são bem definidas em uma narrativa linear. Isso é corroborado pelo fato de que o conto não foi classificado como difícil em nenhuma das 10 avaliações, mas por outro lado, foi considerado muito fácil por duas delas. Essa avaliação pode ser explicada pela simplicidade da trama e pelo uso de uma linguagem acessível, com frases curtas e baixa complexidade sintática, o que torna a leitura mais fluida e compreensível para diferentes públicos.

Por fim, o conto intitulado "O purgativo", escrito por Humberto de Campos, foi classificado como o mais simples de todos os contos analisados pelo NILC-Metrix. Além disso, de acordo com os resultados obtidos a partir do questionário, este conto foi considerado o mais fácil pelos alunos, em comparação aos outros quatro contos avaliados. Conforme os dados do questionário, 80% dos alunos acharam o conto

muito fácil, enquanto 20% o classificaram como fácil, representando o mais alto índice de correlação entre o NILC-Metrix e o questionário. Características como a extensão da obra, história simples e sem reviravoltas com poucos personagens, e uma linguagem clara e acessível são alguns dos motivos que tornam o conto "O purgativo" fácil de ser compreendido. Como pode ser observado no trecho abaixo:

Mãe: Ainda não fez efeito, doutor! Nunca vi uma coisa assim! Médico: Nenhum efeito? Mãe: Absolutamente nenhum. Eu chego a pensar que o farmacêutico talvez se tenha enganado ao preparar a receita. (CAMPOS, 1922)

Acreditamos que os indicadores apresentados sugerem que "O purgativo" é uma excelente escolha para leitores iniciantes em busca de uma leitura agradável e de fácil compreensão, que também auxilie no desenvolvimento de habilidades linguísticas. Além disso, é uma opção interessante no contexto educacional, onde os professores buscam incentivar a leitura entre os alunos. A linguagem simples e a trama descomplicada do conto podem torná-lo menos intimidante para os que estão começando a se aventurar na ficção, e sua brevidade e facilidade de compreensão podem auxiliar no desenvolvimento de habilidades de interpretação textual e análise literária, bem como no enriquecimento do vocabulário dos estudantes. Por isso, é uma opção atraente para despertar o interesse inicial pela literatura entre os jovens estudantes, devendo ser destacado durante as aulas de literatura, como defende Zafalon (2010).

Com base nos dados apresentados na tabela 3, podemos responder à terceira questão estabelecida nesta pesquisa: há correlação entre os dados do software NILC-Metrix e as percepções dos alunos obtidas a partir do questionário sobre os contos. Analisando primeiramente os contos, temos a obra "Cabeça e coração" de Visconde de Taunay é considerado muito difícil e esperávamos que cada aluno atribuísse a ele uma nota 5, totalizando 50 pontos, porém a pontuação total foi de 47 pontos, uma diferença de 3 pontos. Já o conto "O estaqueado" de Alcides Maya, esperávamos uma pontuação de 40, mas a pontuação total foi de 46 pontos, uma diferença de 6 pontos a mais em relação ao NILC-Metrix.

A obra "Quem tudo quer, tudo pode" de Coelho Neto recebeu uma pontuação de 32 pontos, superando a expectativa de 30 pontos. Já a obra "Vista Cansada" de Otto Lara Resende obteve uma pontuação de 22 pontos, ultrapassando a expectativa de 20 pontos, enquanto "O purgativo" de Humberto de Campos recebeu 12 pontos,

excedendo a expectativa de 10 pontos. Como resultado, as três obras obtiveram uma pontuação 2 pontos acima do esperado pelos avaliadores.

Embora as notas não correspondam exatamente à classificação obtida pela abordagem empírica do NILC-Matrix, a avaliação subjetiva baseada em questionários gerou pontuações próximas às da pesquisa. Essa discrepância pode ter sido causada por vários fatores, incluindo diferenças no nível de leitura, dificuldades na interpretação das obras, familiaridade com o contexto cultural e variações nas expectativas dos alunos. Isso fica evidente quando examinamos as pontuações individuais dos alunos, que variaram de 12 a 19, apesar da expectativa de uma pontuação total de 15 para cada aluno após a avaliação dos cinco contos. Essa variação sugere a existência de diferentes perfis de alunos, o que gera também múltiplas percepções.

Com base em nossas análises, concluímos que, apesar das diferenças identificadas, o itinerário apresenta um grande potencial para orientar as escolhas dos professores na seleção de contos literários a serem desenvolvidos em sala de aula. Os resultados promissores obtidos a partir dessa abordagem sugerem que a ferramenta NILC-Matrix pode ser uma adição útil ao conjunto de recursos disponíveis para leitores e professores que desejam ressignificar suas práticas e atividades leitoras, tornando o trabalho com o texto literário mais significativo e enriquecedor para seus alunos.

Levando em consideração a discussão dessa seção, alinhamos nossas observações à BNCC, que destaca a importância da literatura e do uso de textos literários nas aulas de língua portuguesa, sendo fundamental abordar as habilidades necessárias para a formação literária ao longo dos anos finais do ensino fundamental, do 6º ao 9º ano.

Nesse sentido, o itinerário de leitura de contos literários elaborado na pesquisa oferece um caminho claro para ser aplicado nas escolas, uma vez que a BNCC prevê em suas habilidades o trabalho com diferentes autores e obras literárias. As categorias de complexidade presentes no itinerário, que variam de muito fácil a muito difícil, são outro aspecto positivo, pois permitem ao professor selecionar obras adequadas ao nível de conhecimento dos alunos em diferentes estágios de sua trajetória escolar. Dessa forma, é possível estimular o hábito da leitura e o desenvolvimento de habilidades linguísticas ao longo do tempo, de maneira progressiva e adequada às necessidades de cada turma.

Embora a BNCC possua uma abordagem tímida em relação ao léxico, em comparação com as expectativas de uma abordagem ideal, o professor pode preencher esse espaço propondo atividades que sejam contextualizadas, adequadas e desafiadoras para seus alunos. Uma maneira eficaz de fazer isso é utilizar contos literários, considerando os diferentes níveis de complexidade da linguagem presente nesses textos.

Ao propor atividades contextualizadas que envolvam o uso do léxico, o professor pode auxiliar os alunos a expandir seu vocabulário e compreender as nuances da língua de forma mais efetiva. Para alunos iniciantes, é adequado selecionar obras de baixa complexidade, como contos da categoria "muito fácil" ou "fácil", como por exemplo o conto "O purgativo" de Humberto de Campos. No entanto, se a turma for mais avançada, o professor pode buscar obras que desafiem a capacidade leitora e interpretativa dos estudantes, recorrendo às categorias "difícil" ou "muito difícil", como por exemplo o conto "Cabeça e coração" de Visconde de Taunay.

Dessa forma, ao utilizar contos literários em suas atividades, o professor pode promover o desenvolvimento integrado e abrangente das habilidades de leitura, escrita e compreensão da língua, em conjunto com a literatura.

Por fim, o elemento intercultural é considerado pela BNCC como um eixo central no desenvolvimento de habilidades linguísticas e culturais. Os contos literários são uma forma de expressão cultural que reflete a diversidade de experiências humanas ao redor do mundo. Por meio deles, os estudantes têm a oportunidade de se familiarizar com diferentes culturas, tradições, costumes e perspectivas de vida, enriquecendo seu repertório cultural e linguístico.

Ao explorar contos literários de diferentes autores e épocas, os estudantes podem desenvolver habilidades linguísticas, como a compreensão de vocabulário e estruturas gramaticais específicas. Além disso, ao analisar as narrativas, os estudantes podem aprimorar suas habilidades de interpretação de texto, compreensão de contexto, inferência e análise crítica, contribuindo para o desenvolvimento de habilidades de leitura e escrita.

CONSIDERAÇÕES FINAIS

Sob o enfoque das linhas teóricas da Lexicologia e da Linguística de Córpus, a pesquisa em questão buscou explicitar a importância em se considerar a complexidade léxico-sintática para a seleção de textos que fomentem a prática de leitura e conseqüentemente o desenvolvimento de habilidades linguísticas, como a aquisição vocabular.

Além disso, por meio da construção de um córpus composto exclusivamente por contos literários brasileiros, a pesquisa buscou destacar o papel essencial que a literatura brasileira pode desempenhar na formação humanizadora e cultural dos leitores, especialmente no cenário educacional, com a promoção do contato com a literatura.

Para alcançar esse objetivo, primeiramente, foi necessário nos basearmos no contexto educacional atual, no qual os alunos estão cada vez mais desmotivados a praticar o hábito da leitura, o que conseqüentemente os afasta da literatura. Essa falta de contato com a literatura pode resultar em dificuldades em atividades escolares, como a interpretação e produção de textos, uma vez que eles não conseguem expandir seu conhecimento lexical devido à ausência de leitura.

Logo, considerando que a prática de leitura e contato com textos contribuem para o desenvolvimento de competências, foi criado um Itinerário de Leitura de Contos Literários que contendo cinco níveis de complexidade léxico-sintática. Acredita-se que, dessa forma, o objetivo geral da pesquisa foi alcançado. Com o intuito de cumprir nossos objetivos específicos, elaboramos um corpus representativo de contos literários brasileiros. Utilizando o extrator NILC-Matrix, pudemos calcular e identificar o perfil léxico-sintático dos contos. Além disso, foi aplicado questionário de percepção aos alunos da segunda fase do Ensino Fundamental, contemplando os objetivos dessa pesquisa.

Nessa pesquisa, foram definidos três problemas para serem abordados. O primeiro, diz respeito à abordagem da tríade contos, léxico e interculturalidade na Base Nacional Comum Curricular (BNCC), concluiu-se que a BNCC está atenta a esses elementos. Verificou-se que o documento enfatiza a importância da formação literária e da consideração de elementos culturais durante as aulas. Contudo, é necessário ressaltar que é preciso realizar reformulações para que haja uma maior valorização do léxico durante as aulas de língua portuguesa, deixando de ser

considerado de maneira restrita, o que compromete o reconhecimento e valorização dos estudos lexicais.

Com relação ao segundo problema proposto, assumimos a responsabilidade de esclarecer como foi realizada a distribuição dos 878 contos em cinco categorias com base em sua complexidade léxico-sintática. A categoria "Muito Fácil" teve a menor quantidade de contos, apenas 5, correspondendo a 0,6% do corpus. Já a categoria "Fácil" teve 211 contos, representando 24%. A categoria "Moderado" foi a maior, com 484 contos, o que corresponde a 55,1% do corpus analisado. A categoria "Difícil" teve 160 contos, 18,2%, e a categoria "Muito Difícil" obteve 18 contos, o que corresponde a 2,1% do corpus.

Por fim, o terceiro problema proposto buscou identificar se há correlação entre as avaliações parciais obtidas pelo NILC-Matrix, abordagem empírica, e os resultados dos questionários respondidos pelos alunos, que representam uma abordagem subjetiva. Embora os números obtidos não tenham coincidido exatamente, variando em média 4 pontos por conto, existe sim uma correlação entre os dados, uma vez que as pontuações foram próximas

Acreditamos que a realização de uma sequência didática com os contos literários selecionados no questionário poderia levar a resultados mais semelhantes aos apontados pelo NILC-Matrix. Essa abordagem poderia ajudar a aprimorar a compreensão dos alunos sobre as narrativas e desenvolver sua capacidade de análise. Entretanto, para que essa proposta fosse realizada de maneira efetiva, seria necessário um planejamento cuidadoso e tempo adequado para sua execução, o que não foi possível devido ao curto espaço de tempo disponível.

Ainda assim, apesar das dificuldades encontradas, cremos na potencialidade do Itinerário aqui proposto, para orientar professores e leitores na seleção de futuras leituras a serem realizadas.

Destacamos a relevância da Linguística de Córpus no processamento eficiente de grandes volumes de dados linguísticos. A contribuição de suas reflexões teórico-metodológicas foi indispensável para o desenvolvimento e obtenção do itinerário de leitura em nossa pesquisa. Por meio de ferramentas como o NILC-Matrix, conseguimos processar um corpus com mais de um milhão de palavras de maneira ágil e confiável, o que possibilitou uma análise mais rápida e precisa dos dados.

Acreditamos ter contribuído para enriquecer as discussões acerca do tratamento do texto em sala de aula e da abordagem da literatura. Felizmente,

constatamos que a Base Nacional Comum Curricular (BNCC) valoriza e reserva um espaço significativo para o trabalho com textos literários. A literatura deve permanecer como protagonista no contexto educacional, pois é por meio dela que os estudantes desenvolvem habilidades críticas, exercitam a imaginação e cultivam a empatia. Além disso, a leitura de obras literárias permite aos estudantes terem contato com diferentes níveis de vocabulário e construções sintáticas, enriquecendo seu repertório, assim como possibilita o contato com diversas culturas e realidades, contribuindo para a formação integral dos estudantes.

Para futuros projetos, planejamos implementar uma sequência didática com o nosso itinerário de leitura. Acreditamos que a prática pode demonstrar a eficácia do trabalho com textos literários em sala de aula, apresentando textos adequados que possam enriquecer o repertório linguístico e cultural dos alunos. Além disso, também desejamos realizar um estudo das palavras-chave do nosso itinerário literário, de forma que os professores ou leitores que desejem escolher um texto com base em nosso itinerário possam considerar não apenas o nível de complexidade, mas também o tema abordado pelo conto, a fim de agradar ao seu gosto pessoal.

REFERÊNCIAS

- ABBADE, C. M. S. A lexicologia e a teoria dos campos lexicais. *Cadernos do CNLF*, v. 15, n. 5, p. 1332-1343, 2011.
- ANTUNES, Irandé. *O território das palavras: estudo do léxico em sala de aula*. São Paulo: Parábola Editorial, 2012.
- ANTUNES, I. O léxico como componente fundamental da língua: implicações pedagógicas. In: COELHO, F. A.; SILVA, J. E. N. (org.). *Ensino de língua portuguesa: teorias e práticas*. Rio de Janeiro: Gramma, 2018. p. 247-259. v.1.
- ARAÚJO, E. *A construção do livro*. Rio de Janeiro/Brasília: Nova Fronteira/Instituto Nacional do Livro, 2015
- ASSUNÇÃO, C. Ampliação vocabular: glossário de textos do livro didático de língua portuguesa “Vontade de Saber Português” do 9º ano. Dissertação (Mestrado Profissional em 25 Letras) - Instituto de Letras e Linguística, Universidade Federal de Uberlândia, Uberlândia, 2015.
- ASSUNÇÃO, C. ARAÚJO, C. Linguística de corpus: teoria, perspectivas metodológicas e ensino das línguas. *Filologia e Linguística Portuguesa*, São Paulo, v. 21, n. 2, p. 271-288, jul./dez. 2019.
- BARBOSA, L. M. A. O conceito de lexicultura e suas implicações para o ensino-aprendizagem de português língua estrangeira. *Filologia Linguística Portuguesa*, n. 10-11, p. 31-41, 2009.
- BERBER SARDINHA, T. *Linguística de Corpus*. Barueri: Manole, 2004.
- BERBER SARDINHA, T. *Linguística de Corpus*. In: GONÇALVES, A. V.; GÓIS, M. L. S.(org.). *Ciências da linguagem: o fazer científico?*. v 1. p. 321-347, 2012
- BEZERRA, L. L. A. Suplementação de vocábulos e sua importância para análise textual e evolução da compreensão leitora. Monografia de Especialização (Ensino de Língua Portuguesa e Literatura) Universidade Aberta do Brasil (UAB), Curitiba, 2008.
- BRASIL. MINISTÉRIO DA EDUCAÇÃO. *Base Nacional Comum Curricular*. 2018. Disponível em: <http://basenacionalcomum.mec.gov.br/>. Acesso em: 15 de jun. 2022.
- BIDERMAN, Maria Tereza Camargo. Dimensões da Palavra. In. *Revista Filologia e Linguística Portuguesa*, n. 2, p. 81-118, 1998.
- BIDERMAN, M. T. C. As ciências do léxico. In: *As ciências do léxico: lexicologia, lexicografia, terminologia*. 2 ed. Campo Grande: Ed. UFMS, 2001.
- BORGES, J. M. M.; SILVA, E. B. Análise Lexical do escritor goiano José J. Veiga:

Contribuições da Lexicologia e Linguística de Córpus para o ensino de Língua Portuguesa. In: BONAFIM, A.; OLIVEIRA, B. H. M.; SILVA, E. B.; TRINDADE, M. R. *(Des) centralidades Epistemológicas: Estudos em Língua, Literatura e Interculturalidade*. 1. ed. São Paulo: Todas as Musas, 2022, p. 129-169.

CANDIDO, Antonio. O direito à Literatura. In: *Vários escritos*. 4. ed. São Paulo: DuasCidades, 2004.

CASTRO, C. S. Uso de termos hiperônimos em textos dissertativos-argumentativos de estudantes do ensino fundamental de educação de jovens e adultos (EJA). Dissertação (MESTRADO PROFISSIONAL EM LETRAS - PROFLETRAS). Universidade Federal de Pernambuco (UFPE), Recife, 2015.

CHAMBERS, A. Towards the Corpus Revolution: Bridging the Research-Practice Gap, *Language Teaching. Cambridge University Press*. 460–475, 2019.

COSSON, R. A literatura em todo lugar. In: *Círculos de Leitura e Letramento Literário*. São Paulo: Contexto, 2014.

COLLIE, J. & SLATER, S. *Literature in the Language Classroom*. Avon: Longman, 1992.

Dias, J. A. F. R. Menezes, T. S. A. R. Reflexões sobre o ensino da literatura na sala de aula: entraves e possibilidades. *CIFEFIL*, 2014.

Douglas Biber and Randi Reppen (eds.), *The Cambridge handbook of English corpus linguistics*. Cambridge: Cambridge University Press, 2015.

EGBERT, J. LARSSON, T. BIBER, D. Doing Linguistics With A Corpus Methodological Considerations For The Everyday User. *Cambridge University Press*. 2020.

FAILLA, Z. (org.). *Retratos da Leitura no Brasil 5*. São Paulo: Imprensa Oficial do Estado de São Paulo/Instituto Pró-Livro, 2021.

FRANÇA, T. M. Por um ensino de literatura discursivo e decolonial. *Heterotópica*, v.4; n. 1, p. 85-102, jan.-jun. 2022.

GIL, A. C. *Como elaborar projetos de pesquisa*. São Paulo: Atlas, 2002.

Gil, Beatriz Daruj. Recursos léxico-gramaticais no ensino de língua portuguesa. *Revista do Gel*, v. 19, n. 1, p. 100-118, 2022.

GUERRA, M. M.; ANDRADE, K. S. O léxico sob perspectiva: contribuições da Lexicologia para o ensino de línguas. *Revista eletrônica de Linguística, Uberlândia*, v. 6, n. 1, 1. sem./2012. p. 226-241.

HUNSTON, S. *Corpora in applied linguistics (2nd ed.) Cambridge University Press*. 2022.

Jackson, H., & Amvela, E. Z. (2021). *An Introduction to English Lexicology (3rd ed.)*. Bloomsbury Publishing. Retrieved from <https://www.perlego.com/book/2879409/an-introduction-to-english-lexicology-words-meaning-and-vocabulary-pdf> (Original work published 2021)

MAGALHAES JUNIOR, Raimundo. *A arte do conto: sua história, seus gêneros, sua técnica, seus mestres*. 1. ed. Rio de Janeiro: Bloch Editores S. A., 1972.

MARCUSCHI, L. A. *Produção textual, análise de gêneros e compreensão*. São Paulo: Parábola Editorial, 2008.

O'KEEFEE, A. MCCARTHY, M. J. *The routledge handbook of Corpus Linguistics*. Routledge. 2022

ORSI, Vivian. Lexicologia: o que há por trás do Estudo das Palavras. In: GONÇALVES, Adair Vieira; GÓIS, Marcos Lúcio de Sousa. *Ciências da Linguagem: o fazer científico?*. Campinas, SP: Mercado de Letras, 2012, p.164-167.

RODRIGUES, M. I M. Vocabulário e (des)motivação para a leitura. Dissertação (Mestrado em Português Língua Não Materna) Universidade Aberta, Lisboa, 2014.

ROMER, U. Pedagogical Applications of Corpora: Some Reflections on the Current Scope and a Wish List for Future Developments. In : GAST, V. (Org.). *The Scope and Limits of Corpus Linguistics - Empiricism in the Description and Analysis of English*. Special Issue: Zeitschrift für Anglistik und Amerikanistik, v. 54, n. 2, p. 121-134, 2006.

SARMENTO, Simone. Ensino de Cultura na Aula de Língua Estrangeira. *Revista Virtual de Estudos da Linguagem – ReVEL*. V. 2, n. 2, março de 2004.

SARMENTO, S. Linguística de corpus: histórico, metodologia, campos de aplicação. *Trama*, Marechal Cândido Rondon, v. 6, n. 2, p. 87-107, 2010.

SERRA, L. H. O ensino de vocabulário na sala de aula: reflexões e práticas para a produção de textos na educação básica. *Afluente*, São Luís, v. 1, n. 1, p. 1-19, 2019.

SHEPHERD, T. M. G. O estatuto da Linguística de Corpus: metodologia ou área da linguística? *Matraga*, Rio de Janeiro, v. 16, n. 24, p. 150-172, jan./jun. 200

SIMÕES, D; ASSIS, E. F. Ler os clássicos para ampliar domínio lexical. In: SIMÕES, D; ASSIS, E. F.; OSÓRIO, P. (org.). *Léxico: investigação e ensino*. Rio de Janeiro: Dialogarts, 2014. p. 20-30.

SILVA, Maria Cristina Parreira. Lexicografia Bilíngue: Uma verificação dos substantivos mais frequentes em dicionários bilíngues francês –português e português –francês. In: LONGO, Beatriz Nunes de Oliveira; SILVA, Bento Carlos Dias da. *A construção dos dicionários e de bases de conhecimento lexical*. Araraquara: Cultura Acadêmica Editora, 2006.

SZUDARSKI, P. *Corpus Linguistics for Vocabulary: A Guide for Research*. London/New York: Routledge, 2018.

VILELA, Mário. *Estudos de lexicologia do português*. Coimbra: Almedina, 1994.

ZAFALON, Míriam. *Leitura e ensino da literatura: reflexões*. Artigo. Educadores. UEM 2010.

ZILBERMAN, Regina. A escola e a leitura da literatura. In: ZILBERMAN, Regina; ROSING, Tania M. K. (Org.). *Escola e Leitura: velhas crises, novas alternativas*. São Paulo: Global, 2009.

APÊNDICE A – Quadro com as categorizações dos contos literários

Categoria – Muito fácil

#	Autor	Título	Pontuação
1	Humberto de Campos	O purgativo	4,2
2	Fernando Sabino	O gato sou eu	4,9
3	Olavo Bilac	Como os cães	5,1
4	Lima Barreto	Na janela	5,1
5	Alcântara Machado	O tímido José	5,3

Fonte: Dados da presente pesquisa.

Categoria – Fácil

#	Autor	Título	Pontuação
1	Olavo Bilac	O pecado	5,4
2	Olavo Bilac	Como a pescada	5,4
3	Otto Lara Resende	Vista cansada	5,4
4	Clarice Lispector	Sobre a escrita	5,4
5	Dalton Trevisan	Arara bêbada	5,4
6	Cora Coralina	Última Impressão	5,4
7	Machado de Assis	Um Apólogo	5,5
8	Marques Rebelo	A morta	5,5
9	Ricardo Ramos	O Terceiro Irmão	5,5
10	Humberto de Campos	Roubado e contente	5,6
11	Ferreira Gullar	A estante	5,6
12	Cora Coralina	Um Vencedor	5,6
13	Clarice Lispector	Meu Deus, me dê coragem	5,6
14	Clarice Lispector	Por não estarem distraídos	5,6
15	Lygia Fagundes Telles	Apenas um saxofone	5,6
16	João Antônio	Um preso	5,6
17	Coelho Neto	O lenhador	5,6
18	Lygia Fagundes Telles	O jardim selvagem	5,6
19	Carlos Drummond de Andrade	Governar	5,6
20	Cora Coralina	Ideal de Moça	5,6
21	Olavo Bilac	O velho rei	5,6
22	Olavo Bilac	A borboleta negra	5,6
23	Amadeu Amaral	Sanfona	5,6
24	Clarice Lispector	Atenção ao sábado	5,7
25	Marques Rebelo	Caprichosos da Tijuca	5,7
26	Alcântara Machado	O Aventureiro Ulisses	5,7
27	Alcântara Machado	A apaixonada Elena	5,7
28	Carlos Drummond de Andrade	A bailarina	5,7

29	Artur Azevedo	Denúncia Involuntária	5,7
		Apólogo brasileiro sem véu de	
30	Alcântara Machado	alegoria	5,7
31	Xavier Marques	A vida do homem	5,7
32	José J. Veiga	Diálogo da relativa grandeza	5,7
33	João Antônio	Torcedor	5,7
34	Humberto de Campos	O ladrão honesto	5,7
	Carlos Drummond de		
35	Andrade	Lavadeiras de Moçoró	5,7
36	Olavo Bilac	Mãe Maria	5,7
37	Humberto de Campos	O papagaio traído	5,7
38	Artur Azevedo	Como o diabo as arma!	5,7
39	João Simões Lopes Neto	Artigos de fé do gaúcho	5,7
40	Olavo Bilac	Pátria nova	5,7
41	Lima Barreto	Uma conversa	5,7
42	Marques Rebelo	Serrana	5,7
		Maravilhas da Casa Velha da	
43	Cora Coralina	Ponte	5,7
44	Humberto de Campos	O ladrão arrependido	5,8
45	Alcântara Machado	A isigne Cornélia	5,8
46	Humberto de Campos	A despedida	5,8
47	Raul Pompéia	Antes e depois	5,8
48	João Antônio	Dádiva	5,8
49	Humberto de Campos	Número, faz favor	5,8
50	João Guimarães Rosa	A terceira margem do rio	5,8
51	Valdomiro Silveira	Por mexericos	5,8
52	Dalton Trevisan	Questão de família	5,8
53	Dalton Trevisan	Angústia do Viúvo	5,8
54	Marques Rebelo	A derrota	5,8
55	Coelho Neto	O rato	5,8
56	Dalton Trevisan	Menino caçando passarinho	5,8
57	Humberto de Campos	João, o bom jornalista	5,8
58	Coelho Neto	A partilha	5,8
59	Arthur Azevedo	Os compadres	5,8
60	Arthur Azevedo	O meu criado João	5,8
61	Humberto de Campos	Sinceridade	5,8
62	João Antônio	Uma força	5,8
63	Humberto de Campos	A mulher forte	5,8
64	Marques Rebelo	Quatro momentos de um Idílio	5,8
	Carlos Drummond de		
65	Andrade	Um livro e sua lição	5,8
66	Olavo Bilac	Uma vida	5,8
67	Humberto de Campos	Uma rapariga apressada	5,8
68	Marques Rebelo	Almas no jardim	5,8
69	Humberto de Campos	O relógio preguiçoso	5,9
70	José J. Veiga	Uma pedrinha na ponte	5,9

71	João Guimarães Rosa	Soroco, sua mãe, sua filha	5,9
72	Alcântara Machado	A piedosa Tereza	5,9
73	Fernando Sabino	O canto do galo	5,9
74	Valdomiro Silveira	Esperando	5,9
75	Valdomiro Silveira	Constância	5,9
76	Alcântara Machado	O Inteligente Cícero	5,9
	Carlos Drummond de		
77	Andrade	A cor de cada um	5,9
78	Arthur Azevedo	A conselho do marido	5,9
79	João Simões Lopes Neto	O mate do João Cardoso	5,9
80	Alcântara Machado	Amor e sangue	5,9
81	Lima Barreto	O desconto	5,9
82	Valdomiro Silveira	Missa da páscoa	5,9
83	Valdomiro Silveira	As frutas	5,9
84	Lygia Fagundes Telles	Lua Crescente em Amsterdã	5,9
85	Álvares de Azevedo	Último beijo de amor	5,9
86	Rubem Fonseca	Deus e o Diabo	5,9
87	Dalton Trevisan	Chove Chuva	5,9
88	Rubem Fonseca	Jardim de Flores	5,9
89	Lima Barreto	História de um soldado velho	5,9
	Carlos Drummond de		
90	Andrade	A condição geral	5,9
91	Clarice Lispector	Dá-me a tua mão	5,9
92	Humberto de Campos	Bebida para Viúvo	5,9
93	Alcântara Machado	Carmela	5,9
94	Aluísio Azevedo	Heranças	5,9
95	João Antônio	Bolo na garganta	5,9
96	Amadeu Amaral	Justiça	5,9
97	Machado de Assis	Cantiga de esponsais	5,9
98	Rubem Fonseca	O pudico	5,9
99	Monteiro Lobato	Cabelos compridos	5,9
100	Amadeu Amaral	Lenço perdido	5,9
101	Álvares de Azevedo	Uma noite do século	5,9
102	Olavo Bilac	Um homem	5,9
103	Marques Rebelo	A mudança	5,9
104	Cora Coralina	Medo	5,9
105	Clarice Lispector	As águas do mundo	5,9
106	Fernando Sabino	A companheira de viagem	6
107	Humberto de Campos	O perigo das profecias	6
108	Humberto de Campos	As pesquisas de Moisés	6
109	Alcântara Machado	A sociedade	6
110	Ferreira Gullar	Quisera ser um gato	6
111	Alcântara Machado	O Ingênuo Dagoberto	6
		Notas Biográficas do Novo	
112	Alcântara Machado	Deputado	6
113	Marques Rebelo	Uma senhora	6

	Carlos Drummond de		
114	Andrade	Rick e a girafa	6
115	Arthur Azevedo	Uma por outra	6
116	João Antônio	Fujie	6
	Carlos Drummond de		
117	Andrade	Flor, telefone, moça	6
118	João Antônio	Natal na gafua	6
119	Dalton Trevisan	O herói perdido	6
120	Valdomiro Silveira	Última vez	6
121	Raul Pompéia	Tilburi de praça	6
122	Olavo Bilac	O vaso	6
123	Dalton Trevisan	Clínica de repouso	6
124	Lima Barreto	Uma loteria com que sonho	6
125	Valdomiro Silveira	Velha dor	6
126	Olavo Bilac	A costura	6
127	Arthur Azevedo	As asneiras do Guedes	6
128	Cora Coralina	O Corpo de Delito	6
129	Artur Azevedo	Chico	6
130	Marques Rebelo	Stela me abriu a porta	6
	Carlos Drummond de		
131	Andrade	A beleza total	6
132	João Antônio	Sete mungangas	6
		Retalhos de fome numa tarde de G	
133	João Antônio	C	6
134	Olavo Bilac	O diabo	6
135	Lima Barreto	O anel de perdicas	6
136	João Simões Lopes Neto	A mboitatá	6
137	Alcântara Machado	Gaetaninho	6
	Ignácio de Loyola		
138	Brandão	Os músculos	6
139	Marques Rebelo	Labirinto	6
140	Clarice Lispector	Melhor do que arder	6
141	Coelho Neto	O perna de pau	6
142	Olavo Bilac	A civilização	6
143	Cora Coralina	Oração de Natal	6
144	Jorge Amado	Nem a rosa, nem o cravo	6
	Carlos Drummond de		
145	Andrade	O nome	6
146	Humberto de Campos	O cavalo	6
147	Cora Coralina	Quem Foi Ela	6
148	Alcântara Machado	O Monstro de Rodas	6,1
149	Alcântara Machado	Armazém Progresso de São Paulo	6,1
150	Alcântara Machado	O Filósofo Platão	6,1
151	Monteiro Lobato	O casamento da Emília	6,1
152	Dalton Trevisan	Dois velhinhos	6,1
153	Monteiro Lobato	Cavalinhos	6,1

154	José J. Veiga	Entre Irmãos	6,1
155	Humberto de Campos	O duelo	6,1
156	Dalton Trevisan	Uma vela para Dário	6,1
157	Machado de Assis	Terpsícore	6,1
158	Lygia Fagundes Telles	Antes do baile verde	6,1
159	Valdomiro Silveira	Hora quieta	6,1
160	Ingles de Souza	O gado do Valha-me Deus	6,1
161	Clarice Lispector	Banhos de mar	6,1
162	Artur Azevedo	Conjugo Vobis	6,1
163	Valdomiro Silveira	Valentia	6,1
164	Machado de Assis	Conto de Escola	6,1
		Considerações em torno das aves-	
165	Ivan Ângelo	balas	6,1
166	Artur Azevedo	Assunto para um conto	6,1
167	Lygia Fagundes Telles	Os objetos	6,1
168	Arthur Azevedo	As cerejas	6,1
169	Valdomiro Silveira	Saudades do natal	6,1
170	Alcântara Machado	O revoltado Robespierre	6,1
171	Clarice Lispector	Escrever, humildade, técnica	6,1
172	Monteiro Lobato	Os pequeninos	6,1
173	Valdomiro Silveira	Salvação	6,1
174	Clarice Lispector	Mas vai chover	6,1
175	Olavo Bilac	A pátria	6,1
176	João Antônio	Busca	6,1
177	Lygia Fagundes Telles	O menino	6,1
178	Dalton Trevisan	Contos dos bosques de Curitiba	6,1
179	Carlos Heitor Cony	O pai, as filhas e o pinto	6,1
180	Lima Barreto	Uma vagabunda	6,1
181	Machado de Assis	Flor anônima	6,1
182	Dalton Trevisan	O espião	6,1
183	Marques Rebelo	Na rua dona Emerenciana	6,1
184	Humberto de Campos	O inglês tal... qual se o pinta	6,1
185	Humberto de Campos	O Bellerophon	6,1
186	João Antônio	Eguns	6,1
187	Clarice Lispector	A repartição dos pães	6,1
188	Dalton Trevisan	O vampiro de Curitiba	6,1
189	Machado de Assis	Adão e Eva	6,1
190	Humberto de Campos	O café	6,1
191	Olavo Bilac	Sumé	6,1
		Afinação da arte de chutar	
192	João Antônio	tampinhas	6,1
193	Humberto de Campos	A geografia	6,1
194	Machado de Assis	A Ela	6,1
195	Olavo Bilac	O recruta	6,1
196	João Antônio	Televisão	6,1
197	Lima Barreto	A cartomante	6,1

198	Cora Coralina	As Cocadas	6,1
199	Marques Rebelo	História de abelha	6,1
200	Marina Colassanti	A moça tecelã	6,1
201	Lima Barreto	A indústria da caridade	6,1
202	Cora Coralina	As Almofadas de Dona Lu	6,1
203	Érico Veríssimo	O navio das sombras	6,1
204	Clarice Lispector	Precisa-se	6,1
205	Lima Barreto	Arte de governar	6,1
206	Marques Rebelo	História	6,1
207	Alcântara Machado	O Lírico Lamartine	6,1
	Carlos Drummond de		
208	Andrade	A melhor opção	6,1
209	Amadeu Amaral	Brinquedo	6,1
210	Olavo Bilac	O defunto	6,1
211	Machado de Assis	O sermão do diabo	6,1

Fonte: Dados da presente pesquisa.

Categoria – Moderado

#	Autor	Título	Pontuação
		O negrinho do	
1	João Simões Lopes Neto	pastoreio	6,2
2	Alcântara Machado	Lisetta	6,2
3	Dalton Trevisan	Penélope	6,2
4	Lygia Fagundes Telles	O moço do saxofone	6,2
5	Rubem Fonseca	Feliz ano novo	6,2
6	Aluísio Azevedo	Uma lição	6,2
7	Monteiro Lobato	O resto de onça	6,2
8	Valdomiro Silveira	História antiga	6,2
9	João Antônio	Gente do erro	6,2
10	Valdomiro Silveira	Desespero de amor	6,2
11	Alcântara Machado	Trio de Guerra N 35	6,2
12	Cora Coralina	Um Carnaval Antigo	6,2
		Jeca Tatu, a	
13	Monteiro Lobato	Ressurreição	6,2
14	Arthur Azevedo	O último palpite	6,2
		Chasque do	
15	João Simões Lopes Neto	imperador	6,2
16	Lygia Fagundes Telles	A Testemunha	6,2
17	Lygia Fagundes Telles	A medalha	6,2
18	Arthur Azevedo	Uma embaixada	6,2
19	Cora Coralina	O Boi Balão	6,2
20	Dalton Trevisan	Tio Galileu	6,2
21	Arthur Azevedo	Caiporismo	6,2
22	Arthur Azevedo	Puelina	6,2
23	Clarice Lispector	O ovo e a galinha	6,2
24	João Antônio	Frio	6,2

	Como se inventaram os almanaques	6,2
25 Machado de Assis	Ponto de vista	6,2
26 Machado de Assis	O filho de Maria	
27 José Geraldo Vieira	Bárbara	6,2
28 Lima Barreto	Despesa filantrópica	6,2
29 Lygia Fagundes Telles	Eu era mudo e só	6,2
30 Monteiro Lobato	Bucólica	6,2
31 João Guimarães Rosa	Sequência	6,2
	Por que Lopes se casou	6,2
32 Monteiro Lobato	Conto de Amor	6,2
33 Rubem Fonseca	Candoca	6,2
34 Cora Coralina	O Brinco de Pérola	6,2
35 Rubem Fonseca	Natal na barca	6,2
36 Lygia Fagundes Telles	O Patriota Washington	6,2
37 Alcântara Machado	Beira-rio	6,2
38 Carlos Drummond de Andrade	A taça de champagne	6,2
39 José Geraldo Vieira	A prima Lúcia	6,2
40 José Geraldo Vieira	O regenerado	6,2
41 Humberto de Campos	De madrugada.	6,2
42 Raul Pompéia	Uma história de mil anos	6,2
43 Monteiro Lobato	Nossa amiga	6,2
44 Carlos Drummond de Andrade	Oscar Wilde	6,2
45 José Geraldo Vieira	O manteau	6,2
46 Humberto de Campos	O segredo	6,2
47 José Geraldo Vieira	Viver!	6,2
48 Machado de Assis	Helga	6,2
49 Lygia Fagundes Telles	Mãe	6,2
50 Amadeu Amaral	Boa prosa	6,2
51 Amadeu Amaral	Espelho	6,2
52 Marques Rebelo	Moleques da Minha Terra	6,2
53 Cora Coralina	A cozinha feliz, a grandeza da sinceridade	6,2
54 Clarice Lispector	Quando as ciganas passam	6,2
55 José Geraldo Vieira	Um morto	6,2
56 Marques Rebelo	Um homem perfeito	6,2
57 Amadeu Amaral	Uma rosa	6,2
58 Amadeu Amaral	O duelo dos farrapos	6,3
59 João Simões Lopes Neto	Súplica inútil	6,3
60 Humberto de Campos	Corinthians 2 vs Palesta 1	6,3
61 Alcântara Machado		

62	Valdomiro Silveira	Avinha má	6,3
63	Lygia Fagundes Telles	A caçada	6,3
64	Arthur Azevedo	A pequetita	6,3
65	Dalton Trevisan	Na pontinha da orelha	6,3
66	Mario de Andrade	O besouro e a Rosa	6,3
67	Machado de Assis	Trina e una	6,3
68	Artur Azevedo	A dona Branca	6,3
69	Monteiro Lobato	A colcha de retalhos	6,3
70	Artur Azevedo	Uma embaixada	6,3
71	João Guimarães Rosa	Fatalidade	6,3
72	Lima Barreto	O reconhecimento	6,3
73	Lima Barreto	Uma conversa vulgar	6,3
74	João Antônio	Cadelo	6,3
75	João Guimarães Rosa	A menina de lá	6,3
76	Machado de Assis	Vidros quebrados	6,3
77	Marina Colassanti	Um espinho de marfim	6,3
78	Machado de Assis	A cena do cemitério	6,3
79	João Simões Lopes Neto	O boi velho	6,3
80	Artur Azevedo	A cozinheira	6,3
81	Valdomiro Silveira	Cena de amor	6,3
82	Raul Pompéia	Estou roubado	6,3
83	Carlos Drummond de Andrade	Câmara e cadeia	6,3
84	Cora Coralina	Vitalinas	6,3
85	João Guimarães Rosa	Os irmãos Dagobé	6,3
86	Lygia Fagundes Telles	As pérolas	6,3
87	Machado de Assis	Aurora sem dia	6,3
88	Aluísio Azevedo	Aos vinte anos	6,3
89	Lygia Fagundes Telles	A chave	6,3
90	Arthur Azevedo	As escuras	6,3
91	Lygia Fagundes Telles	A ceia	6,3
92	Marques Rebelo	Um destino	6,3
93	Dalton Trevisan	Duas Rainhas	6,3
94	Moacyr Scliar	No retiro da figueira	6,3
95	Lima Barreto	A barganha	6,3
96	Arthur Azevedo	Quem ele era	6,3
97	João Antônio	Barra funda	6,3
98	Clarice Lispector	Mineirinho	6,3
		Correspondências	
99	Raul Pompéia	íntimas I	6,3
100	Carlos Drummond de Andrade	A doida	6,3
101	Machado de Assis	A última receita	6,3
		Menino a bico de	
102	Clarice Lispector	pena	6,3
103	Marques Rebelo	Dois pares pequenos	6,3
104	Cora Coralina	O Capitão-mor	6,3

	Dois Raimundos, um	
105	João Antônio	Lourival 6,3
106	Monteiro Lobato	Café! Café! 6,3
	O feiticeiro e o	
107	Lima Barreto	deputado 6,3
	Salteadores da	
108	Machado de Assis	Tessália 6,3
109	Clarice Lispector	Miss Algrave 6,3
	O Cônego ou	
110	Machado de Assis	Metafísica do Estilo 6,3
111	João Antônio	Milagre Chué 6,3
112	Arthur Azevedo	Toc, toc, toc, toc... 6,3
113	Ribeiro Couto	A conquista 6,3
114	Clarice Lispector	O primeiro beijo 6,3
115	Marques Rebelo	Uma véspera de natal 6,3
116	Arthur Azevedo	O velho Lima 6,3
117	Carlos Drummond de Andrade	A escola perfeita 6,3
118	Lima Barreto	Uma noite no lírico 6,3
119	Raul Pompéia	Milina e Turco 6,3
	À sombra de uma	
120	Autran Dourado	castanheira 6,3
121	Amadeu Amaral	De Amicitia 6,3
122	Lima Barreto	Elkazenadji 6,3
123	Ribeiro Couto	Bilu, Carolina e eu 6,3
	Composição de	
124	Marques Rebelo	carnaval 6,3
125	Humberto de Campos	Estética e pontaria 6,3
126	Amadeu Amaral	O bonde e a rua 6,3
127	Olavo Bilac	A enguia 6,3
128	Amadeu Amaral	Camelô 6,3
129	Carlos Drummond de Andrade	Novo dicionário 6,3
130	Cora Coralina	Rio Vermelho 6,3
131	Humberto de Campos	A água 6,3
132	Amadeu Amaral	Um romance 6,3
133	Lygia Fagundes Telles	Que se chama solidão 6,4
134	Machado de Assis	Noite de Almirante 6,4
135	Arthur Azevedo	O Cuco 6,4
136	Hugo de Carvalho Ramos	Caçando perdizes 6,4
137	João Simões Lopes Neto	Contrabandista 6,4
138	Machado de Assis	A carteira 6,4
139	Rachel de Queiroz	O ateu 6,4
	A firmeza de Al-	
140	Lima Barreto	Bandeirah 6,4
141	João Antônio	Água branca 6,4
142	João Simões Lopes Neto	Trezentas onças 6,4
143	Humberto de Campos	Seio da família 6,4

144	Machado de Assis	Singular Ocorrência	6,4
145	Lima Barreto	O juramento	6,4
146	Valdomiro Silveira	Ana Cabriuvana	6,4
147	Alcântara Machado	Guerra Civil	6,4
148	José J. Veiga	Cachimbo	6,4
149	Aluísio Azevedo	O Madeireiro	6,4
150	Machado de Assis	Ex Cathedra	6,4
151	José Geraldo Vieira	O sacrilégio	6,4
152	Clarice Lispector	Cem anos de perdão	6,4
153	Lima Barreto	Li Via	6,4
154	Humberto de Campos	A buzina	6,4
155	Lygia Fagundes Telles	As formigas	6,4
156	Lygia Fagundes Telles	Verde lagarto amarelo	6,4
157	Rachel de Queiroz	A arte de ser avó	6,4
158	Clarice Lispector	Uma amizade sincera	6,4
		A mulher que matou	
159	Clarice Lispector	os peixes	6,4
160	Clarice Lispector	Uma galinha	6,4
161	Arthur Azevedo	A ama seca	6,4
162	Humberto de Campos	O hipnotizador	6,4
163	Adélia Prado	Sem enfeite nenhum	6,4
164	Raul Pompéia	Niente	6,4
		Onofre, o terrível ou a	
165	Marques Rebelo	sede de justiça	6,4
		A bailarina e o	
166	Carlos Drummond de Andrade	morcego	6,4
167	Lima Barreto	Boa medida	6,4
		Fiscador de	
168	Valdomiro Silveira	Carumbé	6,4
169	Arthur Azevedo	A réclame	6,4
		Das Coisas Bem	
170	Cora Coralina	Guardadas e Suas	6,4
171	Monteiro Lobato	As fitas da vida	6,4
172	Carlos Drummond de Andrade	A salvação da alma	6,4
		O entendimento dos	
173	Carlos Drummond de Andrade	contos	6,4
174	Lima Barreto	A mulher do Anacleto	6,4
175	Érico Veríssimo	As mãos de meu filho	6,4
176	Lima Barreto	O caçador doméstico	6,4
177	Monteiro Lobato	Fatia de vida	6,4
178	Carlos Drummond de Andrade	Parabéns por tudo	6,4
179	Aluísio Azevedo	Vícios	6,4
180	Clarice Lispector	Macacos	6,4
		Sua alteza imperial	
181	Lima Barreto	Jan-Ghote	6,4
182	Dalton Trevisan	Visita à professora	6,4

183	Cora Coralina	Zé Sidrach e Dico Foggia	6,4
		Felicidade	
184	Clarice Lispector	Clandestina	6,4
185	Ribeiro Couto	Largo da matriz	6,4
		O casamento e a cegonha	6,4
186	Cora Coralina	A defesa	6,4
187	Olavo Bilac	A última posição	6,4
188	Humberto de Campos	O vitral	6,4
189	Osman Lins	A filosofia do Mendes.	6,4
190	Arthur Azevedo	A coroa de espinhos	6,4
191	José Geraldo Vieira	Conversa de velho com criança	6,4
192	Carlos Drummond de Andrade	A solidariedade de Al- Bandeirah	6,4
193	Lima Barreto	A fronteira	6,4
194	Coelho Neto	Negrinha.	6,4
195	Monteiro Lobato	Na tormenta	6,4
196	Marques Rebelo	As margens da alegria	6,4
197	João Guimarães Rosa	Rendas e Fitas	6,4
198	Aluísio Azevedo	Praça Mauá	6,4
199	Clarice Lispector	A amiguinha Teresa	6,4
200	Ribeiro Couto	O menino e o velh	6,4
201	Lygia Fagundes Telles	Sinos de Goiás	6,4
202	Cora Coralina	Martha das Terras	
203	José Geraldo Vieira	Baixas	6,4
204	Gonzaga Duque	Sapo	6,4
205	Humberto de Campos	Razão poderosa	6,4
206	Humberto de Campos	A dama séria	6,4
207	Humberto de Campos	Justificação	6,4
208	Arthur Azevedo	O retrato	6,5
209	Machado de Assis	Umás férias	6,5
		Onde andam os didangos	6,5
210	José J. Veiga	Os cavaleiros de	
		Platiplanto	6,5
211	José J. Veiga	Leão de chácara	6,5
212	João Antônio	Juca guerra	6,5
213	João Simões Lopes Neto	Penar de velho	6,5
214	João Simões Lopes Neto	Irmã Tereza	6,5
215	Humberto de Campos	Jogo de osso	6,5
216	João Simões Lopes Neto	O meu carnaval	6,5
217	Lima Barreto	Pedro Pichorra	6,5
218	Monteiro Lobato	Missa do galo	6,5
219	Machado de Assis	Marcha fúnebre	6,5
220	Machado de Assis		

221	José J. Veiga	Tapete florido	6,5
222	Lima Barreto	O traidor	6,5
		A parcinômia do	
223	Humberto de Campos	Coronel	6,5
224	Machado de Assis	Tempo de crise	6,5
225	Arthur Azevedo	Uma carga de sono	6,5
226	Arthur Azevedo	As paradas	6,5
227	José J. Veiga	Espelho	6,5
		Um moço muito	
228	João Guimarães Rosa	branco	6,5
		Quem tudo quer, tudo	
229	Coelho Neto	perde	6,5
230	Machado de Assis	Entre Santos	6,5
		Uma noite de chuva	
		ou Simão diletante de	
231	Ribeiro Couto	ambientes	6,5
232	Ribeiro Couto	Milagre de natal	6,5
233	Carlos Drummond de Andrade	Presépio	6,5
		O Tesouro da Casa	
234	Cora Coralina	Velha	6,5
235	Cora Coralina	Minga, zóio de prata	6,5
236	Arthur Azevedo	A doença do Fabrício	6,5
		Paulo Melado do	
		Chapéu Mangueira	
237	João Antônio	Serralha	6,5
238	Arthur Azevedo	Octogenário	6,5
239	Monteiro Lobato	A "Cruz de Ouro"	6,5
		Os dois bonitos e os	
240	Rachel de Queiroz	dois feios	6,5
241	Monteiro Lobato	Barba azul	6,5
		Um músico	
242	Lima Barreto	extraordinário	6,5
243	Osman Lins	A partida	6,5
244	Dalton Trevisan	A noite da Paixão	6,5
		Uma visita de	
245	Machado de Assis	Alcibíades	6,5
		Os desvaneios do	
246	Érico Veríssimo	general	6,5
247	Ribeiro Couto	A denúncia do sangue	6,5
248	Arthur Azevedo	Um capricho	6,5
249	Lygia Fagundes Telles	A chave na porta	6,5
250	Coelho Neto	O ambicioso	6,5
251	Machado de Assis	Uma senhora	6,5
252	Carlos Drummond de Andrade	A baronesa	6,5
253	Machado de Assis	Lágrimas de Xerxes	6,5

	Uma história de tanto amor	6,5
254	Clarice Lispector	
255	José J. Veiga	Vestido de Fustão 6,5
256	João Antônio	Sufoco 6,5
257	Machado de Assis	Trio em Lá Menor 6,5
258	Valdomiro Silveira	Camunhengue 6,5
259	Ivan Ângelo	Talismã 6,5
260	Olavo Bilac	O cabeça de ferro 6,5
261	Raul Pompéia	No mar 6,5
262	Marques Rebelo	Depoimento simplório 6,5
263	Aluísio Azevedo	Fora de horas 6,5
264	João Antônio	O coração na mão 6,5
265	Osman Lins	Menino mau 6,5
266	Carlos Drummond de Andrade	Meu companheiro 6,5
267	Ricardo Ramos	Viva o Rei 6,5
268	Amadeu Amaral	Rufina 6,5
	A última sessão do Grêmio	6,5
269	Marques Rebelo	
270	Cora Coralina	Goiás e Suas Uvas 6,5
271	Humberto de Campos	A Violência 6,5
272	Marques Rebelo	Em maio 6,5
273	Lúcio Cardoso	Simple encontro 6,5
	Congresso	
274	Lima Barreto	Pamplanetário 6,5
275	Marques Rebelo	Episódio coreográfico 6,5
276	Olavo Bilac	Os aneis 6,5
277	Lima Barreto	Conservou o fez 6,5
278	Amadeu Amaral	Ainda a rosa 6,5
	A pomba e a estrumeira	6,5
279	Raul Pompéia	
280	Cora Coralina	No Gosto do Povo 6,5
281	Arthur Azevedo	O Paulo 6,5
282	Amadeu Amaral	O bonde 6,5
283	Amadeu Amaral	Delicadeza 6,5
284	João Antônio	Publicitário do ano 6,5
285	Amadeu Amaral	Canudo-de-pito 6,5
286	José Geraldo Vieira	O enxoval 6,5
287	Raul Pompéia	Comércio de flores 6,5
288	Érico Veríssimo	Chico 6,5
289	Amadeu Amaral	Ruídos e rumores 6,5
290	Arthur Azevedo	A Ritinha 6,5
291	Coelho Neto	O mentiroso 6,6
292	Arthur Azevedo	Barca 6,6
293	João Simões Lopes Neto	Melância-Coco verde 6,6
	A Viagem de Dez Léguas	6,6
294	José J. Veiga	

295	Hugo de Carvalho Ramos	O saci	6,6
296	José J. Veiga	Os cascamorros	6,6
297	Otto Lara Resende	Três pares de patins	6,6
		Debaixo da Ponte	
298	Dalton Trevisan	Preta	6,6
299	Alcântara Machado	O Mártir Jesus	6,6
		O menininho do	
300	João Simões Lopes Neto	presépio	6,6
301	José J. Veiga	Manuscrito perdido	6,6
302	Marina Colassanti	Começou, ele disse	6,6
303	Machado de Assis	D. Paula	6,6
304	Humberto de Campos	O abatimento	6,6
305	José J. Veiga	Cantilever	6,6
306	Aluísio Azevedo	A serpente	6,6
307	Machado de Assis	Vae Soli	6,6
308	João Guimarães Rosa	Famigerado	6,6
309	Machado de Assis	A senhora do Galvão	6,6
310	Machado de Assis	Anedota Pecuniária	6,6
311	Machado de Assis	Fulano	6,6
		Primeira aventura de	
312	Graciliano Ramos	Alexandre	6,6
		Hussein Ben-Ali Al-	
		Balec e Miquéias	
313	Lima Barreto	Habacuc	6,6
314	João Guimarães Rosa	A befazeja	6,6
315	João Simões Lopes Neto	Deve um queijo	6,6
316	Autran Dourado	A Ilha Escalvada	6,6
317	Machado de Assis	Mariana	6,6
318	Valdomiro Silveira	Mamãe	6,6
		A Desejada das	
319	Machado de Assis	Gentes	6,6
		De como quebrei a	
		cabeça à mulher do	
320	Monteiro Lobato	Melo	6,6
321	Lygia Fagundes Telles	Herbarium	6,6
322	Marques Rebelo	Felicidade	6,6
323	Lygia Fagundes Telles	Venha ver o pôr do sol	6,6
324	Arthur Azevedo	O telefone	6,6
325	Monteiro Lobato	Sorte grande	6,6
		O bloco de mimosas	
326	Ribeiro Couto	borboletas	6,6
327	Cora Coralina	Os meninos verdes	6,6
328	Cora Coralina	As Capas do Diabo	6,6
329	José Geraldo Vieira	A mandíbula	6,6
330	Machado de Assis	O segredo do bonzo	6,6
331	Machado de Assis	Canção de piratas	6,6

332	Machado de Assis	Manuscrito de um Sacristão	6,6
333	Machado de Assis	Último capítulo	6,6
334	Clarice Lispector	O crime do professor de matemática	6,6
335	Ingles de Souza	Acauã	6,6
336	Coelho Neto	O tesouro	6,6
337	Raul Pompéia	Como nasceu, viveu e morreu	6,6
338	Coelho Neto	O pároco	6,6
339	Lygia Fagundes Telles	Suicídio na granja	6,6
340	Cora Coralina	A Menina, as Formigas e o Boi	6,6
341	Machado de Assis	As academias de São	6,6
342	Machado de Assis	Frei Simão	6,6
343	Monteiro Lobato	Os perturbadores do silêncio	6,6
344	Arthur Azevedo	O espírito	6,6
345	Gonzaga Duque	Miss Fatalidade	6,6
346	Artur Azevedo	O viúvo	6,6
347	Olavo Bilac	O bandeirante	6,6
348	Machado de Assis	A igreja do diabo	6,6
349	Raul Pompéia	O piano	6,6
350	Marques Rebelo	A moça e a primavera	6,6
351	Monteiro Lobato	Meu conto de Maupassant	6,6
352	Ribeiro Couto	Mistério de sábado	6,6
353	Marques Rebelo	Tragédia	6,6
354	Ribeiro Couto	Clube das esposas enganadas	6,6
355	Amadeu Amaral	O gato	6,6
356	Humberto de Campos	O Filho do Comendador	6,6
357	Raul Pompéia	Amor de inverno	6,6
358	Amadeu Amaral	A roupa e o gesto	6,6
359	Raul Pompéia	O perfume dos bolos	6,6
360	Raul Pompéia	O hino auriverde	6,6
361	Humberto de Campos	O inglês exigente	6,6
362	Anibal Machado	O Iniciado do Vento	6,6
363	José J. Veiga	Acidente em Sumaúma	6,7
364	Humberto de Campos	O túmulo de ouro	6,7
365	José J. Veiga	O largo do Mestrevinte	6,7
366	João Simões Lopes Neto	Batendo orelha	6,7
367	Machado de Assis	Uns Braços	6,7

368	Machado de Assis	Diana	6,7
369	Arthur Azevedo	Um desastre	6,7
370	Arthur Azevedo	Sova bem merecida	6,7
		Começos de Uma	
371	Clarice Lispector	Fortuna	6,7
372	Artur Azevedo	O contrabando	6,7
373	Clarice Lispector	Os Obedientes	6,7
374	Carlos Drummond de Andrade	A opinião em palácio	6,7
375	Mario de Andrade	O ladrão	6,7
376	Machado de Assis	Pai contra mãe	6,7
377	Ribeiro Couto	O baiano	6,7
378	Humberto de Campos	Desigualdade	6,7
379	Machado de Assis	Suje-se gordo	6,7
380	Monteiro Lobato	Noite de São João	6,7
381	Machado de Assis	Ideias do canário	6,7
382	Lygia Fagundes Telles	A mão no ombro	6,7
383	Machado de Assis	Primas de Sapucaia	6,7
384	Raul Pompéia	Violeta	6,7
385	Arthur Azevedo	A melhor vingança	6,7
386	Lima Barreto	Um especialista	6,7
387	Clarice Lispector	Restos do carnaval	6,7
388	Carlos Drummond de Andrade	No interior da baleia	6,7
		Á procura de uma	
389	Clarice Lispector	dignidade	6,7
390	Lygia Fagundes Telles	Missa do Galo	6,7
391	Rachel de Queiroz	Um caso obscuro	6,7
		Os macacos me	
392	Fernando Sabino	mordam	6,7
393	Machado de Assis	A inglesinha Barcelos	6,7
394	Arthur Azevedo	A viúva de Estanislau	6,7
		Correspondências	
395	Raul Pompéia	íntimas II	6,7
396	Artur Azevedo	Comes e bebes	6,7
397	Machado de Assis	A Cartomante	6,7
398	Cora Coralina	Minha Irmã	6,7
399	Arthur Azevedo	Sabina	6,7
		A praia de Santa	
400	Artur Azevedo	Luzia	6,7
401	Arthur Azevedo	Vovó Andrade	6,7
402	Lima Barreto	A biblioteca	6,7
403	Monteiro Lobato	O estigma	6,7
404	Carlos Drummond de Andrade	Abotoaduras.	6,7
		Maria Pintada de	
405	Dalton Trevisan	Prata	6,7
406	Ferreira Gullar	Sobre o amor	6,7
407	Aluísio Azevedo	Hamleto	6,7

408	José Geraldo Vieira	A princesa Salomé	6,7
409	Lima Barreto	Hóspede ilustre	6,7
410	Artur Azevedo	Vingança	6,7
411	Ingles de Souza	A feiticeira	6,7
412	Carlos Drummond de Andrade	Miguel e seu furto	6,7
413	Mario de Andrade	O peru de natal	6,7
414	Clarice Lispector	Tentação	6,7
415	José Geraldo Vieira	A IX sinfonia	6,7
416	José Geraldo Vieira	O monte Tabor	6,7
417	Lima Barreto	Dentes negros e cabelos azuis	6,7
418	Machado de Assis	Teoria do medalhão	6,7
419	Cora Coralina	Siá Matilde	6,7
420	Humberto de Campos	Profiteur	6,7
421	Olavo Bilac	Os óculos Extraordinária	6,7
422	Carlos Drummond de Andrade	conversa	6,7
423	Amadeu Amaral	Passeio dominical A sereníssima	6,7
424	Machado de Assis	república	6,7
425	José J. Veiga	Roupa no coradouro	6,8
426	José J. Veiga	Tarde de sábado, manhã de domingo	6,8
427	José J. Veiga	Tia Zi rezando	6,8
428	João Simões Lopes Neto	O anjo da vitória A ilha dos gatos	6,8
429	José J. Veiga	pingados	6,8
430	Ricardo Ramos	Circuito fechado A internada do	6,8
431	José J. Veiga	Sossego	6,8
432	Machado de Assis	Anedota Cabriolet	6,8
433	Humberto de Campos	A queda de Abraão	6,8
434	Machado de Assis	O Caso da Vara	6,8
435	José J. Veiga	Luneta	6,8
436	José J. Veiga	Os noivos	6,8
437	Mario de Andrade	Caso pançudo	6,8
438	José J. Veiga	Fronteira	6,8
439	Lygia Fagundes Telles	Seminário dos Ratos	6,8
440	João Simões Lopes Neto	O negro Bonifácio O Natal de "seu"	6,8
441	Breno Accioly	Hermídio	6,8
442	Arthur Azevedo	Uma aposta	6,8
443	Humberto de Campos	Metamorfose	6,8
444	Raul Pompéia	O tapacurá de cendi	6,8
445	Monteiro Lobato	O mata-pau	6,8

446	Clarice Lispector	Mistério em São Cristóvão	6,8
447	Arthur Azevedo	Os dez por cento	6,8
448	Machado de Assis	Um Homem Célebre	6,8
449	Clarice Lispector	Os laços de família	6,8
450	Monteiro Lobato	O pito do reverendo	6,8
451	Machado de Assis	Longe dos olhos	6,8
452	Machado de Assis	Pobre cardeal	6,8
453	Humberto de Campos	O engolidor de sabre	6,8
454	Lima Barreto	A Nova Califórnia	6,8
455	Aluísio Azevedo	O impenitente	6,8
456	Carlos Drummond de Andrade	Um escritor nasce e morre	6,8
457	Clarice Lispector	Gertrudes pede um conselho	6,8
458	Lima Barreto	O moleque	6,8
459	Machado de Assis	Conto Alexandrino	6,8
460	Ribeiro Couto	O egoísta	6,8
461	Ribeiro Couto	Baianinha	6,8
462	Rubem Fonseca	Incorpóreo	6,8
463	Machado de Assis	Galeria Póstuma	6,8
464	Lima Barreto	Clara dos Anjos	6,8
465	José Geraldo Vieira	Os pombos do claustro	6,8
466	Ribeiro Couto	D. Teodorinha	6,8
467	Carlos Drummond de Andrade	Os privilegiados da Terra	6,8
468	Clarice Lispector	Amor	6,8
469	Mario de Andrade	Conto de Natal	6,8
470	Alcântara Machado	Miss Corisco	6,8
471	Aluísio Azevedo	Polítipo	6,8
472	Ingles de Souza	Amor de Maria	6,8
473	Raul Pompéia	Os parricidas	6,8
474	Lima Barreto	Harakashy e as escolas de Java	6,8
475	Cora Coralina	Trincos, pinos e tramelinhas	6,8
476	José Geraldo Vieira	Carola Marwenga	6,8
477	Arthur Azevedo	Questão de honra	6,8
478	Lima Barreto	Os Kalogheras	6,8
479	José Geraldo Vieira	Baixo-relevo em lápis-lazúli	6,8
480	Raul Pompéia	Conto de fadas	6,8
481	José Geraldo Vieira	Uma operação gratuita	6,8
482	Machado de Assis	Verba testamentária	6,8

483	Hugo de Carvalho Ramos	A madre de ouro	6,8
484	Arthur Azevedo	A Marcelina	6,8

Fonte: Dados da presente pesquisa.

Categoria – Difícil

#	Autor	Título	Pontuação
1	José J. Veiga	Era só brincadeira.	6,9
2	João Simões Lopes Neto	Os cabelos da China	6,9
3	José J. Veiga	Domingo de festa	6,9
4	João Simões Lopes Neto	Correr eguada A espingarda do rei da	6,9
5	José J. Veiga	Síria	6,9
6	Rachel de Queiroz	Viagem de bonde Um capitão de	6,9
7	Machado de Assis	voluntários	6,9
8	Arthur Azevedo	A água de Janos	6,9
9	Lygia Fagundes Telles	Noturno Amarelo	6,9
10	Machado de Assis	Pílades e Orestes	6,9
11	Hugo de Carvalho Ramos	O poldro picaço	6,9
12	Monteiro Lobato	Os faroleiros O tal negócio de	6,9
13	Lima Barreto	“Prestações” Nhola dos Anjos e a	6,9
14	Bernardo Élis	cheia do Corumbá	6,9
15	Monteiro Lobato	Bugio Moqueado A quadrilha de Jacob	6,9
16	Ingles de Souza	Patacho	6,9
17	Clarice Lispector	A partida do trem	6,9
18	José Geraldo Vieira	João lágrima	6,9
19	Lima Barreto	Lourenço, o Magnífico	6,9
20	Mario de Andrade	Vestida de preto O comprador de	6,9
21	Monteiro Lobato	fazendas	6,9
22	Mario de Andrade	Primeiro de maio Como é mesmo o	6,9
23	Marina Colassanti	nome	6,9
24	João Antônio	Excelentíssimo A sensação do	6,9
25	João Simões Lopes Neto	passado Das notas de uma	6,9
26	Aluísio Azevedo	viúva	6,9
27	Machado de Assis	Garnier	6,9
28	Lygia Fagundes Telles	Geléia de maçã	6,9
29	Lima Barreto	Cló	6,9
30	Carlos Drummond de Andrade	O sorvete	6,9

31	Mario de Andrade	Tempo de camisolinha	6,9
32	Aluísio Azevedo	Último lance	6,9
33	José Geraldo Vieira	O milagre do gesso	6,9
34	Marques Rebelo	Caso de mentira	6,9
35	Cora Coralina	O Velho Telhado	6,9
36	Raul Pompéia	Idílio retrospectivo	6,9
37	Monteiro Lobato	A vida em Oblivion	6,9
38	Monteiro Lobato	Velha praga	6,9
39	Monteiro Lobato	Cidades mortas	6,9
40	Amadeu Amaral	Modéstia	6,9
41	Gonzaga Duque	Sob a Estola da Morte	6,9
42	Raul Pompéia	Olhos	6,9
43	José J. Veiga	A máquina extraviada	7,0
44	José J. Veiga	Pasta de couro de búfalo	7,0
45	Mario de Andrade	Nelson	7,0
46	João Antônio	Lapa	7,0
47	Arthur Azevedo	O sonho do conselheiro	7,0
48	José J. Veiga	Cadeira	7,0
49	Ribeiro Couto	O primeiro amor de Antônio Maria	7,0
50	José J. Veiga	Os do outro lado	7,0
51	Arthur Azevedo	Paga ou morre	7,0
52	Olavo Bilac	Os anjos	7,0
53	José J. Veiga	Caderno de endereços	7,0
54	Aluísio Azevedo	Resposta	7,0
55	José J. Veiga	O cachorro canibal	7,0
56	Machado de Assis	Goivos e Camélias	7,0
57	Machado de Assis	Capítulo dos chapéus	7,0
58	Machado de Assis	Evolução	7,0
59	Arthur Azevedo	A tia Aninha	7,0
60	Machado de Assis	Eterno	7,0
61	Hugo de Carvalho Ramos	Pelo caiapó velho	7,0
62	Cora Coralina	O lampião da Rua do Fogo	7,0
63	Clarice Lispector	Evolução de uma miopia	7,0
64	Hugo de Carvalho Ramos	A beira do pouso	7,0
65	Machado de Assis	O sainete	7,0
66	Raul Pompéia	Quase tragédia	7,0
67	Raul Pompéia	História Cãndida	7,0
68	Raul Pompéia	Maladetto Francesco	7,0
69	Cora Coralina	Sequeira versus Siqueira	7,0

70	Lima Barreto	Uma academia de roça	7,0
71	Clarice Lispector	Feliz aniversário	7,0
72	Arthur Azevedo	Os dois andares	7,0
73	José Geraldo Vieira	O triste epigrama	7,0
74	Autran Dourado	Homem cavalo e praia	7,0
75	Raul Pompéia	O modelo do anjo	7,0
76	Lima Barreto	Miss Edith e seu tio	7,0
77	Monteiro Lobato	Era no Paraíso	7,0
		Hóspedes e	
78	Cora Coralina	Hospitalidade	7,0
79	Dyonelio Machado	Ronda das Gotas	7,0
		Necessidade de	
80	Carlos Drummond de Andrade	alegria	7,0
		As festas de reis de	
81	Raul Pompéia	minha prima	7,0
82	Humberto de Campos	O Barbadão	7,0
83	José Geraldo Vieira	A casula de ouro	7,0
84	João Simões Lopes Neto	Laurinda Bellfort	7,0
85	Humberto de Campos	A Judia	7,0
86	João Guimarães Rosa	Desenredo	7,0
87	Hugo de Carvalho Ramos	Caminho das tropas	7,1
88	Machado de Assis	D. Jucunda	7,1
89	Arthur Azevedo	A dívida	7,1
90	Raul Pompéia	A cruz da matriz	7,1
		As Bodas de Luís	
91	Machado de Assis	Duarte	7,1
92	Rachel de Queiroz	O amistoso	7,1
		O donativo do capitão	
93	Ingles de Souza	Silvestre	7,1
94	Machado de Assis	Papéis velhos	7,1
95	Rachel de Queiroz	Amor de acidentado	7,1
96	Machado de Assis	Letra vencida	7,1
97	Hugo de Carvalho Ramos	Mágoa de vaqueiro	7,1
98	Arthur Azevedo	Cavação	7,1
99	Monteiro Lobato	Pollice verso	7,1
100	Monteiro Lobato	Urupês	7,1
101	Hugo de Carvalho Ramos	Ninho de periquitos	7,1
102	Monteiro Lobato	O plágio	7,1
103	Hugo de Carvalho Ramos	A alma das aves	7,1
104	Machado de Assis	O dicionário	7,1
		Madame Renon &	
105	Ribeiro Couto	Sobrinhas Modistas	7,1
106	José Geraldo Vieira	Á maneira de Verlaine	7,1
107	Raul Pompéia	Rogério, o rude	7,1
108	Clarice Lispector	Os bonecos de barro	7,1

109	Raul Pompéia	O mal de D. Quixote	7,1
110	Gonzaga Duque	Confirmação	7,1
111	Graciliano Ramos	Baleia	7,2
112	Machado de Assis	A vida eterna	7,2
113	Otto Lara Resende	Mater dolorosa	7,2
114	Machado de Assis	Um erradio	7,2
115	Murilo Rubião	Botão de rosa	7,2
116	Raul Pompéia	A mona do sapateiro	7,2
117	Xavier Marques	A noiva do golfinho	7,2
118	Monteiro Lobato	Um suplício moderno	7,2
		O engraçado	
119	Monteiro Lobato	arrependido	7,2
120	Graciliano Ramos	Um cinturão	7,2
121	Hugo de Carvalho Ramos	Nostalgias	7,2
122	Raul Pompéia	Os gatos e os cães	7,2
123	Raul Pompéia	Fora de horas	7,2
124	Gonzaga Duque	Ciúme póstumo	7,2
125	Raul Pompéia	Um vizinho original	7,2
126	Amadeu Amaral	Escoteiro	7,2
127	José J. Veiga	Cinzeiro	7,3
128	Mario de Andrade	O poço	7,3
129	Hugo de Carvalho Ramos	Peru de roda	7,3
130	Arthur Azevedo	O palhaço	7,3
		O colocador de	
131	Monteiro Lobato	pronomes	7,3
132	Monteiro Lobato	O espião alemão	7,3
133	Alcântara Machado	Nacionalidade	7,3
134	Monteiro Lobato	O bocatorta	7,3
135	José Geraldo Vieira	Van Dongen	7,3
136	Raul Pompéia	Mocinha	7,3
		A andorinha	
137	José Geraldo Vieira	crucificada	7,3
138	Gonzaga Duque	Aquela Mulher	7,3
139	Hugo de Carvalho Ramos	Dias de chuva	7,3
140	Gonzaga Duque	Benditos olhos!	7,3
		A usina atrás do	
141	José J. Veiga	morro	7,4
142	José J. Veiga	Professor Pulquério	7,4
143	Aluísio Azevedo	Músculos e nervos	7,4
144	Monteiro Lobato	O fígado indiscreto	7,4
145	Raul Pompéia	É morto Pulcinella	7,4
	Dalton Trevisan	Dinorá, moça de	7,4
146		prazer	
	Marina Colassanti	E tinha a cabeça	7,4
147		cheia deles	
148	Raul Pompéia	14 de julho na roça	7,4

149	Raul Pompéia	O fruto da formosura	7,4
150	Arthur Azevedo	O galã	7,5
151	Graciliano Ramos	Um amigo em Talas	7,5
152	Mario de Andrade	Frederico paciência	7,5
153	Raul Pompéia	A andorinha da torre	7,5
154	Lima Campos	A Tia Martinha	7,5
155	Gonzaga Duque	Ruínas	7,5
156	José J. Veiga	O galo impertinente	7,5
157	Arthur Azevedo	Black	7,6
158	Alcides Maya	Estaqueado	7,6
159	Visconde de Taunay	O estorvo	7,6
160	Gonzaga Duque	Idílio Roxo	7,6

Fonte: Dados da presente pesquisa.

Categoria – Muito difícil

#	Autor	Título	Pontuação
1	Afonso Arinos de Melo Franco	Assombramento	7,7
2	Machado de Assis	Henriqueta Renan	7,7
3	Machado de Assis	O velho senado	7,7
4	Visconde de Taunay	Pobre menino	7,7
5	Mário de Andrade	Atrás da catedral de ruão	7,8
6	Murilo Rubião	Os dragões	7,9
7	Gonzaga Duque	Morte do palhaço	7,9
8	Raul Pompéia	Último castelo	7,9
9	Hugo de Carvalho Ramos	A bruxa dos Marinheiros	7,9
10	Raul Pompéia	Cavaleiros andantes	8,0
11	Gonzaga Duque	Posse Suprema	8,0
12	Alcides Maya	Miguelito	8,1
13	Raul Pompéia	A batalha dos livros	8,1
14	Visconde de Taunay	Uma vingança	8,2
15	Gonzaga Duque	Agonia por semelhança	8,2
16	Euclides da Cunha	Judas-Ahsverus	8,5
17	Visconde de Taunay	Rapto original	8,5
18	Visconde de Taunay	Cabeça e coração	8,9

Fonte: Dados da presente pesquisa.

APÊNDICE B – Dados fornecidos pelo NILC-Metrix

Devido ao grande número de colunas e linhas do arquivo gerado pelo NILC-Metrix, e devido à conseqüente dificuldade de visualização nessa dissertação, optamos por viabilizar o acesso aos dados por meio do seguinte link:

https://www.dropbox.com/s/25dpqr1gack14bk/contos_metrics_v1%20%281%29.xlsx?dl=0