

**UNIVERSIDADE FEDERAL DE UBERLÂNDIA  
INSTITUTO DE LETRAS E LINGUÍSTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS**

**ANA PAULA CORRÊA PIMENTA**

**REPRESENTAÇÕES DO LÉXICO SERTANISTA EM *CORPUS* DA LITERATURA  
REGIONALISTA BRASILEIRA: PROTÓTIPO DE VOCABULÁRIO  
ETNOTERMINOLÓGICO *ONLINE***

**UBERLÂNDIA - MG  
2019**

ANA PAULA CORRÊA PIMENTA

**REPRESENTAÇÕES DO LÉXICO SERTANISTA EM *CORPUS* DA LITERATURA  
REGIONALISTA BRASILEIRA: PROTÓTIPO DE VOCABULÁRIO  
ETNOTERMINOLÓGICO *ONLINE***

Tese apresentada à banca examinadora do Programa de Pós-Graduação em Estudos Linguísticos do Instituto de Letras e Linguística da Universidade Federal de Uberlândia, como requisito parcial para a obtenção do título de Doutora em Estudos Linguísticos.

Área de concentração: Estudos em Linguística e Linguística Aplicada.

Linha de pesquisa: Teoria, descrição e análise linguística.

Orientador: Prof. Dr. Ariel Novodvorski

UBERLÂNDIA - MG  
2019

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU  
com dados informados pelo(a) próprio(a) autor(a).

P644 Pimenta, Ana Paula Corrêa, 1984-  
2019 Representações do léxico sertanista em corpus da Literatura regionalista brasileira [recurso eletrônico] : protótipo de vocabulário etnoterminológico online / Ana Paula Corrêa Pimenta. - 2019.

Orientador: Ariel Novodvorski.  
Tese (Doutorado) - Universidade Federal de Uberlândia, Pós-graduação em Estudos Linguísticos.  
Modo de acesso: Internet.  
Disponível em: <http://dx.doi.org/10.14393/ufu.te.2019.2139>  
Inclui bibliografia.  
Inclui ilustrações.

1. Linguística. I. Novodvorski, Ariel , 1968-, (Orient.). II. Universidade Federal de Uberlândia. Pós-graduação em Estudos Linguísticos. III. Título.

CDU: 801

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:  
Gizele Cristine Nunes do Couto - CRB6/2091  
Nelson Marcos Ferreira - CRB6/3074

Ana Paula Corrêa Pimenta

Representações do léxico sertanista em *corpus* da Literatura regionalista brasileira: protótipo de vocabulário etnoterminológico *online*

Tese aprovada para obtenção do título de Doutora em Estudos Linguísticos no Programa de Pós-Graduação em Estudos Linguísticos (PPGEL) da Universidade Federal de Uberlândia (UFU) pela banca examinadora formada por:

Uberlândia, 11 de julho de 2019.

Banca Examinadora



---

Prof. Dr. Ariel Novodvorski – PPGEL/UFU/MG  
(orientador)



---

Prof. Dr. Guilherme Fromm – PPGEL/UFU/MG



---

Profa. Dra. Cleci Regina Bevilacqua – Instituto de Letras/UFRGS/RS



---

Profa. Dra. Camila Tavares Leite – PPGEL/UFU/MG



---

Profa. Dra. Eliana Dias – PROFLETRAS/UFU/MG

Dedico este trabalho às pessoas que me são tão valiosas, Sr. Luiz, meu pai, Sra. Voner, minha mãe, Adryan e Ariany, meus filhos amados, e Fernando, meu companheiro de todas as horas.

## AGRADECIMENTOS

A Deus, por ter me concedido forças e sabedoria nessa caminhada se fazendo presente em todos os momentos da minha vida. Sem a ajuda Dele, nada disso seria possível.

Aos meus pais, Luiz e Voner, verdadeiros exemplos de coragem, persistência, humildade e sabedoria, que fizeram dos meus sonhos seus próprios objetivos e das minhas metas a sua própria luta.

Aos meus filhos, Adryan e Ariany, pelo apoio incondicional e pela compreensão nos momentos de ausência. Sem vocês eu não teria chegado até aqui e, tampouco, teria motivos para concluir esta jornada. A vocês, nada menos do que o meu mais sincero e puro amor.

Ao meu esposo Fernando, por sempre acreditar nos meus sonhos e no meu sucesso, por caminhar ao meu lado mesmo nos momentos difíceis e por ser um esteio nas horas amargas.

Ao meu irmão Marcelo, que sempre esteve ao meu lado, encorajando-me nos momentos difíceis e aplaudindo-me nos momentos de glória.

Aos meus avós, Manoel e Mariana, que com tanto amor me fizeram trilhar pelos seus caminhos, ensinando-me os valores e os princípios da vida. Da singeleza de suas vidas de roceiros, das suas sábias palavras, herança valiosa que nos fizeram conhecer e compartilhar, surgiu toda a inquietação deste estudo. Em suas pisadas, coloco os meus pés, nos seus rastros, faço minhas páginas.

Aos professores Guilherme Fromm, Evandro Silva Martins e Eliana Dias, pelas valiosas sugestões e observações na qualificação da tese, que tanto contribuíram para o aprimoramento deste trabalho.

Especialmente, ao professor e orientador desta tese, Ariel Novodvorski, que com muita sabedoria, paciência, dedicação e carinho conduziu cada momento deste trabalho, cedendo seu precioso tempo para que hoje fosse concluído e o meu sonho realizado. Mais que orientador, amigo, a quem devo respeito, admiração, carinho e gratidão.

Ao Heitor, pela colaboração e apoio na fase de elaboração do site, contribuindo não apenas como técnico de informática, mas como um amigo. Graças à sua dedicação e empenho, conseguimos disponibilizar a nossa proposta de Vocabulário em uma página *web*. A você, deixo aqui, a minha eterna gratidão.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão de Bolsa de Estudos, imprescindível durante o desenvolvimento e conclusão desta pesquisa.

A todos os amigos, professores e funcionários do PPGEL, os meus sinceros agradecimentos.

*No centro do sertão, o que é doideira às vezes  
pode ser a razão mais certa e de mais juízo!*

*João Guimarães Rosa*

## RESUMO

Partindo de uma perspectiva etnoterminológica, de uma tentativa de aproximação entre Linguística e Literatura, buscamos neste trabalho analisar as unidades lexicais, na qualidade de vocábulos-termos, usadas no discurso literário regionalista sertanista, com o intuito de propor um vocabulário e disponibilizá-lo em uma página da Internet para estudantes e professores de literatura. Por se tratar de um repertório que traz significados do universo literário sertanista, entendemos que a nossa proposta de Vocabulário poderá servir como auxílio a alunos e professores de literatura nas atividades de leitura e interpretação do texto literário e, também, a outros interessados pelo estudo ou pelo produto etnoterminológico, como: críticos literários, linguistas, lexicógrafos, terminógrafos. Para tanto, partimos de um *corpus* de estudo composto por vinte obras literárias regionalistas de temática sertanista de autores consagrados na Literatura brasileira, sendo cinco do Regionalismo pitoresco, dez do Regionalismo crítico e cinco do Super-regionalismo, formando um total de três *subcorpora*. O *corpus* de estudo, quando processado pelo programa *WordSmith Tools* 6.0, especificamente por suas três ferramentas: *Concord*, *KeyWords* e *WordList*, nos permitiu identificar os vocábulos-termos e ter o acesso a seus contextos linguísticos de ocorrências para elaboração das propostas de definições. Assim, elaboramos as fichas etnoterminográficas, construímos os verbetes e descrevemos a microestrutura e a macroestrutura do vocabulário, demonstrando, a cada etapa, a viabilidade de nossa proposta. Concluímos que os vocábulos-termos apresentam especificidades semânticas referentes ao universo sócio-linguístico-cultural do sertanejo, que erigem da própria conceptualização linguística desse grupo. Essas unidades adquirem essa dupla funcionalidade (de vocábulo e termo) no discurso literário regionalista sertanista devido aos seguintes aspectos: elas fazem parte de um sistema conceptual estruturado dentro de uma temática específica, a temática sertanista; atuam na composição de um mundo ficcional semioticamente construído pela força modelizante da linguagem literária; possuem intertextualidade e interdiscursividade intra e interuniverso de discurso com discursos etnoliterários; atualizam um sistema de valores, um sistema de crenças, enfim, a axiologia de um grupo específico, o sertanejo dos sertões brasileiros; designam conceitos formados com semas do universo de discurso em que são usadas; referem-se às particularidades de um mundo real/ficcional quanto à sua função simbólica; atuam no plano da memória, de maneira que é possível resgatar das narrativas elementos da cultura e da linguagem do sertanejo, ainda presentes em nossa sociedade. Desse modo, acreditamos que este trabalho possa resultar em significativas contribuições para os estudos etnoterminológicos e, ainda, servir como subsídio para novos trabalhos etnoterminográficos e pesquisas futuras.

**Palavras-chave:** Etnoterminologia. Vocábulo-termo. Linguística de Corpus. Literatura regionalista sertanista. Vocabulário.



## ABSTRACT

Starting from an ethnoterminological perspective, an attempt to bring Linguistics and Literature together, we seek in this work to analyze the lexical units, in the quality of vocab-terms, used in the sertanista regionalist literary discourse, with the intention of proposing a vocabulary and of making it available on a webpage for students and literature teachers. Because it is a repertoire that brings meanings of the sertanista literary universe, we understand that our proposal of Vocabulary can serve as an aid to students and teachers of literature in the activities of reading and interpretation of the literary text and also to others interested in the study or ethnoterminological product, such as: literary critics, linguists, lexicographers, terminographers. To these ends, we started with a study corpus composed of twenty regionalist literary works of sertanista themes by notable authors in the Brazilian Literature, five of them coming from picturesque Regionalism, ten from Critical Regionalism and five from Super-Regionalism, forming a total of three subcorpora. The study corpus, when processed by the program WordSmith Tools 6.0, specifically by its three tools: Concord, KeyWords and WordList, allowed us to identify the vocab-terms and have access to their linguistic contexts of occurrences to elaborate the definitions proposals. Thus, we elaborate the ethnoterminographic fiches, constructed the vocabulary entries and described the microstructure and the macrostructure of the vocabulary, demonstrating at each stage the feasibility of our proposal. We conclude that the vocab-terms present semantic specificities referring to the socio-linguistic-cultural universe of the sertanejo, that they erect from the linguistic conceptualization of this group. These units acquire this double functionality (vocab and term) in the sertanista regionalist literary discourse due to the following aspects: they are part of a conceptual system structured within a specific theme, the sertanista theme; act in the composition of a fictional world semiotically constructed by the modeling force of literary language; have intertextuality and interdiscursivity intra and interuniverse of discourse with ethnoliterary discourses; they actualize a system of values, a system of beliefs, in short, the axiology of a specific group, the sertanejo of the Brazilian backlands; designate concepts formed with semes of the universe of discourse in which they are used; refer to the particularities of a real/fictional world as to its symbolic function; act on the memory plane, so that it is possible to retrieve from the narratives elements of the culture and language of the sertanejo, still present in our society. Thus, we believe that this work may result in significant contributions to ethnoterminological studies and, also, serve as subsidy for new ethnoterminographic works and future research.

**Keywords:** Ethnoterminology. Vocab-term. Corpus Linguistics. Sertanista Regionalist Literature. Vocabulary.

## LISTA DE FIGURAS

FIGURA 1- Tensão dialética vocábulo x termo.....	80
FIGURA 2- Tela inicial do <i>WordSmith Tools</i> versão 6.0.....	103
FIGURA 3- Visão parcial do armazenamento do <i>corpus</i> .....	111
FIGURA 4- Visão parcial dos textos da subpasta Reg1.....	112
FIGURA 5- Lista das vinte duas primeiras palavras do 1º <i>subcorpus</i> .....	113
FIGURA 6- Lista das vinte duas primeiras palavras do 2º <i>subcorpus</i> .....	113
FIGURA 7- Lista das vinte duas primeiras palavras do 3º <i>subcorpus</i> .....	114
FIGURA 8- Visão parcial da lista de palavras-chave do 1º <i>subcorpus</i> .....	115
FIGURA 9- Visão parcial da lista de palavras-chave do 2º <i>subcorpus</i> .....	116
FIGURA 10- Visão parcial da lista de palavras-chave do 3º <i>subcorpus</i> .....	116
FIGURA 11- Lista das linhas de concordâncias da palavra “sertão” com destaque na posição 2 para um enunciado definatório.....	118
FIGURA 12- Fragmento do texto onde a palavra selecionada aparece no <i>corpus</i> .....	118
FIGURA 13- Visão parcial da lista de palavras-chave do 1º <i>subcorpus</i> no processo de limpeza.....	119
FIGURA 14- Visão parcial da lista de palavras-chave do 2º <i>subcorpus</i> no processo de limpeza.....	119
FIGURA 15- Visão parcial da lista de palavras-chave do 3º <i>subcorpus</i> no processo de limpeza.....	120
FIGURA 16- Modelos de verbetes segundo Barbosa (2001).....	128
FIGURA 17- Verbetes do vocábulo-termo “cangalha” na versão <i>online</i> .....	130
FIGURA 18- Tela de acesso do administrador.....	133
FIGURA 19- Tela de cadastro das obras literárias em PDF.....	134
FIGURA 20- Tela de cadastro do texto literário com as <i>tags</i> .....	134
FIGURA 21- Tela de cadastro dos vocábulos-termos.....	135
FIGURA 22- Tela de cadastro das acepções.....	135
FIGURA 23- Tela da interface Vocabulário.....	136
FIGURA 24- Tela da interface Vocabulário com os textos literários onde o vocábulo-termo aparece.....	136
FIGURA 25- Texto literário “O Sertanejo” com destaque para o vocábulo-termo “aguilhada”.....	137

FIGURA 26- Janela <i>pop-up</i> do vocábulo-termo “aguilhada”.....	137
FIGURA 27- Tela da interface Leitura com destaque nas obras do Regionalismo Pitoresco.....	138
FIGURA 28- Obra “O Sertanejo” no arquivo PDF.....	138
FIGURA 29- Obra “O Sertanejo” em HTML.....	139
FIGURA 30- Tela inicial de consulta.....	140
FIGURA 31- “Dados da pesquisa” com destaque para a tese.....	140
FIGURA 32- Busca pelo vocábulo-termo “jirau” na interface Vocabulário.....	141
FIGURA 33- Verbete do vocábulo-termo “jirau” e os textos literários em que ocorre.....	142
FIGURA 34- Busca pela obra “O Quinze” para leitura e visualização das definições.....	142
FIGURA 35- Visualização da obra “O Quinze” para leitura.....	143
FIGURA 36- Visualização da definição de “jirau” na obra “O Quinze” em janela <i>pop-up</i> ...	143
FIGURA 37- Tela da interface Leitura.....	144
FIGURA 38- Busca pela obra “O Garimpeiro” para leitura em HTML com as <i>pop-up</i> .....	144
FIGURA 39- Busca pela obra “O Garimpeiro” para leitura em PDF.....	145

## LISTA DE QUADROS

QUADRO 1- Linguagem literária.....	28
QUADRO 2- Obras literárias do <i>corpus</i> de estudo.....	65
QUADRO 3- Tipologia de processos de constituição de conjuntos terminológicos e vocabulares.....	81
QUADRO 4- Tipologia de obras lexicográficas e terminográficas conforme Barbosa (2001).....	91
QUADRO 5- <i>Corpus</i> de estudo.....	107
QUADRO 6- Tipologia do <i>corpus</i> de estudo.....	109
QUADRO 7- Lista dos vocábulos-terms selecionados.....	121
QUADRO 8- Modelo de ficha etnoterminográfica.....	123
QUADRO 9- Exemplo de verbete.....	129
QUADRO 10- Caracterização do Vocabulário.....	131
QUADRO 11- Síntese dos resultados obtidos.....	147

## LISTA DE TABELAS

TABELA 1- Dados estatísticos do <i>corpus</i> de estudo.....	109
TABELA 2- Quantidade de itens que restaram nas listas de palavras-chave.....	120

## LISTA DE GRÁFICOS

GRÁFICO 1- Percentual de vocábulos-terms dicionarizados, não dicionarizados e com acepções diferentes.....	168
GRÁFICO 2- Número de vocábulos-terms presentes em cada dicionário.....	169

## LISTA DE ABREVIATURAS E SIGLAS

adj. – adjetivo

B – dicionarizado em Bluteau

exp. – expressão

fig. – sentido figurado

H – dicionarizado no *Houaiss*

n/d – não dicionarizado em nenhuma das obras consultadas

n/e – não encontrado

P – dicionarizado em Silva Pinto

RP- Regionalismo Pitoresco

RC- Regionalismo crítico

SR- Super-Regionalismo

S – dicionarizado em Moraes Silva

s. m – substantivo masculino

s. f. – substantivo feminino

v. – verbo

### **Siglas para os títulos das obras:**

AB – A Bagaceira

AC – Auto da Compadecida

CB – Chapadão do Bugre

CG – Contos Gauchescos

CL- O Coronel e o Lobisomem

GS- Grande Sertão: Veredas

IN - Inocência

LH- Luzia-Homem

ME – Menino de Engenho

OC – O Cabeleira

OG – O Garimpeiro

OQ – O Quinze

OS – O Sertanejo

OS- Os Sertões

PS – Pelo Sertão

SG – Sargento Getúlio

SV – Seara Vermelha

TB - Tropas e boiadas

UR- Urupês

VS - Vidas secas

## SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>16</b>
1.1 Contextualização da pesquisa, estado da arte e justificativa.....	16
1.2 Hipóteses e questões de pesquisa.....	22
1.3 Objetivos.....	23
1.4 Estrutura da tese.....	24
<b>2 UNIVERSO DA PESQUISA.....</b>	<b>26</b>
2.1 A linguagem literária.....	26
2.2 O discurso etnoliterário.....	32
2.3 O universo literário, temático e linguístico do <i>corpus</i> de estudo.....	36
<b>3 FUNDAMENTAÇÃO TEÓRICA.....</b>	<b>68</b>
3.1 Da Terminologia à Etnoterminologia.....	68
3.1.1 O vocábulo-termo: unidade mínima de significação da Etnoterminologia.....	78
3.1.2 A formação do conceito nos discursos etnoliterários.....	82
3.2 As definições e suas especificidades.....	85
3.3 Caracterização de obras Lexicográficas e Terminográficas: dicionário, vocabulário e glossário.....	90
3.4 Linguística de <i>Corpus</i> .....	94
3.4.1 <i>WordSmith Tools</i> : ferramentas e utilitários.....	102
<b>4 METODOLOGIA.....</b>	<b>106</b>
4.1 O <i>Corpus</i> de estudo.....	106
4.2. Procedimentos metodológicos.....	110
4.2.1 Compilação e preparação do <i>corpus</i> de estudo.....	110
4.2.2 Lista de palavras.....	112
4.2.3 Lista de palavras-chave.....	114

4.2.4 Aplicação da ferramenta <i>Concord</i> e consulta aos dicionários para identificação dos vocábulos-terminos.....	117
4.2.5 Seleção dos vocábulos-terminos.....	121
4.2.6 Elaboração da ficha etnoterminográfica.....	123
4.3 O público-alvo.....	125
4.4 Construção dos verbetes e organização do Vocabulário: micro e macroestruturas.....	126
4.4.1 A Macroestrutura: elaboração da página <i>Web</i> .....	132
4.4.2 Elaboração do sistema.....	132
4.4.3 Apresentação da página e o modo de consulta.....	139
<b>5 APRESENTAÇÃO, ANÁLISE E DISCUSSÃO DOS RESULTADOS.....</b>	<b>147</b>
5.1 Análise quantitativa: número de vocábulos-terminos dicionarizados, não dicionarizados e com acepções diferentes dos dicionários.....	167
5.2 Quanto ao número de ocorrências e fases do Regionalismo.....	169
5.3 Apresentação das fichas etnoterminográficas e dos verbetes.....	171
5.4 Análise dos vocábulos-terminos e discussão dos resultados.....	193
<b>CONSIDERAÇÕES FINAIS.....</b>	<b>198</b>
<b>REFERÊNCIAS.....</b>	<b>202</b>



# 1 INTRODUÇÃO

## 1.1 Contextualização da pesquisa, estado da arte e justificativa

Esta tese é resultado de nossa pesquisa em nível de doutorado, realizada no âmbito do Programa Pós-graduação em Estudos Linguísticos (PPGEL) da Universidade Federal de Uberlândia (UFU), vinculada à linha de pesquisa 1 - Teoria, descrição e análise linguística. A pesquisa está atrelada ao projeto *Linguística de Corpus aplicada a pesquisas empírico-descritivas: tradução, interlíngua, fraseologia, discurso referido e transitividade em contraste*, coordenado pelo Prof. Dr. Ariel Novodvorski, que integra projetos de pesquisa em nível de Graduação (iniciação científica) e Pós-graduação (mestrado e doutorado) na linha de pesquisa 1 do Programa de Pós-graduação em Estudos Linguísticos (PPGEL/UFU), e insere-se no âmbito dos grupos de pesquisa GECon (Grupos em Estudos Contrastivos) coordenado pelo respectivo professor e GPELC (Grupo de pesquisa e estudos em Linguística de *Corpus*) coordenado pelo Prof. Dr. Guilherme Fromm com a participação do Prof. Ariel Novodvorski.

Partindo de um diálogo entre Linguística e Literatura pela via etnoterminológica, buscamos, neste trabalho, analisar os vocábulos-termos presentes em obras regionalistas brasileiras de temática sertanista, com o objetivo de propor um modelo de vocabulário *online* na tentativa de auxiliar alunos e professores de literatura<sup>1</sup> nas tarefas de leitura e interpretação do texto literário. Por se tratar de um objeto linguístico pouco estudado no discurso literário regionalista e, ainda, por ser um elemento preponderante na interpretação do texto literário, acreditamos que um vocabulário dessas unidades lexicais não apenas servirá como suporte na compreensão de obras clássicas da nossa literatura, mas também como subsídio para outras propostas ou trabalhos futuros.

Apesar de partirmos de uma hipótese norteadora de que é possível encontrar elementos de natureza especializada em um texto literário, as considerações teóricas resultantes desta investigação foram derivadas de observações de usos das unidades lexicais em contextos presentes no *corpus*, o que configura uma abordagem direcionada pelo *corpus*. Tal abordagem “tem como doutrina a não-categorização *a priori*” (BERBER SARDINHA, 2004, p. 36), isto é, a não formalização de teorias antes de se conhecer previamente as condições

---

<sup>1</sup> Por se tratar de um repertório que traz significados específicos do universo etnoliterário sertanista, presumimos que a nossa proposta de vocabulário poderá servir como auxílio a alunos e professores de literatura na compreensão do texto literário e, também, como subsídio a outros interessados pelo estudo ou pelo produto etnoterminológico, como: linguistas, lexicógrafos; terminógrafos.

linguísticas e textuais do *corpus*. Assim, foi a partir da investigação do nosso *corpus* de estudo e das observações trazidas pela banca na ocasião da qualificação do projeto de tese que pudemos propor uma abordagem teórico-metodológica adequada ao trabalho de descrição e análise dos dados.

Desta feita, o nosso aporte teórico-metodológico é constituído pelos pressupostos da Etnoterminologia em correlação com a Linguística de *Corpus* e com a Terminologia Textual e Terminologia Cultural. A Etnoterminologia (BARBOSA, 2005, 2006, 2007, 2009), por se dedicar ao estudo das unidades multifuncionais dos discursos etnoliterários, não apenas tem o texto como seu objeto primário de análise como também o contexto sociocultural, ou seja, é no contexto imediato (no texto) e no contexto sociocultural (no seio de uma cultura específica) que as unidades lexicais dos discursos etnoliterários adquirem sentido. Assim, partindo de uma abordagem transdisciplinar, da integração entre Linguística e Literatura, que buscamos analisar o léxico sertanista das principais obras regionalistas da Literatura brasileira enfocando os vocábulos-termos como nosso objeto de estudo.

Antes de tratarmos das hipóteses e questões da pesquisa, dos objetivos e da justificativa sobre a escolha do nosso *corpus* de estudo, cumpre assinalar que o estudo do léxico com enfoque nas relações entre língua, cultura e sociedade tem se multiplicado, substancialmente, nos últimos anos, dando origem a várias pesquisas e novas abordagens teóricas. De maneira particular, o estudo das unidades lexicais que congregam essas relações e que denotam funções específicas em um dado universo de discurso tem sido ultimamente objeto de investigação e discussão no âmbito das pesquisas de base terminológica. Isto se deve aos grandes avanços científicos e tecnológicos e ao rápido crescimento da Linguística de *Corpus* no Brasil, a qual vem trazendo grandes contribuições em diversas áreas e, ao mesmo tempo, expressivo desenvolvimento nas pesquisas linguísticas.

Fato a ser considerado é que contamos, hoje, com inúmeros trabalhos que têm se valido de dados de *corpora* para focalizar questões terminológicas por meio de ferramentas computacionais. Com a criação e propagação de variados *software* foi possível não só coletar e armazenar uma grande quantidade de textos instantaneamente, como visualizar novos e variados fenômenos linguísticos, isto é, questões que antes não eram vistas ou que não poderiam ser analisadas manualmente passaram a ser percebidas e estudadas no âmbito das várias áreas da Terminologia, como é o caso das unidades multifuncionais dos discursos etnoliterários.

A multifuncionalidade das palavras em determinados universos de discurso tem conduzido estudiosos a perceberem a tênue fronteira entre a unidade lexical especializada e a não especializada, ou seja, entre o termo e o vocábulo. Foi observando essa multifuncionalidade das palavras nos discursos etnoliterários e a evidente necessidade de uma disciplina que estudasse essas unidades lexicais que Maria Aparecida Barbosa propôs a formalização de uma nova disciplina científica, a Etnoterminologia, subárea da Terminologia que tem o vocábulo-termo como sua unidade mínima de significação. São, portanto, os preceitos teóricos desta disciplina correlacionada à Terminologia Textual e à Terminologia Cultural que apoiamos e sustentamos as discussões apresentadas neste trabalho.

A Etnoterminologia estuda os discursos etnoliterários, como os de literatura oral, literatura popular, literatura de cordel, fábulas, lendas, mitos, folclore e também os discursos das linguagens especiais com baixo grau de tecnicidade e de cientificidade (BARBOSA, 2009). Ao examinar o universo lexical desses discursos, Barbosa (2005) percebeu que as unidades lexicais contidas nos textos possuíam um estatuto diferente, ou seja, elas possuíam qualidades das linguagens de especialidade e qualidades da linguagem literária, adquirindo, desse modo, o estatuto de vocábulo-termo. Segundo a autora, é um vocábulo, nos seus aspectos referenciais, pragmáticos e simbólicos, e é um termo, na medida em que a unidade léxica em questão tem características de uma linguagem de especialidade (BARBOSA, 2005).

Essas unidades representam não somente um contexto linguístico, mas um contexto sociocultural - a axiologia de um grupo, de uma sociedade, ou seja, os valores, os costumes, os usos, as crenças, os hábitos de um grupo de falantes linguística, histórico e socialmente definido. Por essa razão, elas apresentam características peculiares: no nível da norma, exerce duas funções: vocábulo e termo, e possui significados específicos segundo o universo de discurso a que pertencem; no nível de sistema essa unidade é plurifuncional, ou seja, é a norma e o universo de discurso em que se insere que determina o seu estatuto. (BARBOSA, 2006a).

Logo, o vocábulo-termo é uma unidade lexical composta de expressão e conteúdo que designa um conceito específico dentro do universo de discurso etnoliterário. Em outras palavras, é um vocábulo “metasemiótico”, um quase “termo-técnico” que pertence a uma linguagem especial/especializada (BARBOSA, 2005). Isto porque os discursos etnoliterários são altamente significativos, reveladores de uma cultura, de uma identidade específica de um grupo histórico e socialmente definido. Desse modo, a natureza “especializada” do vocábulo-

termo aflora vínculos de pertencimento ao grupo (características de etnicidade), e é nesse ambiente cultural humano que essas unidades lexicais adquirem sentido.

Desse modo, é imprescindível interpretar, com toda a clareza, de um lado, os traços semântico-conceituais das unidades lexicais que integram um discurso etnoliterário e, de outro lado, as relações de significação que estabelecem no interior de um texto, sempre lembrando que cada vocábulo-termo constitui um registro linguístico, resultado das produções culturais humanas, que atrelados a uma identidade cultural corroboram para a construção de sentidos de um mundo literário verbalizado.

É mediante essa capacidade de construir sentidos em um mundo possível, de evocar imagens, ações, objetos e integrá-los em um mundo ficcional modelizado pelo “traçado verbal-estético” e pela “síntese criativa (racional e emotiva)” do sujeito enunciador, que a linguagem literária se torna o agente responsável pela multiplicidade de ‘mundo literários’ à nossa volta (STEGGER, 1987). Mas não se trata de uma linguagem nova alheia às demais variedades linguísticas, tampouco de uma imitação da realidade (*mimesis*), mas de uma linguagem que se utiliza de outras para construir seu próprio universo, seu próprio contexto referencial que engendram significados específicos em um universo literário particular.

É essa concepção de linguagem dada por Steger (1987) que adotamos neste trabalho. Trata-se de uma linguagem com função pragmática e social que modula e constrói seu próprio universo, gerando especificidades semântico-conceituais no interior de um discurso, de um texto literário. Tal concepção reforça o *princípio do poeta creator* e retoma a noção de *síntese criativa (racional e emotiva) do sujeito enunciador*, ou seja, o autor-criador, ao construir um mundo literário possível, pode fazer uso de um léxico específico para designar elementos próprios desse mundo. Em outros termos, a autonomia da criação literária dá ao autor-criador liberdade para expressar-se de diferentes formas, por meio de vários recursos linguísticos, podendo até adotar um modo de dizer próximo de uma linguagem especializada, por exemplo.

Isto nos leva a inferir e a enfatizar o entendimento de Barbosa (2005, 2006, 2009) de que é possível encontrar elementos de natureza especializada na composição sintética da obra literária. Partindo dessa perspectiva, propomos um estudo do léxico sertanista em um *corpus* literário composto por obras regionalistas de autores consagrados na Literatura brasileira, com o propósito de elaborar um modelo de vocabulário que pudesse servir como auxílio a alunos e professores de literatura na compreensão do texto literário e, também como respaldo a outros interessados pelo estudo ou produto etnoterminológico.

Assim, com sustento nos recursos instrumentais e metodológicos da Linguística de *Corpus*, mais especificamente no programa *WordSmith Tools 6.0* (SCOTT, 2012), buscamos compilar as obras mais representativas de cada fase do regionalismo, circunscritas em universo de discurso etnoliterário de temática sertanista, e que tivessem um repertório lexical diversificado e amplo. À vista disso, escolhemos os romances e contos em virtude da extensão linguística e lexical que possuem, com exceção da peça teatral *Auto da Compadecida* de Ariano Suassuna, por encontrarmos nela características marcantes da terceira fase do Regionalismo.

Investigando as obras que marcaram as fases do Regionalismo literário, selecionamos: *O Sertanejo* de José de Alencar, *O Garimpeiro* de Bernardo Guimarães, *Inocência* de Visconde de Taunay, *O Cabeleira* de Franklin Távora e *Pelo Sertão* de Afonso Arinos, da primeira fase – **Regionalismo Pitoresco** (enaltece o homem do sertão e a terra em que vive); *Os Sertões* de Euclides da Cunha, *Luzia-Homem* de Domingos Olímpio, *Contos Gauchescos* de João Simões Lopes Neto, *Tropas e Boiadas* de Hugo de Carvalho Ramos, *Urupês* de Monteiro Lobato, *A Bagaceira* de José Américo de Almeida, *O Quinze* de Raquel de Queiroz, *Menino de Engenho* de José Lins do Rego, *Vidas Secas* de Graciliano Ramos e *Seara Vermelha* de Jorge Amado, da segunda fase – **Regionalismo Crítico** (denúncia dos problemas econômicos e sociais das regiões interioranas do Brasil); *Grande Sertão: Veredas* de Guimarães Rosa, *O Coronel e o Lobisomem* de José Cândido de Carvalho, *Chapadão do Bugre* de Mário Palmério, *Sargento Getúlio* de João Ubaldo Ribeiro e *Auto da Compadecida* de Ariano Suassuna, da terceira fase- **Super-Regionalismo** (transcendência do regional para o universal através da exploração dos grandes problemas que envolvem o ser humano em qualquer lugar). Nesse ponto é válido destacar que a segunda fase do Regionalismo, denominada por Antonio Candido (1989) de *Regionalismo problemático ou crítico*, foi um período de grande produtividade literária em virtude da propagação da famosa geração do romance regionalista de 30, caracterizado como romance neorrealista moderno da segunda fase do Modernismo, também chamado de “Romance social Nordeste”. Por essa razão, selecionamos um número maior de textos, sendo 10 desta fase e 5 das demais, somando um total de 20 textos. Mesmo sabendo da existência de outras obras de um mesmo autor, escolhemos apenas uma de cada escritor devido às semelhanças estilísticas e lexicais adotadas por estes ao representarem a sua vertente regionalista.

Em resumo, utilizamos as ferramentas do *WordSmith Tools* (doravante WST) para gerar as listas de palavras, as palavras-chave e concordâncias e também obras de referência

para verificar a existência da forma coletada, o seu conceito e a sua relação com o universo sertanejo. Dentre os dicionários mais conhecidos e renomados, optamos pelo Dicionário *Houaiss* (2009) em sua versão eletrônica, por ser uma obra de referência nos estudos lexicográficos de língua portuguesa com um vasto repertório lexical, incluindo regionalismos, expressões e modismos usados no Brasil; e pelos dicionários *online* da Biblioteca Brasileira da USP: *Vocabulario Portuguez e Latino*, de Raphael Bluteau (1712), *Diccionario da Lingua Portuguesa*, de Antonio de Moraes Silva (1789) e *Diccionario da Lingua Brasileira*, de Luiz Maria da Silva Pinto (1832), por apresentarem formas antigas da língua portuguesa e diversos arcaísmos constantes no nosso *corpus* de estudo, nos auxiliando na elaboração das definições.

Escolhemos essas obras regionalistas como *corpus* de análise, em primeiro lugar, por focalizarem o universo sertanejo em seus diferentes aspectos, permitindo-nos conhecer o homem do sertão, suas características físicas, comportamentais e, especialmente linguísticas. Em segundo lugar, por estarem no escopo dos discursos etnoliterários *lato sensu* – aqueles que utilizam elementos provenientes de discursos etnoliterários *stricto sensu* (como elementos da literatura oral, do folclore, das lendas e mitos de uma dada cultura), isto é, os discursos-ocorrência do *corpus* de análise por utilizarem, com frequência, elementos da cultura sertaneja, estabelecendo relações intertextuais e interdiscursivas com o universo de discurso etnoliterário, não apenas se tornam objeto de pesquisa legítimo em Etnoterminologia, como apresentam um material linguístico suficiente e adequado para esta pesquisa.

Investigando o estado da arte deste estudo no que tange ao seu campo teórico e temático, percebemos que a literatura regionalista é um universo de discurso ainda pouco explorado em Etnoterminologia. Dos trabalhos encontrados (citados no terceiro capítulo deste trabalho, p.75), apenas um (1) tem como objeto de investigação um texto regionalista, abordando também o léxico, mas por um viés diferente, a saber: *Uma abordagem etnoterminológica de Grande Sertão: Veredas* (2011) de Vanice Ribeiro Dias Latorre (dissertação de mestrado orientada por Maria Aparecida Barbosa).

Utilizando um *corpus* de estudo e uma abordagem diferente da que propusemos no nosso trabalho, Latorre analisa nessa pesquisa apenas as denominações constantes em *Grande Sertão: Veredas*, ou seja, a natureza dos semas conceptuais formadores do vocábulo-termo. Portanto, o nosso trabalho é inédito, tanto na abordagem, como na proposta de disponibilização do produto em formato de vocabulário *online*<sup>2</sup>.

---

<sup>2</sup> Até o momento da pesquisa, não encontramos nenhum dicionário, vocabulário ou glossário com essa temática, tampouco alguma proposta etnoterminográfica, o que justifica a pertinência de nosso trabalho.

Por essas razões, acreditamos que este estudo pode resultar em contribuições para a área e, ainda, servir como subsídio para pesquisas futuras. Destarte, pretendemos contribuir para uma melhor compreensão do vocábulo-termo no tocante ao seu funcionamento linguístico e aos papéis cultural e social que desempenha, e, por outro lado, também colaborar para o reconhecimento de novas possibilidades de investigação na esfera dos estudos etnoterminológicos.

Além disso, também esperamos que nossa proposta analítico-descritiva possa ser replicada e, ainda, que a partir dela surjam novas pesquisas e propostas etnoterminográficas que levem em conta outros textos literários e outros universos de discurso passíveis de serem abordados etnoterminologicamente.

Ademais, apesar de reconhecer as limitações de nossa proposta, acreditamos em sua validade e pertinência. Por um lado, por ser um trabalho inédito que trata de um léxico específico do universo etnoliterário sertanista ainda pouco explorado e, até então, não registrado e viabilizado sob a forma de vocabulário ou glossário; por outro, por estar em formato eletrônico, o que propicia rapidez e facilidade no acesso, bem como a redução de custos; e, também, por apresentar uma estrutura simples de fácil acesso e consulta; e por último, por apresentar duas interfaces, uma de Vocabulário e outra de Leitura, que permite ao consulente ter acesso tanto ao Vocabulário com os vocábulos-terminos e suas respectivas definições quanto às obras literárias para leitura.

Feitas essas considerações e esclarecimentos sobre o nosso trabalho, convém agora tratar das hipóteses, dos objetivos e da estrutura da tese como um todo.

## **1.2 Hipóteses e Questões de Pesquisa**

Considerando que as obras regionalistas de temática sertanista se configuram como um discurso etnoliterário passível de investigação em Etnoterminologia, partimos do princípio de que as unidades lexicais presentes nos textos têm um estatuto particular e significados específicos dentro do contexto em que estão inseridas, quer no contexto linguístico, quer no contexto sociocultural.

Pressupomos que tais unidades, apesar de serem ficcionais, refletem um sistema de valores culturais do mundo sertanejo ainda presentes na memória coletiva da nossa sociedade. E, mais do que isso, presumimos que elas constituem um registro linguístico de uma

identidade cultural que existiu e que ainda existe não apenas em nossa memória, mas, sobretudo, em nossa língua.

Tais hipóteses nos levam às seguintes questões de pesquisa:

1. Quais seriam os procedimentos metodológicos pertinentes à identificação dos vocábulos-termos num *corpus* da Literatura Regionalista sertanista, a partir dos princípios e ferramentas próprias da Linguística de *Corpus*?
2. Os vocábulos-termos presentes nas obras literárias apresentam características linguísticas e socioculturais do universo sertanejo?
3. Qual seria a distribuição dos vocábulos-termos no *corpus*, conforme cada fase da Literatura Regionalista, em razão de suas ocorrências?
4. Quais as diferenças lexicais nos discursos-ocorrência de uma determinada fase do Regionalismo em contraste com as demais fases?
5. Como disponibilizar o produto final desta pesquisa para o acesso, seja de estudantes, docentes, leitores ou pesquisadores?

### 1.3 Objetivos

Na tentativa de buscar possíveis respostas a essas questões propomos uma base teórico-metodológica para a análise e descrição dos vocábulos-termos que integram o universo de discurso literário regionalista, utilizando como *corpus* as obras de temática sertanista que marcaram as fases do Regionalismo literário. Assim, esta pesquisa tem como objetivo principal:

- Propor um modelo de Vocabulário *online* dos vocábulos-termos presentes em obras regionalistas sertanistas que compõem o *corpus* de estudo, com a finalidade de auxiliar alunos e professores de literatura nas tarefas de leitura e interpretação do texto literário.

Tendo em vista as diferentes etapas da pesquisa para o alcance do objetivo geral, elencamos os seguintes objetivos específicos:

- Identificar os vocábulos-termos presentes no *corpus* de estudo, utilizando as ferramentas do WST;
- Analisar os candidatos a vocábulos-termos por meio das palavras-chave, das linhas de concordância (contexto linguístico) e de consultas aos dicionários;



- Descrever a distribuição dos vocábulos-terminos selecionados, conforme suas ocorrências, nas diferentes fases da Literatura Regionalista;
- Analisar e descrever em fichas etnoterminográficas, em contraste entre as diferentes fases do Regionalismo, as diferenças lexicais nos discursos-ocorrência, para posterior elaboração dos verbetes.
- Propor um modelo de vocabulário eletrônico para o acesso *online* com duas interfaces: uma de leitura, contendo todas as obras; a outra do vocabulário, para disponibilização dos vocábulos-terminos e suas respectivas definições.

#### 1.4 Estrutura da tese

Dada a proposta de pesquisa apresentada acima, a presente tese está organizada nos seguintes capítulos:

Capítulo 1 – **Introdução**; O capítulo 2, **Universo da pesquisa**, apresenta uma breve contextualização da pesquisa, mostrando o universo literário, temático e linguístico do *corpus* de estudo (ALMEIDA, 1981; ATHAYDE, 2004; BILAC, 1934; CANDIDO, 1989, 1997, 2000, 2002; COUTINHO, 2001; GALVÃO 1976; LEITÃO JÚNIOR, 2012; TUFANO, 1988); o conceito e as características do discurso etnoliterário (BARBOSA, 2004, 2005, 2009; PAIS; BARBOSA, 2004); e, ainda, a linguagem literária enquanto funcioleto (LUCKMANN, 1966; NOMURA, 1993, 1996; STEGER, 1987).

O capítulo 3, **Fundamentação teórica**, traça um breve panorama dos estudos em Terminologia, abordando as principais teorias e perspectivas recentes com enfoque na Etnoterminologia (ALMEIDA, 2006; BARBOSA, 2005, 2006, 2007, 2009; BARROS, 2004; CABRÉ, 1999; DIKI-KIDIRI, 2009; FINATTO, 2011; GAUDIN, 2014; KRIEGER, 2011 KRIEGER; FINATTO, 2004; LATORRE, 2013; TEMMERMAN, 2000). O capítulo também trata da constituição do vocábulo-termo, da formação do conceito nos discursos etnoliterários, das definições e suas especificidades, da caracterização de obras lexicográficas e terminográficas, bem como dos dicionários, vocabulários e glossários (ANDRADE, 1998; BARBOSA, 1991, 1994, 2001, 2009; BARROS, 2004; BIDERMAN, 2001; FROMM, 2002; KRIEGER; FINATTO, 2004; LATORRE, 2011; PAIS; BARBOSA, 2004; SAGER, 1990; CARDOSO, 2017). A última seção do capítulo versa sobre a Linguística de *Corpus* e as ferramentas do WST, demonstrando suas principais funções e utilidades (BERBER SARDINHA, 2004, 2009; BERBER SARDINHA ; ALMEIDA, 2008; BEVILACQUA, 2013; BIBER, 1998; BOWKER ; PEARSON, 2002; LEECH, 1992; NOVODVORSKI; FINATTO,

2014; PARODI, 2010; SVARTVIK, 1992; TOGNINI-BONELLI, 2001; SCOTT, 2012; SINCLAIR, 1991; VIANA; TAGNIN, 2011).

O capítulo 4, **Metodologia**, caracteriza o nosso *corpus* de estudo e descreve os procedimentos metodológicos utilizados na realização desta pesquisa. As seções relatam, de forma pormenorizada e sequencial, todos os passos realizados durante as várias etapas da pesquisa que vão desde a compilação e preparação do *corpus* de estudo ao levantamento, seleção, descrição e análise dos dados. As seções também descrevem a composição e a elaboração das fichas etnoterminográficas, bem como a constituição e construção dos verbetes.

O capítulo 5, **Apresentação, análise e discussão dos resultados**, exhibe a síntese dos resultados obtidos a partir do nosso *corpus* de estudo e demonstra suas principais características por meio do preenchimento de fichas etnoterminográficas, de análises quantitativas, comparativas e contrastivas.

Por fim, as **Considerações Finais** apresenta uma apreciação final do percurso investigativo, incluindo um resumo dos resultados obtidos, dificuldades, limitações da pesquisa e indicações de futuras investigações.

## 2 UNIVERSO DA PESQUISA

Este capítulo apresenta uma breve contextualização da pesquisa para melhor compreensão do nosso objeto de estudo. Na primeira seção, traçamos algumas considerações sobre a linguagem literária como *funcioleto*, segundo a teoria de Steger (1987); na segunda falamos sobre o conceito e as características do discurso etnoliterário conforme as proposições de Barbosa (2004, 2005, 2009); e por último discutimos sobre o universo literário, temático e linguístico do *corpus* de estudo, mostrando suas características e especificidades.

### 2.1 A Linguagem literária

Embora existam várias propostas de definições para a linguagem literária (como as dos estruturalistas de Praga, dos formalistas russos, dos estruturalistas franceses e soviéticos, dos gerativistas e dos sociolinguistas), não existe um consenso unívoco sobre o conceito relativo à linguagem da literatura, em razão da divergência de ideias e concepções. Por isso, buscamos amparar as nossas discussões na proposta de Hugo Steger (1929-2011), linguista alemão e estudioso literário que propôs a Teoria Funcional da Comunicação.

Nomura (1993, p. 12) ressalta que “nessa teoria, o discurso é considerado em sua ‘função pragmática’, isto é, em sua função social de uso dentro de determinado contexto comunicativo e voltado para atingir determinados fins práticos”; e a linguagem literária, por sua vez, é “definida como um ‘funcioleto’, isto é, ela é caracterizada pelo seu lugar e função social na vida de uma comunidade linguística” (NOMURA, 1993, p. 77).

Nesse sentido, a linguagem literária é definida como a variante linguística da comunicação literária, cuja função é realizar “a síntese criativa, racional e emotiva, do mundo novo modelizado pelo traçado verbal-estético” (STEGER, 1987, p. 110). Não se trata de uma linguagem nova alheia às demais variedades linguísticas, mas de uma linguagem que se utiliza de outras para construir um mundo ficcional baseado em interpretações do mundo real, modelizado pelos recursos próprios da literatura e pela intervenção artística e estilística do autor. Dito de outro modo:

Através de seu próprio universo modelizado pela linguagem literária, a literatura interpreta o mundo à sua maneira, e explica-o, na medida em que, construindo um universo próprio, confronta-o com a percepção social do mundo, apontando eventualmente as falhas nele existentes, ou, então, criando o modelo de um mundo melhor (NOMURA, 1993, p. 77).

Assim, sob uma nova roupagem, a linguagem literária apresenta uma nova realidade a partir de uma percepção própria de mundo que pode ser caracterizada como “síntese criativa (racional e emotiva)” do sujeito enunciador, reportando a noção de autor como “poeta criador” em substituição ao “poeta imitador” da velha arte poética. Trata-se de um mundo possível com verdades provenientes de uma concepção de mundo estética/linguística intimamente ligada aos contextos funcionais comunicativos.

Nessa definição funcional de linguagem literária, Steger (1987) retoma conceitos básicos vigentes em épocas anteriores, como arte, poesia, mundo e realidade; e, através da expressão “traçado verbal-estético”, retoma, também, o conceito de capacidade de modelização da linguagem literária do estruturalista soviético Jurij Lotman. (NOMURA, 1993). Ao propor sua teoria, Steger (1987) retoma basicamente os seguintes conceitos legados do passado:

a) O conceito de *poeta creator*, da Antiguidade, retomado na Renascença, no século XVI:

A noção de autor como criador de seu próprio mundo teria sido contraposta, na Renascença, à noção de autor como *poeta imitador*, e da arte como imitação da natureza (= mundo). Essa concepção de poeta criador deu origem ao conceito moderno de poesia (= literatura) e possibilitou a distinção teórica entre literatura e filosofia, literatura e história, isto é, entre literatura e não-literatura (NOMURA, 1993, p. 78).

Para Steger (1987, p. 107), o conceito de *poeta-creator* “[...] enfatiza o fato de o autor como criador não imitar a criação divina, mas criar seu próprio mundo como Deus no Gênesis”.

b) O conceito dos *mundos possíveis*, de Leibniz, filósofo alemão do século XVIII:

A concepção de poeta criador combina-se com a teoria dos *mundos possíveis*, de Leibniz. Para este, o mundo da fantasia poética é um mundo possível, pois o poder de evocação da imaginação poética tem a sua própria lógica. Quanto ao conteúdo e à função desta proposta, Leibniz afirma que “a história deve tratar do mundo real, e a poesia (= literatura), de um mundo possível” (NOMURA, 1993, p. 78).

Retomando a noção de mundos possíveis de Leibniz, Steger (1987, p. 108) afirma que “[...] o mundo da fantasia poética é um mundo possível; a fantasia poética tem portanto sua própria lógica. A história trata do mundo real, a literatura de um mundo possível [...]”

c) As concepções históricas de *arte* (séculos XVI, XVII, XVIII):

A concepção antiga de arte como imitação da natureza (mimese) tinha, como objetivo intencional, *ensinar e divertir* (“*docere et delectare*”); entretanto, considerados sob a perspectiva de conteúdo e função, os produtos culturais da época não possuíam uma marca explícita interna que permitisse sua identificação como sendo obra religiosa, filosófica, poética, histórica, ou empírico-científica. Um exemplo dessa mescla e superposição de campos funcionais da comunicação é uma obra do séc. XVI, intitulada “Syphilis sive de morbo gallico”: é uma poesia didática de um médico de nome Gerolamo Fracastoro. Nessa obra, o autor descreve, em forma de poema elaborado segundo modelo virgiliano, os sintomas e a cura da doença, dando a um rapaz, tomado por blasfêmia vítima da doença, o nome de Sífilis (NOMURA, 1993, p. 79).

d) A *separação metodológica* dos campos funcionais das ciências empíricas, da filosofia e da história (século XVIII):

Hamann (1730, 1788), um dos teóricos do Romantismo na Alemanha [...] procurou operacionalizar os diferentes modos de acesso ao mundo aplicando às ciências do conhecimento as operações lógicas de *análise* e *síntese*: a filosofia e as ciências empíricas foram classificadas como sendo “operações metodológicas de *análise* geral do mundo feitas mediante o uso da linguagem” [...]; a literatura, ao contrário, exigiria uma operação de síntese: a poesia (= literatura) seria em seu entender, “um projeto sintetizante de um mundo ficcional” moldado pela linguagem. Com a linguagem da poesia ligar-se-ia o espírito de *profecia*: a ficção teria o poder mágico de evocar e tornar concretas as expressões configuratórias do passado e do futuro (NOMURA, 1993, p. 80, grifo do autor).

Segundo Steger (1987, p. 109), Hamann entende que a ciência atua por meio de “operações metodológicas de uma análise generalizante do mundo real através da linguagem”, enquanto a literatura traça “um projeto sintetizante de um mundo ficcionável a partir da linguagem”.

Com essa separação metodológica em operações de análise e síntese, buscou-se considerar os produtos culturais a partir da noção de campos funcionais, assim teria: o campo funcional das ciências empíricas, o campo funcional da filosofia, o campo funcional da literatura e assim por diante.

Sintetizando e esquematizando todas essas proposições e conceitos abordados por Steger (1987) em sua teoria, temos:

QUADRO 1 – Linguagem literária

Linguagem literária	Literatura
Síntese criativa, racional e emotiva de um mundo novo	O autor, como agente responsável pela construção desse modelo de mundo, tem o poder

	de colocar dentro dele um jogo múltiplo de intenções
Criação	Criação
Ficção	Ficção
Modelo de um mundo ficcional com leis próprias	A literatura é a criação de um modelo de mundo possível, regido por leis próprias
A verdade advinda do poder de persuasão/convencimento de uma concepção estético-verbal	A verdade desse mundo emana do poder de convencimento do traçado verbal estético
Objetos e ações evocados e montados pela linguagem no espaço-tempo ficcionais	A operação de acesso a esse mundo, a operação de síntese, cabe à linguagem literária.

Fonte: Elaboração da autora com base em Nomura (1993) e Steger (1987)

A noção semiótica de capacidade de modelização da linguagem literária completa a série de conceitos que fundamentam a “tese de literatura como funcioleto”. É mediante essa capacidade de “evocar imagens, ações e objetos e de integrá-los em um modelo de mundo verbalizado de forma estética e sintética, que a linguagem literária se torna o agente gerador responsável pela multiplicidade de ‘mundo literários’ à nossa volta” (NOMURA, 1993, p. 82).

A noção de função, no contexto dessa teoria, refere-se tanto aos modos de uso da linguagem em campos específicos (campo funcional da literatura, da filosofia, etc.), como a função social que ela exerce na vida de uma comunidade linguística; e a representação dessa função relaciona-se à noção sociológica de “motivação pragmática” “(abordagem fenomenológica, que remete ao filósofo austríaco Schütz, definida como um ‘aglomerado de significados que diz respeito diretamente a ações presentes e futuras’)” (BERGER; LUCKMANN, 1966, p. 41 apud NOMURA, 1996, p. 199). Quer isto dizer que o modo particular de uso da língua é determinado por intenções pragmáticas, ou seja, pelos propósitos comunicativos em uma dada situação discursiva, e pelo campo específico da referência “(o campo dos objetos do mundo real e dos conteúdos), onde tem validade determinado tipo de comunicação” (NOMURA, 1996, 199). No campo funcional da literatura, a motivação pragmática é a de modelizar uma concepção intencional de um mundo possível, e a função exercida pela linguagem é a de sintetizar criativa, racional e emotivamente um mundo novo modelizado pela comunicação verbal-estética.

Segundo Nomura (1996), ao atribuir uma “função social”, uma “motivação pragmática” à literatura, Steger (1987) possibilita sua integração no quadro geral da comunicação e da ação social. E ao considerar a linguagem literária como forma de expressão da criatividade humana, Steger (1987) estabelece relações íntimas entre literatura e sociedade, mostrando que tanto os fatores socioculturais quanto os estético-literários dependem da concepção social de arte e de cultura vigente no recorte espaço-tempo vivido pelas comunidades que os produzem e os recebem. Nas palavras de Nomura (1996):

A tese da linguagem literária como funcioleto teve o mérito de possibilitar integrar a literatura no quadro geral da comunicação, ao levar em conta a produção e a recepção de textos e ao atribuir à literatura uma motivação pragmática, uma função social definida, na vida da comunidade de falantes. Com a atribuição de uma função social à literatura, estaria afastada a hipótese de uma comunicação “disfuncional” ou idealizada, sem vínculo, portanto, com a sociedade. A representação da linguagem literária com função social dá a entender que a literatura depende do conceito social de arte e cultura vigente em um determinado recorte espaço-temporal (NOMURA, 1996, p. 200).

Entendemos, portanto, que, para Steger (1987) a linguagem literária é uma forma de expressão da criatividade humana que tem como função conduzir e renovar a relação do homem com a realidade, com a sociedade. Nessa definição de linguagem literária como “funcioleto” o autor deixa claro que: a) “o texto literário refere-se a si mesmo, isto é, ao mundo construído dentro dele e que se refere indiretamente ao mundo real”; b) “o texto literário refere-se ao mundo da literatura, colocado contra um painel de fundo sociocultural e literário, formado por um contexto múltiplo de interação desses fatores” c) “a autonomia da criação literária dá ao autor-criador de mundos possíveis liberdade total para exprimir-se em prosa, verso ou drama, utilizando qualquer variedade linguística”; d) “a linguagem literária tem a responsabilidade de modelizar uma concepção de mundo”, com isso, o plano de expressão adquire uma nova dimensão, “visto que a força de verdade dessa concepção depende do poder de convicção e de convencimento da linguagem que a modeliza”; e) a concepção de “síntese criativa do universo literário” permite à literatura “conceber seus próprios *mundos verbais em contraste*, na medida em que integra os modos de falar do cotidiano e das ciências e os desenvolve em um terceiro elemento novo na obra literária”; f) esses “mundos verbais em contraste” são constituídos “pelas variantes linguísticas: por um lado, pela linguagem cotidiana e, por outro, pelas linguagens culturais” (como as linguagens das artes, das ciências e das técnicas, da religião, das instituições políticas, da jurisprudência, etc.), enfim, “todas as manifestações discursivas dos campos funcionais da comunicação

moderna”, g) “a imbricação dos diversos códigos, no mundo cotidiano e nas artes em geral, tem seus reflexos na constituição do texto literário” (NOMURA, 1996, p. 201 - 203).

Diante de toda essa explanação, podemos afirmar que encontramos na teoria de Steger (1987) subsídios suficientes para a compreensão das especificidades linguísticas da linguagem literária que reverbera no discurso etnoliterário. Através do seu próprio universo modelizado pela linguagem literária, o discurso etnoliterário apresenta um mundo ficcional construído a partir de dados do mundo real e materializado pelos recursos próprios da literatura e pelas intenções estético-linguísticas do autor.

Essa capacidade que a literatura tem de interpretar o mundo à sua maneira, de construir seu próprio universo, “seu próprio contexto referencial, que por si só permite que venha à tona o novo sentido linguístico” (STEGER, 1987, p. 129), é também explicada por Pais e Barbosa (2004, p. 84) ao tratarem do discurso etnoliterário: “os discursos etno-literários mostram uma visão de mundo, apresentam as grandes linhas de um *mundo semioticamente construído*”, e por meio de uma “linguagem especial” engendram novos significados, representam aspectos semântico-conceptuais característicos de um grupo ou comunidade linguística, “incorporam, sustentam e caracterizam uma identidade cultural”.

Assim sendo, é possível encontrar elementos de natureza especializada provenientes de outros universos na composição sintética da obra literária. Na verdade, a autonomia da criação literária dá ao autor liberdade para expressar-se de diferentes formas, por meio de vários recursos linguísticos, podendo até adotar um modo de dizer próximo de uma linguagem especializada, por exemplo.

Em outros termos, para verbalizar um mundo ficcional, o autor faz uso de uma linguagem especial do universo da literatura, uma linguagem com função pragmática e social que legitimiza o saber-fazer de uma comunidade linguística. Ela é especial não por propriedades que lhes são intrínsecas, mas pelo uso que se faz dela em um dado universo de discurso, ou seja, não se trata de uma linguagem com características linguísticas exclusivas e distantes de outras variantes linguísticas, mas de uma linguagem que se constitui a partir de outras (língua geral; linguagem especializada, por exemplo) para expressar mundos literários diversos e diferentes recortes e identidades culturais. E é exatamente essa linguagem com propriedades funcionais que percebemos no discurso etnoliterário, discurso este constante no *corpus* de estudo do nosso trabalho e que agora convém discutir.



## 2.2 O discurso etnoliterário

Examinando o universo lexical dos discursos etnoliterários, Barbosa (2005) verificou que as unidades lexicais contidas nos textos possuíam um estatuto diferente, ou seja, elas possuíam qualidades das linguagens de especialidade e qualidades da linguagem literária, adquirindo, desse modo, o estatuto de vocábulo-termo. Isto fez com que a autora criasse uma nova disciplina científica, a Etnoterminologia, para designar “o estudo das unidades multifuncionais dos discursos etnoliterários” (BARBOSA, 2005, p. 106). Trataremos desta disciplina no próximo capítulo para focarmos especificamente no discurso etnoliterário.

Segundo Pais e Barbosa (2004, p. 84), os discursos etnoliterários “apresentam um mundo sermioticamente construído”, mostram uma visão de mundo construída a partir de recortes culturais que “integram o *imaginário coletivo* de uma comunidade humana”, “incorporam, sustentam, caracterizam uma *identidade cultural*”, de modo que são considerados “*documentos* altamente significativos, reveladores de uma cultura e do seu processo histórico”.

São considerados discursos etnoliterários: “literatura oral, literatura popular, literatura de cordel, fábulas, lendas, mitos, folclore e também os discursos das linguagens especiais com baixo grau de tecnicidade e de cientificidade” (BARBOSA, 2009a, p. 1). A esses, damos o nome de discursos etnoliterários *stricto sensu*. Porém, há também os discursos etnoliterários *lato sensu* – “aqueles discursos que mantêm relações intertextuais e interdiscursivas com os discursos etnoliterários *stricto sensu*”, isto é, “o leitor, por meio dos discursos etnoliterários *lato sensu*, tem acesso indireto aos discursos etnoliterários *stricto sensu* que dialogam com o texto literário” (CARNEIRO, 2016, p. 103).

Esse esclarecimento permite-nos incluir manifestações literárias, como as obras regionalistas sertanistas no escopo dos discursos etnoliterários *lato sensu*, visto que os seus discursos-ocorrência utilizam, com frequência, elementos provenientes de discursos etnoliterários *stricto sensu* (como elementos da literatura oral, do folclore, das lendas e mitos da cultura sertaneja), estabelecendo relações intertextuais e interdiscursivas. Com efeito, assim como os discursos etnoliterários *strictu sensu*, os discursos etnoliterários *lato sensu* são também considerados objetos da Etnoterminologia.

Conforme Barbosa (2009b), os discursos etnoliterários sustentam-se numa tensão dialética entre dois termos: documentais x ficcionais, visto que reúnem características tanto dos discursos literários como dos discursos sociais não literários. De certo ângulo, os

discursos etnoliterários podem ser considerados ficcionais “na medida em que os ‘eventos’ narrados são ou parecem ser inverossímeis, se tomados denotativa-mente, e não correspondem a fatos historicamente comprovados” (BARBOSA, 2009b, p. 41). De outro ângulo, esses discursos, “revelam e sustentam sistemas de valores, sistemas de crenças, um ‘saber compartilhado sobre o mundo’, que integram um imaginário coletivo de uma cultura, de uma sociedade”. Nesse sentido, além de ficcionais são também documentais, pois “contribuem para o sentimento de sua permanência no eixo da História e para a configuração de uma identidade cultural”. Nas palavras da autora, “parece legítimo vê-los como documentos do pensamento e dos valores coletivos, imprescindíveis, portanto, para a compreensão do processo histórico da cultura” (BARBOSA, 2009b, p. 41).

Isso nos permite inferir que os sememas que constituem as unidades lexicais do discurso etnoliterário são derivados do mundo real e reconstruídos em um mundo ficcional possível. Assim, “mundo real e mundo ficcional constituem, sincreticamente, a base de referência semântica para a lexemização<sup>3</sup> de traços semântico-conceptuais em unidades lexicais ficcionais, que integram os discursos literários ficcionais” (CARNEIRO, 2016, p. 56). Tais unidades portam verdades oriundas do conhecimento acumulado sobre o mundo, sobre a natureza humana.

Porém, não se trata de uma imitação da realidade (*mimesis*), mas de uma construção do próprio universo, do próprio contexto referencial que engendram significados específicos de um mundo literário particular circunscrito em universo de discurso etnoliterário. Isto nos remete à concepção de linguagem literária dada por Steger (1987). Uma linguagem com função pragmática e social que modula e constrói seu próprio universo, gerando especificidades semântico-conceptuais no interior de um discurso, de um texto literário.

Como vimos, as unidades linguísticas que integram a cadeia figurativa de um mundo ficcional semiotizado apresentam características linguísticas diversas a depender das normas discursivas, do universo de discurso em que se inserem e da intenção estético-linguística do autor. Isso nos leva a inferir que é possível encontrar elementos de natureza especializada provenientes de outros universos na composição sintética da obra literária. No caso dos discursos etnoliterários, as unidades lexicais não só apresentam dupla funcionalidade: a de vocábulo e termo simultaneamente, como também revelam elementos próprios de uma cultura presentes na memória coletiva de um grupo, de uma comunidade linguística. Nesse sentido, é pertinente considerar que:

---

<sup>3</sup> “configuração do conceito em grandeza-signo, no próprio ato de instaurar a significação” (BARBOSA, 2004, p. 57).

[...] a semiose literária é caracterizada pelo caráter relacional do signo que permite que relações diversas sejam estabelecidas no interior da linguagem, propiciando a criação de particularidades ficcionais [...]. Percebe-se que termos ficcionais devem ser analisados a partir da esfera cultural; é no ambiente cultural humano, nas produções culturais humanas que essas unidades lexicais adquirem sentido. Elas constituem um registro linguístico, um documento cultural relativo às crenças e ao imaginário coletivo de uma cultura (CARNEIRO, 2016, p. 58).

Todas essas considerações nos conduzem a observar que os discursos etnoliterários exercem importantes funções culturais e sociais, pois além de revelar elementos de uma dada cultura, “ensinam ao sujeito-enunciário individual e/ou coletivo elementos cruciais da natureza humana, da alma, dos impulsos, da afetividade, em suma, da *psyché* humana”. Tais discursos “desempenham, com certeza, um papel na socialização dos membros da comunidade e, além disso, uma função estética, uma função didática, uma função mítica” (BARBOSA, 2009b, p. 40, 41).

Para Barbosa (2009b), os discursos etnoliterários assemelham-se, em muitos aspectos, ao *mýthos* da cultura grega antiga, pois “da mesma forma que os mitos gregos, os mitos indígenas da América do Norte, esses discursos e seus textos oferecem subsídios importantes para os estudos antropológicos e para as reflexões psicanalíticas” na medida em que ensinam ao sujeito-enunciário elementos cruciais da natureza humana (BARBOSA, 2009b, p. 40). Daí resultam as funções “didática” e “mítica”, além da “estética”

De fato, é um discurso com traços distintos e particulares, que é preciso aprender, para interpretar corretamente os seus textos, ou seja, “é preciso estar familiarizado com as histórias, conhecer o pensamento e o sistema de valores da cultura em questão, para poder compreendê-los bem” (BARBOSA, 2005, p. 105). Por vezes, essa não é uma tarefa fácil, segundo Barbosa (2009b), porque ao assimilar o universo semiótico-linguístico de um discurso etnoliterário, o sujeito-enunciário precisará assimilar e construir o saber e o saber-fazer específicos da cultura de uma dada comunidade linguística, para poder entender, “rediscursar” e realimentar não só os modelos figuratizados, como também a sua própria “visão do mundo” anterior, num processo de amadurecimento intelectual e pessoal.

Nesse sentido, é imprescindível interpretar, com toda a clareza, de um lado, os traços semântico-conceituais das unidades lexicais que integram um discurso etnoliterário e, de outro lado, as relações de significação que estabelecem no interior de um texto, sempre lembrando que cada vocábulo-termo constitui um registro linguístico, resultado das produções

culturais humanas, que atrelados a uma identidade cultural corroboram para a construção de sentidos de um mundo literário verbalizado.

De acordo com Pais e Barbosa (2004), o estudo do modo de existência e de produção dos discursos etnoliterários, o exame das normas frásticas e transfrásticas que os caracterizam indicam que o universo de discurso etnoliterário se submete a uma tipologia discursiva única, uma “classe discursiva específica”, podendo ser considerados como um “universo de discurso”. Compreendendo o discurso como o “*processo discursivo de produção* - que compreende uma enunciação de codificação e uma enunciação de decodificação, e o *texto*, enquanto *produto, enunciado*”, entende-se por universo de discurso:

[...] um conjunto não finito, ou que tendem *ad infinitum*, de todos os discursos manifestados que apresentam certas características comuns e constantes, assim como certas coerções suscetíveis de configurar uma norma [...] que compreende, por sua vez, uma série de normas frásticas, lexicais, semântico-sintáticas, às vezes fonético-fonológicas, e outras tantas normas transfrásticas, narrativas, discursivas, dependentes da argumentação, da veridicção, da verossimilhança, da eficácia, ou aquelas concernentes aos mecanismos da persuasão/interpretação, da manipulação/contramanipulação, da sedução, às formulações específicas das relações intersubjetivas, espaciais e temporais de enunciação e enunciado e, ainda, às que dizem respeito às modalidades, às modalizações discursivas dominantes, e, enfim, aos processos de produção e reiteração da ideologia, próprios de um determinado universo de discurso (PAIS; BARBOSA, 2004, p. 85).

Nesse fragmento, Pais e Barbosa (2004) destacam a norma que se refere ao estatuto semântico, sintático e funcional do conjunto de unidades lexicais que caracterizam os universos de discurso etnoliterários. Trata-se, portanto, de um discurso regido por uma norma<sup>4</sup> particular e exclusiva, própria do universo etnoliterário, que estabelece com outros textos e discursos relações intertextuais e interdiscursivas.

Em outras palavras, essas características do modo de existência e de produção dos discursos etnoliterários, assim como sua função mítica e pedagógica, refletem-se nas estruturas lexicais de modo que “as unidades lexicais atualizadas nos textos mantêm uma rede de relações semânticas específicas - intra-universo - e têm funções peculiares, quanto à designação e à referência, e por isso são *multifuncionais*” (PAIS; BARBOSA, 2004, p. 98).

Essa multifuncionalidade das unidades lexicais – vocábulos/termos – nos conduz a refletir sobre a semiótica do discurso etnoliterário (mais especificamente a semiótica dos

---

<sup>4</sup> Norma esta que compreende outras normas frásticas e transfrásticas provindas de outros universos de discurso (PAIS; BARBOSA, 2004).

discursos etnoliterários *lato sensu* em que as obras regionalistas sertanistas estão inseridas), bem como sobre a importância da Etnoterminologia no estudo dessas unidades. Além de apresentarem características exclusivas reunindo qualidades da linguagem literária e das linguagens de especialidade, essas unidades são reveladoras de vários aspectos da cultura sertaneja, da geografia física e humana do sertão brasileiro.

Vale lembrar que língua e cultura se inter-relacionam e mantêm um vínculo de interdependência, ou seja, é por meio da língua que os traços culturais de um povo se evidenciam, é através dela que os grupos sociais representam o mundo e expressam suas ideias e experiências (LYONS, 2009). Desse modo, a linguagem funciona como “um instrumento social e cultural, uma vez que viabiliza a interação do homem e a sua atuação sobre a realidade, possibilitando a vida em comunidade” (PIMENTA, 2013, p. 23, 24).

Enquanto alguns discursos literários pretendem apoiar-se eminentemente na ficcionalidade, na imaginação, na fantasia, distanciando-se do mundo real, os discursos etnoliterários sustentam-se a partir de uma linguagem que incorpora aspectos do mundo real, “assumindo o importante papel de *sustentar, conservar, atualizar, transmitir* aspectos relevantes de um saber compartilhado sobre o mundo”, de um saber-fazer específico de uma cultura, de um grupo, de uma comunidade linguística. (PAIS, BARBOSA, 2004, p. 98).

Por tudo isso entendemos que os discursos etnoliterários, apesar de também construírem um mundo ficcional, assim como os discursos literários de fantasia, aproximam-se mais do mundo real do que estes. Isto se deve ao fato de estabelecerem um recorte semântico-conceitual de maneira a caracterizar uma identidade cultural presente no mundo real. Prova disso, encontramos nas obras regionalistas sertanistas elementos linguísticos e culturais do sertanejo que evidenciam realidades dos sertões do Brasil (aspectos geográficos, históricos, políticos e socioculturais), permitindo-nos conhecer o homem do sertão, suas características físicas, comportamentais e, especialmente, linguísticas. Sobre esses aspectos que iremos tratar a seguir.

### **2.3 O universo literário, temático e linguístico do *corpus* de estudo**

Não é novidade o fato de que o termo regionalismo possui ressonâncias conceituais exponenciais dentro do amplo painel de estudos da literatura. Trata-se de uma tendência, um fenômeno histórico-cultural concreto que se manifestou em vários momentos da história do

sistema literário brasileiro de formas diversas, razão pela qual possui diferentes concepções, graus ou fases que compreende desde o Romantismo ao Modernismo.

Diante da multiplicidade de definições propostas pela crítica literária procuramos nos ater, especialmente, às proposições de Antonio Candido, crítico literário e professor, um dos grandes expoentes da crítica literária brasileira. Para ele, o regionalismo é uma tendência literária que busca descrever e representar as peculiaridades regionais, os costumes, as tradições e falares locais (dialetos), “principalmente, do mundo sertanejo, documentando e retratando os tipos humanos, paisagens e costumes considerados tipicamente brasileiros.” (CANDIDO, 2002, p. 87). É nessa noção de regionalismo que apoiamos e sustentamos as discussões deste trabalho.

Segundo Candido (2000), a literatura regionalista germinou no Romantismo em um dramático cenário de busca de uma identidade nacional, em meados do século XIX. Aguçados por um sentimento nacionalista de representação da matéria local, da verdadeira forma física e humana da nação recém-fundada, os românticos foram pouco a pouco construindo uma tradição literária própria, embora ainda apoiada em modelos estrangeiros. Nas palavras do crítico:

No Brasil o romance romântico, nas suas produções mais características (em Macedo, Alencar, Bernardo Guimarães, Franklin Távora, Taunay), elaborou a realidade graças ao ponto de vista, à posição intelectual e afetiva que norteou todo o nosso Romantismo, a saber, o nacionalismo literário. Nacionalismo, na literatura brasileira, consistiu basicamente, como vimos, em escrever sobre coisas locais; no romance, a consequência imediata e salutar foi a descrição de lugares, cenas, fatos, costumes do Brasil. É o vínculo que une as *Memórias de um Sargento de Milícias* ao *Guarani* e a *Inocência*, e significa, por vezes, menos o impulso espontâneo de descrever a nossa realidade, do que a intenção programática, a resolução patriótica de fazê-lo. Esta tendência naturalizou a literatura portuguesa no Brasil, dando-lhe um lastro ponderável de coisas brasileiras. E como além de recurso estético foi um projeto nacionalista, fez do romance verdadeira forma de pesquisa e descoberta do país [...] (CANDIDO, 2000, p. 99).

Inserido em um projeto nacionalista de criar uma expressão nova de um país novo, o ideal romântico concentrava-se, principalmente, na supervalorização das coisas locais, no interesse por costumes e regiões, no gosto pela expressão local, pela natureza, pelo pitoresco e exótico, pela descrição dos tipos humanos, resultando em obras idealistas e sentimentais de considerável valor estético-literário. A princípio, a primeira forma de expressão desse nacionalismo romântico foi o indianismo, fundado na supervalorização do índio como símbolo da identidade nacional. A título de exemplificação, podemos citar as obras de José de Alencar: *O Guarani*, *Iracema* e *Ubirajara*. Na medida em que o índio, enquanto potencial de

expressão mítico-heroica e de representação autóctone começa a se enfraquecer, surge outro tipo humano: o sertanejo, o homem do campo, das regiões interioranas do país, tomando lugar a temática sertanista.

Ora, em um país constituído por uma população mestiça, com pouquíssimas povoações e, ainda, predominantemente rurais, nada mais natural do que recorrer à figura do índio e do sertanejo como símbolos autênticos de uma identidade nacional, como observa Candido (2000):

Quanto à matéria, o romance brasileiro nasceu regionalista e de costumes; ou melhor, pendeu desde cedo para a descrição dos tipos humanos e formas de vida social nas cidades e nos campos. O romance histórico se enquadrou aqui nesta mesma orientação; o romance indianista constitui desenvolvimento à parte do ponto de vista da evolução do gênero, e corresponde não só à imitação de Chateaubriand e Cooper, como a certas necessidades [...], poéticas e históricas, de estabelecer um passado heróico e lendário para a nossa civilização, a que os românticos desejavam, numa utopia retrospectiva, dar tanto quanto possível traços autóctones (CANDIDO, 2000, p. 101).

Nessa mesma linha de raciocínio, Coutinho (2001) argumenta:

Cria-se, inclusive, um tipo de herói – o herói regional – de estatura quase épica em seus aspectos de super-homem, em luta contra um destino fatal, traçado pelas forças superiores do ambiente. [...] O Regionalismo foi uma das formas que assumiu o nacionalismo literário brasileiro, a partir do Romantismo. É uma das respostas à pergunta do séc. XIX sobre como deveria ser a literatura para ter caráter e identidade nacionais, isto é, para ser brasileira (COUTINHO, 2001, p. 237).

Com isso, podemos inferir que o surgimento da literatura regionalista no Brasil está, portanto, ligado ao nacionalismo literário que, por sua vez, é fruto do romance romântico de cunho sertanista que se baseia na supervalorização dos aspectos regionais, voltando-se para as descrições físicas e humanas do mundo sertanejo: a paisagem, os costumes, as tradições, a linguagem, o perfil físico e psicológico do sertanejo de diversas regiões do Brasil. Segundo Almeida (1981):

O sertanejo tinha a seu favor vários elementos que o recomendavam para a função. Via de regra é um mestiço do branco com o índio (não com o negro, raro nas áreas mais pobres do sertão) [...] Metaforicamente poder-se-ia afirmar que o sertanejo é o descendente direto de Peri e Ceci, de Martim e Iracema. Vivendo em regiões isoladas, sem grande contato como os centros litorâneos tem a evolução cultural relativamente autônoma [...] Na obra *O sertanejo*, mais do que em qualquer outra, do próprio romancista ou de terceiros, transparece o desejo de substituir o mito indianista, então em acentuado processo de desgaste, pelo mito sertanista na busca de arquétipos

com que se pudessem identificar as aspirações nacionalistas tão atuantes no Romantismo brasileiro (ALMEIDA, 1981, p. 35; 49).

Cercado por um forte desejo de compensação e representação da identidade nacional do país, a figura do sertanejo tornou-se um elemento central de expressão do espírito da época. Em pleno século XIX, onde havia apenas alguns pequenos povoados em regiões isoladas, afastadas do litoral, o sertão era algo real, notável em extensão na imensidão da terra brasileira. Não por acaso é ao sertanejo e à sua terra que são dedicadas as páginas mais intensas das variadas histórias da ficção romanesca regionalista.

Importa mencionar que o sertão, de início, era considerado como uma área vasta e indefinida do interior do Brasil, onde habitava o homem bom, o sertanejo das terras longínquas e desabitadas. Para Galvão (1976):

O que se chama sertão, no Brasil, é toda a região interiorana do país, abrangendo mais da metade de seu território. Sua determinação é mais histórico-econômica que geográfico-política. Num país criado por determinação do mercado externo, e criado para suprir esse mercado de bens de consumo que lhe interessavam [...] as terras boas eram naturalmente reservadas para a produção desses bens [...]. Mas era preciso, sem desperdiçar as terras boas, prover o sustento das pessoas necessárias a essa produção, de um lado; e de outro, entra o fator histórico-político de tratar de assegurar a maior extensão de terra possível para a coroa portuguesa. Foi assim que, aos poucos, os territórios interiores – mais distantes do mar e também menos rentáveis para a produção agro-industrial – foram sendo ocupados pela criação de gado. [...] Desse modo, no Brasil, desde o primeiro século da colonização, vai sendo chamado de *sertão* o interior e a palavra carrega consigo os significados de *interior*, *indesbravado*, *selvagem*, *desconhecido*, *não-urbano* (GALVÃO, 1976, p. 36).

Assim, coube à literatura regionalista retratar a realidade da sociedade brasileira e isto não só contribuiu para a construção de uma literatura própria com elementos tipicamente brasileiros através do Romantismo, como a transformou no principal veículo de informação e expressão da realidade do país através do sertanismo. Em outras palavras, além de um recurso estético, o regionalismo literário foi um projeto nacionalista de construção e conhecimento da recém-nação que contribuiu, sobremaneira, para a formação e desenvolvimento da cultura brasileira.

Esse regionalismo de vertente romântica-sertanista, marcado pela supervalorização dos aspectos regionais, pela exaltação do homem do interior, pela descrição e representação do mundo sertanejo, é denominado por Candido (2000) de *Regionalismo pitoresco*. De acordo com o crítico, os autores mais notáveis dessa fase são: José de Alencar (*O Sertanejo*), Bernardo Guimarães (*O Garimpeiro*), Visconde de Taunay (*Inocência*) e Franklin Távora (*O*



*cabeleira*). As obras em destaque foram selecionadas para compor o *corpus* de estudo primeiro, por focalizarem o mundo sertanejo, segundo, por serem as mais representativas dessa fase e, também, por apresentarem um material linguístico suficiente e adequado para esta pesquisa.

Nessa primeira fase, “o sertão era visto somente no seu aspecto róseo, o sertão bom e saudável povoado de criaturas boas, sadias e vigorosas, de almas puras” (COUTINHO, 2001, p.237). Sob um ângulo pitoresco, sentimental e idealista o sertanejo era retratado nos romances como um herói, forte, corajoso e generoso, habitante de uma terra abençoada e de uma natureza exuberante, como se pode notar em *O Garimpeiro* (1872) de Bernardo Guimarães:

As regiões que formam os municípios de Araxá, Patrocínio e Bagagem, na província de Minas, encerram paisagens as mais risonhas e encantadoras que se podem imaginar, e quem uma vez tem percorrido esses férteis e pitorescos sertões nunca mais os perde da lembrança. [...] Tudo é belo e grandioso, é risonho e enlevador por aquelas imensas solidões. Inúmeras manadas de gado e de éguas, mugindo e relinchando pelos vargedos de viço perenal, bandos de emas e siriemas vagando pelos camegais, alegram a solidão daqueles sertões abençoados.[...] Os habitantes dessas regiões são notáveis pela amenidade dos costumes e pela amabilidade do trato. Nessas paragens os homens são robustos, ativos e inteligentes, as moças são bem feitas, meigas e formosas (GUIMARÃES, 1997, p. 1).

Essa valorização do sertão e da tipologia humana do sertanejo é também percebida em *O Sertanejo* (1875) de José de Alencar e *Inocência* (1872) de Visconde de Taunay. Em Arnaldo, protagonista de *O Sertanejo* e Pereira (pai de Inocência) podemos ver a criatividade dos escritores em construir personagens realistas em um cenário onde tudo parece ser verossímil, com numerosas descrições da paisagem, da fauna e flora dos sertões do Nordeste (em *O Sertanejo*) e do Mato Grosso do Sul (em *Inocência*):

Esta imensa campina, que se dilata por horizontes infínidos, é o sertão de minha terra natal. Aí campeia o destemido vaqueiro cearense, que à unha de cavalo acossa o touro indômito no cerrado mais espesso, e o derriba pela cauda com admirável destreza. Aí, ao morrer do dia, reboa entre os mugidos das reses, a voz saudosa e plangente do rapaz que abóia o gado para o recolher aos currais no tempo da ferra (ALENCAR, 1995, p. 1).

Em *Inocência*:

Corta extensa e quase despovoada zona da parte sul-oriental da vastíssima Província de Mato Grosso a estrada que da Vila de Sant'Ana do Paranaíba vai ter ao sítio abandonado de Camapuã. Desde aquela povoação, assente próximo ao vértice do ângulo em que confinam os territórios de São Paulo, Minas Gerais, Goiás e Mato Grosso até ao Rio Sucuriú, afluente do

majestoso Paraná, isto é, no desenvolvimento de muitas dezenas de léguas, anda-se comodamente, de habitação em habitação, mais ou menos chegadas umas às outras, rareiam, porém, depois as casas, mais e mais, caminham-se largas horas, dias inteiros sem se ver morada nem gente até ao retiro de João Pereira, guarda avançada daquelas solidões, homem chão e hospitaleiro, que acolhe com carinho o viajante desses alongados páramos, oferece-lhe momentâneo agasalho e o provê da matalotagem precisa para alcançar os campos de Miranda e Pequiri, ou da Vacaria e Nioac, no Baixo Paraguai. Ali começa o sertão chamado bruto (TAUNAY, 1998, p. 1).

Embora a linguagem culta seja também empregada nessas narrativas, percebemos que o linguajar sertanejo é predominante tanto nos discursos diretos como nos discursos indiretos. Os diálogos são vivos, naturais, reproduzem o falar do sertanejo dessas regiões:

- Que bichão! murmura ele contemplando um rasto mais fortemente impresso no solo; com um bom onceiro não se me dava de acuar este diabo e meter-lhe uma chumbada no focinho. [...] - Neste rumo daqui a 20 léguas, fica o espigão mestre de uma serra *braba*, depois um rio grosso; dali a cinco léguas outro mato sujo que vai findar num brejal. Se *vassuncê* frechar direitinho assim umas duas horas, topa com o pouso do Tatu, no caminho que vai a Cuiabá (TAUNAY, 1998, p. 4).

Em *O Sertanejo*:

[...] Mas o cavalo estacou espavorido, com o pêlo hispido e as narinas insufladas pelo terror [...]. Seu cavalo cardão rompeu o mato a galope, como quem estava acostumado a campear o barbatão no mais espesso bamburral; e com pouco o sertanejo, atalhando a distância, avistou D. Flor parada além, no caminho. [...] — Lá o grito da doninha, não digo nada, mas o rincho do cavalo chega longe; e então quando o fogo começasse a chamuscar-lhe a pele! (ALENCAR, 1995, p. 5 - 9).

Em *O Garimpeiro*, a presença da linguagem sertanista é tão marcante que podemos comparar Bernardo Guimarães a um exímio contador de causos, como declara Candido (2000):

[...] os romances deste juiz, Bernardo Joaquim da Silva Guimarães, parecem boa prosa da roça, cadenciada pelo fumo de rolo que vai caindo no côncavo da mão ou pela marcha das bestas de viagem, sem outro ritmo além do que lhes imprime a disposição de narrar sadiamente, com simplicidade, o fruto de uma pitoresca experiência humana e artística. [...] e quase todos os outros livros não deixam de apresentar essa tonalidade de conversa de rancho. Conversa de bacharel bastante letrado para florear as descrições e suspender a curiosidade do ouvinte, mas bastante matuto para exprimir fielmente a inspiração do gênio dos lugares (CANDIDO, 2000, p. 212).

Segundo o crítico, valendo-se dos recursos próprios da narrativa oral, Bernardo enreda a história do garimpeiro Elias em um tom romântico-regionalista surpreendente. Com base na

realidade mineira da época, Bernardo retrata o cotidiano dos sertanejos com maestria, apresentando com uma riqueza de detalhes, a paisagem do sertão, os hábitos, os costumes e tradições da região (como, por exemplo, os garimpos e as festas da Vila do Patrocínio, as chamadas cavalhadas). A título de exemplificação, citamos o seguinte trecho:

Cavalhadas! Ainda mais isso! Que bom! E eu que sou doida por cavalhadas! Não pode haver brinquedo mais bonito. Há quanto tempo não há disso por aqui! Esta terra já não é o que era dantes. No meu tempo, ah! Sinhazinha! Se Você visse! Que bonitas cavalhadas não se corriam aqui e no Araxá! Era um gosto! Hoje isto já não presta para nada. Que é dos corredores de fama que então havia? Já morreu tudo. Agora isso há de ser alguma coisa à-toa. - Estás enganada, Joana, estas vão ser muito boas. Aquele moço que aqui passou outro dia, não te lembrás? Aquele moço alto, de cabelo preto e anelado. . . - Ah! Já sei. . . O Sr Elias, aquele moço de Uberaba. . . - Isso mesmo, Joana; ele também vai correr, e pediu a meu pai o cavalo rosilho. - Oh! Aquele sim, que bonito cavaleiro não há de ser! É um mocetão sacudido e muito bem parecido (GUIMARÃES, 1997, p. 2).

O êxito de Bernardo em construir narrativas com elementos tão verossímeis parece residir, diante da reflexão de Candido (2000), no amor que tinha por Minas Gerais, seu estado natal, e Goiás, onde exerceu o cargo de juiz municipal da cidade de Catalão e viveu por alguns anos (de 1832 até 1854). Segundo ele, “essa disposição de narrar sadiamente, com simplicidade, é fruto de uma pitoresca experiência humana e artística” (CANDIDO, 2000, p. 212).

De modo um pouco diferente, Franklin Távora enreda *O Cabeleira* (1876), romance que abriu caminho a uma nova linhagem regionalista, “um regionalismo de cunho social, voltado para a interpretação social de uma determinada zona”, um regionalismo que “culminou no chamado Regionalismo Nordestino ratificado pela geração de 1930” (CANDIDO, 2000, p. 268). De acordo com Candido (2000), Távora foi um dos primeiros “romancistas do Nordeste” a se preocupar com os problemas sociais da região e fazer da literatura um instrumento para exprimir a realidade local: a vida difícil do sertanejo e a sua luta pela sobrevivência, o cangaço e suas consequências, a violência naquelas regiões esparsas e isoladas, o atraso da região em relação às demais.

Embora ainda revestida por um tom sentimental e idealista em relação à figura do sertanejo, como prova os seguintes fragmentos: “A história de Pernambuco oferece-nos exemplos de heroísmo e grandeza moral que podem figurar nos fastos dos maiores povos da antigüidade sem desdourá-los”/ “A sua audácia e atrocidades deve seu renome este herói legendário para o qual não achamos par nas crônicas provinciais” (TÁVORA, 1973, p. 2),

nessa narrativa o autor descreve o sertão não como um “lugar bom para se viver” (característica marcante nos romances anteriores), mas como um lugar inóspito, perigoso, cujo atraso e carência vitimam transformar os bons em pessoas rudes e cruéis, como é o caso de o Cabeleira, um homem naturalmente bom e corajoso que é corrompido pelo pai (velho cangaceiro, “sujeito de más entranhas, dado à prática dos mais hediondos crimes” (TÁVORA, 1973, p.1) e pelo meio em que vive, tornando-se um assassino cruel. Os trechos a seguir revelam tais características:

Entra neste número o protagonista da presente narrativa, o qual se celebrou na carreira do crime, menos por maldade natural, do que pela crassa ignorância que em seu tempo agrilhoava os bons instintos e deixava soltas as paixões canibais. [...] As populações do interior, em sua maioria destituídas de bens da fortuna, e então muito mais espalhadas do que atualmente, pouco tinham já com que cevar a voracidade dos três aventureiros a quem desde muito pagavam um triplo imposto consistente em víveres, dinheiro e sangue. O assalto foi resolvido em secreto conciliábulo dentro das matas de Pau d'Alho onde mais uma vez se haviam reunido para concertos idênticos. [...] A notícia da sua aproximação a maior parte dos moradores, deixando os povoados, então muito fracos por não terem ainda a densidade que só um século depois tornou alguns deles respeitáveis, emigrou para os matos, único abrigo com que lhes era permitido contar, embora se achassem a poucas léguas do Recife. [...] Contavam-se então as casas por aquelas paragens. Em torno delas o deserto começava a aumentar antes de pôs-se o sol. Uma lei cruel, a lei da necessidade, obrigava os moradores a trancar-se cedo por bem da própria conservação. Os roubos e assassinatos reproduziam-se com incrível frequência nos caminhos e até nas beiradas dos sítios (TÁVORA, 1973, p. 1- 4).

Segundo Candido (2000, p. 271), “para Távora [...], a região não era apenas motivo de contemplação, orgulho ou enlevo: mas um complexo de problemas sociais, sobrelevando (não custa repisar) a perda de hegemonia político-econômica”. Sua intenção era desvelar aquilo que até então outros escritores românticos não tiveram a audácia de mostrar: o lado rude e difícil naqueles sertões abandonados e carentes. Para o crítico, Távora conhecia bem a região que retratava, por isso a descrevia com fidelidade e respeito, mas sem praticar o pitoresco e o exotismo. Nas palavras de Candido (2000):

A virtude maior de Távora foi sentir a importância literária de um levantamento regional; sentir como a ficção é beneficiada pelo contacto de uma realidade concretamente demarcada no espaço e no tempo, que serviria de limite e em certos casos, no Romantismo, de corretivo à fantasia. Ora, para ele este contacto se funda na experiência direta da paisagem, que o romancista deve conhecer e descrever precisamente (CANDIDO, 2000, p. 269).

De modo sucinto, podemos afirmar, com base em Candido (2000), que o regionalismo de Távora funda-se em três elementos, “que ainda hoje constituem, em proporções variáveis, a principal argamassa do regionalismo literário do Nordeste”: primeiro, “o senso da terra, da paisagem que condiciona tão estreitamente a vida de toda a região, marcando o ritmo da sua história pela famosa “intercadência” de Euclides da Cunha”; segundo, “o que se poderia chamar patriotismo regional, orgulhoso das guerras holandesas, do velho patriarcado açucareiro, das rebeliões nativistas; por fim, “a disposição polêmica de reivindicar a preeminência do Norte, reputado mais brasileiro, onde abundam os elementos para a formação de uma literatura propriamente brasileira, filha da terra”. (CANDIDO, 2000, p. 268).

Esse senso da terra, esse patriotismo regional é também percebível em *Pelo Sertão* (1898) de Afonso Arinos, um livro de contos que reúne histórias do interior do estado de Minas Gerais e Goiás, mostrando a vida rural do sertanejo, os mitos e superstições, os costumes e tradições, a beleza selvagem da paisagem e, principalmente, a variedade linguística utilizada pelos sertanejos dessas regiões. A fidelidade e verossimilhança são notáveis na narrativa não apenas porque o autor conhecia muito bem essas regiões, mas porque nutria uma paixão por sua cidade Natal (Paracatu), como bem observa Silva:

O sertão de Paracatu tocou a sensibilidade artística de Arinos, contribuindo com elementos da vida rural e iletrada para sua obra ficcional. Os temas do sertão, a caracterização dos personagens e tipos populares, a descrição do espaço em contos como “Assombramento”, “Joaquim Mironga” e “Pedro barqueiro”, provêm da observação e da vivência no meio da região rural de sua cidade natal. Arinos percebeu, também, as credences populares e o comportamento religioso do homem do sertão (SILVA, 2008, p. 18).

A valorização do sertão e da tipologia humana do sertanejo, o culto à natureza e aos aspectos sentimentais, a evocação dos aspectos regionais, como as tradições, mitos e superstições (características dessas típicas dos romances sertanistas), são também percebíveis em alguns contos, como em: “Paisagem alpestre”, “Buriti perdido”, “Desamparados” e “A árvore do pranto” (exaltações alusivas à natureza do sertão, sua fauna e flora); em “Manuel Lúcio”, “A fuga”, “O contratador dos diamantes”, “A velhinha”, “A cadeirinha”, “O Mão pelada”, “Feiticeira” e “A rola encantada”, há uma evocação do passado histórico, das lendas e tradições; e em “Assombramento”, “Joaquim Mironga”, “Pedro barqueiro”, e “A garupa”) há uma valorização da tipologia humana do sertanejo, heroicizado por sua coragem, lealdade e bravura.

Já nos contos “A esteireira” e “O contratador dos diamantes”, diferentemente dos outros contos, a representação do sertanejo é mais objetiva e fiel, sem intenções idealizadoras ou sentimentais. Em “A esteireira” Ana, a protagonista, é retratada como uma mulher fria e cruel que, depois de matar Candinha, a suposta rival, comete inesperado gesto de vampirismo: “Ana [...] encostou os lábios ao lugar de onde irrompia aos cachões, e, carnívora esfaimada, chupou, chupou por muito tempo, carregando, depois, o corpo da desventurada para bem longe [...]” (ARINOS, 1947, p. 97). Em “O contratador dos diamantes”, Afonso Arinos nos mostra o lado hostil do ser humano, a ambição, o interesse, a inveja, os prazeres da carne. O autor constrói personagens fidedignos à época a partir de uma perspectiva realista-naturalista, cujos comportamentos são impulsionados por seus instintos e pelo meio em que vivem. A exploração das minas de diamantes no interior mineiro não só teve lugar na literatura de Afonso Arinos, como fez parte de um projeto estético-literário na busca por um nacionalismo renovado, mais preciso e objetivo, despido da roupagem idealista romântica.

Além desse contraste na representação regional, a linguagem é também um elemento bem marcante nessa obra. Em uma parte dos contos, a linguagem culta, por vezes poética, é predominante: “Velha palmeira solitária, testemunha sobrevivente do drama da conquista, que de majestade e de tristura não exprimes, venerável epônimo dos campos” (Buriti perdido, p. 67)/ “Ninguém pôde, ninguém que tenha alma sensível aos espetáculos da natureza ou à Poesia das eras já mortas, poderá deixar de recolher-se, de concentrar-se em fundas cogitações ou em caroáveis devaneios” (Paisagem Alpestre, p. 105);/ “Pelas fraldas dos morros, cingindo-os, bordando os vales, em cujo fundo se espreguiçavam paus sonolentos, o buritizal erguia suas verdes frondes” (Desamparados, p. 115); em outra parte, é a linguagem sertanista que predomina: “À beira do caminho das tropas, num taboleiro grande, onde cresciam a canela d'ema e o pau santo, havia uma tapera. A velha casa assobradada, com grande escadaria de pedra levando ao alpendre” (Assombramento, p. 9)/ “— Ali tão perto ? Que tem isso ? Vamos no mouro de papai ; é um cavallinho muito bom e dá garupa. Antes do sol alto lá estalemos” (A esteireira, p. 77)/ “Trabalhado de dores, com a ruga da fronte aprofundada pelo pensar fixo — refolho onde descobria o observador talvez a magua dorida” (Manuel Lúcio, p. 91)

Diante dessas observações, fica evidente que Afonso Arinos, ao compor *Pelo Sertão*, adota duas diferentes frentes de abordagem para representar o mundo sertanejo: por um lado temos contos com características do Romantismo e, por outro, contos com marcas do

Realismo. Isto se deve ao fato de que os contos foram escritos em momentos diferentes, como o próprio autor esclarece na introdução do seu livro:

O livro que ora se apresenta ao público devia ter sido publicado há cerca de três anos. O leitor descobrirá nele a falta de unidade, quer na maneira ou na execução, quer no estilo propriamente. A razão disto é que os contos foram escritos em épocas diversas, num período que medeia entre os 19 e os 26 anos. Os primeiros datam de 1888 e 1889; os últimos de 1895 (ARINOS, 1968, p. 47).

Tudo isso nos leva a crer que *Pelo Sertão* é uma obra de grande destaque na história do regionalismo brasileiro não só porque marca a transição do Regionalismo pitoresco para o chamado Regionalismo problemático ou crítico, mas por sua avultada riqueza estético-linguística. Nas palavras de Olavo Bilac (1934):

Com que entusiasmos, com que admiração comovida, com que energia de pincel, com que colorido intenso de estilo, contaís a nobreza de alma, a coragem heróica, os amores brandos ou impetuosos, os fogosos ciúmes, a abnegação rara, a paciente resignação, e também as grandes cóleras desses homens fortes e simples, que vivem para amar a vida e o trabalho, a natureza e a liberdade de Deus... e da faca que trazem à cinta! Estas poucas novelas, que enfeixastes em livro, são os Fastos da Alma Sertaneja... (BILAC, 1934, p. 184).

Afonso Arinos relatou o homem do sertão e a sua terra de forma tão expressiva e particular que lhe rendeu um lugar privilegiado na Academia Brasileira de Letras em 1903, alavancando definitivamente a trajetória artística do escritor. É por essa razão, ou seja, por sua riqueza linguística, e também por estar numa linha de transição entre a primeira fase do regionalismo (Regionalismo pitoresco) e a segunda (Regionalismo problemático ou crítico) que agora convém discutir, é que selecionamos essa obra para compor o nosso *corpus* de estudo.

Essa preocupação em produzir uma literatura mais consciente, mais próxima da realidade local (princiada por Távora e Arinos), intensificou-se eminentemente no fim do século XIX e início do século XX com o advento do Realismo e, posteriormente do Modernismo. No Brasil, a partir da década de 80, grandes transformações foram ocorrendo nas esferas econômica, política e social, o que acabou determinando mudanças significativas no regionalismo brasileiro, como nos esclarece Tufano (1988):

A civilização burguesa, industrial e mecânica, começa a se firmar; as ideias de liberalismo e democracia ganham dimensões cada vez maiores; as ciências naturais desenvolvem-se e os métodos de experimentação e

observação da realidade passam a ser encarados como os únicos capazes de explicar racionalmente o mundo físico [...]. No aspecto social, ocorriam no Brasil grandes transformações, pois com o desenvolvimento das cidades surgiu uma significativa população urbana, marcada por desigualdades econômicas que provocaram o aparecimento de uma pequena massa proletária (TUFANO, 1988, p. 139).

Desse modo, o regionalismo foi se modificando e desvestindo-se do saudosismo e do idealismo romântico, passando a se concentrar na relação homem-região. É sobre a tensão dessa relação que os escritores agora se concentravam e procuravam desvelar: a vida miserável do sertanejo e a sua luta pela sobrevivência, a seca, a rusticidade e a precariedade da região, a desumanização do ser humano no isolamento do sertão, a exploração dos trabalhadores rurais dominados pelos poderosos coronéis, a força arbitrária dos soldados, o cangaço e a emigração dos retirantes para o litoral.

Sobrepondo-se ao otimismo romântico, esse Regionalismo, calcado na pré-consciência do subdesenvolvimento (do atraso econômico e social) é chamado por Candido (1989) de *Regionalismo problemático*, fase em que se acentua o interesse por um sertanismo mais renovado, comprometido com as questões sociais. “Trata-se, portanto, de uma literatura comprometida com o nacional, ou seja, de uma literatura que traduz a realidade de uma determinada região sob a forma de denúncia social como meio de ação para a transformação” (PIMENTA, 2011, p.58). Nas palavras de Candido (1989):

Na fase de pré-consciência do subdesenvolvimento [...] tivemos o regionalismo problemático, que se chamou de “romance social”, “indigenismo”, “romance do Nordeste” [...]. Ele nos interessa mais, por ter sido um precursor da consciência de subdesenvolvimento [...] (CANDIDO, 1989, p.159).

O marco dessa fase do Regionalismo se deu, primeiramente, com a publicação de *Os Sertões* de Euclides da Cunha, em 1902, uma obra pré-modernista que trata da guerra de Canudos, no sertão da Bahia, revelando um Brasil problemático: a situação miserável do sertanejo nordestino, abandonado por um governo que age com extrema crueldade e violência; o sofrimento e a luta dos sertanejos refugiados (liderados por Antônio Maciel, o Conselheiro); os graves desequilíbrios econômicos; as grandes diferenças (sociais, econômicas, culturais) entre os homens do litoral e os homens do sertão.

Unindo a palavra rebuscada ao rigor científico e à preocupação social, Euclides da Cunha compõe as três partes de *Os Sertões*, baseando-se nas pesquisas e reportagens feitas para o jornal paulista, “O Estado de S. Paulo”, na campanha do exército contra Canudos. A obra divide-se em: 1) “A terra” – parte em que escritor apresenta os caracteres geológicos e



topográficos das regiões que medeiam entre o Rio Grande do Norte e o Sul de Minas Gerais, de modo particular a bacia do rio São Francisco, e também, com grande abundância de pormenores, as regiões sertanejas de Monte Santo (Canudos), que abrangem os rios Vaza-Barris e Itapicurus; 2) “O homem” – nesta segunda parte, Euclides da Cunha procura apresentar, num estudo genérico, os elementos étnicos do meio brasileiro, os caracteres de sua índole e sua distribuição pelo território nacional. O autor fala das raças (índio, português, negro) e de sub-raças (que nomeia de “mestiço”), das características peculiares dos sertanejos, em especial, o do norte (o vaqueiro) e o do sul (o gaúcho), e também, da genealogia de Antônio Conselheiro; 3) “A luta”- parte mais extensa que narra as quatro expedições que se investiram contra Canudos, mostrando com uma riqueza de detalhes os combates ocorridos entre as tropas do governo e os sertanejos. Para exemplificar, citamos os seguintes fragmentos:

O sertão de Canudos é um índice sumariando a fisiografia dos sertões do Norte. Resume-os, enfeixa os seus aspectos predominantes numa escala reduzida. É-lhes de algum modo uma zona central comum. De fato, a inflexão peninsular, extremada pelo cabo de S. Roque, faz que para ele convirjam as lindes interiores de seis Estados — Sergipe, Alagoas, Pernambuco, Paraíba, Ceará e Piauí — que o tocam ou demoram distantes poucas léguas. Desse modo é natural que as vicissitudes climáticas daqueles nele se exercitem com a mesma intensidade, nomeadamente em sua manifestação mais incisiva, definida numa palavra que é o terror máximo dos rudes partícios que por ali se agitam — a seca. [...] Ao passo que a caatinga o afoga; abrevia-lhe o olhar; agride-o e estonteia-o; enlaça-o na trama espinescente e não o atrai; repulsa-o com as folhas urticantes, com o espinho, com os gravetos estalados em lanças; e desdobra-se-lhe na frente léguas e léguas, imutável no aspecto desolado: árvores sem folhas, de galhos estorcidos e secos, revoltos, entrecruzados, apontando rijamente no espaço ou estirando-se flexuosos pelo solo, lembrando um bracejar imenso, de tortura, da flora agonizante [...] O sertanejo é, antes de tudo, um forte. Não tem o raquitismo exaustivo dos mestiços neurastênicos do litoral. A sua aparência, entretanto, ao primeiro lance de vista, revela o contrário. Falta-lhe a plástica impecável, o desempenho, a estrutura corretíssima das organizações atléticas. É desgracioso, desengonçado, torto. Hércules-Quasímodo, reflete no aspecto a fealdade típica dos fracos. O andar sem firmeza, sem aprumo, quase gigante e sinuoso, aparenta a translação de membros desarticulados. Agrava-o a postura normalmente abatida, num manifestar de displicência que lhe dá um caráter de humildade deprimente [...]. O sertanejo é, antes de tudo, um forte. Não tem o raquitismo exaustivo dos mestiços neurastênicos do litoral. A sua aparência, entretanto, ao primeiro lance de vista, revela o contrário. Falta-lhe a plástica impecável, o desempenho, a estrutura corretíssima das organizações atléticas. É desgracioso, desengonçado, torto. Hércules-Quasímodo, reflete no aspecto a fealdade típica dos fracos. O andar sem firmeza, sem aprumo, quase gigante e sinuoso, aparenta a translação de membros desarticulados. Agrava-o a postura normalmente abatida, num manifestar de displicência que lhe dá um caráter de humildade deprimente. [...] Quando se tornou urgente pacificar o sertão de Canudos, o governo da

Bahia estava a braços com outras insurreições. A cidade de Lençóis fora investida poratrevida malta de facínoras, e as suas incursões alastravam-se pelas Lavras Diamantinas; o povoado de Brito Mendes caíra às mãos de outros turbulentos; e em Jequié se cometiam toda a sorte de atentados (CUNHA, 1984, p. 17- 91).

Em 1903, seguindo essa mesma linhagem de Euclides da Cunha, Domingos Olímpio publica *Luzia-Homem*, romance considerado naturalista que retrata, em um linguajar regional, a vida difícil do sertanejo e a sua luta pela sobrevivência, a seca no Nordeste, a escassez de alimentação e o abandono das propriedades que se tornam improdutivas. Marcada por características naturalistas – “do homem como um produto biológico, cujo comportamento é resultado da pressão do ambiente social e da hereditariedade psicofisiológica” (TUFANO, 1988, p. 140), a sertaneja Luzia é apresentada nessa obra como uma mulher forte e corajosa que é desumanamente embrutecida pelo meio em que vive, pelas necessidades e sofrimentos trazidos por aquele sertão rústico, improdutivo e abandonado. Constituída por uma alma feminina e de um corpo quase másculo, essa retirante sertaneja foi então apelidada de Luzia-Homem:

Era Luzia, conduzindo para a obra, arrumados sobre uma tábua, cinquenta tijolos. Viram-na outros levar, firme, sobre a cabeça, uma enorme jarra d'água, que valia três potes, de peso calculado para a força normal de um homem robusto [...]. Em plena florescência de mocidade e saúde, a extraordinária mulher, que tanto impressionara o francês Paul, encobria os músculos de aço sob as formas esbeltas e graciosas das morenas moças do sertão. Trazia a cabeça sempre velada por um manto de algodãozinho, cujas curelas prendia aos alvos dentes, como se, por um requinte de casquilhice, cuidasse com meticoloso interesse de preservar o rosto dos raios do sol e da poeira corrosiva. Pouco expansiva, sempre em tímido recato, vivia só, afastada dos grupos deconsortes de infortúnio, e quase não conversava com as companheiras de trabalho, cumprindo, com inalterável calma, a sua tarefa diária, que excedia à vulgar, para fazer jus a dobrada razão — É de uma soberbia desmarcada — diziam as moças da mesma idade, na grande maioria desenvoltas ou deprimidas e infamadas pela miséria. — A modos que despreza de falar com a gente, como se fosse uma senhora dona — murmuravam os rapazes remordidos pelo despeito da invencível recusa, impassível às suas insinuações galantes. — Aquilo nem parece mulher fêmea — observava uma velha alcoveta e curandeira de profissão. Reparem que ela tem cabelos nos braços e um buço que parece bigode de homem [...] (OLÍMPIO, 1983, p. 4).

Podemos perceber nesses dois romances que o sertão já não é mais visto como “um lugar bom para se viver” e, por outro lado, o sertanejo não é mais um personagem heroicizado por suas qualidades, mas é visto como um ser humano comum, dotado de qualidades e também de defeitos, que, muitas vezes, mostra-se corrompido pelas circunstâncias ou pelo

meio em que vive. O sertão, portanto, começa a ser desconstruído e a figura do sertanejo é mais parecida com a imagem de um “anti-herói”: um homem rústico, pobre, extremamente castigado pelo flagelo das secas e, ao mesmo tempo, pelos donos das terras e do poder, um ser humano real apanhado em suas fraquezas e virtudes que luta pela sua própria sobrevivência.

Influenciados pelos mesmos ideais de denúncia e crítica dos problemas econômicos e sociais e desejosos em colaborar para uma mudança social, vários escritores começam a representar as suas regiões entre os anos de 1905 a 1922 (período em que surgem as primeiras inquietações modernistas), dentre eles destacam-se: João Simões Lopes Neto, Hugo de Carvalho Ramos e Monteiro Lobato. Em *Contos Gauchescos* (1912), obra composta por dezenove contos, João Simões revela a vida do sertanejo gaúcho, os costumes, as tradições, a subjugação do humilde trabalhador rural, a violência na região, a carnificina motivada por diferentes situações naquele sertão do Rio Grande do Sul. Hugo de Carvalho Ramos, em *Tropas e boiadas* (1917), descreve a realidade do sertanejo goiano, mais especificamente o cotidiano dos tropeiros, suas tradições, seus costumes, suas condições de vida naquele sertão inóspito e abandonado. Em *Urupês*<sup>5</sup> (1918), também um livro de contos, Monteiro Lobato se preocupa em mostrar, através do personagem Jeca Tatu, o estado de miséria do sertanejo, o descaso do governo em relação aos sertões do Brasil, o atraso econômico e social em que vivia o homem do interior, a decadência da zona rural e de seus habitantes, as queimadas praticadas pelos caboclos na Serra da Mantiqueira e os problemas por elas causados. Ao construir o personagem Jeca Tatu, um dos personagens mais icônicos da literatura brasileira, Monteiro Lobato tinha uma intenção muito clara: revelar a miséria, a ignorância, o sofrimento do sertanejo alheio a tudo que acontecia naquele sertão decadente, atrasado, esquecido, menosprezado. Não por acaso esse personagem é apelidado de “urupê”, uma espécie de fungo parasita, nome também dado ao título da obra.

Contrapôs-lhe a cruel etrologia dos sertanistas modernos um selvagem real, feio e brutesco, anguloso e desinteressante, tão incapaz. muscularmente, de arrancar uma palmeira, como incapaz, moralmente, de amar Ceci [...] O indianismo está de novo a deitar copa, de nome mudado. Crismou-se de

---

<sup>5</sup> “*Urupês*, obra publicada originalmente em 1918, reúne ao todo 14 contos de Monteiro Lobato. Segundo o prefácio da 2ª. edição do livro, esta obra surgiu do artigo “Velha praga”, publicado originalmente no jornal O Estado de São Paulo no ano de 1914. Na época, Monteiro Lobato dedicava-se ao trabalho na fazenda que recebeu como herança de seu avô e amargava um ano terrível por conta da seca. Além do problema causado pela seca do inverno, Monteiro Lobato estava exausto das constantes queimadas praticadas pelos caboclos. Por conta disso, ele resolveu escrever uma carta de indignação ao jornal, que viu naquele texto algo muito valioso e o publicou fora da seção de cartas dos leitores. “Velha praga” causou grande impacto e polêmica, fazendo com que Monteiro Lobato publicasse outros textos que dariam origem ao livro *Urupês*”. Disponível em: <<https://guiadoestudante.abril.com.br/estudo/urupes-resumo-e-analise-da-obra-de-monteiro-lobato/>>. Acesso em: 20 jul. 2018.

"caboclismo". O cocar de penas de arara passou a chapéu de palha rebatido à testa; o ocará virou rancho de sapé: o tacape afilou, criou gatilho, deitou ouvido e é hoje espingarda troxada; o boné descaiu lamentavelmente para pio de inambu; a tanga ascendeu a camisa aberta ao peito. Mas o substrato psíquico não mudou: orgulho indomável, independência, fidalguia, coragem, virilidade heróica, todo o recheio em suma, sem faltar uma azeitona, dos Peris e Ubirajaras [...] Jeca Tatu é um piraquara do Paraíba, maravilhoso epítome de carne onde se resumem todas as características da espécie. Ei-lo que vem falar ao patrão. Entrou, saudou. Seu primeiro movimento após prender entre os lábios a palha de milho, sacar o rolete de fumo e disparar a cusparada d'esguicho, é sentar-se jeitosamente sobre os calcanhares [...] Pobre Jeca Tatu! Como és bonito no romance e feio na realidade! Jeca mercador, Jeca lavrador, Jeca filósofo... Quando comparece às feiras, todo o mundo logo adivinha o que ele traz: sempre coisas que a natureza derrama pelo mato e ao homem só custa o gesto de espíchar a mão e colher - cocos de tucum ou jicara, guabirobas, bacuparis, maracujás, jataís, pinhões, orquídeas; ou artefatos de taquarapoca- peneiras, cestinhas, samburás, tipitis, pios de caçador; ou utensílios de madeira mole - gamelas, pilõezinhos, colheres de pau. Nada mais (LOBATO, 1994, p. 90-92).

Nesse fragmento, percebemos como a figura do sertanejo vai sendo construída ao longo da narrativa: um homem pobre, ignorante (desprovido de educação), xucro, desleixado com a sua aparência e higiene pessoal, acomodado com a sua situação de miséria e de total abandono; um verdadeiro retrato do sertanejo brasileiro de então. A visão crítica da realidade brasileira, a denúncia dos problemas econômicos e sociais de seu tempo, a linguagem empregada (a oralidade, recurso utilizado para representar a fala do sertanejo brasileiro da zona rural)<sup>6</sup> revelam, sem dúvida, a face moderna de Lobato, assegurando-lhe um lugar de destaque não só na literatura, mas na história da nossa cultura.

De igual modo, ou seja, em um tom crítico e caricatural, e utilizando também os recursos da narrativa oral, Hugo de Carvalho Ramos e João Simões Lopes Neto enredam suas histórias. De um lado, Hugo nos apresenta a fala do sertanejo goiano:

[...] a rosilha não ia bem, sentida dos rins, o que o punha preocupado, tanto mais que o macho mascarado, trazido à escoteira, sempre à mão, para atalhar tropeços e incômodos de montar na necessidade lombo chucro de animal ruim ou passarinho, aguara dos cascos na subida da serra do Corumbá [...] (RAMOS, 1917, p. 9).

De outro lado, Simões Neto nos apresenta a fala do sertanejo gaúcho:

---

<sup>6</sup> “A partir do livro surgiram diversas palavras e expressões que hoje são dicionarizadas, como por exemplo o termo “jeca”, que vem da personagem Jeca Tatu e passou a ser sinônimo de “caipira”, “morador da zona rural” ou ainda “pessoa de hábitos rudimentares”. Disponível em: <<https://guiadoestudante.abril.com.br/estudo/urupes-resumo-e-analise-da-obra-de-monteiro-lobato/>>. Acesso em: 20 jul. 2018.

Genuíno tipo — crioulo — rio-grandense (hoje tão modificado), era Blau o guasca sadio, a um tempo leal e ingênuo, impulsivo na alegria e na temeridade, precavido, perspicaz, sóbrio e infatigável; e dotado de uma memória de rara nitidez brilhando através de imaginosa e encantadora loquacidade servida e floreada pelo vivo e pitoresco dialeto gauchesco [...] - Eu tropeava, nesse tempo. Duma feita que viajava de escoteiro, com a guaiaca empanzinada de onças de ouro, vim varar aqui neste mesmo passo, por me ficar mais perto da estância da Coronilha, onde devia pousar (LOPES NETO, 1976, p. 2, 3).

Nas obras desses três autores observamos que o emprego do discurso direto é recorrente. Além de ser uma técnica para representar a fala do sertanejo, esse recurso também foi utilizado com o propósito de dar voz ao homem do interior que, na época, era vigorosamente marginalizado, sem direitos e proteção, isto é, sem “vez e nem voz”. Sobre essa condição de hilota, Leitão Júnior (2012) reitera:

Ademais, o resultado da “marcha civilizatória e modernizadora” não é a “anexação” permanente dos sertões aos domínios da soberania hegemônica, mas sim a (re)criação perpétua dos espaços sertanejos no âmbito da formação territorial e social brasileiras: sempre sob os signos da necessidade de sintonia aos centros político-econômicos hegemônicos do capitalismo mundial, as supostas condições do “atraso” do Brasil na concertação das nações alimentou constantemente um imaginário do território a conquistar, motivando uma vitalícia “marcha modernizadora”. Nesse processo, tornou-se recorrente o reconhecimento de porções territoriais apartadas e desintegradas aos núcleos hegemônicos do poder, reproduzindo incessantemente os sertões e, como corolário, as preocupações quanto à necessidade de alteração (superação) das suas condições sertanejas no âmbito da consolidação da brasilidade – é nessa perpetuação dos espaços de “lógicas atrasadas”, assim qualificadas quando comparadas à lógica hegemônica, que reside um dos cerne da formação territorial brasileira (LEITÃO JÚNIOR, 2012, p. 5).

Na verdade, naquele momento histórico se conclamava a revelação da realidade legítima do Brasil: uma nação geográfica bipartida entre as porções territoriais do Sertão e do Litoral, constituída pelo sertanejo miserável/atrasado/arcaico/ignorante e pelo homem civilizado/moderno/instruído/. A intenção era anular ou neutralizar o que havia de ruim no país para que pudessem seguir rumo à (re)construção de um Brasil melhor. Por essa razão, o sertão se tornou o foco das atenções, figurando como protagonista nas obras literárias de então.

Assim, preocupados com os problemas humanos e sociais de seu tempo, esses escritores, além de intensificar ainda mais essa linhagem regionalista, abrem caminho para uma nova fase na literatura brasileira: a famosa geração do romance regionalista de 30, romance neorrealista moderno da segunda fase do Modernismo, caracterizado como “romance

social nordestino” que consolidou definitivamente o Regionalismo na prosa literária, passando a vigorar como uma tendência literária de grande valor. Trata-se de uma arte predominantemente social, neorrealista, fundamentada na denúncia dos problemas econômicos e sociais do Nordeste, a saber: a realidade do sertão nordestino, a rusticidade e precariedade da região, os dramas vivenciados pelos sertanejos retirantes, o flagelo das secas, a desumanização do ser humano frente ao primitivismo agreste do sertão, a exploração do trabalhador rural em um sistema social injusto. Para Bechara (1991, p. 9), esse regionalismo enquadra-se “no que se costuma considerar literatura de protesto pelo que contribuem em no chamar a atenção para os dramas dos nossos sertões no afã de que, com a denúncia, se minorem a miséria”.

O marco inicial dessa linha neorrealista se deu com a publicação de *A Bagaceira* de José Américo de Almeida, em 1928. A partir daí surgiram várias obras que marcaram definitivamente essa fase do Regionalismo, dando continuidade ao gênero “Ciclo das Secas” da Literatura Nordestina. Dentre as obras mais importantes dessa fase destacam-se: *O Quinze* (1930) de Raquel de Queiroz, *Menino de Engenho* (1932) de José Lins do Rego, *Vidas secas* (1938) de Graciliano Ramos, *Seara Vermelha* (1946) de Jorge Amado.

De acordo com Leitão Júnior (2012), o sertão é elencado nessas obras como “região-problema”, sob as bandeiras do “atraso” e da “obsolescência”. Esse sertão representa:

[...] o ambiente das secas e da Caatinga, semiárido, desértico e marrom, [...] é essa abrangência que permite caracterizar e unificar minimamente os ambientes literários de Rachel de Queiroz, Graciliano Ramos, José Lins do Rego e Jorge Amado na qualidade de sertões, porque, embora heterogêneos, estão todos ligados à condição de não responderem às lógicas hegemônicas – sequer enquanto periferias – da política, da cultura, da sociedade e da economia capitalista, em desacordo que estão com tais diretrizes ao serem estruturados em torno de poderes personificados, violências naturais e sociais, arbitrariedades políticas e fragilidades político-econômicas (LEITÃO JÚNIOR, 2012, p. 5).

Em “A Bagaceira”, romance que abriu uma nova fase na história da literatura brasileira, José Américo nos apresenta em um tom de denúncia: os dramas vivenciados pelos sertanejos nordestinos nos períodos de seca, o autoritarismo e a violência praticados pelos senhores de engenho, os sofrimentos dos “brejeiros” (homens que trabalham nos engenhos) e a sua desumanização, os assassinatos motivados pela chamada “honra sertaneja”, enfim, a luta constante do sertanejo pela sua sobrevivência.

Utilizando uma linguagem regional, o escritor caracteriza o sertanejo e o seu sertão de forma surpreendente, retratando verdadeiros tipos humanos e paisagens realistas do sertão nordestino. Para exemplificar, citamos os seguintes fragmentos:

[...] A casa-grande, situada numa colina, sobranceava o caminho apertado, no trecho fronteiro, entre o cercado e o açude. Num repentino desenfado, Dagoberto estirou o olhar, por cima das mangueiras meãs enfileiradas ladeira abaixo, para a estrada revolta. Parecia a poeira levantada, a sujeira do chão num pé de vento. Era o êxodo da seca de 1898. Uma ressurreição de cemitérios antigos – esqueletos redivivos, com o aspecto terroso e o fedor das covas podres. [...] Meninotas, com as pregas da súbita velhice, caretavam, torcendo as carinhas decrépitas de ex-voto. Os vaqueiros másculos, como titãs alquebrados, em petição de miséria. Pequenos fazendeiros, no arremesso igualitário, baralhavam-se nesse anônimo aniquilamento. [...] Párias da bagaceira, vítimas de uma emperrada organização de trabalho e de uma dependência que os desumanizava, eram os mais insensíveis ao martírio das retiradas (ALMEIDA, 2004, p. 4, 5)

Sobre esse estilo admirável do escritor, Cavalcanti Proença, na introdução do livro, argumenta:

Escritor que sabe escutar-se e escutar, José Américo tira do primeiro saber a justeza da frase, no sentido de ajustamento ao assunto, à circunstância, ao personagem; do segundo, a fidelidade na transcrição da fala regional e a precisão das imagens com que traduz as vozes da natureza e das coisas, muito frequentemente animizadas. [...] O que se lê não é, apenas, a fala rural nordestina, ou, mesmo, paraibana, mas, a fala particular da bagaceira, no timbre e no entôo do Marzagão. Os diálogos conferem com os personagens; o alongamento de vogais (*preee-tinho*, etc.) acentua os superlativos; alterações de letras ou de timbres (*sé-vergonho*, *felha da pota*, etc.) e o reforço de consoantes (*cabra sssfado*, etc.) trazem ao leitor o tom e o andamento da linguagem do eito (PROENÇA, 2004, s/p, introdução).

Complementando as afirmações de Cavalcanti Proença, Tristão de Athayde reitera: “este livro pode ser colocado ao lado dos maiores romances brasileiros, pois não é apenas um grande livro nosso: é um grande livro humano” (ATHAYDE, 2004, s/p, parte introdutória). Cumpre dizer que, além de nos apresentar uma visão realista do sertão, esse romance nos revela (ainda que em uma linguagem artística) aspectos da nossa cultura e da nossa língua antes não observados.

Valendo-se também de um linguajar regional nordestino e de uma percepção objetiva da realidade, Raquel de Queiroz focaliza em *O Quinze* a seca no nordeste, descrevendo vários aspectos da vida do interior cearense: o drama do sertanejo devastado pela seca, o desespero, a fome, os crimes cometidos, as doenças, a falta de trabalho. A título de ilustração citamos os seguintes excertos:

[...] Encostado a uma jurema seca, defronte ao juazeiro que a foice dos cabras ia pouco a pouco mutilando, Vicente dirigia a distribuição de rama verde ao gado. Reses magras, com grandes ossos agudos furando o couro das ancas, devoravam confiadamente os rebentões que a ponta dos terçados espalhava pelo chão [...] Na serra, também, o recurso falta... Também o pasto seca... Também a água dos riachos afina, afina, até se transformar num fio gotejante e transparente. Além disso, a viagem sem pasto, sem bebida certa, havia de ser um horror, morreria tudo (QUEIROZ, 2012, p. 2, 3).

Em um tom de narrativa oral, José Lins do Rego narra em *Menino de engenho* as experiências do personagem Carlos Melo, um menino de 8 anos que, após a morte trágica do pai e da mãe, foi morar no engenho Santa Rosa (região entre Pernambuco e Paraíba) de seu avô José Paulino, o maior proprietário de terras da região. Inspirado em suas memórias de infância, época em que passou no engenho do avô e nos contadores de causos da época, José Lins do Rego retrata nesse romance: a vida do sertanejo na época dos engenhos de cana de açúcar; a exploração do trabalhado rural, o sofrimentos dos escravos e o castigo do “tronco”; a enchente do rio e suas consequências; as superstições e crendices populares, como a do “lobisomem”; os bandidos, os cangaceiros e a violência exercida por eles.

[...] TRÊS DIAS depois da tragédia levaram-me para o engenho de meu avô materno. Eu ia ficar ali a morar com ele. Um mundo novo se abria para mim. Lembro-me da viagem de comboio e de uns homens que iam conosco no mesmo carro. O tio Juca, que fora buscar-me, contava a história, afirmando que o meu pai estava doido. Todos olhavam para mim com um grande pesar. — Eu avalio como deve estar o coronel Cazuza — dizia um deles. — Naquela idade, a sofrer destas coisas! Compreendi que falavam do meu avô [...]. Quando cheguei, com o meu tio Juca, ao pátio da casa, o alpendre estava cheio de gente. Desapeamo, e uma mulher muito parecida com a minha mãe foi logo me abraçando e beijando [...] (REGO, 2001, p. 21, 22).

Seguindo os trilhos abertos por Raquel de Queiroz, Graciliano Ramos também expõe em *Vidas secas* os efeitos do flagelo das secas na região nordestina, como: a fuga dos sertanejos retirantes massacrados pelas agruras do sertão, a desonestidade do patrão e a subjugação do humilde trabalhador rural, a arbitrariedade da classe dominante que deixa o sertanejo numa situação degradante de miséria e total abondo. A seca em Graciliano Ramos é tão marcante que reduz os personagens a uma condição de “bicho”. Essa desumanização, coisificação ou animalização é visível nos fragmentos:

[...] Chegara naquele estado, com a família morrendo de fome, comendo raízes. Caíra no fim do pátio, debaixo de um juazeiro, depois tomara conta da casa deserta. Ele, a mulher e os filhos tinham-se habituado à camarinha escura, pareciam ratos — e a lembrança dos sofrimentos passados esmorecera. P. 2 [...]— Você é um bicho, Fabiano. Isto para ele era motivo de orgulho. Sim senhor, um bicho, capaz de vencer dificuldades [...]Apossara-



se da casa porque não tinha onde cair morto, passara uns dias mastigando raiz de imbu e sementes de mucunã. Viera a trovoada. E, com ela, o fazendeiro, que o expulsara. [...] Aparecera como um bicho, entocara-se como um bicho [...]. Chape-chape. As alpercatas batiam no chão rachado. O corpo do vaqueiro derreava-se, as pernas faziam dois arcos, os braços moviam-se desengonçados. Parecia um macaco [...] (RAMOS, 2006, capítulo 2, p. 1).

Brutalizados e animalizados pelo ambiente e, sobretudo, pela miséria e exploração, o vaqueiro Fabiano e sua família fogem daquela região (no interior de Pernambuco) em busca do mínimo para sobreviver. Para potencializar ainda mais o sofrimento da família e a sua condição de hilota, Graciliano além de não dar nome aos filhos, utiliza uma linguagem seca, rústica, objetiva, com raríssimos diálogos:

[...] Vivia longe dos homens, só se dava bem com animais. Os seus pés duros quebravam espinhos e não sentiam a quentura da terra. Montado, confundia-se com o cavalo, grudava-se a ele. E falava uma linguagem cantada, monossilábica e gutural, que o companheiro entendia. A pé, não se agüentava bem. Pendia para um lado, para o outro lado, cambaio, torto e feio. As vezes utilizava nas relações com as pessoas a mesma língua com que se dirigia aos brutos – exclamações, onomatopéias. Na verdade falava pouco. Admirava as palavras compridas e difíceis da gente da cidade, tentava reproduzir algumas, em vão, mas sabia que elas eram inúteis e talvez perigosas [...] (RAMOS, 2006, capítulo 2, p. 1).

De temática semelhante à obra de José Lins do Rego, só que agora representando o ciclo do cacau no Sul da Bahia e utilizando também uma linguagem regional, Jorge Amado nos apresenta *Seara Vermelha*. A obra conta a história de Zé trovoada, jagunço temível, que por uma ironia do destino mata seu próprio irmão. Nesse romance de cunho neorrealista o autor destaca: o coronelismo na região, os jagunços, a impunidade e a violência, banditismo, as mortes dos trabalhadores e as tocais, a obediência e a “animalização” dos homens daqueles sertões, a exploração do trabalho, os jogos de interesses dos políticos, a luta dos fazendeiros pela fixação e expansão das terras do cacau, o “atraso” expressivo da região em comparação aos grandes centros político-econômicos. Nas palavras de Leitão Júnior (2012):

Portanto, o emblema do “atraso” da zona cacauzeira também pesa na descrição regionalista de Jorge Amado, assim como o fizera José Lins do Rego para a zona açucareira. Na obra desses literatos, há notáveis aproximações que decantam características denunciadoras da “obsolescência” dos ambientes privilegiados em suas tramas romanescas: a violência, o coronelismo, a perversidade sexual e a prostituição são máculas sociais e políticas que irrevogavelmente traçam o perfil regional em sua condição sertaneja, em falta de sintonia com a moralidade e os códigos éticos hegemônicos (LEITÃO JÚNIOR, p. 15, 16).

De modo geral e sucinto, podemos afirmar, consoante Leitão Júnior (2012), que, seja por meio da violência natural – em que pesa a imagem generalizada da seca, representada por Raquel de Queiroz e Graciliano Ramos, ou por meio da violência social – em que pesam as características da miséria social, do atraso, da ignorância, da exploração, da subjugação do pobre sertanejo nordestino - representada por José Lins do Rego e Jorge Amado, todos eles contribuíram significativamente para a legitimação da “eficácia política”<sup>7</sup> e, principalmente, para a constituição de uma literatura propriamente brasileira amplamente impregnada na nossa cultura e cristalizada no imaginário coletivo da nossa sociedade.

Desse modo, calcados em um Regionalismo<sup>8</sup>, consciente dos problemas que agitavam o Brasil na época e mergulhados em projeto nacionalista de denúncia social e crítica como meio de ação para a transformação, esses escritores produziram obras de grande valor estético e linguístico, razão por que possuem um *status* privilegiado na nossa literatura e merecida consagração na história do regionalismo literário brasileiro.

Importa observar que essa fase do Regionalismo (Regionalismo problemático ou crítico) foi um dos períodos de maior produtividade literária regionalista, razão pela qual escolhemos um número maior de obras para compor o nosso *corpus* de estudo. Mesmo apresentando características diferentes do Regionalismo pitoresco, esse Regionalismo se insere numa linha de continuidade artística da tendência anterior, fortificando-se nessa segunda fase e consolidando-se dentro da conjuntura literária com Guimarães Rosa.

De acordo com Candido (1989), a partir de 1945 (3ª geração modernista), mais especificamente com a publicação de *Grande Sertão: Veredas* (1956) de Guimarães Rosa, nasce a terceira vertente do Regionalismo, o *Super-regionalismo*: um regionalismo calcado na consciência “dilacerada” do subdesenvolvimento, uma tendência firmada “no que se poderia chamar de a universalidade da região”. Nas palavras do crítico:

Iniciado o *super-regionalismo*, a partir de 45, a ficção brasileira passou por grandes experiências com o aparecimento de Guimarães Rosa, autor que alcança êxito universal ao recriar a linguagem regional de forma extremamente elaborada, dando-nos mostra palpável de como a palavra é

---

<sup>7</sup> “Transformados em canônicos, Rachel de Queiroz, Graciliano Ramos, José Lins do Rego e Jorge Amado tiveram os seus discursos alçados à condição de *eficácia política*, de modo que, mesmo que incorporados flagrantemente de modo parcial, participaram significativamente na legitimação da necessidade de superação dos sertões nacionais por meio da ampla repercussão de suas retóricas no imaginário geográfico e na “alma nacional”, denunciadoras das características perniciosas dos sertões que divulgaram. Quer isto dizer, que esses escritores, contribuíram, ainda que de forma parcial, para a implantação de projetos e ações estatais em favor das comunidades carentes dos sertões brasileiros de então” (LEITÃO JÚNIOR, 2012, p.17).

<sup>8</sup> Por todas essas razões é que optamos pelo termo “Regionalismo crítico”, em vez de “Regionalismo problemático”, como denomina Candido (1989).

flexível e a língua, consequentemente, maleável [...] (CANDIDO, 1989, p.162).

Para Candido (1989), a universalidade dos temas, a transcendência do regional através da exploração dos grandes problemas que envolvem o ser humano em qualquer lugar, a criação de uma nova linguagem fizeram a obra de Rosa ultrapassar todos os limites até então impostos pelo próprio regionalismo, obrigando o crítico a criar uma nova categoria: o *super-regionalismo*. No seu ponto de vista:

Descartando o sentimentalismo e a retórica; nutrida de elementos não-realistas, como o absurdo, a magia das situações; ou de técnicas antinaturalistas, como o monólogo interior, a visão simultânea, o escorço, a elipse – ela implica não obstante em aproveitamento do que antes era a própria substância do nativismo, do exotismo e do documentário social. Isto levaria a propor a distinção de uma terceira fase, que se poderia (pensando em surrealismo, ou super-realismo) chamar de *super-regionalista*. Ela corresponde à consciência dilacerada do subdesenvolvimento e opera uma explosão do tipo de naturalismo que se baseia na referência a uma visão empírica do mundo; naturalismo que foi a tendência estética peculiar a uma época onde triunfava a mentalidade burguesa e correspondia à consolidação das nossas literaturas (CANDIDO, 1989, p. 160,161).

Assim, o *super-regionalismo* instituiu-se como uma tendência renovada, desprendendo-se do pitoresco e de uma visão superficial sobre a região, para uma visão universalista. Trata-se de um Regionalismo novo, comprometido com as questões sociais e com as problemáticas que envolvem o ser humano em qualquer lugar do mundo, ou seja, agora a região deixa de ser particular e passa a ser universal. Além de *Grande Sertão: Veredas*, outras obras ganham destaque nessa fase, como: *O Coronel e o Lobisomem* (1964) de José Cândido de Carvalho, *Chapadão do Bugre* (1965) de Mário Palmério, *Sargento Getúlio* (1971) de João Ubaldo Ribeiro e a peça *Auto da Compadecida* (1955) de Ariano Suassuna.

Em um estilo único, totalmente inovador, em uma linguagem regional reelaborada - “exploração de aspectos sonoros da linguagem (aliterações, onomatopeias); criação de palavras; aproveitamento do linguajar regionalista pleno de arcaísmos, adaptação de termos e expressões extraídas de várias línguas modernas e antigas” (TUFANO, 1988, p. 280) - Guimarães Rosa nos apresenta *Grande sertão: veredas*, um romance regionalista que retrata o sertão como o próprio mundo com seus mistérios e desencontros, como um lugar vasto e sem limites (fronteiras), como bem diz o personagem Riobaldo: “o Sertão está em toda a parte”, “o Sertão é sem lugar” (ROSA, 1994, p. 5), sobretudo, como um pano de fundo para extrair do regional matéria para atingir o universal, ou seja, o escritor faz do sertão cenário para

mostrar os problemas vitais que existem no interior do homem de qualquer região, como: o sentido da existência humana, o destino do ser humano, a existência de Deus e do diabo, o significado da vida e da morte.

Essas indagações religiosas e metafísicas estão presentes em toda a narrativa por meio dos relatos de Riobaldo, narrador-protagonista que conta a sua vida de jagunço a um suposto doutor, demonstrando seus medos, anseios, desejos, reflexões e angústias naquele sertão difícil e perigoso. Em busca de respostas para o mistério de sua própria condição, Riobaldo nos apresenta o sertão-mundo para revelar a si próprio: um homem atormentado com várias dúvidas e incertezas sobre a existência de Deus ou do diabo; a relação entre o Bem e o Mal; o sentimento que experimentou por Diadorim – uma mulher que se passa por homem para vingar a morte de seu pai; o sentido de sua vida como jagunço; o significado da vida e da morte; o destino do ser humano. A título de exemplificação citamos os seguintes fragmentos:

–NONADA. TIROS QUE O SENHOR ouviu foram de briga de homem não, Deus esteja. Alvejei mira em árvores no quintal, no baixo do córrego. Por meu acerto. Todo dia isso faço, gosto; desde mal em minha mocidade. Daí, vieram me chamar. Causa dum bezerro: um bezerro branco, eroso, os olhos de nem ser –se viu –,e com máscara de cachorro. Me disseram; eu não quis avistar. Mesmo que, por defeito como nasceu, arrebitado de beiços, esse figurava rindo feito pessoa. Cara de gente, cara de cão: determinaram – era o demo. Povo prascóvio. Mataram [...] p. 3º senhor tolere, isto é o sertão. Uns querem que não seja: que situado sertão é por os campos-gerais a fora a dentro, eles dizem, fim de rumo, terras altas, demais do Urucuia [...]Esses gerais são sem tamanho. Enfim, cada um o que quer aprova, o senhor sabe: pão ou pães, é questão de opiniões... O sertão está em toda a parte. Do demo? Não glosa. Senhor pergunte aos moradores. Em falso receio, desfalcam no nome dele. [...]Pois, hem, que, despontar o Rio pelas nascentes, será a mesma coisa que um se redobrar nos internos deste nosso Estado nosso (ROSA, 1994, 3- 6).

Além dos conflitos de Riobaldo, podemos perceber que a linguagem de Rosa é inovadora, cheia de neologismos e arcaísmos (“nonada”, “opiniões”, “prascóvio”; “eroso”); é bela, eloquente, por vezes simples e rústica, beirando à oralidade, por vezes culta ou poética. Isto nos revela o talento surpreendente de Rosa, a sua criatividade e simplicidade em fazer do simples, do comum, do rústico, ou seja, da “pedra bruta” meramente regional - o sertão, uma verdadeira obra-prima que atingiu um veio universal, como bem pontua Candido (2002b):

Na extraordinária obra-prima *Grande Sertão: Veredas* há de tudo para quem souber ler, e nela tudo é forte, belo impecavelmente realizado. Cada um poderá abordá-la a seu gosto, conforme o seu ofício; mas em cada aspecto aparecerá o traço fundamental do autor: a absoluta confiança na liberdade de inventar [...]. A experiência documentária de Guimarães Rosa, a observação da vida sertaneja, a paixão pela coisa e pelo nome da coisa, a capacidade de

entrar na psicologia do rústico, - tudo se transformou em significado universal graças à invenção, que subtrai o livro à matriz regional para fazê-lo exprimir os grande lugares comuns, sem os quais a arte não sobrevive: dor, júbilo, ódio, amor, morte, - para cuja órbita nos arrasta a cada instante, mostrando que o pitoresco é acessório e que na verdade o Sertão é o Mundo (CANDIDO, 2002b, p. 121-122).

Portanto, o valor de *Grande Sertão: Veredas* não se resume simplesmente em sua riqueza estética, mas, principalmente em sua riqueza linguística que nos permite enxergar as realidades expressivas e humanas não apenas do sertão mineiro, mas a realidades que existem no interior do homem de qualquer região. Não por acaso Guimarães foi considerado um dos escritores mais notáveis da nossa literatura, assumindo em 1967 a cadeira 2 (dois) na Academia Brasileira de Letras. Além de receber vários prêmios como: o Prêmio Machado de Assis, do Instituto Nacional do Livro, em 1961; o Prêmio Carmem Dolores Barbosa, de São Paulo, em 1957; e o Prêmio Paula Brito, do Rio de Janeiro, Rosa encabeçou a tríplice, composta por Clarice Lispector e João Cabral de Melo Neto, como os melhores escritores da terceira geração modernista brasileira.

Nessa mesma linhagem regionalista de Rosa, José Cândido de Carvalho em *O coronel e o lobisomem* (1964), romance também em primeira pessoa, nos mostra os conflitos vivenciados pelo coronel Ponciano de Azeredo Furtado, homem do sertão (dos Campos dos Goitacazes) que luta contra seus próprios medos aguçados pelas lendas e superstições (lobisomem, ururau<sup>9</sup>, sereia, onças) e contra as mais variadas formas de injustiça: contra o valente do circo (Vaca-Braba), contra o cobrador de impostos, contra o suposto agiota. Seu caráter de personagem lendário, de herói picaresco, se solidifica no momento em que é vitorioso contra as forças sobrenaturais do lobisomem. Para bem elucidar essas características, citamos o trecho a seguir:

A bem dizer sou Ponciano de Azeredo Furtado, coronel de patente, do que tenho honra e faço alarde. Herdei do meu avô Simeão terras de muitas medidas, gado do mais gordo, pasto do mais fino. Leio no corrente da vista e até uns latins arranhei em tempos verdes da infância, com uns padres-mestres a dez tostões por mês. Digo, modéstia de lado, que já discuti e joguei no assoalho do Foro mais de um doutor formado. Mas disso não faço glória, pois sou sujeito lavado de vaidade, mimoso no trato, de palavra educada. Já morreu o antigamente em que Ponciano mandava saber nos ermos se havia um caso de lobisomem a sanar ou pronta justiça a ministrar. Só de uma regalia não abri mão nesses anos todos de pasto e vento: a de falar alto, sem

---

<sup>9</sup> Lenda de um rapaz muito pobre que se apaixona pela filha de um coronel muito rico. Esse coronel, insatisfeito como o namoro joga o rapaz no rio Paraíba do Sul que, por sua vez, é transformado pelo Deus das águas em um enorme jacaré de papo amarelo. Louco de saudades de sua amada, o ururau vem à beira do rio encontrar sua amada e como não a encontra acaba por tragar alguém que passa perto dele. Informação obtida no seguinte endereço eletrônico: <<https://pt.wikipedia.org/wiki/Ururau>>. Acesso em: 23 jul. 2018.

freio nos dentes, sem medir consideração, seja em compartimento do governo, seja em sala de desembargador. Trato as partes no macio, em jeito de moça. Se não recebo cortesia de igual porte, abro o peito: — Seu filho de égua, que pensa que é? [...]. De noite, era aquela algazarra de lobisomem, pio de coruja, asa de caburé, fora outros atrasos dos ermos [...] sujeito especial em lobisomem como eu não ia cair na armadilha de pouco pau (CARVALHO, 2014, p. 6).

Aproximando-se de Guimarães Rosa quanto à recriação da linguagem, embora sem a complexidade e eloquência rosiana, podemos perceber a grande capacidade imaginativa e estilística de José Cândido em representar os tipos humanos do sertão, seus costumes, suas tradições e superstições, sua variedade linguística, em um universo mítico-popular, senão fantasmagórico.

Seguindo os mesmos passos de Rosa e José Cândido, Mário Palmério nos apresenta em *Chapadão do Bugre* (1965) o falar do sertanejo dos “campos gerais”, reproduzindo em um tom de oralidade as preocupações e astúcias de José de Arimatéia, um dentista ambulante do interior de minas que mata o amante de sua noiva Maria do Carmo, dando início a uma sucessão de tragédias. Para exemplificar, eis alguns fragmentos:

A montaria mal se encostara à cerca de limão brabo, e o cavaleiro já desapareava. Vazou a tronqueirinha, rumando direito para a janela de frente da meiágua. Chamou: — João, ô João da Preta! Teve de bater m ais uma vez com o argolão da t aca e chamar de novo: — João, ô João da Pret a! Sou eu, é o José de Arimateia! A mulher foi quem veio ver à porta, mesmo assim sem arredar de todo a escora. Meia tirinha só de confiança na fresta mal aberta: — Seu Isé? — Sou eu sim . Anda logo, que a friagem aqui no tempo... Siá Preta acabou de abrir a porta, e o cavaleiro entrou . Treme-soprava: — Viajinh a danada de fora de horas... Qu ' é de o João? Recebeu o aviso? Descalço, encolhido num resto de poncho campanheiro, João da Preta apareceu . Mal salvou, e foi dizendo: — Recebi de tardezinha. Seu Valério f oi quem trouxe. Quer baldear agora, ou dilata um tico pro café? E antes que o outro resolvesse: — Vigia a mulher já na cozinha. Fogo de sabugo esquentado de vereda... Não, não carecia tanta pressa — José de Arimateia calculava. Havia saído com muita folga do Sassafrás, e a besta rompera bem o trecho até o porto, apesar do nevoeiro que fechava o rio e a Serra (PALMÉRIO, 2006, p. 8).

É importante observar que, além de utilizar vocábulos e expressões sertanejas Mário Palmério também realiza criações pessoais, porém menos radicais que os neologismos de Rosa, como por exemplo: “como em-antes”, “xibiu”, “um pimpingar”; “remolgueia”; “macioso”. Isso mostra o talento do escritor em transpor para a literatura a linguagem e os costumes do homem do interior, do sertanejo mineiro.

Baseado em uma história real e misteriosa ocorrida no interior de Minas Gerais, mais especificamente uma chacina ocorrida em Passos (MG), esse romance altamente dramático revela os desmandos de um patriarcalismo ainda imperante: a força arbitrária dos coronéis e jagunços; a violência e a falta de segurança naqueles sertões onde opera a lei de talião: “olho por olho, dente por dente”; os costumes de reparar com sangue a honra manchada e as perseguições sangrentas da família em busca do assassino.

Valendo-se também de uma linguagem sertanista, repleta de arcaísmos e invenções próprias (neologismos)<sup>10</sup>, João Ubaldo Ribeiro nos mostra em *Sargento Getúlio* (1971) o banditismo que imperava no sertão de Sergipe: as mortes sangrentas, a violência empregada pelos policiais e jagunços, os políticos corruptos, os pactos de lealdade e as consequências trágicas das traições. O trecho a seguir é uma amostra da sua criatividade linguística:

A gota serena é assim, não é fixe. Deixar, se transforma-se em gancho e se degenera em outras mazelas, de sorte que é se precatar contra mulheres de viagem. Primeiro preceito. De Paulo Afonso até lá, um esticão, inda mais de noite nessas condições. Estrada de carroça, peste. Temos Canindé de São Francisco e Monte Alegre de Sergipe e Nossa Senhora da Glória e Nossa Senhora das Dores e Siriri e Capela e outros mundãos, sei quantos. Própria e Maruim, já viu, poeiras e caminhãos algodoados, a secura fria. E sertão do brabo: favelas e cansações, tudo ardiloso, quipás por baixo, um inferno. Plantas e mulheres reimosas, possibilitando chagas, bichos de muita aleiva, potós, lacraias, piolhos de cobra, veja. Matei uns três infelizes assim, pelo cima de uns quipás, sendo que um chegou devagar no chão, receando os espinhos sem dúvidas. Assunte se quem vai morrer se incomoda com conforto. Fosse dado a sangria, terminava o vivente no ferro, porém faz um barulho esquisito e não é asseado por causo de todo aquele esguincho que sai. E dessa forma acertei um disparo no cachaço, procurando atitude para não desperdiçar munição. Inda xinguei por me obrigar a caçar por essas catingas, arremetido naquela soaeira, estropeando as reiúnas novas naquelas catanduvras embaracentas [...] (RIBEIRO, 2010, p. 3).

Inspirado nas histórias que seu pai Manoel Ribeiro, então chefe da Polícia Militar de Sergipe, contava em sua época de infância, João Ubaldo narra a história de Getúlio Santos Bezerra, sargento da polícia militar de Sergipe, um homem “bruto”, “violento” assim como o sertão, mas também fiel e corajoso, que recebe uma ordem para prender um adversário de um importante chefe político e levá-lo para Aracaju. Nessa trajetória, Getúlio enfrenta várias dificuldades para levar o preso até o destino, como: a resistência do preso, as perseguições dos jagunços contratados por Acrísio Nunes (o tal político) para tentar impedi-lo de seguir com a

---

<sup>10</sup> Eis algumas criações do escritor: “bexiguento”, “almospenados”, “sinfetar”, “consumições”, “soaeira”, dentre outras.

sua missão, a corrupção de um político que decide revogar a ordem original, a violência imperante naquela região. Altamente dramático e sanguinolento, o romance termina com a morte de Getúlio em confronto com os correligionários de Acrísio Nunes. De matéria regional, esse drama vivido por Getúlio torna-se universal na medida em que focaliza o estado de abatimento físico e moral do homem subjugado por uma lei que não é válida no mundo corrupto onde vive.

Incorporando os mesmos elementos regionalistas dos escritores anteriores (quanto à temática e a linguagem) só que agora em um estilo bem diferente Ariano Suassuna compõe o Auto da compadecida, uma peça teatral em forma de auto<sup>11</sup> que foi encenada pela primeira vez 1956 em Recife e adaptada para o cinema em 1969, alavancando definitivamente a trajetória artística do escritor. Em um tom satírico-humorístico, com elementos da literatura de cordel, do gênero comédia e do barroco católico brasileiro, a peça retrata o drama vivido pelos sertanejos nordestinos castigados pela seca, pela fome e miséria e sua constante luta pela sobrevivência. Nesse contexto de opressão e subjugação diante da dominação dos poderosos coronéis e da corrupção política e religiosa, o personagem João Grilo, é a representação própria dos sertanejos que tentam de todas as formas escapar dos flagelos do sertão. Por serem carentes de recursos muitas vezes utilizam as próprias armas e, no caso do personagem, a única arma que ele possui é a esperteza, a inteligência e a astúcia, como se pode observar nos diálogos a seguir:

MANUEL- E agora? Que é que você diz em sua defesa? Sei que você é astuto, mas não pode negar o fato de que foi acusado.

JOÃO GRILO- O senhor vai me desculpar, mas eu não fui acusado de coisa nenhuma.

MANUEL- Não?

ENCOURADO- Foi mesmo não. Começou com uma confusão tão grande que eu me esqueci de acusá-lo. Vou começar.

JOÃO GRILO- Você não vai começar coisa nenhuma, porque a hora de acusar já passou [...]

JOÃO GRILO- É, é verdade, mas do jeito que eles me pagavam, o jeito era eu me virar. Além disso eu estava com pena do gato, tão abandonado, e queria que ele passasse bem .

MULHER- É, e nessa pena levou meus quinhentos mil-réis.

ENCOURADO- Depois, foi ele quem matou Severino e o cabra dele, com uma história de gaita, Padre Cícero e não sei que mais.

JOÃO GRILO- Legítima defesa, Nosso Senhor! [...] (SUASSUNA, 2014, p. 113, 114).

O brilhantismo do autor está no humor que provoca ao abordar temáticas tão sérias, como: a corrupção na igreja; a exploração do trabalho; a desumanização do ser humano

---

<sup>11</sup> Gênero da literatura que trabalha com elementos cômicos e tem intenção moralizadora.



dominados por seus patrões; as desigualdades sociais e os privilégios dos mais favorecidos; a perspicácia do ser humano para se salvar. Em outras palavras, para tratar dos problemas que envolviam os sertanejos naquela região, o escritor não apenas utilizou um linguajar do sertanejo nordestino como se serviu do humor para atenuar os dissabores do sertão. Nos diálogos a seguir podemos identificar essas características:

ENCOURADO - Ele e a mulher foram os piores patrões que Taperoá já viu.

MULHER- É mentira!

JOÃO GRILO- É não, é verdade. Três dias passei... MANUEL- Em cima de uma cama, com febre, e nem um copo d'água lhe mandaram. Já sei, João, todo mundo já sabe dessa história, de tanto ouvir você contar.

JOÃO GRILO- Mas eu posso? Me diga mesmo se eu posso! Bife passado na manteiga pra o cachorro e fome pra João Grilo. É demais!

ENCOURADO- Avareza do marido, adultério da mulher. Bem-medido e bem-pesado, cada um era pior do que o outro.

JOÃO GRILO- Está aí Chicó que o diga.

MANUEL- Chicó?

JOÃO GRILO- Ah, é verdade, Chicó ficou. Já estava tão acostumado a aperrear pobre de Chicó que me esqueci de que ele tinha ficado. É um amigo meu.

MANUEL- Eu o conheço, estou até de olho nele por causa das histórias que vive contando.

JOÃO GRILO- Aquilo é o sol. Não vá ligar isso não. O sol do sertão é quente e Chicó começa a ver demais. É o sol [...] (SUASSUNA, 2014, p.109,110).

Além de tratar das mazelas do sertão de uma forma cômica, Suassuna constrói personagens picarescos: O padeiro e a mulher são avarentos, exploram o pobre Chicó, tratando-o como bicho enquanto o cachorro é por eles tratado como gente; Dorinha, a mulher do padeiro, além de avarenta é uma adúltera que se diz santa; o padre, o bispo e o sacristão são gananciosos, utilizam da autoridade religiosa para enriquecerem. Todos eles são mandados para o purgatório com a interseção de Nossa Senhora (A compadecida). Já Severino e o cangaceiro, apesar de todos os crimes cometidos em vida, são poupados por serem considerados vítimas daquela situação e a João Grilo lhe é dada uma nova chance, retornando a seu corpo para a alegria de seu amigo Chicó.

Por toda essa riqueza estética e linguística e também por tratar de temas universais assim como fizera Guimarães Rosa (o significado da vida e da morte, o destino do ser humano, a existência de Deus e do diabo, o pecado e a salvação), o Auto da Compadecida foi considerado o texto mais popular do moderno teatro brasileiro, tornando-se um clássico da nossa literatura. Não por acaso essa peça foi apresentada em 1999 (minissérie) pela Rede Globo e também exibida nos cinemas, saindo do regional para o nacional.

O QUADRO 2, a seguir, sintetiza as obras que compõem o nosso *corpus* de estudo.

QUADRO 2: Obras literárias do *corpus* de estudo

<b>REGIONALISMO PITORESCO</b>					
NOME DA OBRA	GÊNERO	AUTOR	DATA	REGIÃO	CORRENTE LITERÁRIA
O Sertanejo	Romance	José de Alencar	1875	Ceará (Nordeste)	Romantismo
O Garimpeiro	Romance	Bernardo Guimarães	1872	Minas Gerais	Romantismo
Inocência	Romance	Visconde de Taunay	1872	Mato Grosso do Sul	Romantismo
O Cabeleira	Romance	Franklin Távora	1876	Pernambuco (Nordeste)	Romantismo
Pelo Sertão	Contos	Afonso Arinos	1898	Minas Gerais/Goiás	Romantismo/Realismo
<b>REGIONALISMO CRÍTICO</b>					
Os Sertões	Romance	Euclides da Cunha	1902	Bahia (Nordeste)	Pré-Modernismo
Luzia-Homem	Romance	Domingos Olímpio	1903	Ceará (Nordeste)	Realismo-Naturalismo
Contos Gauchescos	Contos	João Simões Lopes Neto	1912	Rio Grande do Sul	Pré-Modernismo
Tropas e boiadas	Contos	Hugo de Carvalho Ramos	1917	Goiás	Pré-Modernismo
Urupês	Contos	Monteiro Lobato	1918	São Paulo	Pré-Modernismo
A Bagaceira	Romance	José Américo de Almeida	1928	Paraíba (Nordeste)	2ª fase do Modernismo
O Quinze	Romance	Raquel de Queiroz	1930	Ceará (Nordeste)	2ª fase do Modernismo
Menino de Engenho	Romance	José Lins do Rego	1932	Paraíba (Nordeste)	2ª fase do Modernismo

Vidas secas	Romance	Graciliano Ramos	1938	Alagoas (Nordeste)	2ª fase do Modernismo
Seara Vermelha	Romance	Jorge Amado	1946	Bahia (Nordeste)	2ª fase do Modernismo
<b>SUPER-REGIONALISMO</b>					
Grande Sertão: Veredas	Romance	Guimarães Rosa	1956	Minas Gerais/Bahia	3ª fase do Modernismo
O coronel e o Lobisomem	Romance	José Cândido de Carvalho	1964	Rio de Janeiro	3ª fase do Modernismo
Chapadão do Bugre	Romance	Mário Palmério	1965	Minas Gerais	3ª fase do Modernismo
Sargento Getúlio	Romance	João Ubaldo Ribeiro	1971	Sergipe (Nordeste)	3ª fase do Modernismo
Auto da compadecida	Peça teatral	Ariano Suassuna	1955	Paraíba (Nordeste)	3ª fase do Modernismo

Fonte: Elaboração da autora

Diante de todas essas considerações, pode-se inferir que o Regionalismo, além de ser uma tendência literária que contribuiu decisivamente para a constituição da literatura brasileira, foi um projeto nacionalista de valorização e expressão dos aspectos físicos e humanos do Brasil, ou seja, foi uma forma de identificar e apresentar o “tipicamente brasileiro”, como aponta Candido (2002a):

Trata-se de um caso privilegiado para estudar o papel da literatura num país em formação, que procura a sua identidade através da variação dos temas e da fixação da linguagem, oscilando para isto entre a adesão aos modelos europeus e a pesquisa de aspectos locais. [...] O Regionalismo, que o sucedeu e se estende até os nossos dias, foi uma busca do *tipicamente brasileiro* através das formas de encontro, surgidas do contato entre o europeu e o meio americano. Ao mesmo tempo documentário e idealizador, forneceu elementos para a auto-identificação do homem brasileiro e também para uma série de projeções ideais (CANDIDO, 2002a, p. 86, grifo do autor).

Ao mesmo tempo em que a literatura desempenha uma função social - seja por meio de um regionalismo pitoresco, de idealização do homem do sertão e da sua terra, de um regionalismo crítico, de denúncia e crítica dos problemas regionais, ou de um super-regionalismo, em que os temas locais se transfiguram em universais, ela também desempenha

uma função humanizadora e transformadora na medida em que atua na própria formação do homem, na constituição da sua cultura e transformação da sua língua.

Assim, baseado na realidade física e humana do país, o homem do sertão tornou-se tema central do regionalismo brasileiro como representação humana autêntica. A cada fase, a figura do sertanejo e de sua terra foi sendo moldada e representada de acordo com as aspirações da época e, assim, o regionalismo foi adquirindo novas roupagens até se consolidar dentro da conjuntura literária. A esse respeito, Candido (2002a) argumenta:

Mas é forçoso convir que, justamente porque a literatura desempenha funções na vida da sociedade, não depende apenas da opinião crítica que o regionalismo exista ou deixe de existir. Ele existiu, existe e existirá enquanto houver condições como as do subdesenvolvimento, que forcem o escritor a focalizar como tema as culturas rústicas mais ou menos à margem da cultura urbana. O que acontece é que ele se vai modificando e adaptando, superando as formas mais grosseiras até dar a impressão de que se dissolveu na generalidade dos temas universais, como é normal em toda obra bem feita. E pode mesmo chegar à etapa onde os temas rurais são tratados com um requinte que em geral só é dispensado aos temas urbanos, como é o caso de Guimarães Rosa, a cujo propósito seria cabível falar num super-Regionalismo [...] (CANDIDO, 2002a, p. 86, 87).

Por exercer uma função social, Candido (2002a) nos deixa claro que o regionalismo não “morreu”, pelo contrário, ele ainda está vivo e presente em nossa literatura, ainda que, de forma sutil, discreta. Fato a ser considerado é que ele nos trouxe importantes contribuições não só em termos literários com obras grandiosíssimas (em nível estético e linguístico), mas também em termos políticos, econômicos e sociais a partir do momento em que se propôs revelar o atraso de algumas regiões (interioranas) em relação às demais (próximas ao litoral) e, por conseguinte, o atraso do Brasil em relação aos países desenvolvidos.

Portanto, devemos procurar entender o regionalismo não como uma tendência anacrônica ou como sinônimo de literatura menor, como considerava alguns críticos, mas como um fenômeno literário dinâmico e significativo que vai se enriquecendo com as peculiaridades de cada escritor à medida que produzem obras que se complementam e se continuam.

O capítulo seguinte apresenta os pressupostos teóricos que fundamentam a nossa pesquisa, incluindo um breve panorama da Linguística de *Corpus* e algumas considerações e esclarecimentos sobre as ferramentas do *WordSmith Tools*®, versão 6,0 (SCOTT, 2012), programa escolhido para esta análise.

### **3 FUNDAMENTAÇÃO TEÓRICA**

Sabemos que o estudo das relações entre língua, cultura e sociedade tem se multiplicado nos últimos anos dando origem a várias pesquisas com diferentes abordagens linguísticas. De maneira particular, o estudo das unidades lexicais que congregam essas relações tem se constituído em um objeto substancial no âmbito das pesquisas de base terminológica. A multifuncionalidade das palavras em determinados universos de discurso tem conduzido os estudiosos a perceberem a tênue fronteira entre a unidade lexical especializada e a não especializada, ou seja, entre o termo e o vocábulo. Prova disso, contamos hoje com uma nova disciplina científica, a Etnoterminologia, subárea da Terminologia, criada por Maria Aparecida Barbosa para designar o estudo das unidades lexicais dos discursos etnoliterários. São, portanto, os preceitos teóricos desta disciplina correlacionada à Terminologia Textual e Terminologia Cultural que apoiamos e apresentamos neste capítulo.

#### **3.1 Da Terminologia à Etnoterminologia**

Falar em Etnoterminologia implica falar, antes de tudo, em Terminologia, já que a primeira é parte constitutiva da segunda, ou seja, uma subárea, uma disciplina teoricamente circunscrita nas instâncias desse campo das ciências da linguagem.

A Terminologia, “disciplina científica que estuda as chamadas línguas (ou linguagens) de especialidade e seu vocabulário” (BARROS, 2004, p. 21), também compreendida como o campo de estudos que tem como objetos o termo técnico-científico, a fraseologia especializada e a definição terminológica (KRIEGER; FINATTO, 2004), tem se desenvolvido consideravelmente nas últimas décadas, corroborando para o crescimento da produção científica e técnica em diferentes áreas.

Segundo Barros (2004), o desenvolvimento da investigação científica no âmbito da Terminologia e a atuação prática dos terminólogos em criar mecanismos mais eficientes para a disseminação da informação, fizeram com que os estudos terminológicos se expandissem para além dos domínios da pesquisa e chegassem em outros setores da vida social, como por exemplo, nas empresas, “deixando de ser uma atividade restrita a grupos de cientistas altamente especializados e se tornando cada vez mais uma necessidade na formação de profissionais de inúmeras áreas” (BARROS, 2004, p. 22).

Notadamente, os avanços técnicos e científicos têm provocado profundas transformações na sociedade e, paralelo a essas mudanças, os estudos terminológicos também têm sofrido modificações. A cada situação nova o universo lexical das línguas renova-se, ajusta-se, amplia-se substancialmente para satisfazer as necessidades humanas e, de igual modo, sucede com o seu conjunto terminológico. Verificamos, assim, uma evolução da Terminologia, dos seus pressupostos teóricos e metodológicos que a cada momento são revisados, reformulados, e novas propostas e diferentes campos de atuação são criadas no âmbito desta disciplina.

Desde a promulgação da Teoria Geral da Terminologia (TGT), proposta por Wüster nos anos de 1930, novas teorias e vertentes têm surgido, como a Teoria Comunicativa da Terminologia (TCT), a Socioterminologia (ST), a Teoria Sociocognitiva da Terminologia (TSCT), a Terminologia Textual (TT), a Terminologia Cultural (TC) e a Etnoterminologia (ET). Cada uma delas com enfoques diferentes no tratamento terminológico, divergindo em alguns pontos, mas todas, de um modo geral, contrárias aos pressupostos da TGT.

Antes de nos atentarmos especificamente para as discussões sobre a Etnoterminologia, convém apresentar brevemente as características principais dessas teorias e vertentes que contribuíram para o surgimento e desenvolvimento da Terminologia.

De fato, a Terminologia surgiu a partir dos trabalhos do engenheiro austríaco Eugen Wüster (1898-1977), professor da Universidade de Viena que instituiu a chamada Escola de Viena e elaborou a Teoria Geral da Terminologia (TGT). Segundo Barros (2004), Wüster concebia a Terminologia como “uma ciência de caráter filosófico” que mantinha relações privilegiadas com a Lógica, a Teoria da Classificação e a Linguística. Para a autora, esta relação com a Linguística proposta por Wüster é ambígua, “uma vez que se interessava apenas pelos termos, dissociando o léxico da gramática, do contexto e do discurso, vendo-os como unidades que existem e têm vida independente”, não havendo, desse modo, termos polissêmicos, sinônimos ou homônimos (BARROS, 2004, p. 55).

Na teoria wüsteriana há uma valorização do termo em seu aspecto conceitual, sendo, portanto, compreendido não como uma unidade lexical sujeita a interferências do sistema linguístico e das situações comunicativas, mas como um nódulo cognitivo universal e imutável, como uma unidade monorreferencial, invariável e absoluta (KRIEGER; FINATTO, 2004). Isso significa dizer que, nessa proposta, os termos eram considerados apenas como uma representação ontológica de uma área do conhecimento e estudados alijados de seus contextos de uso.

De base normativa e prescritiva (TGT), a Terminologia assume contornos linguístico-descritivos com o surgimento da Teoria Comunicativa da Terminologia (TCT) proposta por Maria Teresa Cabré, no início da década de 1990. Trata-se de uma teoria com princípios distintos aos da TGT que revolucionou os paradigmas epistemológicos em relação ao estudo dos termos e redimensionou a compreensão dos fenômenos terminológicos tornando-se “referência teórica em grande parte das pesquisas terminológicas realizadas no Brasil<sup>12</sup>” (ALMEIDA, 2006, p. 86).

Para Cabré (1999a), a TCT se fundamenta entre três teorias: a) a teoria do conhecimento, que explica como se conceptualiza a realidade, “os tipos de conceptualização e relação dos conceitos entre si e suas possíveis denominações”; b) a teoria da comunicação, que descreve os tipos de situações, “e o tipo de comunicação em toda a sua amplitude e diversidade, [...] as características, possibilidades e limites dos diferentes sistemas de expressão de um conceito e de suas unidades”; c) e a teoria da linguagem, que dá conta das unidades terminológicas propriamente ditas, “que fazem parte da língua natural e participam de suas características, mas singularizando sua especificidade significativa e explicando como se atua na comunicação” (CABRÉ, 1999a, p. 122,123).

Percebe-se que a TCT é uma abordagem de caráter linguístico-descritivo que visualiza o termo além da sua dimensão cognitiva, ou seja, analisa e descreve os aspectos semânticos, gramaticais, textuais e pragmáticos das unidades terminológicas levando em conta os seus contextos reais de ocorrência. De modo contrário à teoria tradicional, reconhece-se a existência de variação e de sinonímia dos termos e se aceita que “o conteúdo de um termo nunca é absoluto, mas relativo, segundo cada âmbito e situação de uso”, e também “que os termos não pertencem a um âmbito, mas são usados em um âmbito com um valor singularmente específico” (CABRÉ, 1999b, p. 132-133). Ainda segundo Cabré (1999b), não há termos nem palavras, mas apenas unidades lexicais que vão se tornar termo ou palavra a partir do contexto em que estão inseridos.

Com contornos ainda mais expressivos a Terminologia passa por uma nova fase a partir da formalização da Socioterminologia, instituída na década de 1990 por François Gaudin com a colaboração de Y. Gambier, e J.-C. Boulanger. Conforme Gaudin (2014), a Socioterminologia propõe o estudo do termo em uma situação de interação social, ou seja,

---

<sup>12</sup> “Essa grande influência da TCT no Brasil não é sem razão, haja vista que uma teoria descritiva de base linguística parece ser muito mais adequada ao contexto brasileiro: país monolíngüe com grande variedade dialetal. Só mesmo uma teoria descritiva para dar conta das especificidades das terminologias aqui praticadas” (ALMEIDA, 2006, p. 86).

procura analisar o termo levando em conta as situações discursivas e sociais concretas. Para o autor, “a circulação dos termos é projetada sob o ângulo da diversidade dos usos sociais, o que engloba o estudo das condições de circulação e de apropriação dos termos, considerados como signos linguísticos, e não como etiquetas de conceitos” (GAUDIN, 2014, p. 304).

Nesse sentido, a Socioterminologia privilegia o caráter linguístico e social do termo, situando a “comunicação especializada em seu lugar social” (BARROS, 2004, 69). Ela trata da diversidade de usos sociais dos termos e reconhece a existência da variação linguística nos discursos especializados, bem como da sinonímia, homonímia e polissemia. “A Socioterminologia ocupa-se, ainda, de estudos acerca dos processos de banalização da linguagem” e considera que as línguas de especialidade, assim como os conjuntos terminológicos a elas pertencentes, “devem ser analisados em sua completude, incluindo-se aí todos os aspectos sociais que os envolvem e as variações que deles decorrem” (BARROS, 2004, p.70).

Com o avançar dos estudos em Terminologia surge a Teoria Sociocognitiva da Terminologia, proposta por Temmerman (1997), ancorada nos pressupostos da Semântica Cognitiva. A autora defende a ideia de que os termos são unidades de compreensão, são produtos da cognição humana que funcionam como modelos cognitivos (ao invés de conceitos padronizados) estruturalmente organizados na mente a partir de diferentes conjuntos de informações (provenientes do mundo interior e exterior). Em outras palavras, “os termos são motores no processo de compreensão, visto que eles conectam novas compreensões a compreensões anteriores” (TEMMERMAN, 2000, p. 228), propiciando a formação de conceitos e denominações. Temmerman (2000) também valoriza o papel da metáfora nesse processo de construção e constituição das unidades terminológicas, assim como a sinonímia e polissemia como elementos funcionais nessa compreensão.

De acordo com Barros (2004, p. 56), as últimas décadas foram marcadas por intensa produção no campo da Terminologia, “sobretudo no que concerne à descrição de conjuntos terminológicos, elaboração de obras terminográficas, desenvolvimento de tecnologia adequada ao tratamento computacional de terminologias e ao planejamento linguístico”. Para a autora, essa intensa atividade promoveu reflexões sobre o ser e o fazer da Terminologia, conduzindo a novas observações. É nesse bojo que surgem a Terminologia Textual, Terminologia Cultural e a Etnoterminologia.

Com propósitos mais adequados ao estudo da Terminologia, essas vertentes congregam aspectos já vistos nas teorias anteriores sobre o caráter linguístico, cognitivo, comunicativo e



social do termo, mas de forma mais contundente, aprofundada e ampliada, “levando a Terminologia a contar com um corpo teórico, com maior poder descritivo e explicativo para dar conta da complexidade envolvida na constituição e funcionamento das terminologias” (KRIEGER, 2011, p. 447).

O objeto de estudo da Terminologia, prioritariamente o termo, tem sido observado sob diferentes ângulos, analisado sob vários pontos de vista e tratado de maneira diversa conforme a perspectiva teórica adotada. No entanto, com o avançar das pesquisas, têm-se considerado também a “fraseologia especializada, a definição terminológica, e ainda, o texto especializado como objetos de estudo da Terminologia” (KRIEGER, 2011, p. 448). Isto implica dizer que os estudos terminológicos têm avançado significativamente nos últimos anos, abrindo caminhos para novos olhares sobre o componente lexical especializado.

É nessa perspectiva que a Terminologia Textual (TT) nasce para enfatizar que “os termos estão nos textos e a terminologia torna-se um estudo textual” (CLAS, 2004, p. 238). Portanto, o texto especializado, consoante Finatto (2011), tornou-se um dos objetos centrais de estudo da Terminologia, uma vez que é nele que se abrigam e se constituem funcionalmente as unidades lexicais, ou seja, é através dele que se pode determinar o estatuto dessas unidades, se vocábulos ou termos, ou ainda se vocábulos-termos (unidade mínima de significação da Etonoterminologia).

Dessa forma, têm se considerado cada vez mais o estudo *in vivo* das unidades lexicais, ou seja, o estudo dos termos no lugar em que ocorrem - “nos textos especializados, sobretudo, porque eles são o *habitat* natural das terminologias” (KRIEGER, 2011, p. 447). De acordo com Krieger e Finatto (2004), o termo deve ser visto como um componente ativo da tessitura do texto, como uma unidade léxico-semântica portadora de semas temáticos e de conteúdo que se se particulariza em um texto dependendo do uso e da situação discursiva e se constrói socialmente e culturalmente de forma múltipla e dinâmica.

Observando esse caráter sociocultural dos termos e a evidente necessidade de uma proposta que abordasse de maneira mais profícua esse aspecto, Diki-Kidiri (2009) propõe a Terminologia Cultural (TC) como alternativa para essa necessidade eminente. Como se pode deduzir a partir do próprio nome, o cerne do enfoque terminológico está na cultura, concebida por Diki-Kidiri (2009, s/p) como “o conjunto das experiências vividas, das produções realizadas e dos conhecimentos gerados por uma comunidade humana que vive em um mesmo espaço e em uma mesma época”. Tal enfoque se deve porque é através da cultura que entendemos a linguagem e por ela é construída. Não existe cultura sem a linguagem,

tampouco a linguagem sem a cultura, isto é, sem o povo que a usa. Portanto, é por meio da língua que os grupos sociais representam o mundo e expressam suas ideias e experiências, é através dela que perpassa todo o conhecimento adquirido socialmente que, por sua vez, é repassado de geração em geração. Nas palavras de Diki-Kidiri (2009, s/p), “A visão de mundo de uma cultura determina a sua forma de classificar, ordenar, nomear e categorizar tudo o que os seus membros percebem e concebem”.

Entendendo a cultura como um processo cumulativo de conhecimentos e práticas que os indivíduos de uma comunidade compartilham com seus pares por meio de unidades lexicais elaboradas e memorizadas ao longo de suas vidas, podemos compreender claramente a origem dos termos, sua formação e funcionamento segundo a proposta de Diki-Kidiri (2009). Para o autor, o termo é um produto das linguagens culturalmente integradas e, como tal, seu funcionamento depende das representações que se constituíram no percurso de evolução histórica da cultura (DIKI-KIDIRI, 2009). Assim, os termos se constituem não por uma propriedade que lhes é intrínseca, mas mediante os interesses e necessidades de uma comunidade linguística que definem e redefinem os seus significados a um conhecimento especializado. Isto implica dizer que estatuto das unidades lexicais é definido em um contexto cultural e não pelo pensamento universalizante especializado, ou seja, os vocábulos só assumem valor de termo no seio de uma cultura específica.

Tanto as proposições da TT como as da TC estão no bojo dos pressupostos da Etnoterminologia, uma subárea da Terminologia que busca estudar as unidades lexicais dentro do contexto em que estão inseridas, quer em um contexto imediato (no texto), quer e em um contexto sociocultural. É nessa vertente que fundamentamos o nosso trabalho e que agora convém discutir.

A Etnoterminologia é uma disciplina científica instituída por Maria Aparecida Barbosa para designar “o estudo das unidades multifuncionais dos discursos etnoliterários” (BARBOSA, 2005, p. 106). “Ela estuda os discursos etnoliterários, como os de literatura oral, literatura popular, literatura de cordel, fábulas, lendas, mitos, folclore e também os discursos das linguagens especiais com baixo grau de tecnicidade e de cientificidade” (BARBOSA, 2009a, p. 1).

Examinando o universo lexical dos discursos etnoliterários, Barbosa (2005) verificou que as unidades lexicais contidas nos textos possuíam um estatuto diferente, ou seja, elas possuíam qualidades das linguagens de especialidade e qualidades da linguagem literária, adquirindo, desse modo, o estatuto de vocábulo-termo. “É um vocábulo, nos seus aspectos

referenciais, pragmáticos e simbólicos, e é um termo, na medida em que a unidade léxica em questão tem características de uma linguagem de especialidade” (BARBOSA, 2005, p. 105). Foi observando essa multifuncionalidade das palavras nos discursos etnoliterários e a evidente necessidade de uma disciplina que estudasse essas unidades lexicais que Barbosa (2005) propôs a Etnoterminologia, disciplina científica que tem o vocábulo-termo como sua unidade mínima de significação.

Nas palavras de Barbosa (2006a):

Essas unidades lexicais têm sememas muito especializados, construídos com semas específicos do universo de discurso em causa, provenientes das narrativas, cristalizados, de modo a tornar-se verdadeiros símbolos dos temas envolvidos. Consta-se então, que sustentam o pensamento e o sistema de valores da cultura que configuram uma axiologia. Assim, as unidades lexicais do universo de discurso etno-literário têm um estatuto próprio e exclusivo. Nos níveis da norma e da fala, subsumem duas funções, vocábulo e termo (BARBOSA, 2006a, p.48).

Nessas condições, percebe-se que as unidades lexicais do discurso etnoliterário têm características específicas, têm significados particulares, próprios do universo de discurso a que pertencem. Elas representam não somente um contexto linguístico, mas um contexto sociocultural - a axiologia de um grupo, de uma sociedade, ou seja, os valores, os costumes, os usos, as crenças, os hábitos de um grupo de falantes linguística, histórico e socialmente definido. Por isso, elas adquirem um estatuto diferente, correspondente a uma norma discursiva e a um texto-ocorrência.

Em outras palavras, a Etnoterminologia estuda a unidade léxica (vocábulo-termo), “que representa um grupo de falantes que detém os valores conceptuais próprios de uma realidade” que são transformados em signos e incorporados a um sistema de significação, condicionando o processo de denominação (LATORRE, 2013, p. 75). Desse modo, “para usar o vocábulo-termo é preciso conhecer a axiologia do grupo em que teve origem, do grupo que conceptualizou seus signos-símbolos” (LATORRE, 2013, p.76).

Nessa perspectiva, “a Etnoterminologia está intimamente associada ao sentido de etnia e etnismo na formação social e cultural de um grupo, e às interferências históricas e geográficas que subordinam o processo de conceptualização dos seus sujeitos” (LATORRE, 2013, p. 76). A natureza “especializada” do vocábulo-termo aflora vínculos de pertencimento ao grupo, características de etnicidade, o que o torna um documento autêntico de identidade, tal como concebe Pais e Barbosa (2004, p. 84): essas unidades “incorporam, sustentam,

caracterizam uma identidade cultural; representam um saber compartilhado sobre o mundo, traduzido em amplas sucessões de metáforas”. Isto porque:

Enquanto expressão cultural, a língua é a mais importante forma de representação de um contexto, e com os grupos evolui, simbolizando suas realidades. É o elo histórico do passado e presente que une todo homem às suas tradições remotas e às suas origens. Desse modo, as diferentes formas de conceber a realidade, que linguisticamente sustentam as manifestações socioculturais, identificam cada grupo humano, afastando-os em especificidades profundas e aproximando-os em identidades não menos importantes, registrando perenemente os valores, crenças e práticas no tempo e espaço em que se produzem o saber e o saber-fazer (LATORRE, 2013, p. 65).

Assim, a Etnoterminologia estuda o vocábulo-termo em uma perspectiva *in vivo*, isto é, estuda os aspectos linguísticos desse componente considerando não somente o contexto linguístico de uso, mas, também, o contexto sociocultural em que está inserido. Por isso, a função mítica, a memória social, o conhecimento compartilhado sobre o mundo e a natureza humana, os sistema de valores dos grupos humanos e o sistema de crenças são também aspectos tratados por essa disciplina.

Diante disso, podemos afirmar que, além de considerar as proposições da TT e da TC, a Etnoterminologia também mantém conexões com a TCT, ST, TSCT para dar conta da complexidade envolvida na constituição e funcionamento desse novo componente lexical - o vocábulo-termo. Além desses suportes, esta disciplina também conta com um “número crescente de produtos terminológicos resultantes de programas informatizados, revelando uma interface produtiva da Terminologia com a Linguística de *Corpus*” (KRIEGER, 2011, p. 449).

Vale salientar que com o advento da Linguística de *corpus* as pesquisas terminológicas têm avançado significativamente. Hoje, no Brasil, contamos com inúmeros trabalhos que têm se valido de dados de *corpora* para focalizar questões terminológicas por meio de ferramentas computacionais. Com isso, questões que antes não eram vistas ou que não poderiam ser analisadas manualmente passaram a ser percebidas e estudadas no âmbito das várias áreas da Terminologia, como é o caso das unidades multifuncionais dos discursos etnoliterários, objeto de estudo da Etnoterminologia.

Paralelo a esse avanço nos estudos terminológicos, a Etnoterminologia vem ganhando cada vez mais destaque. Resultado disso são os vários trabalhos que têm surgido nesse campo com temáticas e abordagens diferentes. A título de exemplificação, podemos citar: *As denominações no léxico de Grande Sertão: Veredas sob uma perspectiva Etnoterminológica* (2010), *Uma abordagem etnoterminológica de Grande Sertão: Veredas* (2011), *A dialética*

entre os extremos: da Terminologia à Etnoterminologia (2013), *A Etnoterminologia no âmbito dos estudos da tradução* (2016) de Vanice Ribeiro Dias Latorre; *A Socioterminologia e Etnoterminologia das plantas medicinais do Nordeste* de Maria do Socorro Silva de Aragão (2010); *A etno-terminologia da culinária baiana na obra Dona Flor e seus dois maridos: análise dos aspectos do discurso etno-literário na versão para o inglês* de Manoel Messias Alves da Silva e Jonathas de Paula Chaguri (2010); *A transcodificação de textos científicos em textos etnoliterários, o cordel: o desenvolvimento da cognição com reflexão crítica* de Albelita Lourdes Monteiro Cardoso (2011); *O dicionário do folclore brasileiro: um estudo de caso da etnoterminologia e tradução etnográfica* de Flávia Medeiros de Carvalho (2013); *Estudo etnoterminológico preliminar do sistema de cura e cuidados do povo Mundurukú (Tupi)* de Nathalia Martins Peres Costa (2013); *A Etnoterminologia da língua Mundurukú-Tupi e as contribuições da Ecolinguística* de Nathalia Martins Peres Costa e Dione M. Gomes (2013); *O conhecimento etnobotânico dos Kalunga: uma relação entre língua e meio ambiente* de Gilberto Paulino de Araújo (2014); *Discurso literário de fantasia infantojuvenil: proposta de descrição terminológica direcionada por corpus* de Raphael Marco Oliveira Carneiro (2016).

A Etnoterminologia vem conquistando terreno pouco a pouco ao lado da Terminologia Aplicada. O seu reconhecimento, ainda que a passos lentos, se justifica pelo fato de congregar aspectos já observados nas teorias e vertentes anteriores, sintetizando questões terminológicas oriundas de décadas de trabalho sobre a unidade lexical especializada; e também pelo fato de suprir uma necessidade teórica e epistemológica, instrumentalizando os especialistas na matéria com uma proposta mais adequada ao tratamento terminológico (especificamente, o estudo do vocábulo-termo). Além disso, “esta disciplina também conta com os modelos teóricos da Lexicologia, da Semiótica e da Semântica cognitiva” (BARBOSA, 2009a, p. 1)

A explicação dessa cooperação mútua entre as ciências é dada por Barbosa (2006b) da seguinte maneira:

[...] Toda ciência ou tecnologia, seja, do ponto de vista epistemológico, seja do metodológico, seja, ainda, daquele da construção do seu saber metalingüístico, estabelece estreitas relações de cooperação - interdisciplinares, no nível das ciências básicas, ou no nível das ciências aplicadas, e de alimentação e realimentação entre estas e aquelas -, com outras ciências básicas, ciências aplicadas e/ou tecnologias. Esse processo de contribuição recíproca, entre tais disciplinas, não lhes retira, contudo, a especificidade do objeto de estudo, campo, métodos e técnicas e, até mesmo, de modelos e de metalinguagem. De fato, sustentando-se todas nesse relacionamento complexo e dinâmico de interdisciplinaridade, multidisciplinaridade, alimentação e realimentação, intra e inter-áreas do

conhecimento humano, perseguem, efetivamente, objetivos comuns: a busca da verdade, a análise e descrição do seu objeto, a redução dos fatos a modelos, a construção do saber, o aprimoramento da qualidade de vida, a construção de um discurso metalingüístico específico (BARBOSA, 2006b, p. 1).

Os conhecimentos resultantes dessas relações interdisciplinares e multidisciplinares dão sustentação teórica a trabalhos de diversas ciências aplicadas. No caso da Terminologia, Barros (2004, p. 34, 35) afirma que “os estudos terminológicos fornecem as bases teóricas e metodológicas para inúmeras pesquisas, tendo aplicações no ensino das línguas (materna e estrangeira), na tradução, na elaboração de obras terminográficas (dicionários especializados)”, e também no ensino de disciplinas técnicas e científicas, na documentação, “no jornalismo científico, nas ciências sociais, na transferência do saber técnico e científico, na produção industrial e nas políticas linguísticas”.

Sabemos que existem determinadas áreas do conhecimento que prevalece o saber exclusivo, mas no caso Terminologia, em especial da Etnoterminologia os dois saberes, o compartilhado e o exclusivo coexistem e articulam-se necessariamente. Apesar da importância dessas correlações e de todo esse desenvolvimento nos estudos terminológicos, ainda há o que ser estudado e analisado, isto porque a “identificação terminológica é sempre complexa, considerando que o termo técnico-científico tende a ser multivocabular em detrimento de sua constituição por meio de um único item lexical” (KRIEGER, 2011, p. 450). Por outro lado, “os termos apresentam problemas de graus de lexicalização para alcançar sua realização sintagmática”, e essa ordem de problemas ainda não é equacionada satisfatoriamente pelos programas computacionais. Dessa maneira, “compreender o fenômeno terminológico, hoje, significa compreender também outros elementos que integram e corroboram o estabelecimento e a divulgação dos saberes especializados” (KRIEGER, 2011, p. 450).

Diante do exposto, podemos dizer que o percurso da Terminologia ainda está sendo escrito, pois se trata de um campo de interfaces e com diferentes subáreas, cada qual com focos distintos decorrentes de interesses e visões de linguistas. Então, dizer que área está consolidada é contraditório, senão um tanto imaturo, pois o léxico é dinâmico e multifacetado, e como a Etnoterminologia, vão surgindo novas vertentes e teorias à medida que se percebem novos aspectos associados ao fenômeno terminológico.

### 3.1.1 O vocábulo-termo: a unidade mínima de significação da Etnoterminologia

Como vimos, o vocábulo-termo são unidades multifuncionais presentes nos discursos etnoliterários que congregam aspectos tanto das linguagens de especialidade como da linguagem literária. Por essa razão, ele apresenta características peculiares: no nível da norma, exerce duas funções: vocábulo e termo, e possui significados específicos segundo o universo de discurso a que pertencem; no nível de sistema essa unidade é plurifuncional. Isto implica dizer que o estabelecimento preciso de sua função depende de sua inserção em uma norma discursiva e em um universo de discurso, ou seja, é a norma e o universo de discurso em que se insere que determina o seu estatuto de vocábulo-termo.

Sobre esse caráter monofuncional e plurifuncional do vocábulo-termo, Barbosa (2006) afirma que:

[...] Uma unidade lexical não é termo ou vocábulo, em si mesma, mas, ao contrário, está em função “termo” ou em função “vocábulo”, ou seja, o universo de discurso em que se insere determina o seu estatuto, em cada caso. Assim não é possível estabelecer uma taxionomia paradigmática dos conjuntos termos e dos conjuntos vocábulos, pois toda a classificação resulta dos entornos discursivos e dos condicionamentos das normas discursivas, dependente, portanto, dos universos de discurso e das situações de discurso. Concebe-se um percurso possível de uma “unidade lexical”, ao longo de um eixo *continuum*, do mais alto grau de banalização ao mais alto grau de cientificidade e vice-versa. Em suma, toda unidade lexical é plurifuncional, no nível de sistema, e monofuncional, no nível de uma norma ou do falar concreto (BARBOSA, 2006a, p. 50).

Vale lembrar que o termo é uma unidade lexical (ou unidade terminológica) com conteúdo específico dentro de um domínio específico, também entendido como uma “unidade padrão das línguas de especialidade”; e o conjunto de termos de uma área especializada chama-se conjunto terminológico ou terminologia (BARROS, 2004). Por outro lado, o vocábulo é um “modelo de realização das palavras que o representam no texto”, ou seja, um modelo de realização lexical no texto. Isto significa dizer que a palavra é “uma unidade do texto” e o vocábulo é uma “unidade do léxico; e o conjunto de vocábulos de um texto é chamado de conjunto vocabular” (BARROS, 2004, p. 41).

Compreendendo a palavra como uma unidade léxica (ou unidade lexical), - “um signo linguístico, composto de expressão e de conteúdo, que pertence a uma das grandes classes gramaticais (substantivo, verbo, adjetivo ou advérbio)”, podemos afirmar que “o termo e o vocábulo também são palavras ou unidades lexicais” (BARROS, 2004, p. 40, 41). Segundo Cabré (1999b, p. 123), o termo é uma palavra “ativada singularmente por suas condições

pragmáticas de adequação a um tipo de comunicação” e, no âmbito da análise quantitativa de um texto, é também um vocábulo, “uma vez que é um modelo de realização lexical no texto”. Seu caráter de termo se dá quando “designa um conceito específico de um domínio especializado”, ou seja, quando é “definido e empregado em textos de especialidade” (BARROS, 2004, p. 41,42).

Logo, o vocábulo-termo é uma unidade lexical composta de expressão e conteúdo que designa um conceito específico dentro do universo de discurso etnoliterário. Em outras palavras, é um vocábulo “metasemiótico”, um quase “termo-técnico” que pertence a uma linguagem especial/especializada (BARBOSA, 2005). É também um componente lexical semanticamente representativo de uma área temática que se constitui a partir das dimensões cognitiva, linguística, comunicativa e sociocultural. Por outro prisma, é também considerado como um “documento autêntico de identidade cultural”; um “patrimônio cultural” que preserva um valor semântico social e se configura como um “documento do processo histórico da cultura” (PAIS; BARBOSA, 2004, p. 84).

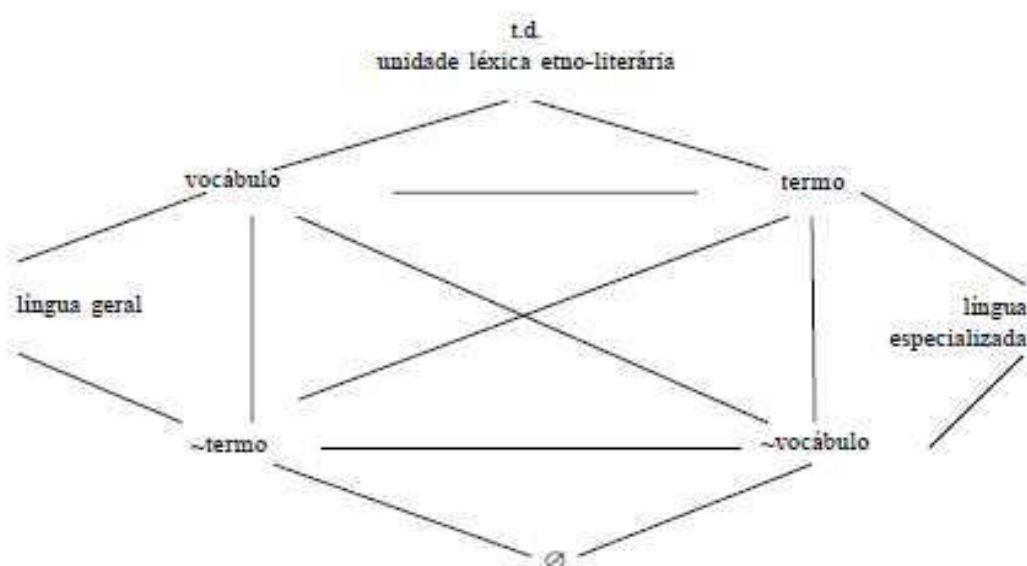
Então, se considerarmos dois universos de discurso, o da língua comum e o das linguagens de especialidade ou especializadas, temos, no primeiro, os vocábulos e, no segundo, os termos. Mas se considerarmos o universo dos discursos etnoliterários, temos a presença da linguagem literária (os vocábulos) e das linguagens de especialidade (os termos) e, ao mesmo tempo características de discursos ficcionais e de discursos documentais, conforme explica Barbosa (2009b):

De certo ângulo, esses discursos etnoliterários poderiam ser considerados ficcionais, na medida em que os “eventos” narrados são ou parecem ser inverossímeis, se tomados denotativa-mente, e não correspondem a fatos historicamente comprovados. Aproximar-se-iam, então, da fábula. De outro ângulo, porém, esses discursos, como vimos, revelam e sustentam sistemas de valores, sistemas de crenças, um “saber” compartilhado sobre o “mundo”, que integram o imaginário coletivo de uma cultura, de uma sociedade [...]. Documentais e ficcionais são termos de metalinguagem, devem ser lidos aqui como duas tendências contrárias. Os discursos sociais não literários têm um estatuto sociosemiótico, conferido pela sociedade que os caracteriza como documentais x não ficcionais, de acordo com o seu modo de existência e produção socialmente aceito, de forma que constituem a dêixis positiva do modelo. Os discursos literários *stricto sensu* são vistos pela sociedade como aqueles que tendem a ser combinação de ficcionais x não documentais, o que lhes dá a posição da dêixis negativa, no mesmo modelo. Nessa perspectiva, **os discursos etnoliterários sustentam-se numa tensão dialética entre os dois termos, documentais x ficcionais**, por todas as razões acima expostas. Confirma-se, uma vez mais, a sua função mítica e a sua função pedagógica. (BARBOSA, 2009b, p. 41, grifo nosso).



Notadamente, com características muito específicas, esse tipo de discurso confere às suas unidades lexicais dupla funcionalidade: a função vocábulo e ao mesmo tempo a função de termo. Segundo Latorre (2011, p. 68), o vocábulo-termo é uma unidade de conhecimento com forte marca de conceptualização da realidade fenomênica composta de duas faces distintas e indissociáveis: “a face vocábulo, uma das suas funções, que pertence à língua geral e porta elementos paradigmáticos, referenciais e simbólicos”, e sua face termo, das linguagens terminológicas, sua outra função, “que porta verdades gerais e universais, registros que são do conhecimento acumulado sobre o mundo, sobre a natureza humana, e a axiologia de um povo”. Instaura-se assim, em semiótica profunda, uma tensão dialética vocábulo x termo, como Barbosa (2005) mostra na figura a seguir:

FIGURA 1- Tensão dialética vocábulo x termo



Fonte: Barbosa (2005, p. 105)

Vale lembrar que são os entornos discursivos e o condicionamento das normas discursivas que determinam a funcionalidade das unidades lexicais. São eles que, conforme Barbosa (2009a), situam os vocábulos-termos no mais elevado patamar de densidade terminológica ou em baixo grau de especialidade.

Nessa dinâmica de processos que constituem o vocábulo-termo ainda convém mencionar que essa unidade “resulta do cruzamento de um processo de metaterminologização e de metavocabularização” (BARBOSA, 2006a, p. 48), o que significa dizer que esses movimentos (entre vocábulo e termo) podem tanto ser horizontais - de um universo de

discurso para outro, quanto verticais - quando ocorre a passagem do nível conceptual para o nível terminológico. Para explicar melhor esses fenômenos, elaboramos um quadro com a tipologia dos processos de constituição de conjuntos terminológicos e vocabulares proposta por Barbosa (2005, 2006a):

QUADRO 3- Tipologia de processos de constituição de conjuntos terminológicos e vocabulares

TIPOLOGIA DE PROCESSOS DE CONSTITUIÇÃO DE CONJUNTOS TERMINOLÓGICOS E VOCABULARES			
Vocabularização	Terminologização <i>stricto sensu</i>	Terminologização <i>lato sensu</i>	Metaterminologização
Passagem da terminologia para a língua comum, ou seja, é a transformação do termo em vocábulo. (ex: “entrar em órbita” – transposto da área técnico-científica para a língua geral, por um processo de metaforização).	Passagem da língua comum para a terminologia. Refere-se à transposição de uma unidade lexical da língua geral para uma linguagem de especialidade, ou seja, a transformação do vocábulo em termo. (ex: “peregrinismo”, na língua comum, significava “ir em romaria” e, nas ciências da linguagem, passou a significar “emprego de vocábulo estranho à língua vernácula, estrangeirismo”).	Passagem do nível conceptual para o nível terminológico. Trata-se de uma criação <i>ex-nihilo</i> , que tem graus diferentes de motivação mas que não resulta da transposição de um universo de discurso para o outro e, sim, da instauração de um nova grandeza signica – numa combinatória inédita, no caso do processo fonológico e sintagmático – e de uma função metassemiótica – no caso do processo semântico. Diferentemente dos outros casos, a relação aqui é vertical. (Ex.: lembremos do nome da marca “Omo”, como designativo de sabão em pó)	1. Passagem da terminologia para a terminologia com a manutenção de um núcleo sêmico comum aos termos das diferentes áreas (ex.: “estrutura” e “função”, em diferentes áreas). 2. Passagem da terminologia para a terminologia sem a manutenção de núcleo sêmico. (ex.: arroba – “medida de peso” e “arroba”- @, como símbolo de endereço eletrônico).

Fonte: Elaboração da autora com base em Barbosa (2005, 2006a)

Essa tipologia de processos demonstra que os conjuntos terminológicos e vocabulares podem ser formados por diferentes processos tanto em nível horizontal como vertical. Daí compreende os percursos realizados pelo vocábulo-termo e o modo como é engendrado.

De maneira sucinta, é preciso conhecer a natureza e o funcionamento do vocábulo-termo para interpretá-lo corretamente nas reais situações discursivas. É incoerente, senão

inadequado, apresentar uma análise lexical subsidiada pela Etnoterminologia sem compreender essas questões.

Propor uma análise lexical em uma perspectiva etnoterminológica requer, portanto, o conhecimento e a compreensão do vocábulo-termo em seus diferentes aspectos: linguístico, cognitivo, comunicativo e sociocultural, inserido em um universo de discurso e circunscrito em uma norma discursiva. Assim é possível realizar uma análise coerente e válida das unidades lexicais de modo a evidenciar as suas reais funções nos discursos etnoliterários, o que elas significam e o que verdadeiramente representam (como documentos de identidade cultural de um povo).

### **3.1.2 A formação do conceito nos discursos etnoliterários**

A construção do conceito assume características semânticas, sintáticas, semióticas, pragmáticas diversas dependendo do universo de discurso e de sua ocorrência quer nas linguagens de especialidade, nos discursos literários, etnoliterários ou em outros discursos sociais não-literários. Isto significa dizer que em cada universo de discurso o modo de engendramento de um conceito tem aspectos específicos ao lado de outras marcas que caracterizam o discurso.

De acordo com Barbosa (2004), é na instância discursiva que se produz a cognição e a semiose, ou seja, é nela e através dela que se instaura a conceptualização de um fato e se engendra um conceito. Nas palavras da autora, “é no discurso manifestado que se presentificam os traços conceptuais, num procedimento de codificação; e é dele que se extraem, num procedimento de investigação, esses mesmos traços” (BARBOSA, 2004, p. 56).

Esse processo de construção de um conceito, o *modus operandi* conceptual, pode ser segundo Barbosa (2004), “vertical” – do “fato” para o patamar cognitivo –, ou “pode ser desencadeado nas relações sintagmáticas de um discurso manifestado, em que o autor vai pouco a pouco construindo, no seu texto, um conceito qualquer” (BARBOSA, 2004, p. 79). O primeiro, chamado de onomasiológico, é caracterizado pela autora da seguinte forma:

Observe-se que os atos de conceituar, ou de engendrar um discurso manifestado qualquer – são processos *onomasiológicos* – tomam como ponto de partida o *continuum* amorfo dos dados da experiência, passam pelo nível noêmico e chegam ao nível lexemático, que vai do fato ao nome, e cujo produto é a *denominação*. É o percurso do fazer persuasivo do *sujeito de enunciação de codificação*, desencadeado por quem fala, quem escreve. Esse sujeito de enunciação de codificação, tendo uma intenção de comunicação de determinado esquema lógico conceptual, pode selecionar diferentes formas

lingüísticas, suscetíveis de representá-lo, para engendrar o seu discurso enfim manifestado. Essa escolha integra o processo de modalização do discurso, enquanto competência e desempenho do sujeito enunciador. Desse percurso resultam: conceitos, seus representantes semiotizados – grandezas-signos – e, em etapas posterior, presentificados em discursos manifestados (BARBOSA, 2004, p. 60).

O segundo, chamado de semasiológico, é assim explicado por Barbosa (2004):

De outro ângulo, tem-se o percurso que toma como ponto de partida o discurso manifestado, para chegar novamente ao nível conceptual, que caracteriza o *fazer interpretativo* do *sujeito enunciatário*, ou, noutras palavras, um processo *semasiológico*, do signo para o *conceito*, realizado por quem ouve ou quem lê; qualifica-se, assim também, o percurso lexicográfico-terminográfico, como processo que parte da manifestação do nível lexemático, com as seleções, restrições e combinatórias sêmicas estabelecidas em discurso, para, num *metadiscurso* igualmente configurado como *fazer interpretativo*, articular semas representados por *metatermos lexemáticos*, operação de que resulta a *definição* (BARBOSA, 2004, p. 79, grifo do autor).

Daí resultam interpretações sobre as diferenças entre conceituar e definir. Baseando-nos na compreensão de Barbosa (2004) podemos dizer que conceituar é o processo de construção de um modelo mental a partir da percepção dos fatos naturais e/ou culturais que, por sua vez, desencadeia em escolhas lexicais para a transmissão da mensagem ou do discurso. Por outro lado, definir é o processo de interpretação do semema linguístico a partir da estrutura léxica manifestada.

Então, podemos inferir que o conceito estabelece-se num nível pré-linguístico, enquanto a definição no nível semiótico, pois resulta de uma análise e descrição de grandezas sígnicas para a reconstrução do modelo mental. De modo sucinto, o “conceito é o resultado de uma interpretação de fatos naturais e/ou culturais”, e a “definição é o resultado de uma interpretação de unidades lexicais” (BARBOSA, 2004, p. 59).

Considerando essas reflexões, Barbosa (2004) ressalta, com base em Pottier (1981), que, para realizar um estudo das unidades lexicais dos discursos etnoliterários, é necessário, pelo menos, quatro níveis de análise: a) nível referencial, “do mundo real ou imaginário (objetos, sonhos, memória do discurso de outrem, foto...)”; b) nível conceptual, “de apreensão mental, da representação construída a partir do referencial, ao mesmo tempo tributária dos hábitos sociais e das necessidades criativas individuais”; c) nível da língua natural, “da competência linguística em seu conjunto”; d) nível do texto produzido, “resultado de diversos componentes (linguístico, cognitivo, contextual, situacional, intencional), em correlação

possível com outros sistemas semióticos (gestos, proxêmica, ilustrações...)” (BARBOSA, 2004, p. 58, 59).

Notadamente, ao criar o próprio campo referencial, a obra literária pode construir conceitos de formas diversas: a partir de referências externas do mundo real; a partir de referências internas (imaginação) construídas pelo próprio autor no âmbito da ficção; a partir de referências exclusivas do universo de discurso a que pertence; e/ou a partir de outros textos ou outros discursos (relações intertextuais e interdiscursivas). Por essa razão, é importante adotar critérios e níveis de análise (com maior rigor e precisão metodológica) ao estudar os itens lexicais dos discursos literários ou etnoliterários.

Sobre isso, Barbosa (2004) afirma que:

No discurso literário, uma obra pode ser auto-suficiente, no engendramento de um conceito, numa intertextualidade intra e interdiscursiva. No discurso técnico-científico, teórico e/ou prático, assim como no discurso literário, o engendramento do conceito é sintagmático, narrativo, transfrástico; no discurso terminológico, é eminentemente paradigmático, como processo e produto final, embora resulte de extrações de contextos de natureza transfrástica (BARBOSA, 2004, p. 79).

Podemos depreender disso que o modo de engendramento de um conceito nunca é uniforme ou simétrico, dado os inúmeros fatores que constituem e caracterizam um discurso manifestado. A cada ato de fala há uma reorganização de fenômenos linguísticos e conceptuais por parte do sujeito enunciador e, de outro lado, também pelo enunciatário, que é determinada pela situação de discurso e de enunciação, configurando-se num processo de codificação e decodificação. Segundo Barbosa (2004), constitui-se, assim, o caráter instável e dinâmico do nível cognitivo, de forma que a estrutura do conhecimento pode mudar fundamentalmente por duas razões: “seja porque aparece um novo eixo, levando todos os conceitos anteriores a redefinir-se em função dessa nova dimensão”, ou porque surge uma nova forma de ver as coisas, “de sorte que os conceitos pré-existentes não mudam de eixo mas mudam de posição no seu eixo” (BARBOSA, 2004, p. 59, 60).

Vale ainda lembrar que existem três subconjuntos que formam o conceito *latu sensu* (modelo mental/*conceptus*) conforme Barbosa (2004, p. 63): a) o subconjunto de noemas biofísicos ou universais chamado de **conceito *stricto sensu*** ou **arquiconceito**, primeiro nível de formação, resultante das latências (“atributos semânticos possíveis dos ‘objetos’ e ‘processos’ da semiótica natural”) e saliências (“os atributos que se destacam, na estrutura, funcionamento e hierarquia dos ‘fatos naturais’ - o perceber”); b) o subconjunto de traços semânticos conceptuais ideológicos, culturais - o **metaconceito**, segundo nível de formação

resultante das pregnancies “(constituem o resultado da atividade do homem, das escolhas que faz na apreensão daqueles ‘fatos’, -‘o conceber’); c) o subconjunto de traços semânticos conceptuais ideológicos, intencionais, modalizadores - o **metametaconceito**.

Assim, nesse percurso que vai da conceptualização à semiotização, há uma seleção de traços semântico-conceptuais específicos conforme o universo de discurso e as pregnancies do autor (no caso do texto literário). Em outros termos, “cada universo de discurso apreende e reelabora certos traços semântico-conceptuais, deixando outros traços latentes”. Consequentemente, “o conceito vai ser ‘tematizado’ e ‘figurativizado’<sup>13</sup> de acordo com as pregnancies do sujeito enunciador” (BARBOSA, 2004, p. 81).

Nessas condições, podemos dizer que os discursos etnoliterários tendem a dar ênfase ao *metaconceptus* – “subconjunto dos traços semântico-conceptuais culturais, produzindo, simultaneamente, uma modificação do recorte cultural, própria de uma reconstrução particular do mundo semioticamente construído” (BARBOSA, 2004, p. 82). Isto porque mostram uma visão de mundo construída pelo sujeito enunciador (autor) a partir de recortes culturais pré-existentes em um universo específico, figuratizando um conjunto lexical altamente significativo, revelador de uma identidade cultural.

Em suma, no discurso etnoliterário a construção do conceito se dá de maneira peculiar, podendo uma obra ser autossuficiente em uma intertextualidade intradiscursiva e interdiscursiva quanto ao *modus operandi* conceptual. Na verdade, o *conceptus* assume características semântico-conceptuais específicas em função do universo de discurso, das pregnancies do autor e das situações discursivas e, de igual modo, a unidade lexical como portadora desse traço linguístico. Por pertencerem a esse universo de discurso etnoliterário elas apresentam características semânticas, sintáticas, semióticas, pragmáticas exclusivas de um universo cultural semioticamente reinventado.

### 3.2 As definições e suas especificidades

A discussão acerca da definição se justifica pela importância que desempenha na descrição dos vocábulos-termos e, sobretudo, por ser um componente indispensável na construção do vocabulário, objetivo primordial deste trabalho.

---

<sup>13</sup> Entende-se por tema a semiotização do conceito, por tematização, o processo de construção de ideias abstratas e, por figurativização, o processo de corporificação dessas ideias. Daí resultam as isotopias temáticas e as isotopias figurativas, como processos de redundância sêmica (BARBOSA, 2004, p. 80).

Como vimos no item anterior, várias são as etapas no percurso que vai da conceptualização à semiotização. Na verdade, a primeira etapa inicia-se com a percepção dos fatos naturais – “que são substâncias estruturáveis, como informação potencial, para os homens, mas que se convertem em substâncias estruturadas, quando apreendidas pelos grupos lingüísticos e socioculturais”. Esse primeiro momento, o da **percepção**, desencadeia um segundo momento, o de início do processo de **conceptualização** que, por sua vez, compreende três tipos de atributos semânticos: o das **latências** – “em que os fatos observáveis têm os seus traços identificadores em estado potencial, como substâncias de conteúdo”; o das **saliências** – “em que certas características dos fatos se destacam por si mesmas, na semiótica natural”; e o das **pregnâncias** – “em que o sujeito enunciador individual e/ou coletivo seleciona e escolhe os traços que irão configurar o conceito que têm do fato em questão”. O terceiro momento consiste na produção dos modelos mentais (os *conceptus*), momento em que se conclui o processo de conceptualização, e o quarto é o da **lexemização** ou semiotização – “corresponde à conversão do conceito em grandeza-signo, em que se deixa o nível cognitivo, para se passar ao nível semiótico propriamente dito”. Tem-se, ainda, um quinto momento, o da **contextualização** em que há uma escolha, determinada por uma situação de discurso e de enunciação, ou seja, uma seleção possível no sistema linguístico (metassistema lexemático ou terminológico) “que precede e autoriza a atualização da lexia num discurso concretamente realizado” (BARBOSA, 2004, p. 56 - 58).

Importa observar nesse percurso que a configuração do conceito é um processo anterior ao processo da definição, uma vez que o primeiro resulta de uma interpretação dos fatos naturais, enquanto o segundo, de uma interpretação de unidades lexicais. Assim sendo, “o conceito, conjunto ordenado de traços conceptuais, está contido na definição, e nela adequa-se à estrutura sintático-semântica, sua forma de conteúdo e expressão [...]”, de modo que “os traços conceptuais são organizados em forma de frase, ou seja, manifestados como metatermos” (BARBOSA, 2004, p. 76, 77). Logo, o enunciado definitório é um metatermo na medida em que retoma o *definiendum* (termo a ser definido) numa espécie de paráfrase que busca descrever os traços semântico-conceptuais de um conceito.

Dessa maneira, “um conceito pode ser representado, nessa instância de semiotização, por uma ou várias unidades lexicais, respectivamente, campos lexicais unitário ou múltiplo, num mesmo texto ou em textos distintos” (BARBOSA, 2004, p. 83). De outro ângulo, uma unidade lexical pode ser representada por diferentes tipos de definições e pode, também, sofrer alterações sintático-semânticas nos diferentes universos de discurso.

“Enquanto a designação é um modo sucinto de se fazer referência a um conceito, a definição permite separá-lo da extensão e distingui-lo de outros conceitos no domínio” (CARDOSO, 2017, p. 41). Isto implica dizer que não se pode haver vários conceitos em uma mesma definição. A definição descreve unicamente um conceito diferenciando-o de outros conceitos associados. A respeito disso, Sager (1990) destaca que:

1. A definição é necessária para situar o termo em sua posição na estrutura de conhecimento apropriada. Uma vez que esta é uma atividade puramente terminológica, nós chamamos este processo de definição terminológica. Ela pressupõe um entendimento da intenção do termo que é adquirida de definições existentes, de contextos, de consultas a especialistas e por meio de conhecimento da área.
2. A definição é necessária para fixar o significado especializado do termo. Esta é a definição “intencional” usada por especialistas para determinar a referência precisa de um termo. Essa definição será flexível e será menos rigorosa em certas áreas do conhecimento. Variações pequenas em designação e desempenho são geralmente adicionadas à intenção de um termo sem levar a outra redefinição ou redesignação.
3. A definição é necessária para dar aos não-especialistas algum grau de entendimento do termo, e este tipo pode ser chamado de “enciclopédico” (SAGER, 1990 apud CARDOSO, 2017, p. 40).

A definição é, portanto, a responsável pela descrição de um conceito, e é feita através do enunciado definatório – “um tipo de enunciado que expressa um segmento de relações de significação de uma dada área do saber [...]” (KRIEGER; FINATTO, 2004, p. 95). Trata-se de [...] um enunciado-texto que dá conta de significados de termos ou de expressões de uma área do conhecimento (KRIEGER; FINATTO, 2004, p. 93).

Então “a definição é o enunciado que descreve e explica o termo, fazendo parte de um texto maior, ou seja, de uma predicação definicional composta de um sujeito (a entrada) e de um predicado (enunciado definicional, definição)” (BARROS, 2004, p. 163). Elas podem ser de diferentes tipos dependendo da natureza linguística da palavra descrita e do tipo de informação que elas transmitem.

Em relação aos tipos de definições a norma ISO (International Organization for Standardization, 704:2009) propõe os seguintes: extensionais, intencionais, ostensivas, lexicais, estipulativas e de precisão.

As definições extensionais contêm uma lista de conceitos subordinados, que podem ser individuais ou gerais. “Essa lista representa conceitos correspondentes aos objetos que compõem a extensão, não os objetos em si. Com isso, a definição extensional apenas sugere a intensão (ideia que uma expressão linguística referente a um objeto transmite) de um objeto”



(CARDOSO, 2017, p. 42), ou seja, ela descreve o conceito pela enumeração exaustiva dos conceitos aos quais se aplica.

Por outro lado, as definições intencionais são as mais precisas na definição de um conceito, uma vez que oferece a menor quantidade de informação possível para formar a base de abstração, sendo possível reconhecer e diferenciar o conceito de outros conceitos afins. Em outras palavras, “ela define o conceito como uma unidade, com uma intensão despida de ambiguidade, refletida por uma extensão única. A combinação exclusiva de características que criam a intensão é suficiente para identificar o conceito e diferenciá-lo de outros” (CARDOSO, 2017, p. 43)

Já as definições ostensivas, também chamadas de demonstrativas, é uma espécie de definição que exhibe um ou mais objetos representativos na extensão do conceito, ou seja, exhibe representações não lexicais do conceito. Assim, a definição acontece “por meio do apontar o objeto ou pela apresentação de desenhos, figuras, vídeos, sons, animações” (CARDOSO, 2017, p. 42). Este tipo de definição se aplica melhor como complemento às definições intensionais, porque nem sempre deixa claro o objeto particularmente exibido ou o que está sendo referido (ISO, 2009).

De outro ângulo, as definições lexicais são aquelas encontradas em dicionários gerais da língua. Elas descrevem o conceito de forma generalizada e não como um conceito específico de um campo ou área. Por outro lado, as definições estipulativas se prestam a atender necessidades específicas, sendo, portanto, limitadas a situações únicas. Conforme a ISO (2009), é uma definição que resulta da adaptação de uma definição léxica a uma situação única para um propósito específico e que não é uso padrão.

Por fim, as definições de precisão são aplicadas quando se quer dar uma explicação de um termo inerente a um campo de conhecimento específico. Nesse processo, restringe a extensão do conceito, acrescentando a ele características mais precisas. “A definição de precisão tem sua aplicação em definições lexicais e é precedida da indicação do campo de estudo específico a que se refere. Assim, também possui aplicabilidade limitada a hipóteses em que o campo de estudos é mais específico” (CARDOSO, 2017, p. 42).

Barros (2004) propõe alguns outros tipos: as **definições substanciais** – “exprimem a substância do termo definido e se aplicam a quatro categorias (gramaticais), mas sobretudo ao substantivo e ao verbo, que nunca têm outro tipo de definição”; as **definições relacionais** – “exprimem a relação que une o termo definido a uma outra palavra que o qualifica (os advérbios e os adjetivos recebem, em geral, definições relacionais)”; as **definições**

**morfossemânticas** – “apenas reproduzem os elementos da unidade lexical definida, ou seja, os elementos morfológicos constitutivos do termo-entrada”; as **definições nominais** – espécie de definição que “não respeita o princípio da não-circularidade que orienta a não se definir uma unidade lexical ou terminológica por meio de unidades linguísticas parecidas, tão desconhecidas quanto a entrada”; as **definições etimológicas** – “se aproxima formalmente da “definição morfossemântica, mas com enfoque no significado original da palavra”; as **definições por compreensão** – descrevem “o termo por meio de traços distintos (características), seguindo o modelo clássico *gênero próximo + diferenças específicas*”, ou seja, “a definição descreve única e exclusivamente uma dada unidade terminológica e não outra”. Esse tipo de definição é considerada “ideal para a elaboração dos vocabulários técnicos, científicos e especializados”; e as **definições por extensão** – “consistem em enumerar todas as espécies que estão no mesmo nível de abstração ou todos os objetos individuais que pertencem ao conceito definido” (BARROS, 2004, p. 168- 171).

Segundo Krieger e Finatto (2004), existem três tipos essenciais de definição: definições lexicográficas, definições terminológicas e definições enciclopédicas. As definições lexicográficas descrevem as unidades lexicais de uma língua ou discurso; as definições terminológicas tratam de termos especializados de um domínio específico; e as definições enciclopédicas “contêm explicações variadas sobre um dado objeto da realidade” (KRIEGER; FINATTO, 2004, p. 92).

Considerando as diferenças entre uma unidade lexical e uma unidade terminológica na descrição de um conceito, podemos fazer distinções entre as definições terminológicas das definições lexicográficas. A primeira é utilizada pelos dicionários de especialidade, já a segunda é própria dos dicionários de língua geral. No caso do vocábulo-termo não há um tipo de definição específica, já que reúne características tanto da primeira quanto da segunda. Por essa razão, buscamos utilizar um tipo de definição que melhor atendesse aos propósitos deste trabalho e às necessidades do público alvo, a saber: **a definição por compreensão**, esta porque descreve o conceito de forma precisa, sem fazer referências a outros conceitos associados, ou seja, apresenta o conceito exatamente como é e o que significa em um dado universo de discurso.

Como já mencionamos, o discurso etnoliterário é um tipo de discurso com características peculiares, posto que reúne qualidades da linguagem literária e das linguagens de especialidade, bem como apresenta características dos discursos ficcionais e dos discursos documentais. Com efeito, as unidades lexicais dos discursos etnoliterários possuem

características particulares e apresentam conceitos específicos próprios do universo de discurso a que pertencem. De igual modo as definições (apresentadas pelo enunciador-autor) também possuem traços distintos, isto porque (os vocábulos-terminos) mostram uma visão de mundo construída pelo sujeito enunciador (autor) a partir de recortes culturais pré-existentes em um universo específico, figuratizando um conjunto lexical altamente significativo, revelador de uma identidade cultural. Desse modo, ao construir uma definição, é preciso antes de tudo ter um conhecimento prévio desses aspectos, bem como de outros elementos linguísticos e extralinguísticos que caracterizam e acompanham o texto literário, a fim de que se possa exprimir com exatidão os atributos semântico-conceptuais não dando margem a ambiguidades.

### **3.3 Caracterização de obras Lexicográficas e Terminográficas: dicionário, vocabulário e glossário.**

Sabemos que a Terminografia é uma face aplicada da Terminologia que se dedica à prática de elaboração de dicionários, vocabulários (técnicos, científicos e especializados) e glossários, incluindo o estudo de metodologias e princípios teóricos para essa construção. Já a Lexicografia é parte da Lexicologia que se dedica à construção dessas mesmas obras, mas ocupa-se da descrição de unidades lexicais, diferentemente da Terminografia que se ocupa do inventário de termos de diferentes domínios de especialidade.

Como são tênues as fronteiras entre os tipos de obras lexicográficas e terminográficas, existem divergências quanto à classificação e nomenclatura dessas obras, ou seja, obras da mesma natureza ou função são classificadas ou definidas de modo diverso, “segundo os critérios adotados por este ou aquele autor, fato que conduz à existência de numerosas denominações para o mesmo núcleo conceptual ‘obra lexicográfica / terminográfica’ (dicionário técnico, dicionário terminológico, dicionário de língua científica, dicionário especial, vocabulário, vocabulário técnico-científico, glossário, dentre outras denominações) (BARBOSA, 2001, p. 32). Dada essa divergência de tipologias e classificações, buscamos restringir-nos apenas à caracterização de três tipos de obras, **dicionário de língua, vocabulário e glossário**, de acordo com as proposições de Barbosa (2001).

Segundo Barbosa (2001, p. 33), os chamados dicionários de língua “processam as unidades lexicais da língua geral”, já os vocabulários, dicionários terminológicos, dicionários técnicos e glossários, “processam vocábulos representativos de uma norma linguística,

inclusive as das línguas de especialidade”; e, ainda, os glossários ou vocabulários “processam o vocabulário de um texto-ocorrência”. De maneira mais clara e detalhada:

Assim, por exemplo, ao nível do sistema corresponde a unidade padrão lexical chamada lexema; o dicionário de língua tende a reunir o universo dos lexemas, apresentando, para cada um deles, os vocábulos que representam suas diferentes acepções. Os vocabulários técnico-científicos e especializados buscam situar-se ao nível de uma norma lingüística e sociocultural, têm como unidade-padrão o vocábulo, constituindo-se como conjuntos vocabulares, representativos de universos de discurso. O vocabulário fundamental, por sua vez, procura reunir os elementos constitutivos da intersecção dos conjuntos vocabulários de uma comunidade ou de um segmento social, elementos esses que são selecionados pelo duplo critério de alta freqüência e distribuição regular entre os sujeitos falantes-ouvintes envolvidos; de maneira geral, o glossário *lato sensu* resulta do levantamento das palavras-ocorrências e das acepções que têm num texto manifestado (BARBOSA, 2001, p. 35).

Esquematizando o exposto, temos:

QUADRO 4 – Tipologia de obras lexicográficas e terminográficas conforme Barbosa (2001)

<b>Dicionário</b>	<b>Vocabulário</b>	<b>Glossário</b>
<i>Nível do sistema</i>	<i>Nível da norma</i>	<i>Nível da fala</i>
Trabalha com todo o léxico disponível e o léxico virtual	Trabalha com conjuntos manifestados dentro de uma área especializada ou universo de discurso	Trabalha com conjuntos manifestados em um determinado texto
Unidade: lexema (significado abrangente; freqüência regular)	Unidade: vocábulos/termos (significado restrito; alta freqüência)	Unidade: palavras (significado específico; única aparição)
Apresenta (teoricamente) todas as acepções de um mesmo verbete Perspectivas: diacrônica, diatópica, diafásica, diastrática	Apresenta todas as acepções de um verbete dentro de uma área especializada ou universo de discurso Perspectivas: sincrônica e sinfásica	Apresenta uma única acepção do verbete (dentro de um contexto determinado) Perspectivas: sincrônica, sintópica, sinstrática, sinfásica

Fonte: Adaptado pela autora com base em Barbosa (2001 apud FROMM, 2002, p. 17).

Depreendemos, então, que o dicionário de língua é um repertório que faz uma descrição do vocabulário de uma língua, “buscando registrar e definir os signos lexicais que referem os conceitos elaborados e cristalizados na cultura” (BIDERMANN, 2001a, p. 18). Já o Vocabulário é um repertório de unidades lexicais (vocábulos) ou terminológicas (termos) com significado restrito, pertencentes a uma determinada área ou a um dado universo de discurso. Por outro lado, o Glossário é um repertório de palavras-ocorrências com significado específico em um texto manifestado.

Como vimos, alguns aspectos são fundamentais para determinar a natureza e a tipologia de um dado repertório, como: os níveis de abstração da linguagem (sistema, norma e fala); o recorte feito na língua (todo o léxico ou parte dele); o tipo de unidade lexical; a quantidade de acepções e as perspectivas linguísticas (diacrônica, diatópica, diafásica, diastrática, sincrônica, sintópica, sinfásica, sinstrática).

É com base nessas questões estabelecidas por Barbosa (2001) que classificamos o produto final deste estudo como vocabulário, uma vez que se trata de um repertório de vocábulos-termos circunscritos em uma norma linguística, inseridos em um universo de discurso etnoliterário e atualizados no conjunto de textos-ocorrência (obras literárias regionalistas sertanistas), revelando aspectos socioculturais de um grupo específico: o sertanejo.

Em relação à organização interna, as obras lexicográficas e terminográficas possuem três componentes estruturais: a macroestrutura, a microestrutura e o sistema de remissivas. Por macroestrutura “entende-se a organização interna de uma obra lexicográfica ou terminográfica”, que está relacionada “às características gerais do repertório, ou seja, à estruturação das informações em verbetes, à presença ou não de anexos, índices remissivos, ilustrações, setores temáticos, mapa conceptual e outros” (BARROS, 2004, p. 151). Esta organização é dividida em três partes: a) as páginas iniciais que apresentam e introduzem a obra; b) o corpo do dicionário, isto é, o número de entradas lexicais ou lemas, geralmente organizados em ordem alfabética, denominado de nomenclatura (conjunto de verbetes); c) e finalmente as páginas finais.

A microestrutura diz respeito ao modo pelo qual os verbetes são organizados. De acordo com Barros (2004, p. 152), “os verbetes reúnem os dados relativos à unidade lexical ou terminológica descrita e compõem-se de pelo menos dois elementos: **entrada e o enunciado lexicográfico/terminográfico**”. Entende-se por entrada a unidade lexical ou terminológica a ser explicada (palavra-entrada que encabeça um verbete) e por enunciado lexicográfico/terminográfico o conjunto de informações fornecidas sobre ela. Para Barros (2004, p. 154), a definição é um componente importantíssimo e indispensável na construção de uma obra lexicográfica/terminográfica e, por isso, “deve veicular as informações necessárias para a total compreensão do conteúdo semântico-conceptual da entrada”.

Contudo, nem sempre as definições são capazes de suprir todas as necessidades para a compreensão da palavra-entrada, já que muitas vezes nelas são empregadas unidades lexicais desconhecidas do leitor. É nesse ponto que entram as remissivas – “mecanismo estrutural que

procura resgatar as relações semântico-conceituais existentes entre as unidades lexicais ou terminológicas que compõem a nomenclatura de uma obra lexicográfica ou terminográfica” (BARROS, 2004, p. 174). Elas, geralmente, remetem o a outro termo ou verbete, onde se encontra a informação completa. De um modo geral, toda obra lexicográfica/terminográfica comporta um sistema de remissivas em sua macro e microestrutura. A função delas é orientar o leitor/consulente “sobre o percurso a seguir para obter as informações procuradas, permitindo uma ampliação do conhecimento, dos pontos de vista do conteúdo e das funções do termo consultado.” (BARROS, 2004, p. 174).

Desse modo, o enunciado lexicográfico/terminográfico se organiza em três macroparadigmas: I) o paradigma informacional (PI) - constituído de elementos como abreviaturas, categoria gramatical, transcrição fonética, conjugação, gênero, número, pronúncia, etc; II) o paradigma definicional (PD) - que faz a descrição dos semas ou unidades de significação (acepção discursiva); III) e o paradigma pragmático (PP) - formado de informações contextuais, tais como exemplos de uso, abonações, remissivas, notas, variantes e outros (BARBOSA, 1991).

Nas palavras de Andrade (1998, p. 5):

O número de “informações” sobre uma entrada pode ampliar-se indefinidamente. Os macroparadigmas podem se subdividir em microparadigmas, variáveis em quantidade e qualidade, conforme a natureza da obra. Significa que outros paradigmas podem ser acrescentados ao artigo mínimo, ampliando as informações da microestrutura: índices da frequência, nível de rapidez da difusão de uma palavra, emprego preferencial por um autor, relações de significação (sinonímia, hiperonímia, antonímia, homonímia, analogias, ilustrações, etc.).

Podemos inferir, então, que conforme a natureza da obra, os objetivos do terminógrafo/especialista e o público-alvo a quem se destina há variações na macro/microestrutura das obras lexicográficas e terminográficas. É sabido que a tarefa prioritária na prática lexicográfica/terminográfica é a construção de uma boa definição. Mas, para que isso ocorra, é importante e necessária a adoção de critérios na elaboração de qualquer tipo de obra, para que o produto final não somente esteja adequado às particularidades do repertório em questão, mas, principalmente, atenda com exatidão as necessidades dos usuários.

### 3.4 Linguística de *Corpus*

A prática de compilação de *corpora* é antiga, mas a compilação e análise de *corpora* eletrônicos para a pesquisa de uma língua ou variedade linguística é recente. Com a popularização dos computadores e a criação de vários *corpora* eletrônicos e ferramentas computacionais vimos surgir, ao longo das últimas cinco décadas, um grande volume de pesquisas que se valem de *corpus* para a descrição dos mais variados aspectos da linguagem. Testemunhamos, assim, o (re)surgimento da Linguística de *Corpus* (doravante LC) e também o renascimento do empirismo.

Dizemos ressurgimento porque antes do computador passar a integrar a rotina de pesquisa dos linguistas já se compilava *corpus*, ou seja, antes da invenção do computador em 1960, já se fazia o uso de *corpus*, como por exemplo, o *Corpus* Helenístico criado na Grécia Antiga e os *corpora* de citações da Bíblia produzidos na Antiguidade e na Idade Média. Nessa época, os dados eram coletados, armazenados e analisados manualmente (PARODI, 2010).

Apesar das dificuldades, os estudos baseados em *corpora* não pararam. No final dos anos de 1950 com a publicação de *Syntactic Structures* de Chomsky, houve uma mudança de “paradigma na linguística: saía de cena o empirismo e a sustentação dos trabalhos baseados em *corpora*, tomando lugar central as teorias racionalistas da linguagem, notadamente a linguística gerativa” (BERBER SARDINHA, 2004, p. 4). Desse período ganham destaque os estudos de Firth (1957) e o *corpus* SEU (*Survey of English Usage*), compilado e etiquetado manualmente em 1959.

Com o advento do computador nos anos de 1960 e a queda de prestígio das pesquisas puramente racionalistas, novos *corpora* são compilados, dentre os principais estão: Brown, Birmingham e BNC. Segundo Berber Sardinha (2004), o *Brown University Standard Corpus of Present-day American English*, primeiro *corpus* linguístico eletrônico, lançado em 1964, com um milhão de palavras, alavancou substancialmente os estudos em LC e se tornou uma marco na história da LC em nível mundial, mesmo apesar de ter nascido em um período ainda desfavorável para os estudos empiristas e pela dificuldade de compilação em computadores *mainframe*.

De acordo com Leech (1992), mesmo diante desse grande feito (a compilação do primeiro *corpus* linguístico eletrônico - o *Brown Corpus*) o uso da expressão *corpus linguistics* só aparece em 1984 em um livro publicado por Aarts e Meijs, década em que, segundo Svartvik (1992), a LC atingiu a maioria tornando-se um campo de grande

importância científica e de grande relevância para a sociedade. Isto porque com o progresso da ciência nos anos de 1980, os computadores se estenderam aos lares, passando também a equipar grandes centros de pesquisa universitários, o que facilitou a criação e a popularização de *corpora* e de ferramentas de processamento, “contribuindo decisivamente para o reaparecimento e fortalecimento da pesquisa linguística baseada em *corpus*” (BERBER SARDINHA, 2004, p. 5).

Desde então, a LC vem ganhando cada vez mais destaque nas análises linguísticas nos seus mais diversos ramos, trazendo grandes contribuições para os estudos linguísticos em diferentes áreas. Mas não é só nos centros acadêmicos que a LC se destaca, como assevera Berber Sardinha (2004, p. 6), “também no âmbito empresarial há um interesse crescente nas aplicações comerciais de estudos baseados em corpora”. Segundo o autor, muitas empresas (como por exemplo, as empresas de telecomunicações) utilizam pesquisas baseadas em *corpus* para vários fins comerciais, “como o processamento automático de textos, informatização de grandes bases de dados e a montagem de sistemas inteligentes de reconhecimento de voz e gerenciamento de informação” (BERBER SARDINHA, 2004, p. 6).

No Brasil, desde o final do século XX, é cada vez maior o número de pesquisadores nessa área, o que denota “crescimento qualitativo e quantitativo das pesquisas realizadas [...], bem como a existência de *corpora* e ferramentas para pesquisa em várias línguas” (BERBER SARDINHA; ALMEIDA, 2008, p. 17). Apesar da propagação de diversos eventos científicos desde 1999, o marco considerado relevante para o desenvolvimento da LC no Brasil foi a publicação do livro *Linguística de Corpus* em 2004, de Tony Berber Sardinha, um dos maiores expoentes da LC no Brasil, hoje professor associado do Departamento de Linguística e do Programa de Estudos Pós-Graduados em Linguística Aplicada e Estudos da Linguagem da Pontifícia Universidade Católica de São Paulo (PUC-SP).

Assim, em vista de meio século de desenvolvimento em nível mundial e um pouco mais de uma década no Brasil, já podemos dizer que a LC se consolidou como uma aventura mais do que adequada no âmbito dos estudos linguísticos, como afirmam Novodvorski e Finatto (2014, p. 15): “a aventura tem sido, sim, adequada e, mais do que isso, já muito bem-sucedida”, haja vista a “diversidade das pesquisas desenvolvidas e publicações feitas no país”. Segundo os autores:

Atualmente, a expansão do uso dos termos *corpus* e *corpora*, além da menção a muitas das ferramentas e princípios caros à LC, alcança áreas que poderiam parecer, num primeiro momento, incompatíveis ou inimagináveis. Assim, a alusão às terminologias típicas de LC (como *types*, *tokens* e *concordâncias*) vem se tornando cada vez mais recorrente. Em eventos



científicos, em publicações, em nomes de disciplinas, teses e dissertações, a recorrência com que aparecem referências ou vestígios da LC denotam já uma presença marcada no plano acadêmico e servem como um bom termômetro do estado da arte (NOVODVORSKI; FINATTO, 2014, p. 8, 9).

Fato a ser considerado é que contamos hoje com um grande número de trabalhos e publicações na área. A título de exemplificação, podemos citar: Berber Sardinha (2004, 2007, 2009, 2012), Tagnin e Vale (2008), Frankenberg-Garcia (2008), Orenha-Ottaiano (2008), Gonçalves (2008) Viana e Tagnin (2010), Pires e Salomão (2012), Raso e Mello (2012), Dutra e Mello (2012), Shepherd, Berber Sardinha e Pinto (2012), Canever (2012), Ferregueti, Pagano e Figueredo (2012) Tagnin e Bevilacqua (2013); Viana e Tagnin (2010, 2015), dentre outros.

De acordo com Berber Sardinha (2004, p. 3), a LC contemporânea caracteriza-se como “uma abordagem empirista de exploração da linguagem por meio de evidências empíricas, extraídas por computador” e, como tal, “ocupa-se da coleta e da exploração de *corpora*, ou conjuntos de dados linguísticos textuais coletados criteriosamente, com o propósito de servirem para a pesquisa de uma língua ou variedade linguística”. De forma similar Tognini-Bonelli (2001, p. 2) afirma que a LC “[...] é uma abordagem empírica para a descrição da língua em uso; opera dentro de uma teoria contextual e funcional do significado; faz uso de novas tecnologias.”<sup>14</sup>

Em Linguística, entende-se por empirismo a abordagem que confere “primazia aos dados provenientes da observação da linguagem, em geral reunidos sob a forma de um *corpus*” (BERBER SARDINHA, 2004, p. 30). Isso significa dizer que na LC a linguagem é observada em seu contexto de uso a partir da exploração sistemática de um *corpus*, ou seja, por meio de uma investigação empírica é possível ver através de um *corpus* os traços linguísticos (lexicais, sintáticos, semânticos, discursivos) que caracterizam uma língua ou variedade linguística. Em suma, trata-se de uma abordagem que tem como foco: o desempenho linguístico, em vez de competência; a descrição linguística, em vez de universais linguísticos; e uma visão mais empirista do que racionalista da pesquisa linguística.

Dessas considerações, fica evidente não apenas a conexão existente entre a Linguística de *Corpus* e a linguística hallidayana, como também o contraste entre a tradição empirista de Halliday e a tradição racionalista de Chomsky. Enquanto Halliday vê a linguagem como probabilidade, Chomsky a enxerga como possibilidade, isto é:

---

<sup>14</sup> No original: “[...] it is an empirical approach to the description of language use; it operates within the framework of a contextual and functional theory of meaning; it makes use of the new Technologies” (TOGNINI-BONELLI, 2001, p. 2).

A linguística chomskyana gerativista enfatiza a determininação de quais agrupamentos sintáticos são possíveis (permissíveis) dado o conhecimento que um falante nativo possui de sua língua. Já a linguística hallidayana descreve a probabilidade dos sistemas linguísticos, dados os contextos em que os falantes os empregam (BERBER SARDINHA, 2004, p. 30).

A LC baseia-se, portanto, numa visão probabilística da linguagem que pressupõe que os traços linguísticos não se apresentam de forma aleatória, ou seja, embora muitos traços sejam teoricamente possíveis, eles não ocorrem com a mesma frequência. Isto porque as regularidades (padrões) estão associadas às características linguísticas e situacionais (os contextos de uso), ou seja, “há um mapeamento regular entre a frequência maior ou menor de um traço em um contexto de ocorrência” (BERBER SARDINHA, 2004, p. 31)

Por isso é comum dizer que a linguagem é padronizada, ou seja, há uma regularidade expressa na recorrência sistemática de unidades coocorrentes de várias ordens: lexical, gramatical, semântica, dentre outras. Nas palavras de Berber Sardinha (2004, p. 31), “a linguagem forma padrões que apresentam regularidade”, e essa padronização se evidencia pela “recorrência, isto é, uma colocação, coligação<sup>15</sup> ou estrutura que se repete significativamente e mostra sinais de ser, na verdade, um padrão lexical ou léxico-gramatical”

Isso mostra ainda mais a ligação íntima entre a Linguística de *Corpus* e linguística hallidayana. Apesar de Halliday não se denominar linguista de *corpus*, Berber Sardinha (2004) afirma que a sua visão de linguagem se encaixa perfeitamente nos preceitos da LC e serve como arcabouço teórico para qualquer trabalho que se dedica à análise de *corpora*.

Por isso, muito mais do que um recurso metodológico, o uso de *corpus* na pesquisa linguística implica uma forma diferente de se estudar uma língua. Traços que antes não eram possíveis ver a olho nu em uma análise manual, hoje, com o desenvolvimento das tecnologias e ferramentas computacionais, é possível não só coletar e armazenar uma grande quantidade de textos em questão de segundos, como visualizar novos e variados fenômenos linguísticos antes não observados (BIBER et al, 1998).

Entretanto, um *corpus* “não é apenas um repositório de exemplos usados como suporte para validar uma teoria pré-existente ou um modo de se calcular estatisticamente um sistema já bem definido”(TOGNINI-BONELLI, 2001, p. 84). Mais do que isso, um *corpus* deve contemplar textos (orais ou escritos) que tenham sido efetivamente produzidos por falantes de determinada língua, ou seja, deve ser representativo de uma linguagem, e também deve ser

---

<sup>15</sup> “Colocação: associação entre itens lexicais, ou entre o léxico e campos semânticos. [...] Coligação: associação entre itens lexicais e gramaticais” (BERBER SARDINHA, 2004, p.40).

adequado (em extensão e conteúdo) à investigação do fenômeno linguístico, ou seja, deve estar afinado com os objetivos da análise, com os propósitos da pesquisa.

Existem várias definições de *corpus* na literatura atual, contudo convém apresentar apenas as que melhor se aplicam neste trabalho. Entre elas, está a proposta por Sinclair (1991, p. 10) que define *corpus* como “uma coleção de textos em formato eletrônico, selecionados de acordo com critérios externos para representar, na medida do possível, uma língua ou variedade de língua como fonte de dados para a pesquisa linguística”<sup>16</sup>. A definição de Biderman (2001) que considera *corpus* como “[...] um conjunto homogêneo de amostras da língua de qualquer tipo (orais, escritos, literários, coloquiais, etc.) ”; e o *corpus* linguístico informatizado como “como uma coletânea de textos selecionados segundo critérios linguísticos, codificados de modo padronizado e homogêneo”, que pode ser tratada mediante processos informáticos (BIDERMAN, 2001b, p. 79). E também a de Parodi (2010) que define *corpus* como: “un conjunto amplio de textos digitales de naturaleza específica y que cuenta con una organización predeterminada em torno a categorías identificables para la descripción y análisis de una variedad de lengua” (PARODI, 2010, p. 25).

Incorporando as características principais mencionadas por esses autores, Berber Sardinha (2004) define *corpus* como:

Um conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou em 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 (BERBER SARDINHA, 2004, p. 18).

Segundo o autor, essa definição é mais completa porque menciona vários pontos importantes: a) a origem: “os dados devem ser autênticos”; b) o propósito: “o *corpus* deve ter a finalidade de ser um objeto de estudo linguístico”; c) a composição: “o conteúdo do *corpus* deve ser criteriosamente escolhido”; d) a formatação: “os dados do *corpus* devem ser legíveis por computador”; e) a representatividade: “o *corpus* deve ser representativo de uma língua ou variedade”; f) e a extensão: “o *corpus* deve ser vasto para ser representativo” (BERBER SARDINHA, 2004, p. 18, 19).

Sobre essa última característica importa salientar que, embora seja um critério fundamental na representatividade, não há regras consistentes para determinar o tamanho

---

<sup>16</sup> No original: “[...] a collection of pieces of language text in electronic form, selected according to external criteria to represent, as far as possible, a language or language variety as a source of data for linguistic research” (SINCLAIR, 1991, p. 10)

ideal de um *corpus*. Essa é uma decisão que envolve diversos fatores, como as necessidades do projeto, a disponibilidade de dados, o tempo disponível para o desenvolvimento da pesquisa, o conhecimento acerca da probabilidade de ocorrência dos traços linguísticos que se pretende investigar, os parâmetros a utilizar para saber que quantidade de dados é suficiente para evidenciar certas características linguísticas, dentre inúmeros outros.

Sobre isso, Bowker e Pearson (2002, p. 45) ressaltam que “não se suponha que maior é sempre melhor”<sup>17</sup>, pois é possível que um *corpus* pequeno e melhor planejado seja mais representativo do que um *corpus* com um zilhão de palavras (tal como propõe Charles Fillmore), sem personalização e adequação aos interesses da pesquisa. Para Bowker e Pearson (2002) antes de uma preocupação com tamanho ou quantidade, deve-se buscar a verificação de qualidade do *corpus* em termos de organização, planejamento e representatividade, ou seja, a adequação do conteúdo de um *corpus* deve prevalecer sobre as questões acerca de seu tamanho.

Quanto à tipologia do *corpus*, convém também apresentar somente as que se aplicam neste trabalho, a saber: a) *corpus* de estudo: aquele em que se baseia a pesquisa a ser desenvolvida; b) *corpus* de referência: serve de termo de comparação, contraste com o *corpus* de estudo (deve ter três a cinco vezes o tamanho do *corpus* de estudo); c) *corpus* escrito: composto de textos escritos; d) *corpus* diacrônico: compreende vários períodos de tempo; e) *corpus* de língua nativa: os autores são falantes nativos (BERBER SARDINHA, 2004).

Todas essas observações nos permite inferir que a história da LC está intimamente relacionada ao desenvolvimento das tecnologias e ferramentas computacionais que se propagaram com o advento do computador. Tais recursos não só permitiram a criação e o armazenamento de *corpora*, como também a sua exploração. Atualmente, contamos com um número considerável de *corpora* eletrônicos disponíveis *online*, dentre os quais convém destacar:

- 1) o **Corpus Brasileiro**, *corpus* de português brasileiro com aproximadamente 1 bilhão de palavras, disponível nas plataformas *Sketch Engine* e *Compara*, compilado por Tony Berber Sardinha (com financiamento da Fapesp) e etiquetado com a ferramenta *Tree Tagger*.
- 2) o **Lingueca**<sup>18</sup>, é um centro de recursos distribuído para o processamento computacional da língua portuguesa, que tem como objetivo servir a comunidade que se dedica ao processamento da nossa língua. Principais recursos: **AC/DC** (acesso a *corpora*/disponibilização de *corpora*/com mais de 250 milhões de palavras em português, nos

<sup>17</sup> No original: “[...] not to assume that bigger is Always better” (BOWKER; PEARSON, 2002, p. 45).

<sup>18</sup> Disponível em: < <https://www.lingueca.pt/>>. Acesso em 7 de maio de 2018.

registos jornalístico, literário, didático e correio eletrônico); **COMPARA** (é um *corpus* paralelo bidirecional com textos em português e inglês e as suas traduções); **CETEMPúblico** e **CETENFolha** (Dois corpora de texto jornalístico de grandes dimensões, separados em extractos, e integralmente disponíveis); **Corpógrafo** (conjunto de ferramentas disponíveis *online* que permite colecionar textos em vários formatos, formar e analisar *corpora*, extrair terminologia e criar bases de dados terminológicas).

3) o **Lacio-Web**, projeto que visa compilar *corpora* de livre acesso. Esse projeto compreende seis corpora: a) um *corpus* de referência chamado Lacio-Ref composto de textos em português brasileiro; b) Mac-Morpho, *corpus* etiquetado compreendendo 1,1 milhões de palavras que foi validado manualmente para *tags* morfossintáticas; c) uma parte automaticamente anotada do Lacio-Ref com lemas, POS e *tags* sintáticas que são utilizadas pelo analisador Curupira desenvolvido no Núcleo Interinstitucional de Linguística Computacional (NILC/São Carlos); d) um *corpus* de desvio composto por textos não revisados (Lacio-EV); e) o Par-C, *corpus* paralelo inglês-português; e f) o Comp-C, *corpus* comparável com amostras do gênero jurídico em português e inglês;

4) o **COMET**, *corpus* multilíngue para Ensino e Tradução, desenvolvido pela Universidade de São Paulo, que tem por objetivo servir de suporte a pesquisas linguísticas, principalmente nas áreas de Tradução, Terminologia e Ensino e Aprendizagem de Línguas. É composto por três *subcorpora*: Corpus Técnico-Científico (**CorTec**); Corpus de Tradução (**CorTrad**) e Corpus Multilíngue de Aprendizizes (**CoMAprend**).

5) o **Corpus do Português**, *corpus* linguístico e textos da língua portuguesa, compilado e mantido pelos pesquisadores Mark Davies (Universidade Brigham Young) e Michael J. Ferreira (Universidade de Georgetown), com suporte financeiro proveniente do US National Endowment for the Humanities, além de suas respectivas instituições de ensino. O *corpus* original (criado em 2004) contém uma base de dados com 45 milhões de palavras, extraídas de quase 57.000 textos em português dos séculos XIII ao XX. Em 2016, o *Corpus* passou por uma nova atualização, contendo uma base de dados com cerca de um bilhão de palavras coletadas a partir de textos mais recentes (2013-14). Em recente atualização (2018), o *Corpus* do Português passou a contar com mais de 1,1 bilhão de palavras de quatro países de língua portuguesa diferentes. Além disso, o *Corpus* também contém o **WordAndPhrase-Portuguese** que permite pesquisar pelas 40.000 palavras principais em português (com base na frequência do *corpus*)<sup>19</sup>

---

<sup>19</sup> Disponível em: <<https://www.corpusdoportugues.org/xp.asp>>. Acesso em: 7 de maio de 2018.

Atualmente, a expansão do uso dos termos *corpus* e *corpora*<sup>20</sup>, além da menção a muitas ferramentas (como *types*, *tokens* e concordâncias) e princípios caros à LC, alcança áreas que pareciam, num primeiro momento, inimagináveis (NOVODVORSKI; FINATTO, 2014). Em eventos científicos, em publicações, em nomes de disciplinas, teses e dissertações, e também no âmbito empresarial, a presença da LC é cada vez mais recorrente, o que denota um grande avanço da área, tanto em nível nacional quanto internacional.

Hoje, a LC exerce grande influência em diversas áreas, em especial, no campo da Terminologia, área que integra o aporte teórico deste trabalho. Sobre essa importante disciplina científica convém mencionar que, paralelo ao desenvolvimento da LC, surgem novos paradigmas teóricos e metodológicos voltados para o uso real do termo, isto é, a sua utilização nos diferentes contextos e situações comunicativas. Segundo Bevilacqua (2013), essas mudanças teóricas e metodológicas, bem como a criação e popularização de vários *corpora*, proporcionaram grandes avanços na prática terminológica, a saber:

Uma das principais mudanças que se observa é que, a partir do pressuposto de que é preciso identificar os termos *in vivo*, não se pode mais seguir o método onomasiológico, ou seja, não se pode ir mais do conceito (significado) ao termo (significante), conforme propunha a TGT. Parte-se da premissa de que o termo é um signo linguístico e que, portanto, não é possível separar significante e significado. Contudo, ao partir os textos, começa-se a identificação dos termos pelo significante e, após a investigação de seu uso em contexto, se estabelece o seu significado, ou, mais ainda, seu valor especializado. Começa a predominar, então, o método semasiológico. A segunda mudança, relacionada à anterior, é que para que se siga esse método, é preciso descrever os vários aspectos que caracterizam os textos especializados, considerando os diferentes níveis de análise textual: funcional, isto é, a função do texto (informar, argumentar etc.), situacional [...], de conteúdo semântico [...] e linguístico (constituição os termos – simples, sintagmáticos, sua morfologia -, fraseologias, fórmulas retóricas, adjetivos frequentes etc.) [...] Em relação à prática, isto é, ao fazer terminográfico, uma inovação é o uso de *corpora* textuais e ferramentas computacionais de extração de informação linguística (BEVILACQUA, 2013, p. 11 - 13).

Em outras palavras, de uma perspectiva prescritivista passa-se para uma perspectiva descritivista, que prevê que se identifiquem, analisem e descrevam os termos, considerando não apenas os seus aspectos linguísticos, comunicativos e cognitivos, mas o contexto em que estão inseridos.

---

<sup>20</sup>Além dessas formas latinas, existe ainda a forma “córpus”, que começou a ser usada em português para indicar tanto a forma singular quanto a plural (BERBER SARDINHA, 2009).

De modo geral, o uso de corpora na Terminologia/Terminografia trouxe importantes mudanças e significativas contribuições para os estudos terminológicos, principalmente no que se refere ao trabalho terminográfico. Nas palavras de Bevilacqua (2013):

Sem dúvida, a Linguística de *Corpus*, além de estabelecer os princípios e critérios para a compilação de *corpora* [...], também oferece recursos e ferramentas que auxiliam nas diferentes etapas terminográficas: desde a própria compilação de *corpora*, passando pela identificação de candidatos a termos e fraseologias e chegando à identificação de elementos que permitem a elaboração de definição. Dentre alguns dos recursos disponíveis é possível citar os *softwares* *Wordsmith Tools*, *AntConc* e o site *WebCorp*, assim como ferramentas elaboradas por grupos de pesquisa. Alguns grupos brasileiros que disponibilizam recursos desse tipo são: CoMET (USP), GETerm (UFSCAR), NILC (USP-UFSCAR), Termisul (UFRGS), TermNeo (USP), TextQuim (UFRGS), e plataformas como o e-Termos (BEVILACQUA, 2013, p. 17).

Embora se perceba e se conheça as facilidades e possibilidades de análise que a LC tem trazido para as pesquisas linguísticas em Terminologia, é preciso que o pesquisador, além de observar os elementos linguísticos do domínio discursivo estudado, também verifique a existência de novos elementos ou informações que o *corpus* poderá fornecer no momento da aplicação das ferramentas de processamento. É sabido que a própria seleção da composição textual de um *corpora* para a realização de uma pesquisa implica a pré-existência de uma hipótese que se quer ver comprovada. Mas, é possível que o *corpus* mostre outros elementos, talvez de grande importância para a pesquisa, sem que o pesquisador tenha percebido previamente ou pelo menos hipoteticamente imaginado.

### **3.4.1 *WordSmith Tools*: ferramentas e utilitários**

Como vimos, a Linguística de *Corpus* é uma das áreas de pesquisa de linguagem mais ativas nos últimos anos e, notavelmente, por meio de variados programas de computador e ferramentas computacionais, vem auxiliando cada vez mais pesquisadores a lidar com grandes quantidades de dados antes inacessíveis. Dentre os vários *software* existentes atualmente destacamos o *WordSmith Tools*.

A opção por esse programa justifica-se por duas razões: primeiro, por ser um *software* de referência nos estudos linguísticos (particularmente, os lexicais) já consagrado no campo da LC, e utilizado por um grande número de usuários em todo o mundo. Por essa razão, “tem sido um fator de divulgação da Linguística de *Corpus* no Brasil” (BERBER SARDINHA,

2009, p. 8); e segundo, por já possuímos certa habilidade com as ferramentas do programa a partir do desenvolvimento de análises baseadas em *corpus* na ocasião das disciplinas do Doutorado.

Basicamente, o *WordSmith Tools* (doravante WST), criado em 1996 por Mike Scott, da Universidade de Liverpool, Reino Unido, e comercializado pela *Oxford University Press* é um conjunto integrado de programas (‘suíte’) destinado à análises linguísticas. Além de permitir a realização de análises baseadas na frequência e na coocorrência de palavras em corpora, ele também permite “pré-processar os arquivos do *corpus* (retirar partes indesejadas de cada texto, organizar o conjunto de arquivos, inserir e remover etiquetas etc.), antes da análise propriamente dita” (BERBER SARIDINHA, 2009, p. 8). O WST já se encontra em sua sétima versão e possui ferramentas essenciais de análise: a *WordList* (lista de palavras), a *Concord* (concordâncias) e a *KeyWords* (palavras-chave), as quais apresentamos a seguir:

FIGURA 2: Tela inicial do *WordSmith Tools* versão 6.0



Fonte: *WordSmith Tools* 6.0

A *WordList* permite gerar uma lista de palavras presentes no *corpus*. Ela produz duas listas, uma em ordem alfabética e outra por ordem de frequência. Além disso, essa ferramenta possui uma terceira janela, denominada de *Statistics*, onde é feito o levantamento dos dados



estatísticos do *corpus*, ou seja, essa função informa o número de *tokens* (itens), *types* (formas) e *type/token ratio* (razão forma/item).

De acordo com a nomenclatura e explicação feita por Berber Sardinha (2004 apud NOVODVORSKI, 2013, p. 66):

(1) os *itens* (*tokens*) ou *palavras corridas* (*running words*) indicam a totalidade de ocorrências ou palavras contidas, seja no geral (primeira coluna), seja em cada um dos textos (demais colunas);

(2) as *formas* (*types*) indicam a quantidade de palavras diferentes, isto é, computadas uma única vez em cada um dos textos; e

(3) a *Razão forma/item* (*type/token ratio*) é a porcentagem resultante da seguinte fórmula:

$$\frac{\text{formas}}{\text{itens}} \times 100 = \text{razão forma/item}$$

Essa relação também é oferecida pelo programa.

Já a ferramenta *Concord* realiza concordâncias, isto é, “extrai todas as ocorrências de uma palavra de busca num corpus juntamente com seu cotexto, apresentando-as na forma de uma concordância” (TAGNIN, 2011a, p. 358). Além disso, essa ferramenta também oferece listas de “*Collocates*, que apresenta os colocados da palavra de busca; e *Clusters*, que relaciona os agrupamentos em que aparece a palavra de busca” (TAGNIN, 2011a, p. 358).

Por outro lado, a *KeyWords* compara a lista de palavras de um *corpus* de estudo com a lista de palavras de um *corpus* de referência e gera uma lista de palavras-chave, ou seja aquelas palavras consideradas mais chaves no *corpus* de estudo, “cujas freqüências são estatisticamente diferentes (maiores ou menores) do que as freqüências das mesmas palavras num outro corpus (de referência)”. Essa ferramenta mostra a frequência de cada item, bem com a sua chavicidade<sup>21</sup> (*keyness*) e “calcula também palavras-chave chave, que são chave em vários textos” (BERBER SARDINHA, 2009, p. 9)

Sobre os utilitários do programa, Berber Sardinha (2009) diz que o WST possui: o *File Manager*, que “abre uma janela para gerenciamento de arquivos”; o *Splitter*, que “permite dividir um arquivo em vários arquivos menores”; o *Text Converter*, que oferece várias funções para “o pré-processamento de textos, como a substituição de palavras, partes de palavras ou partes de textos, simultaneamente num conjunto de arquivos, a renomeação e a

---

<sup>21</sup>É um termo usado para designar aquelas palavras que caracterizam um *corpus* nos seus mais variados aspectos (a saber: o léxico, o conteúdo semântico, a temática e a variedade linguística predominante nos textos) a partir de cálculos de significância estatística.

mudança de pasta”; e o *Viewer & Aligner*, que “fornece meios para visualização de textos e para o alinhamento de dois textos num só” (BERBER SARDINHA, 2009, p. 10).

Essa explicação detalhada sobre o funcionamento do programa nos permite não apenas compreender o desenvolvimento da análise que realizamos nesta pesquisa, como também vislumbrar a qualidade do WST, tanto no que se refere à organização dos dados gerados quanto à variedade de funções disponibilizadas, o que por sua vez propicia uma análise mais vigorosa e precisa dos dados.

Segundo Berber Sardinha (2009), cada vez mais, o programa torna-se referência para a análise linguística via computador. Isto se deve por várias razões:

A primeira é a facilidade de uso; trata-se de um programa escrito para o ambiente Windows, o ambiente operacional dominante no mundo de hoje, o que significa dizer que a maioria dos interessados já terá alguma familiaridade com operações exigidas pelo *WordSmith Tools* [...]. A segunda razão é decorrência da primeira: devido ao fato de rodar num ambiente gráfico como Windows, o programa oferece uma facilidade maior na utilização dos seus recursos disponíveis, o que por sua vez propicia um aprendizado mais rápido e intuitivo de suas várias funções. A terceira razão é o fato de ser disponibilizado via Internet, o que significa que o usuário não precisa comprá-lo numa loja ou por correio, bastando baixá-lo da rede e encomendar a sua senha pagando com cartão de crédito [...]. A quarta razão do sucesso de *WordSmith Tools* é sua versatilidade. O software consiste na verdade de uma ‘suíte’ de diferentes programas, que se destinam a várias aplicações, que compreendem o pré-processamento, a organização dos dados, e a análise propriamente dita de corpora ou textos isolados (BERBER SARDINHA, 2009, p. 9).

Como vimos, o WST tem sido fundamental para a divulgação da LC, auxiliando cada vez mais pesquisadores de diversos lugares do mundo. Sabemos que LC é um campo que se dedica à criação e análise de *corpora* para fins de exploração da linguagem nos seus mais diversos níveis. Mas para lidar com *corpora*, ou seja, para analisar e descrever determinados fenômenos linguísticos é preciso que o pesquisador lance mão de um programa de computador adequado. Isto significa dizer que escolhemos o WST, não apenas pelas qualidades que possui, mas, sobretudo, por ser o programa mais adequado para a análise que intentamos fazer.

No capítulo a seguir, trataremos de todos os aspectos metodológicos desenvolvidos nesta pesquisa por meio das ferramentas do WST.

## 4 METODOLOGIA

Este capítulo apresenta o *corpus* de estudo e os procedimentos metodológicos adotados e desenvolvidos, ao longo da pesquisa, em consonância com o tipo de aporte teórico que a fundamenta.

A primeira seção exibe uma breve descrição do *corpus*, demonstrando suas principais características. As seções seguintes relatam, de forma pormenorizada e sequencial, todos os passos realizados durante as várias etapas da pesquisa que vão desde a compilação e preparação do *corpus* de estudo ao levantamento, seleção, descrição e análise dos dados.

Por fim, o capítulo mostra o modelo de ficha etnoterminográfica, o modelo de verbete, a organização da macro e microestrutura de nossa proposta de Vocabulário e, também, a construção da página Web. Assim, é possível observar, por um lado, a aplicabilidade e a eficácia dos procedimentos adotados, por outro, as condições linguísticas do *corpus* e a significância dos resultados alcançados.

### 4.1 O *corpus* de estudo

O *corpus* linguístico de estudo adotado para a realização desta pesquisa, conforme mencionado, é um *corpus* literário composto por obras regionalistas de autores consagrados na literatura brasileira. Procuramos compilar as obras mais representativas de cada fase do regionalismo, circunscritas em universo de discurso etnoliterário de temática sertanista, e que tivessem um repertório lexical diversificado e amplo, a fim de obter um material linguístico suficiente e adequado para esta pesquisa. Por essa razão, optamos por romances e contos em virtude da extensão linguística (amplitude e profundidade) que possuem, com exceção a peça teatral *Auto da Compadecida* de Ariano Suassuna, por encontrarmos características marcantes da terceira fase do Regionalismo. Nesse ponto é válido destacar que, mesmo sabendo da existência de outras obras de um mesmo autor, selecionamos apenas uma de cada escritor devido às semelhanças estilísticas e lexicais adotadas por estes ao representarem a sua vertente regionalista.

O QUADRO 5 informa o nome das obras que compõem o *corpus* de estudo segundo as fases do Regionalismo a que pertencem e, ainda, gênero, autor, data da 1ª publicação, região abordada e corrente literária.

QUADRO 5: *Corpus* de estudo

<b>REGIONALISMO PITORESCO</b>					
NOME DA OBRA	GÊNERO	AUTOR	DATA	REGIÃO	CORRENTE LITERÁRIA
O Garimpeiro	Romance	Bernardo Guimarães	1872	Minas Gerais	Romantismo
Inocência	Romance	Visconde de Taunay	1872	Mato Grosso do Sul	Romantismo
O Sertanejo	Romance	José de Alencar	1875	Nordeste (Ceará)	Romantismo
O Cabeleira	Romance	Franklin Távora	1876	Nordeste (Pernambuco)	Romantismo
Pelo Sertão	Contos	Afonso Arinos	1898	Minas Gerais/Goiás	Romantismo/Realismo
<b>REGIONALISMO CRÍTICO</b>					
Os Sertões	Romance	Euclides da Cunha	1902	Nordeste (Bahia)	Pré-Modernismo
Luzia-Homem	Romance	Domingos Olímpio	1903	Nordeste (Ceará)	Realismo-Naturalismo
Contos Gauchescos	Contos	João Simões Lopes Neto	1912	Rio Grande do Sul	Pré-Modernismo
Tropas e boiadas	Contos	Hugo de Carvalho Ramos	1917	Goiás	Pré-Modernismo
Urupês	Contos	Monteiro Lobato	1918	São Paulo	Pré-Modernismo
A Bagaceira	Romance	José Américo de Almeida	1928	Nordeste (Paraíba)	2ª fase do Modernismo
O Quinze	Romance	Raquel de Queiroz	1930	Nordeste (Ceará)	2ª fase do Modernismo
Menino de Engenho	Romance	José Lins do Rego	1932	Nordeste (Paraíba)	2ª fase do Modernismo
Vidas secas	Romance	Graciliano	1938	Nordeste	2ª fase do Modernismo

		Ramos		(Alagoas)	
Seara Vermelha	Romance	Jorge Amado	1946	Nordeste (Bahia)	2ª fase do Modernismo
<b>SUPER-REGIONALISMO</b>					
Auto da Compadecida	Peça teatral	Ariano Suassuna	1955	Nordeste (Paraíba)	3ª fase do Modernismo
Grande Sertão: Veredas	Romance	Guimarães Rosa	1956	Minas Gerais/Bahia	3ª fase do Modernismo
O coronel e o Lobisomem	Romance	José Cândido de Carvalho	1964	Rio de Janeiro	3ª fase do Modernismo
Chapadão do Bugre	Romance	Mário Palmério	1965	Minas Gerais	3ª fase do Modernismo
Sargento Getúlio	Romance	João Ubaldo Ribeiro	1971	Nordeste (Sergipe)	3ª fase do Modernismo

Fonte: Elaboração da autora

A escolha dessas obras como *corpus* de análise foi motivada, em primeiro lugar, por representarem o sertão nos seus mais variados aspectos (geografia física e humana), contemplando um número significativo de unidades lexicais com especificidades semânticas próprias do universo sertanejo; e por outro lado, por estarem atreladas a uma vertente literária que buscava não apenas valorizar e preservar as peculiaridades regionais, sobretudo, as peculiaridades do sertão e as suas manifestações linguísticas, como também minorar, através da denúncia, os problemas sociais, econômicos e políticos das regiões interioranas do país, no afã de um Brasil melhor.

Dessa maneira, analisar o léxico sertanista nessas obras e, ainda, propor um vocabulário dos vocábulos-termos, além de se constituir um material substancial para os estudos linguísticos, nos faz refletir sobre a importância desses itens lexicais na história e na cultura de um povo, sobretudo na história do sertanejo das regiões interioranas do Brasil, povo este que existiu e ainda existe na memória coletiva da nossa sociedade.

Em relação à composição do *corpus* de estudo, procuramos caracterizá-lo a partir das tipologias apresentadas por Berber Sardinha (2004). O QUADRO 6 mostra, com detalhes, as características do *corpus*:

QUADRO 6 – Tipologia do *corpus* de estudo

Critérios	Características
Língua	Monolíngue (Português)
Modo	Escrito (narrativas)
Tempo	diacrônico (textos de 1872-1971)
Seleção	Amostragem (amostra de textos literários ficcionais) Estático (não dinâmico, típico do <i>corpus</i> de amostragem)
Conteúdo	Especializado/Regional (textos da literatura regionalista sertanista)
Autoria	Vários autores; falantes nativos
Finalidade	De estudo (análise etnoterminológica)
Nível de codificação	Com cabeçalhos, sem etiquetas
Tamanho	1.348.165 itens (médio-grande) <sup>22</sup>

Fonte: Elaboração da autora

Na TABELA 1, detalhamos os dados estatísticos do *corpus* geral e de cada *subcorpus* (correspondente a cada uma das fases do Regionalismo), obtidos por meio da ferramenta *WordList* do WST.

TABELA 1 – Dados estatísticos do *corpus* de estudo

Corpus	<i>Tokens</i> (itens)	<i>Types</i> (formas)	Razão forma/item %	Nº de textos
<b>Corpus geral</b>	1.348.165	71.442	5,32	20
<b>Subcorpus 1(RP)<sup>23</sup></b>	271.201	26.921	9,95	5
<b>Subcorpus 2 (RC)</b>	568.178	47.422	8,37	10
<b>Subcorpus 3 (SR)</b>	508.786	33.667	6,64	5

Fonte: Elaboração da autora

As próximas seções descrevem passo a passo os procedimentos metodológicos adotados e desenvolvidos neste trabalho.

<sup>22</sup> De acordo com Berber Sardinha (2004), um dos critérios fundamentais na constituição de um *corpus* é a representatividade. Em razão disso, o autor sugere a seguinte classificação dos corpora segundo suas extensões: pequeno (menos de 80 mil palavras); pequeno-médio (80 a 250 mil palavras); médio (250 mil a 1 milhão de palavras); médio-grande (1 milhão a 10 milhões de palavras); grande (10 milhões ou mais de palavras).

<sup>23</sup> RP – Regionalismo Pitoresco ; RC – Regionalismo Crítico; SR – Super-regionalismo.

## 4.2 Procedimentos Metodológicos

Adotando uma perspectiva de análise direcionada pelo *corpus*, percorremos as seguintes etapas para a realização desta pesquisa:

- 1) Compilação e preparação do *corpus* de estudo: digitalização, revisão, limpeza e armazenamento do *corpus*;
- 2) Levantamento das listas de palavras;
- 3) Levantamento das palavras-chave;
- 4) Aplicação da ferramenta *Concord* e consulta aos dicionários para identificação dos vocábulos-termos;
- 5) Seleção dos vocábulos-termos;
- 6) Elaboração das fichas etnoterminográficas;
- 7) Construção dos verbetes; organização do vocabulário: macro e microestruturas.

Importa destacar que os critérios metodológicos foram passando por alterações no decorrer da pesquisa, em razão de novas necessidades metodológicas que foram surgindo desde o momento em que nos propusemos a acatar as sugestões da banca na instância da qualificação da tese.

### 4.2.1 Compilação e preparação do *corpus* de estudo

Para a coleta do *corpus* desta pesquisa, seguimos os seguintes parâmetros de compilação: disponibilidade dos textos em formato digital, possibilidade de acesso sem custos ou restrições, facilidade de conversão dos textos em PDF para TXT (arquivos sem formatação) e necessidade de digitalização de algum texto na íntegra.

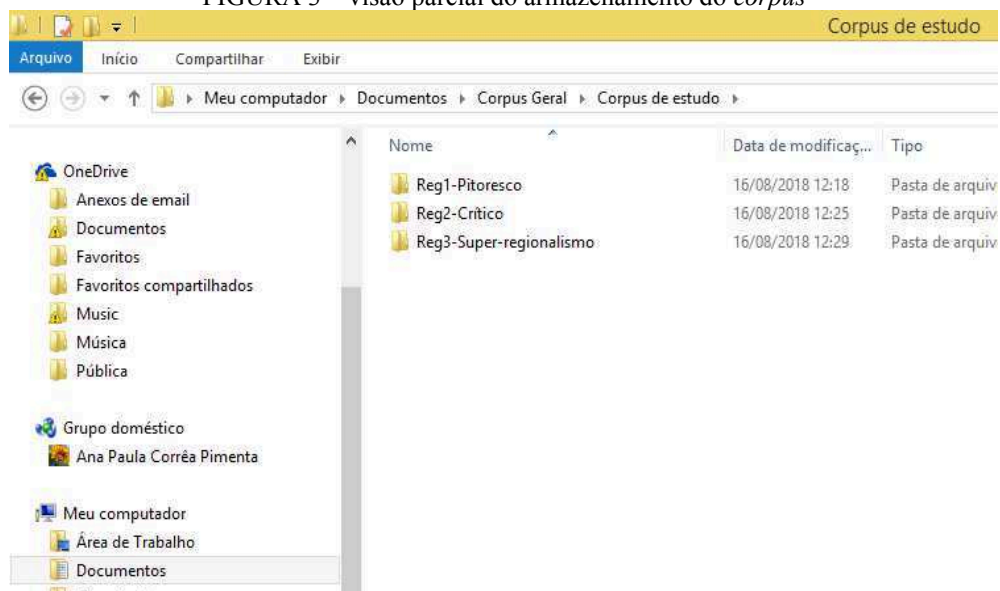
Utilizando o *Google*, encontramos várias obras em PDF provenientes de diversos sítios. Capturamos somente os arquivos que apresentavam o texto completo (legítimo) e uma boa qualidade na digitalização<sup>24</sup>. Apenas a obra “A Bagaceira” não foi encontrada, então compramos o livro e digitalizamos o texto na íntegra. Após compiladas, as obras foram salvas numa pasta denominada *Corpus geral* (sem etiquetas) ainda no formato PDF. Em seguida, fizemos a conversão dos arquivos para TXT, formato adequado para o processamento com o *WordSmith Tools*, e salvamos numa pasta denominada *Corpus de estudo*.

---

<sup>24</sup> Nessa etapa, não tivemos a preocupação em obter autorização para uso dos textos porque as obras já são consideradas de domínio público, ou seja, não são protegidas por direitos autorais.

Nessa instância de preparação do *corpus* realizamos: 1) a limpeza dos textos - retiramos elementos de natureza paratextual e partes dos textos (por exemplo, capa, lombada, dedicatórias, agradecimentos, apresentações, biografias, comentários de outros autores, notas de rodapé) dispensáveis na pesquisa, também alguns caracteres insignificantes (por exemplo: #, letras repetidas e números) e certos tipos de erros que poderiam prejudicar ou atrapalhar o andamento das análises; 2) e a inserção de siglas (títulos dos arquivos) - após a revisão e correção, organizamos os textos em subpastas com siglas correspondentes a cada fase do Regionalismo, sendo a primeira – **Reg1- Pitoresco**, com 5 textos; a segunda- **Reg2- Crítico**, com 10 textos; e a terceira- **Reg3- Super-regionalismo**, com 5 textos. Os nomes dos textos foram inseridos conforme a fase em que pertencem e segundo as iniciais do título. Assim ficou: **Reg1- IN** (correspondente à obra Inocência da fase 1 (Regionalismo Pitoresco); **Reg2- OS** (correspondente à obra Os Sertões da fase 2 (Regionalismo Crítico); e **Reg3- GSV** (correspondente à obra Grande Sertão: Veredas da fase 3 (Super-regionalismo), por exemplo. As demais obras com as suas respectivas abreviaturas constam na “Lista de abreviaturas e siglas” nas páginas iniciais deste trabalho. Nas FIGURAS 3 e 4, podemos ver uma imagem parcial do *corpus* de estudo, armazenado em (sub)pastas, de acordo com os formatos utilizados:

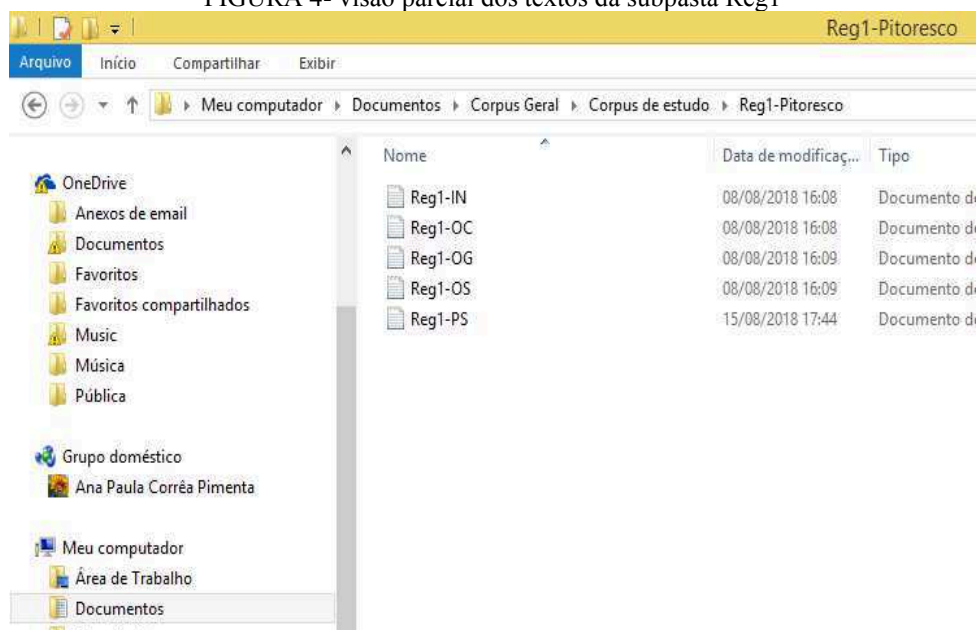
FIGURA 3 – visão parcial do armazenamento do *corpus*



Fonte: Elaboração da autora



FIGURA 4- visão parcial dos textos da subpasta Reg1



Fonte: Elaboração da autora

O *corpus* foi organizado desse modo com o objetivo de contrastar cada *subcorpus* com o *corpus* de referência de modo a gerar listas de palavras-chave separadas e, assim, compararmos as regularidades e diversidades lexicais entre as fases do regionalismo.

Os passos seguintes consistiram no levantamento das listas de palavras, palavras-chave e aplicação da ferramenta *Concord*.

#### 4.2.2 Lista de palavras

Para analisar o *corpus* por meio do programa *WordSmith Tools*, realizamos, inicialmente, o levantamento das listas de palavras por meio da ferramenta *WordList*. Na realização desse procedimento, obtivemos uma lista de todas as palavras de cada *subcorpus* com suas respectivas frequências, porcentagem de ocorrências e o número de textos em que ocorrem as palavras. Este último recurso possibilitou identificar a distribuição das ocorrências no *corpus*. Para exemplificar, fizemos um recorte das listas por ordem alfabética com as vinte e duas primeiras palavras constantes em cada *subcorpus*, como mostram as figuras a seguir:

FIGURA 5- Lista das vinte duas primeiras palavras do 1º *subcorpus*

CorpusEstudo-Fase1.lst							
File	Edit	View	Compute	Settings	Windows	Help	
N	Word	Freq.	%	Texts	%	Lemmas	Set
1	ABA	10		4	80,00		
2	ABADIA	3		1	20,00		
3	ABAFADA	1		1	20,00		
4	ABAFADO	5		3	60,00		
5	ABAFADOS	4		2	40,00		
6	ABAFANDO	6		3	60,00		
7	ABAFAR	3		3	60,00		
8	ABAFASSE	2		2	40,00		
9	ABAFAR	1		1	20,00		
10	ABAFAR	2		1	20,00		
11	ABAIXANDO	8		3	60,00		
12	ABAIXAR	2		2	40,00		
13	ABAIXAVA	1		1	20,00		
14	ABAIXO	26		5	100,00		
15	ABAIXOU	7		3	60,00		
16	ABALA	2		1	20,00		
17	ABALADA	3		3	60,00		
18	ABALADO	6		3	60,00		
19	ABALADOS	1		1	20,00		
20	ABALANÇARAM	1		1	20,00		
21	ABALANÇARIA	1		1	20,00		
22	ABALAR	3		1	20,00		

Fonte: Elaboração da autora por meio da ferramenta *WordList* do *WordSmith Tools* 6.0.

FIGURA 6- Lista das vinte duas primeiras palavras do 2º *subcorpus*

CorpusEstudo-Fase2.lst							
File	Edit	View	Compute	Settings	Windows	Help	
N	Word	Freq.	%	Texts	%	L	Set
1	ABA	10		4	40,00		
2	ABACATEIROS	1		1	10,00		
3	ABADE	11		1	10,00		
4	ABAFADA	6		5	50,00		
5	ABAFADAS	4		2	20,00		
6	ABAFADIÇA	1		1	10,00		
7	ABAFADO	8		4	40,00		
8	ABAFADOR	1		1	10,00		
9	ABAFADOS	7		5	50,00		
10	ABAFAMENTO	2		1	10,00		
11	ABAFANDO	3		1	10,00		
12	ABAFAR	1		1	10,00		
13	ABAFAVA	2		2	20,00		
14	ABAFAR	2		1	10,00		
15	ABAI	1		1	10,00		
16	ABAIXA	2		2	20,00		
17	ABAIXAMENTO	1		1	10,00		
18	ABAIXANDO	1		1	10,00		
19	ABAIXARA	1		1	10,00		
20	ABAIXAVA	1		1	10,00		
21	ABAIXEI	1		1	10,00		
22	ABAIXO	59	0,01	8	80,00		

Fonte: Elaboração da autora por meio da ferramenta *WordList* do *WordSmith Tools* 6.0.

FIGURA 7 - Lista das vinte duas primeiras palavras do 3º *subcorpus*

N	Word	Freq.	%	Texts	%	l	Set
1	ABA	15		4	80,00		
2	ABACATE	1		1	20,00		
3	ABACATEIRO	1		1	20,00		
4	ABADIA	15		2	40,00		
5	ABAETE	1		1	20,00		
6	ABAETÉ	6		1	20,00		
7	ABAF	1		1	20,00		
8	ABAFADA	1		1	20,00		
9	ABAFADIÇO	1		1	20,00		
10	ABAFADO	2		1	20,00		
11	ABAFADOS	1		1	20,00		
12	ABAFAVA	1		1	20,00		
13	ABAF	2		1	20,00		
14	ABAIXA	3		1	20,00		
15	ABAIXADA	1		1	20,00		
16	ABAIXADO	1		1	20,00		
17	ABAIXAMOS	1		1	20,00		
18	ABAIXANDO	6		3	60,00		
19	ABAIXAR	4		2	40,00		
20	ABAIXARAM	1		1	20,00		
21	ABAIXAVA	4		3	60,00		
22	ABAIXEI	2		1	20,00		

Fonte: Elaboração da autora por meio da ferramenta *WordList* do *WordSmith Tools* 6.0.

Apesar de a lista de palavras não nos dar condições de visualizar de maneira mais apurada as particularidades do *corpus* de estudo, podemos perceber possíveis candidatos a vocábulos-termos, como a palavra “abalançaram”, na vigésima posição da 1ª lista do *subcorpus* 1 (Regionalismo Pitoresco).

Contudo, para garantir a legitimidade dos dados e obter resultados mais contundentes realizamos, também, o levantamento das listas de palavras-chave, conforme explicitamos a seguir.

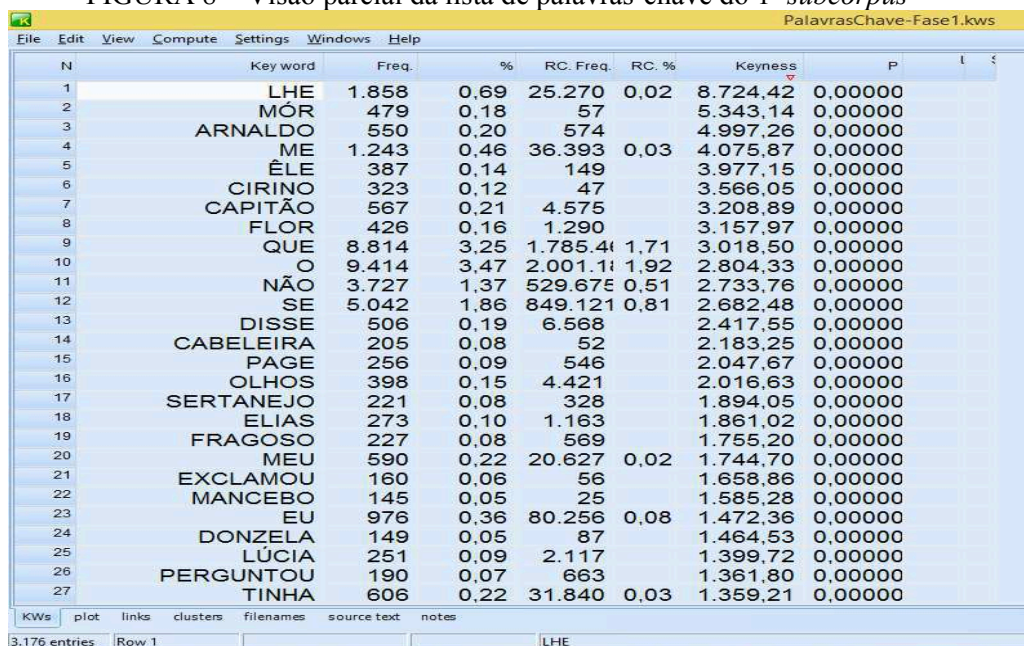
#### 4.2.3 Lista de palavras-chave

Por meio da ferramenta *KeyWords* do WST, realizamos o levantamento das palavras-chave de cada *subcorpus*. Essa ferramenta compara a lista de palavras do *corpus* de estudo com a lista de palavras do *corpus* de referência, gerando uma nova lista com as palavras consideradas chaves no *corpus* de estudo, ou seja, aquelas que caracterizam o *corpus* em seus aspectos linguísticos, temáticos e estilísticos, como observa Viana (2010):

[...] as palavras-chave geralmente são de três tipos: temáticas, gramaticais e/ou identificativas. A primeira categoria abarca todas as palavras que apontam para o assunto abordado no corpus de estudo, as quais seriam inicialmente identificadas por um analista humano. [...] As palavras gramaticais, por serem geralmente empregadas em quantidades semelhantes em ambos os corpora, não aparecem no topo da lista de palavras-chave como ocorre numa lista de palavras regular. Contudo, o surgimento delas – com altos valores de chavicidade – indica uma característica estilística do corpus estudado. [...] Finalmente, as palavras identificativas correspondem a nomes próprios: por serem formas únicas de expressão, é esperado que apareçam entre as palavras-chave (VIANA, 2010, p. 65).

Assim, utilizando como *corpus* de referência o CorpRef-AcadTeses<sup>25</sup>, obtivemos três listas correspondentes a cada *subcorpus*, com um total de 16.395 palavras-chave. As listas foram geradas de acordo com a seguinte configuração: valor de  $p$ <sup>26</sup>: 0,000001 (valor padrão do WST) e frequência mínima de 1 ocorrência. As FIGURAS 8, 9 e 10 apresentam uma visão parcial do início das listas organizadas em função da chavicidade<sup>27</sup>:

FIGURA 8 – Visão parcial da lista de palavras-chave do 1º *subcorpus*



N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P	L	S
1	LHE	1.858	0,69	25.270	0,02	8.724,42	0,00000		
2	MÓR	479	0,18	57		5.343,14	0,00000		
3	ARNALDO	550	0,20	574		4.997,26	0,00000		
4	ME	1.243	0,46	36.393	0,03	4.075,87	0,00000		
5	ÊLE	387	0,14	149		3.977,15	0,00000		
6	CIRINO	323	0,12	47		3.566,05	0,00000		
7	CAPITÃO	567	0,21	4.575		3.208,89	0,00000		
8	FLOR	426	0,16	1.290		3.157,97	0,00000		
9	QUE	8.814	3,25	1.785,41	1,71	3.018,50	0,00000		
10	O	9.414	3,47	2.001,11	1,92	2.804,33	0,00000		
11	NÃO	3.727	1,37	529,67	0,51	2.733,76	0,00000		
12	SE	5.042	1,86	849,12	0,81	2.682,48	0,00000		
13	DISSE	506	0,19	6.568		2.417,55	0,00000		
14	CABELEIRA	205	0,08	52		2.183,25	0,00000		
15	PAGE	256	0,09	546		2.047,67	0,00000		
16	OLHOS	398	0,15	4.421		2.016,63	0,00000		
17	SERTANEJO	221	0,08	328		1.894,05	0,00000		
18	ELIAS	273	0,10	1.163		1.861,02	0,00000		
19	FRAGOSO	227	0,08	569		1.755,20	0,00000		
20	MEU	590	0,22	20.627	0,02	1.744,70	0,00000		
21	EXCLAMOU	160	0,06	56		1.658,86	0,00000		
22	MANCEBO	145	0,05	25		1.585,28	0,00000		
23	EU	976	0,36	80.256	0,08	1.472,36	0,00000		
24	DONZELA	149	0,05	87		1.464,53	0,00000		
25	LÚCIA	251	0,09	2.117		1.399,72	0,00000		
26	PERGUNTOU	190	0,07	663		1.361,80	0,00000		
27	TINHA	606	0,22	31.840	0,03	1.359,21	0,00000		

Fonte: Elaboração da autora por meio da ferramenta KeyWords do WordSmith Tools 6.0.

<sup>25</sup> Trata-se de um dos segmentos do *corpus* acadêmico compilado no âmbito do projeto *Corpus Brasileiro*, de Berber Sardinha, constituído exclusivamente por textos escritos de diferentes áreas extraídos a partir de um banco de teses.

<sup>26</sup> “A coluna “p” registra o valor desse índice estatístico [chavicidade]. O conceito de p indica em que proporção o resultado encontrado é atribuído ao fator chance. Em outras palavras, quanto menor for o valor registrado para p, maior é a probabilidade de o resultado realmente expressar uma diferença entre, nesse caso, os domínios contrastados” (VIANA, 2010, p. 64).

<sup>27</sup> “A chavicidade reporta o resultado de um procedimento estatístico pelo qual a ferramenta levanta o quão importante cada palavra-chave positiva é para o *corpus* de pesquisa em relação ao de referência (e vice-versa no caso das palavras-chave negativas). Quanto maior o valor apresentado nessa coluna, maior a relevância da palavra em questão” (VIANA, 2010, p. 64).



FIGURA 9 – Visão parcial da lista de palavras-chave do 2º *subcorpus*

PalavrasChave-Fase2.kws							
File	Edit	View	Compute	Settings	Windows	Help	
N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	LHE	2.458	0,43	25.270	0,02	9.335,11	0,0000000000
2	ÊLE	640	0,11	149		5.917,51	0,0000000000
3	SE	10.276	1,81	849.121	0,81	5.129,45	0,0000000000
4	O	18.592	3,27	2.001.183	1,92	4.554,40	0,0000000000
5	OLHOS	722	0,13	4.421		3.411,86	0,0000000000
6	ERA	2.573	0,45	116.782	0,11	3.294,11	0,0000000000
7	ME	1.362	0,24	36.393	0,03	2.889,16	0,0000000000
8	ALI	783	0,14	8.954		2.822,01	0,0000000000
9	PAGE	369	0,06	546		2.623,47	0,0000000000
10	JUCUNDINA	251	0,04	0		2.619,84	0,0000000000
11	NÃO	5.983	1,05	529.675	0,51	2.533,47	0,0000000000
12	LUZIA	351	0,06	797		2.258,76	0,0000000000
13	FABIANO	275	0,05	218		2.195,88	0,0000000000
14	NOITE	580	0,10	6.157		2.167,66	0,0000000000
15	OS	7.653	1,35	793.247	0,76	2.094,17	0,0000000000
16	ESTAVA	961	0,17	25.804	0,02	2.030,18	0,0000000000
17	NUM	1.002	0,18	28.397	0,03	2.026,78	0,0000000000
18	JUVÊNCIO	216	0,04	46		2.011,55	0,0000000000
19	IA	600	0,11	8.339		1.953,10	0,0000000000
20	LÁ	750	0,13	15.095	0,01	1.952,85	0,0000000000
21	UM	6.332	1,11	634.216	0,61	1.922,79	0,0000000000
22	CABEÇA	474	0,08	5.007		1.775,62	0,0000000000
23	VELHO	447	0,08	4.334		1.743,31	0,0000000000
24	LUCAS	286	0,05	888		1.691,18	0,0000000000
25	SINHA	184	0,03	48		1.684,46	0,0000000000
26	MAS	2.045	0,36	135.685	0,13	1.542,49	0,0000000000
27	CANUDOS	198	0,03	202		1.514,34	0,0000000000
KW's plot links clusters filenames source text notes							
7.394 entries Row 1 LHE							

Fonte: Elaboração da autora por meio da ferramenta *KeyWords* do *WordSmith Tools* 6.0

FIGURA 10 – Visão parcial da lista de palavras-chave do 3º *subcorpus*

PalavrasChave-Fase3.kws							
File	Edit	View	Compute	Settings	Windows	Help	
N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	EU	5.035	0,99	80.256	0,08	16.218,85	0,00000000
2	ME	2.571	0,51	36.393	0,03	8.817,76	0,00000000
3	VEREDAS	899	0,18	417		7.942,49	0,00000000
4	SENHOR	1.568	0,31	9.204		7.864,67	0,00000000
5	DIADORIM	649	0,13	173		6.072,85	0,00000000
6	PAGE	705	0,14	546		5.805,36	0,00000000
7	SERTÃO	1.068	0,21	5.390		5.643,52	0,00000000
8	ERA	3.158	0,62	116.782	0,11	5.599,68	0,00000000
9	O	17.927	3,52	2.001.183	1,92	5.572,37	0,00000000
10	NÃO	7.165	1,41	529.675	0,51	5.460,62	0,00000000
11	MEU	1.523	0,30	20.627	0,02	5.341,57	0,00000000
12	DELE	1.116	0,22	9.395		4.871,95	0,00000000
13	LÁ	1.277	0,25	15.095	0,01	4.791,36	0,00000000
14	ELE	2.512	0,49	85.767	0,08	4.778,48	0,00000000
15	GENTE	1.495	0,29	25.582	0,02	4.617,90	0,00000000
16	QUE	15.663	3,08	1.785.461	1,71	4.555,22	0,00000000
17	BEBELO	439	0,09	91		4.193,54	0,00000000
18	MAS	2.919	0,57	135.685	0,13	4.124,16	0,00000000
19	GRILLO	432	0,08	149		3.944,03	0,00000000
20	ARIMATEIA	368	0,07	0		3.921,99	0,00000000
21	SÓ	1.675	0,33	46.557	0,04	3.759,34	0,00000000
22	CORONEL	756	0,15	4.799		3.684,66	0,00000000
23	ZÉ	487	0,10	705		3.584,83	0,00000000
24	DISSE	787	0,15	6.568		3.447,58	0,00000000
25	NEM	1.320	0,26	30.376	0,03	3.389,55	0,00000000
26	TINHA	1.324	0,26	31.840	0,03	3.298,93	0,00000000
27	MINHA	1.133	0,22	22.069	0,02	3.239,29	0,00000000
KW's plot links clusters filenames source text notes							
5.825 entries Row 7 SERTÃO							

Fonte: Elaboração da autora por meio da ferramenta *KeyWords* do *WordSmith Tools* 6.0

Convém salientar que o *corpus* de referência contém 96.669.768 itens e 620.068 formas e o *corpus* de estudo geral 1.348.165 itens e 71.442 formas, o que indica que há um equilíbrio em termos quantitativos, tal como sugerido por Berber Sardinha (2004, 2009) que o *corpus* de referência dever ser no mínimo cinco vezes maior do que o *corpus* de estudo.

Assim, fizemos listas separadas para cada *subcorpus* com o intuito de observar: (1) a chavicidade dos itens constantes em cada *subcorpus*, bem como a sua frequência; (2) a diversidade lexical, semântica e estilística de cada fase do Regionalismo, considerando as escolhas lexicais dos autores em correlação com a fase a que pertence e a região abordada na narrativa; (3) e a proeminência do léxico sertanista de um modo geral.

Por meio destas listas, conseguimos não apenas perceber as características e especificidades do *corpus*, como também levantar suposições sobre os possíveis candidatos a vocábulo-termo. Notamos, por exemplo, que as palavras “sertanejo” (na posição 17 da lista de palavras-chave da fase 1) e “sertão” (na posição 7 da lista de palavras-chave da fase 3) além de serem chaves e apresentarem uma alta frequência, apontam para a temática do *corpus* e se configuram como fortes candidatas a vocábulo-termo.

Contudo, sabemos que a lista de palavras-chave por si só não é suficiente para validar nossas hipóteses. Então, verificamos também as linhas de concordância, isto é, os contextos linguísticos e definitórios de cada item; e em seguida, recorremos aos dicionários para elucidação dos sentidos de cada um. A seção seguinte mostra como estes procedimentos foram realizados.

#### **4.2.4 Aplicação da ferramenta *Concord* e consulta aos dicionários para identificação dos vocábulos-termos**

Utilizando a ferramenta *Concord* do WST, tivemos acesso aos contextos linguísticos de ocorrência dos possíveis candidatos a vocábulos-termos. Essa ferramenta nos permite visualizar a palavra de busca no centro e no entorno textual em que é usada, tanto à direita quanto à esquerda, o que corrobora para apreensão do sentido de cada item em seu contexto de uso. Assim, por meio dessa ferramenta conseguimos não apenas verificar em quais textos o item aparecia, como recuperar o seu conceito para posterior elaboração das definições. A FIGURA 11 apresenta uma visão parcial das linhas de concordância da palavra “sertão” correspondente à lista de palavras-chave do 3º *subcorpus*.

FIGURA 11 –Lista das linhas de concordâncias da palavra “sertão” com destaque na posição 2 para um enunciado definatório

Concord													
File Edit View Compute Settings Windows Help													
N	Concordance	Word #	Sent. #	Sent. Pos.	Para.	f	Hez	Sec	Sect	f	File	Date	
1	o medo no mais afundado <b>sertão</b> de Cacimbas, covil	57.458	3.284	63%	0	49%	0	49%	0	49%	Reg3-CL.txt	2018/ago/	
2	que nem um tremer d'água. <b>Sertão</b> foi feito é para ser	164.725	13.931	20%	0	91%	0	91%	0	91%	Reg3-GS.txt	2018/ago/	
3	para se transformar aquele <b>sertão</b> inteiro do interior,	136.468	11.496	69%	0	76%	0	76%	0	76%	Reg3-GS.txt	2018/ago/	
4	universozinho nosso aqui. <b>Sertão</b> . O senhor sabe:	3.995	338	100%	0	2%	0	2%	0	2%	Reg3-GS.txt	2018/ago/	
5	de amanhecer o campo <b>sertão</b> ? A enquanto	161.976	13.674	100%	0	90%	0	90%	0	90%	Reg3-GS.txt	2018/ago/	
6	as pernas de meu cavalo. <b>Sertão</b> velho de idades.	177.660	15.000	40%	0	99%	0	99%	0	99%	Reg3-GS.txt	2018/ago/	
7	velhas, altas cidades... <b>Sertão</b> é o sozinho.	100.800	8.711	40%	0	56%	0	56%	0	56%	Reg3-GS.txt	2018/ago/	
8	de todos os cotovelos. <b>Sertão</b> , – se diz –, o	124.865	10.671	15%	0	69%	0	69%	0	69%	Reg3-GS.txt	2018/ago/	
9	Sertão é o penal, criminal. <b>Sertão</b> é onde homem tem	34.382	3.032	14%	0	19%	0	19%	0	19%	Reg3-GS.txt	2018/ago/	
10	afundava nas safadezas de <b>sertão</b> adentro. Assim, sem	26.153	1.487	96%	0	22%	0	22%	0	22%	Reg3-CL.txt	2018/ago/	
11	requerida, uma terra de <b>sertão</b> adentro, braba como	98.458	5.601	65%	0	84%	0	84%	0	84%	Reg3-CL.txt	2018/ago/	
12	nome, e ele sendo de <b>sertão</b> do mesmo nome,	52.881	4.501	72%	0	29%	0	29%	0	29%	Reg3-GS.txt	2018/ago/	
13	chuvas? Aí este mundo de <b>sertão</b> tinha se perdido –	128.297	10.905	43%	0	71%	0	71%	0	71%	Reg3-GS.txt	2018/ago/	
14	bem! – eu de mim dei. <b>Sertão</b> é isto, o senhor	49.931	4.266	18%	0	28%	0	28%	0	28%	Reg3-GS.txt	2018/ago/	
15	Otacília – a vinda dela, <b>sertão</b> a dentro, por me	185.056	453	39%	0	25%	0	25%	0	25%	Reg3-GS.txt	2018/ago/	
16	desamparadas, lá dentro. <b>Sertão</b> ! Logo que o	44.550	3.799	100%	0	25%	0	25%	0	25%	Reg3-GS.txt	2018/ago/	
17	esconso, nas toesas deste <b>sertão</b> ...” - Grande Sertão:	126.686	10.802	100%	0	70%	0	70%	0	70%	Reg3-GS.txt	2018/ago/	
18	por todos os pousos deste <b>sertão</b> ... Ah, ela vinha,	160.026	13.504	100%	0	89%	0	89%	0	89%	Reg3-GS.txt	2018/ago/	
19	por julgamento, deste <b>sertão</b> . Tudo estava sendo	143.848	12.172	100%	0	80%	0	80%	0	80%	Reg3-GS.txt	2018/ago/	
20	socorrer as infâncias deste <b>sertão</b> !” Eu ia fazer o	130.099	11.048	100%	0	72%	0	72%	0	72%	Reg3-GS.txt	2018/ago/	

Fonte: Elaboração da autora por meio da ferramenta *Concord* do *WordSmith Tools* 6.0

FIGURA 12 – Fragmento do texto onde a palavra selecionada aparece no *corpus*

C:\Users\AnaPaula\Documents\Doutorado\Corpus de estudo\Reg3-Super-regionalismo\Reg3-GS.txt													
File Edit View Compute Settings Windows Help													
<p>uns meses, a igreja, o defunto tinha se secado sozinho... Ao por tanto, que se ia, conjuntamente, Diadorim e eu, nós dois, como já disse. Homem com homem, de mãos dadas, só se a valentia deles for enorme. Aparecia que nós dois já estávamos cavalhando lado a lado, par a par, a vai-a-vida inteira. Que: coragem – é o que o coração bate; se não, bate falso. Travessia – do sertão – a toda travessia.</p> <p>Só aquele sol, a assaz claridade – o mundo limpava que nem um tremer d'água. <b>Sertão</b> foi feito é para ser sempre assim: alegrias! E fomos. Terras muito deserdadas, desdoadas de donos, avermelhadas campinas. Lá tinha um caminho novo. Caminho de gado.</p> <p>Arte que eu achei o meu projeto.</p> <p>Só digo como foi: do prazer mesmo sai a estonteação, como que um perde o bom tino. Porque, viver é muito perigoso... Diadorim, o rosto dele era fresco, a boca de amor; mas o orgulho dele condescendia uma tristeza. Matéria daquilo</p> <p>- Grande Sertão: Veredas</p>													

Fonte: Elaboração da autora por meio da ferramenta *Concord* do *WordSmith Tools* 6.0

Uma vez verificadas as linhas de concordâncias, consultamos também os dicionários para verificar a existência da forma coletada, o seu conceito e a sua relação com o universo sertanejo. Dentre os dicionários mais conhecidos e renomados, optamos pelo *Dicionário Houaiss* em sua versão eletrônica, por ser uma obra de referência nos estudos lexicográficos de língua portuguesa com um vasto repertório lexical, incluindo regionalismos, expressões e modismos usados no Brasil; e pelos dicionários *online* da Biblioteca Brasileira da USP<sup>28</sup>: *Vocabulario Portuguez e Latino*, de Raphael Bluteau (1712), *Diccionario da Lingua*

<sup>28</sup> Os respectivos dicionários podem ser acessados pelo site: <<http://dicionarios.bbm.usp.br/dicionario>>.



*Portuguesa*, de Antonio de Moraes Silva (1789) e *Diccionario da Língua Brasileira*, de Luiz Maria da Silva Pinto (1832), por apresentarem formas antigas da língua portuguesa e diversos arcaísmos constantes no nosso *corpus* de estudo.

Assim, observando as linhas de concordância e os dicionários, conseguimos identificar os vocábulos-termos e selecionar os que realmente seriam incluídos na nossa proposta de vocabulário. A FIGURA 13, 14 e 15 mostram um recorte das listas de palavras-chave no processo de limpeza (exclusão dos itens irrelevantes para a pesquisa).

FIGURA 13- Visão parcial da lista de palavras-chave do 1º *subcorpus* no processo de limpeza

PalavrasChave-Fase1.kws							
N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	LHE	1.858	0,69	25.270	0,02	8.724,42	0,0000000000
2	MÓR	479	0,18	57		5.343,14	0,0000000000
3	ARNALDO	550	0,20	574		4.997,26	0,0000000000
4	ME	1.243	0,46	36.393	0,03	4.075,87	0,0000000000
5	ÊLE	387	0,14	149		3.977,15	0,0000000000
6	CIRINO	323	0,12	47		3.566,05	0,0000000000
7	CAPITÃO	567	0,21	4.575		3.208,89	0,0000000000
8	FLOR	426	0,16	1.290		3.157,97	0,0000000000
9	QUE	8.814	3,25	1.785.461	1,71	3.018,50	0,0000000000
10	O	9.414	3,47	2.001.183	1,92	2.804,33	0,0000000000
11	NÃO	3.727	1,37	529.675	0,51	2.733,76	0,0000000000
12	SE	5.042	1,86	849.121	0,81	2.682,48	0,0000000000
13	DISSE	506	0,19	6.568		2.417,55	0,0000000000
14	CABELEIRA	205	0,08	52		2.183,25	0,0000000000
15	PAGE	256	0,09	546		2.047,67	0,0000000000
16	OLHOS	398	0,15	4.421		2.016,63	0,0000000000
17	SERTANEJO	221	0,08	328		1.894,05	0,0000000000
18	ELIAS	273	0,10	1.163		1.861,02	0,0000000000
19	FRAGOSO	227	0,08	569		1.755,20	0,0000000000
20	MEU	590	0,22	20.627	0,02	1.744,70	0,0000000000

Fonte: Elaboração da autora por meio da ferramenta *KeyWords* do *WordSmith Tools* 6.0


FIGURA 14- Visão parcial da lista de palavras-chave do 2º *subcorpus* no processo de limpeza

PalavrasChave-Fase2.kws							
N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	LHE	2.458	0,43	25.270	0,02	9.335,11	0,00000000
2	ÊLE	640	0,11	149		5.917,51	0,00000000
3	SE	10.276	1,81	849.121	0,81	5.129,45	0,00000000
4	O	18.592	3,27	2.001.183	1,92	4.554,40	0,00000000
5	OLHOS	722	0,13	4.421		3.411,86	0,00000000
6	ERA	2.573	0,45	116.782	0,11	3.294,11	0,00000000
7	ME	1.362	0,24	36.393	0,03	2.889,16	0,00000000
8	ALI	783	0,14	8.954		2.822,01	0,00000000
9	PAGE	369	0,06	546		2.623,47	0,00000000
10	JUCUNDINA	251	0,04	0		2.619,84	0,00000000
11	NÃO	5.983	1,05	529.675	0,51	2.533,47	0,00000000
12	LUZIA	351	0,06	797		2.258,76	0,00000000
13	FABIANO	275	0,05	218		2.195,88	0,00000000
14	NOITE	580	0,10	6.157		2.167,66	0,00000000
15	OS	7.653	1,35	793.247	0,76	2.094,17	0,00000000
16	ESTAVA	961	0,17	25.804	0,02	2.030,18	0,00000000
17	NUM	1.002	0,18	28.397	0,03	2.026,78	0,00000000
18	JUVÊNIO	216	0,04	46		2.011,55	0,00000000
19	IA	600	0,11	8.339		1.953,10	0,00000000
20	LÁ	750	0,13	15.095	0,01	1.952,85	0,00000000

Fonte: Elaboração da autora por meio da ferramenta *KeyWords* do *WordSmith Tools* 6.0



FIGURA 15- Visão parcial da lista de palavras-chave do 3º *subcorpus* no processo de limpeza



N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P
1	EU	5.035	0,99	80.256	0,08	16.218,85	0,0000000000
2	ME	2.571	0,51	36.393	0,03	8.817,76	0,0000000000
3	VEREDAS	899	0,18	417		7.942,49	0,0000000000
4	SENHOR	1.568	0,31	9.204		7.864,67	0,0000000000
5	DIADORIM	649	0,13	173		6.072,85	0,0000000000
6	PAGE	705	0,14	546		5.805,36	0,0000000000
7	SERTÃO	1.068	0,21	5.390		5.643,52	0,0000000000
8	ERA	3.158	0,62	116.782	0,11	5.599,68	0,0000000000
9	O	17.927	3,52	2.001,18	1,92	5.572,37	0,0000000000
10	NÃO	7.165	1,41	529.675	0,51	5.460,62	0,0000000000
11	MEU	1.523	0,30	20.627	0,02	5.341,57	0,0000000000
12	DELE	1.116	0,22	9.395		4.871,95	0,0000000000
13	LÁ	1.277	0,25	15.095	0,01	4.791,36	0,0000000000
14	ELE	2.512	0,49	85.767	0,08	4.778,48	0,0000000000
15	GENTE	1.495	0,29	25.582	0,02	4.617,90	0,0000000000
16	QUE	15.663	3,08	1.785,46	1,71	4.555,22	0,0000000000
17	BEBELO	439	0,09	91		4.193,54	0,0000000000
18	MAS	2.919	0,57	135.685	0,13	4.124,16	0,0000000000
19	GRILO	432	0,08	149		3.944,03	0,0000000000
20	ARIMATEIA	368	0,07	0		3.921,99	0,0000000000

Fonte: Elaboração da autora por meio da ferramenta *KeyWords* do *WordSmith Tools* 6.0

Convém esclarecer que, selecionamos apenas os substantivos, adjetivos e verbos com um sentido específico no *corpus* de estudo referente ao universo sertanejo. Os demais itens foram excluídos por não serem objetos desta investigação, ou seja, por não estarem dentro do escopo desta pesquisa.

As *KeyWords* eram compostas por: 3.176 itens (do 1º *subcorpus*), 7.394 (do 2º *subcorpus*) e 5.825 (do 3º *subcorpus*), somando um total de 16.395 palavras-chave, conforme demonstrado na seção anterior. Após a exclusão das palavras gramaticais e dos itens irrelevantes para a pesquisa, restaram 1.458 itens, conforme mostra a tabela a seguir.

TABELA 2 – Quantidade de itens que restaram nas listas de palavras-chave

Corpus de estudo	Quantidade de itens	Quantidade de itens que restaram
<i>Subcorpus</i> 1	3.176	212
<i>Subcorpus</i> 2	7.394	717
<i>Subcorpus</i> 3	5.825	529
Corpus Geral	16.395	1.458

Fonte: Elaboração da autora

Desses 1.458 itens, chegamos a um total de 987 vocábulos-terminos, em virtude da repetição de palavras e variantes constantes nos três *subcorpora*. Isto revela que há um número considerável de vocábulos-terminos presentes no *corpus* de estudo, o que confirma a nossa hipótese inicial sobre o construto lexical do *corpus*. Esses resultados estão dispostos, com mais detalhes, no capítulo 5, onde foram realizadas análises quantitativas, comparativas e contrastivas dos dados encontrados.

Como se trata apenas de uma proposta de vocabulário, dos 987 itens encontrados, selecionamos somente 40 vocábulos-terminos para compor a nossa amostra. Esta seleção foi proposital, ou seja, escolhemos aqueles que apresentavam características diversas, a saber: substantivo, adjetivo e verbo; baixa, média e alta frequência; presente ora em uma fase, ora em duas, ora nas três fases; dicionarizado e não dicionarizado; com uma ou mais de uma acepção; e com aspectos semânticos variados (características físicas e comportamentais do sertanejo, ocupações, objetos de uso diário, geografia do sertão, fauna, flora, mitos e lendas).

A próxima seção mostra os vocábulos-terminos selecionados para compor os verbetes de nossa proposta de Vocabulário.

#### 4.2.5 Seleção dos vocábulos-terminos

De posse da lista definitiva (das palavras-chave), reordenamos os itens em ordem de frequência, do menos frequente para o mais frequente para, em seguida, verificar quais itens iríamos selecionar para compor a nossa amostra. Como já mencionado, extraímos 40 vocábulos-terminos dos três *subcorpora* (referentes às três fases do Regionalismo) com características diversas, como mostra o quadro a seguir:

QUADRO 7: Lista dos vocábulos-terminos selecionados

Vocábulo- termo	Classe gramatical	Fases do Regionalis- mo	Frequência	Dicionarização	Nº de acepções
1- aboiar	v.	2	4	H (≠ P/S) <sup>29</sup>	1
2- aguilhada	s. f.	1, 2	7	H/B/P/S	1
3- aió	s. m.	2	10	H	1
4- amoitado	adj.	1, 3	13	H	1
5- aporrinhado	adj.	3	8	H	1
6- apostemado	adj.	3	3	n/d	1
7- arção	s. m.	1, 2, 3	28	H/B/P/S	1

<sup>29</sup> Utilizamos o sinal ≠ para indicar os vocábulos-terminos com acepções diferentes das que estão registradas nos dicionários informados.

8- arreio	s. m.	1, 2, 3	94	H/B/P/S	1
9- barbatão	s. m.	1, 2	28	H	1
10- barrigueira	s. f.	2, 3	11	H	1
11- bolandeira	s. f.	2	24	H/P/S	1
12- burgariana	s. f.	2	3	n/d	1
13- caçuá	s. m.	1, 2	22	H	1
14- cacunda	s. f.	2, 3	10	H	2
15- cangalha	s. f.	1, 2, 3	52	H/B/P/S	1
16- capiau	s. m.	3	10	H	1
17- catolé	s. m.	1	3	H	1
18- fiofó	s. m.	3	4	H	1
19- galista	adj. e s. m.	3	12	H	1
20- garajau	s. m.	3	9	H	3
21- garrota	s. f.	3	3	H	1
22- jirau	s. m.	2	30	H	3
23- jururu	adj.	2	3	H	1
24- macambira	s. f.	2, 3	34	H	1
25- maravalhas	s. f.	1	3	B/P/S	1
26- murzelo	adj. e s. m.	3	3	H/B/P/S	1
27- paiol	s. m.	1, 2, 3	78	H (≠B/P/S)	1
28- panasco	s. m.	1, 2	7	H/P/S	1
29- patativa	s. f.	1	3	H	1
30- pesteadó	adj.	3	4	H	1
31- pindaibeira	s. f.	3	4	n/d	1
32- potrilho	s. m.	2	4	H	1
33- quartau	s. m.	1, 2	6	H	1
34- rancho	s. m.	1, 2, 3	166	H/P/S	3
35- rosilho	adj.s. m	1, 2	17	H/B/P/S	1
36- suaçuapara	s. m.	3	4	H	1
37- tiririca	s. f.	3	3	H	2
38- tronqueira	s. f.	3	4	H	1
39- ururau	s. m.	3	30	H	1
40- xexéu	s. m.	2	6	H	1

Elaboração da autora

A partir da seleção feita acima, elaboramos fichas etnoterminográficas para cada vocábulo-termo e, posteriormente, construímos os verbetes, conforme os critérios apresentados nas seções a seguir. No capítulo 5, realizamos análises quantitativas, comparativas e contrastivas de todas essas unidades lexicais.

Em suma, a seleção dos vocábulos-terminos que compõem a nomenclatura de nossa amostra de vocabulário foi realizada com base no critério de que as unidades lexicais trazem significados específicos no universo etnoliterário sertanista que, muitas vezes (por se tratar de arcaísmos, regionalismos ou palavras de pouquíssimo uso), precisam ser esclarecidas para uma melhor compreensão das obras. Assim, priorizamos os itens que são reveladores de uma

identidade cultural, ou seja, aqueles que evidenciam vários aspectos da cultura sertaneja, permitindo-nos conhecer o homem do sertão, suas características físicas, comportamentais e, especialmente linguísticas.

Para fins de sistematização e análise dos itens selecionados, construímos um modelo de ficha etnoterminográfica, que será apresentada a seguir.

#### 4.2.6 Elaboração da ficha etnoterminográfica

Compartilhando do entendimento de Barros (2004) de que a ficha lexicográfica ou terminológica é um elemento de grande importância na organização dos itens fundamentais para a geração de um dado repertório lexicográfico/terminográfico e, ainda, que o modelo de ficha varia de acordo com “a natureza do projeto, as necessidades de registro das informações, a natureza da unidade linguística estudada e as características particulares da pesquisa em questão” (BARROS, 2004, p. 211), buscamos elaborar a nossa proposta de ficha etnoterminográfica.

A ficha etnoterminográfica pode ser definida como um registro completo e organizado de informações referentes a um dado vocábulo-termo para posterior elaboração do verbete. Para a construção do nosso modelo de ficha pautamo-nos na proposta de Barbosa (2004), aliando-a aos nossos objetivos e às características linguístico-textuais do *corpus*. Por se tratar de um modelo de ficha conceptual-terminológica, fizemos algumas adaptações e complementamos com alguns campos necessários à realização da pesquisa, como por exemplo, os campos que indicam a frequência e o registro em dicionários.

Assim, priorizando o número de ocorrências dos vocábulos-termos, as informações gramaticais, as abonações (o contexto linguístico de uso) e as definições trazidas pelos dicionários, chegamos ao seguinte modelo de ficha:

QUADRO 8- Modelo de ficha etnoterminográfica

Número da ficha – <b>vocábulo-termo</b> (Classe gramatical) _____ frequência
<i>Abonações</i>
Registro em dicionários
1. Houaiss:
2. Bluteau
3. Silva Pinto

4. Moraes Silva

Nota: (campo opcional)

Fonte: Elaboração da autora

Constituição da ficha:

- a) Ao lado esquerdo, colocamos o número da ficha, o vocábulo-termo e, entre parênteses, sua classe gramatical; e do lado direito, a frequência, ou seja, o número de ocorrências.
- b) No campo “abonações” (ou contextos de uso) apresentamos, em itálico, os contextos linguísticos de ocorrência da palavra-entrada, ou seja, um recorte do trecho em que o vocábulo aparece no *corpus* seguido do título do conto (abreviado) e da linha de concordância coletada. Não estipulamos uma quantidade exata de abonações. O número de abonações depende do nível de abstração do vocábulo-termo, isto é, nos casos em que não foi possível apreender o seu sentido (no contexto linguístico) com apenas uma abonação, transcrevemos quantas julgamos necessárias para extração do conceito.
- c) Logo abaixo, no campo “registro em dicionários”, inserimos as definições constantes nos dicionários consultados. Para os itens não dicionarizados, empregamos a expressão “n/d” e para aqueles registrados com um sentido diferente do empregado no texto, utilizamos a expressão “n/e” (não encontrado). Para os vocábulos com duas ou mais acepções, construímos fichas separadas para cada conceito.
- d) Por fim, criamos o campo “Nota” (item opcional) para inserir observações e esclarecimentos sobre certos aspectos dos vocábulos-termos.

As fichas dos vocábulos-termos selecionados estão dispostas na seção 5. 3 do capítulo 5.

Importa destacar que a ficha etnoterminográfica foi um elemento de grande importância na elaboração dos verbetes. Além de servir como um instrumento para a sistematização e descrição dos dados, ela nos possibilitou analisar cada item separadamente, se dicionarizado ou não, o número de ocorrências, as acepções de cada um no contexto, enfim, se seria incluído no vocabulário ou não. Fato é que ela foi a base para a elaboração das definições.

As próximas seções tratam do público-alvo, da construção do verbete, da macro e microestrutura da proposta de Vocabulário e da elaboração da página *Web* para disponibilização do produto.

### 4.3 O público-alvo

A produção de toda obra terminográfica ou lexicográfica requer, antes de tudo, que se estabeleçam parâmetros e objetivos que garantam condições de exequibilidade do projeto e a sua eficácia e, principalmente, que se determine o protótipo do público-alvo. Segundo Barros (2004) são os objetivos e o público-alvo que determinam as características essenciais de uma obra lexicográfica/terminográfica: a sua organização interna, a estrutura dos verbetes, o tipo de definição, a linguagem empregada. Nas palavras da autora:

As características essenciais de uma obra terminográfica dependem fundamentalmente de seus objetivos e do público que se deseja atingir [...] Cada obra terminográfica é única; a sua organização interna, as informações que transmite, a linguagem que emprega, tudo depende de seus objetivos e público-alvo, cujos processos de determinação são complementares, um não podendo existir sem o outro (BARROS, 2004, p. 191, 192).

Assim, por se tratar de um repertório que traz significados específicos do universo etnoliterário sertanista, presumimos que este protótipo de vocabulário poderá servir como auxílio a alunos e professores de literatura na interpretação do texto literário e, também, como subsídio a outros interessados pelo estudo ou pelo produto etnoterminológico, como: linguistas, lexicógrafos; terminógrafos.

Apesar de prevermos os usuários de nosso vocabulário não podemos deixar de considerar que ele poderá ser de interesse para outros consulentes, como por exemplo: escritores e críticos literários, estudiosos da indústria cultural, antropólogos, sociolinguistas, dentre outros.

É válido destacar que, dado o escopo da pesquisa, não consultamos o público-alvo em relação aos tipos de informações que ele julgaria importantes em um vocabulário, tampouco o tipo de definição que melhor o auxiliasse no entendimento do texto literário. Assim, julgamos com base em Barbosa (2001) e Barros (2004), que a definição por compreensão seria a mais adequada ao nosso tipo de repertório.

A definição por compreensão, como já mencionamos, é um tipo de definição baseada na compreensão de um conceito, ou seja, no conjunto de características que constitui um conceito, seguindo o modelo clássico *gênero próximo + diferenças específicas* (BARROS, 2004). Por essa razão, é considerada por Barbosa (2001) e Barros (2004) a mais apropriada aos trabalhos terminológicos. Em outras palavras, ela situa o conceito no âmbito de uma classe e adiciona as características que o distingue de outros conceitos situados na mesma

classe, isto é, especifica o conceito genérico mais próximo que já tenha sido definido ou que pode ser assumido como de conhecimento geral e adiciona as características de restrição que delimitam o conceito a ser definido (FROMM, 2007).

Por tudo isso, entendemos que essa espécie de definição é a mais adequada à nossa proposta de descrição etnoterminográfica, uma vez que permite ao consulente ampliar seu horizonte de informações sobre o vocábulo-termo no campo maior de suas associações (linguísticas e socioculturais).

Ademais, ainda que haja algumas limitações, acreditamos que, por se tratar de uma proposta de Vocabulário *online*, este trabalho poderá servir como base para outras pesquisas e trabalhos etnoterminográficos futuros, podendo ser estendido ou replicado à medida que novas investigações surgirem.

#### **4.4 Construção dos verbetes e organização do Vocabulário: micro e macroestruturas**

Antes de apresentar a nossa proposta de verbete, convém retomar algumas informações sobre a estruturação dos repertórios lexicais. Segundo Barbosa (2001), um repertório lexical, seja lexicográfico ou terminográfico, é constituído por dois elementos básicos: a macroestrutura e a microestrutura. Por macroestrutura, entende-se a organização da obra como um todo: “às características gerais do repertório, ou seja, à estruturação das informações em verbetes, à presença ou não de anexos, índices remissivos, ilustrações, setores temáticos, mapa conceptual e outros” (BARROS, 2004, p. 151).

A microestrutura, por sua vez, refere-se à organização interna de um verbete. De acordo com Barros (2004, p. 152), “os verbetes reúnem os dados relativos à unidade lexical ou terminológica descrita e compõem-se de, pelo menos, dois elementos: **entrada e o enunciado lexicográfico/terminográfico**”. Entende-se por entrada a unidade lexical ou terminológica a ser explicada (palavra-entrada que encabeça um verbete) e por enunciado lexicográfico/terminográfico o conjunto de informações fornecidas sobre ela.

Para a construção dos verbetes, os seguintes aspectos foram considerados:

- a) Os objetivos da pesquisa;
- b) A natureza e as características linguístico-textuais do *corpus*;
- c) O protótipo de público-alvo.

Partindo de nossa ficha etnoterminográfica anteriormente caracterizada, construímos um modelo de verbete com base na configuração das entradas apresentadas por Barbosa (2001, p. 39):



Níveis de análise	Conjunto de unidades lexicais	Unidades padrão	Estatuto semiótico -sinótico das unidades padrão	Tipo de obra lexicográfica	Microestrutura	Macroestrutura	Sistema de remissivas
Sistema	Universo léxico	Lexema	Forma semiótico-sintática ampla (sobressensu polissêmico)	Dicionários de língua	Artigo = [ + Enrada (lexema) + Enunciado lexicográfico (+ Par. Inf. 1 (pronúncia, abreviatura, categoria, gênero, número, etimologia, homônimos, campos léxico-semânticos, etc.) + Par. Definicional (acepção <sub>1</sub> , acepção <sub>2</sub> , ... acepção <sub>n</sub> ) + Par. Pragmático, +/- Par. Inf. 2, Par. Inf. n) +/- Remissivas da cadeia interpretante de língua] Lexema <sub>n</sub> = { V <sub>1</sub> , V <sub>2</sub> , ... V <sub>n</sub> (acepções) }	Lexema <sub>1</sub> = { V <sub>1</sub> , V <sub>2</sub> , ... V <sub>n</sub> (acepções com núcleo semântico comum) } Lexema <sub>2</sub> = { V <sub>1</sub> , V <sub>2</sub> , ... V <sub>n</sub> (acepções) } .....	Remissivas da cadeia interpretante de língua
Normas	Conjuntos-vocabulários ou Conjuntos terminológicos	Vocabulo  Termo	Forma semiótico-sintática restrita e caracterizadora de um universo de discurso (sensu de UD)	Vocabulários fundamentais Vocabulários técnico-científicos Vocabulários especializados	Artigo = [ + Enrada (vocabulo) + Enunciado lexicográfico (+ Par. Inform. 1 (pronúncia, abreviatura, categoria, gênero, número, etimologia, área, domínio, subdomínio etc.) + Par. Definicional (acepção específica da área científica, tecnológica ou de um falar especializado) +/- Par. Pragmático (exemplo de emprego específico daquela área) +/- Par. Inform. 2 (frequência, normalização, banalização / vulgarização popularização, etc.) ... +/- Par. Inform. n) + Remissivas (relativas ao Universo de Discurso em questão)].	Vocabulo <sub>1</sub> = acepção restrita e caracterizadora de um UD Vocabulo <sub>2</sub> = acepção restrita e caracterizadora de um UD ..... Vocabulo <sub>n</sub> = acepção restrita e caracterizadora de um UD	Remissivas relativas ao universo de Discurso
Falar	Conjuntos-ocorrência	Palavra	Forma semiótico-sintática específica de um ato de fala, de um discurso manifestado (epissensu)	Glossário	Artigo = [ + Enrada (palavra-ocorrência) + Enunciado Lexicográfico (+ Par. Inform. 1 (categoria, gênero, número, pronúncia, etimologia, etc.) + Par. Definicional (sentido da palavra naquele discurso concreto) - Par. Pragmático, +/- Par. Inform. n, +/- Remissivas (circunscritas ao texto em questão)] de um discurso manifestado	Palavra <sub>1</sub> = acepção específica de um discurso manifestado Palavra <sub>2</sub> = acepção específica de um discurso manifestado ..... Palavra <sub>n</sub> = acepção específica	Remissivas circunscritas ao texto do discurso manifestado

Segundo o modelo proposto por Barbosa (2001), a microestrutura do verbete de um vocabulário é composta por três componentes básicos (ou paradigmas): I) o paradigma informacional (PI) - informações ortográficas, fonéticas, etimológicas, cronológicas, gramaticais, etc. ; II) o paradigma definicional (PD) – definição; ); III) e o paradigma pragmático (PP) - exemplos de uso e abonações.

Seguindo esse modelo, chegamos na seguinte fórmula de composição do nosso verbete:

**Verbetes = [+ entrada + enunciado etnoterminológico (+ PI + PD + PP)]**

Assim temos:

- a) PI: Informacional (dicionarização, classe gramatical, gênero, indicação das obras literárias em que o vocábulo-termo ocorre, fase do Regionalismo, frequência no *corpus*);
- b) PD: Definicional (definição etnoterminológica<sup>30</sup>);
- c) PP: Pragmático (exemplos retirados diretamente do *corpus*).

O QUADRO 9 mostra um exemplo do verbete de acordo com os paradigmas apresentados acima. Com o intuito de apresentar um exemplo mais completo de verbete, escolhemos o vocábulo-termo com o maior número de campos preenchidos nas fichas etnoterminográficas.

QUADRO 9 –Exemplo de verbete

**CANGALHA** (H/B/P/S) s. f. Armação de madeira que se coloca sobre o dorso dos animais para transporte de carga. “*Levantou-se o ofertante com toda a boa vontade e às apalpadelas começou a procurar a cama do patrão, o que só conseguiu depois de ter esbarrado na mesa e numas **cangalhas** velhas atiradas a um canto da sala*” (IN, Reg1; freq. 52).

Fonte: Elaboração da autora

<sup>30</sup> Vale reiterar que, para a elaboração da definição etnoterminológica, partimos do entendimento de Barros (2004) que a definição mais apropriada para elaboração dos vocabulários técnicos, científicos e especializados é a **definição específica** ou **definição por compreensão** – aquela que “descreve o termo por meio de traços distintos (características)” [...], ou seja, “descreve única e exclusivamente uma dada unidade terminológica e não outra” (BARROS, 2004, p. 171).

Assim, para a versão *online*, organizarmos a microestrutura de nossa proposta de Vocabulário da seguinte forma:

- As entradas estão dispostas em ordem alfabética, com letras maiúsculas e em negrito. São apresentadas da mesma forma em que aparecem no *corpus*, salvo os verbos que, por causa da diversidade de formas, foram colocados na forma infinitiva; e os substantivos e adjetivos no singular;
- Após a entrada é indicado, entre parênteses, se o vocábulo é dicionarizado pelo Houaiss (H), pelo Bluteau (B), Pinto (P) Silva (S) ou se não é dicionarizado (n/d);
- Em seguida, é exibida a classe gramatical por uma abreviatura. As abreviaturas e convenções estão dispostas em uma lista nas páginas iniciais deste trabalho;
- Após esses elementos, aparece a definição do vocábulo-termo segundo o seu contexto linguístico de uso;
- Posteriormente, é mostrado o contexto linguístico (as abonações) em que o vocábulo aparece no *corpus*, seguido do título da obra (abreviada) e, ainda, a fase do Regionalismo e o número de ocorrências nos *corpus* de estudo.

Segue o exemplo de verbete na versão *online*:

FIGURA 17 – Verbetes do vocábulo-termo “cangalha” na versão *online*

The screenshot shows the GECON online interface. The top bar is blue with 'GECON' on the left and a search bar 'Pesquisar vocábulo-termo' in the center. On the left side, there is a sidebar with a user icon and three menu items: 'Dados da pesquisa', 'Vocabulário', and 'Corpus: leitura'. The main content area is titled 'Interface Vocabulário' and is divided into three columns.

The first column on the left lists various terms in a vertical list, including ABOIAR, AGUILHADA, AIÓ, AMOITADO, APORRINHADO, APUSTEMADO, ARÇÃO, ARREIO, BARBATÃO, BARRIGUEIRA, BOLANDEIRA, BURGARIANA, CAÇUÁ, CACUNDA, CANGALHA, CAPIAU, CATOLÉ, FIOFÓ, GALISTA, GARAIAU, and GARROTA. The term 'CANGALHA' is highlighted.

The second column in the center is titled 'Acepções' and shows the definition for 'CANGALHA': (H/B/P/S) s. f. Armação de madeira que se coloca sobre o dorso dos animais para transporte de carga. "Levantou-se o ofertante com toda a boa vontade e às apalpadelas começou a procurar a cama do patrão, o que só conseguiu depois de ter esbarrado na mesa e numas **cangalhas** velhas atiradas a um canto da sala" (IV, Reg!; freq.52).

The third column on the right is a table with three columns: 'Obra', 'Sigla', and 'Autor'. It lists several works where the term 'cangalha' appears.

Obra	Sigla	Autor
Inocência	IN	Visconde de Taunay
O Sertanejo	OS	José de Alencar
O Cabeleira	OC	Franklin Távora
Pelo Sertão	PS	Afonso Arinos
Luzia-Homem	LH	Domingos Olímpio
Contos Gauchescos	CG	João Simões Lopes Neto
Tropas e boiadas	TB	Hugo de Carvalho Ramos

Fonte: <https://www.ileel.ufu.br/gecon>

A próxima seção explica, de maneira detalhada, como a página foi elaborada e quais as formas de consulta, isto é, a macroestrutura de nossa proposta de Vocabulário *online*.

Sintetizando e complementando o que foi exposto, reunimos no QUADRO 10 as principais características da nossa proposta de Vocabulário:

QUADRO 10 –Caracterização do Vocabulário

Critérios	Características
Objetivos	Divulgação de produtos etnoterminológicos do universo sertanejo
Público-alvo	estudiosos de literatura (alunos e professores); linguistas, lexicógrafos; terminógrafos
Universo de discurso	Literatura regionalista sertanista
Temática	O sertanejo e o seu sertão
Atitude linguística	Descritiva
Natureza dos dados	Etnoterminológicos
Organização dos dados	Direcionada por um <i>corpus</i> textual monolíngue (português); baseada em fichas etnoterminográficas
Acesso	meio eletrônico ( <i>online</i> )

Fonte: Elaboração da autora com base em Rondeau (1984 apud KRIEGER; FINATTO, 2004)

A seguir, o passo a passo de como elaboramos a página *Web* para a disponibilização do produto e como foi realizada a inserção dos dados para habilitação das duas interfaces: Vocabulário e Leitura.

#### 4.4.1 A Macroestrutura: elaboração da página *Web*

Ao propor qualquer obra lexicográfica, terminográfica ou como no nosso caso, etnoterminográfica, o pesquisador deve pensar previamente no tipo de saída ou meio pelo qual o produto será publicado/acessado. Levando em consideração que, nos dias de hoje, o meio eletrônico vem substituindo o meio impresso, com a propagação da Internet e os diversos serviços que ela tem oferecido, decidimos disponibilizar a nossa proposta em formato eletrônico para o acesso *online*.

Apesar de reconhecer as limitações<sup>31</sup> de nossa proposta acreditamos que, uma vez concretizada, ela possa servir como auxílio a alunos e professores de literatura nas atividades de leitura e interpretação do texto literário e, também, como subsídio a outros interessados, como linguistas, lexicógrafos e terminógrafos. Em primeiro lugar, por ser um trabalho inédito que trata de um léxico específico do universo etnoliterário sertanista, ainda pouco explorado e até então não registrado e viabilizado sob a forma de vocabulário ou glossário; em segundo lugar, por estar em formato eletrônico, o que propicia rapidez e facilidade no acesso, bem como a redução de custos; em terceiro lugar, por apresentar uma estrutura simples de fácil acesso e consulta; e por último, e talvez a mais importante, por apresentar duas interfaces uma de vocabulário e outra de leitura que permite ao consulente ter acesso tanto ao vocabulário quanto às obras literárias para leitura, como será explicado a seguir.

Apresentamos, a seguir, o passo a passo de como elaboramos o sistema (ambiente *web*) para disponibilização do produto, compreendendo a inserção dos dados feita pelo administrador para habilitação das duas interfaces (vocabulário e leitura), a descrição de seu funcionamento e os modos de consulta.

#### 4.4.2. Elaboração do sistema

Como qualquer outro produto eletrônico (dicionário, vocabulário ou glossário), percorremos várias etapas para a construção de um modelo de consulta que atendesse aos objetivos da pesquisa e ao protótipo do público-alvo. Entre essas etapas, uma das mais importantes foi decidir o design de *layout web* para disponibilização do produto. Pensando em

---

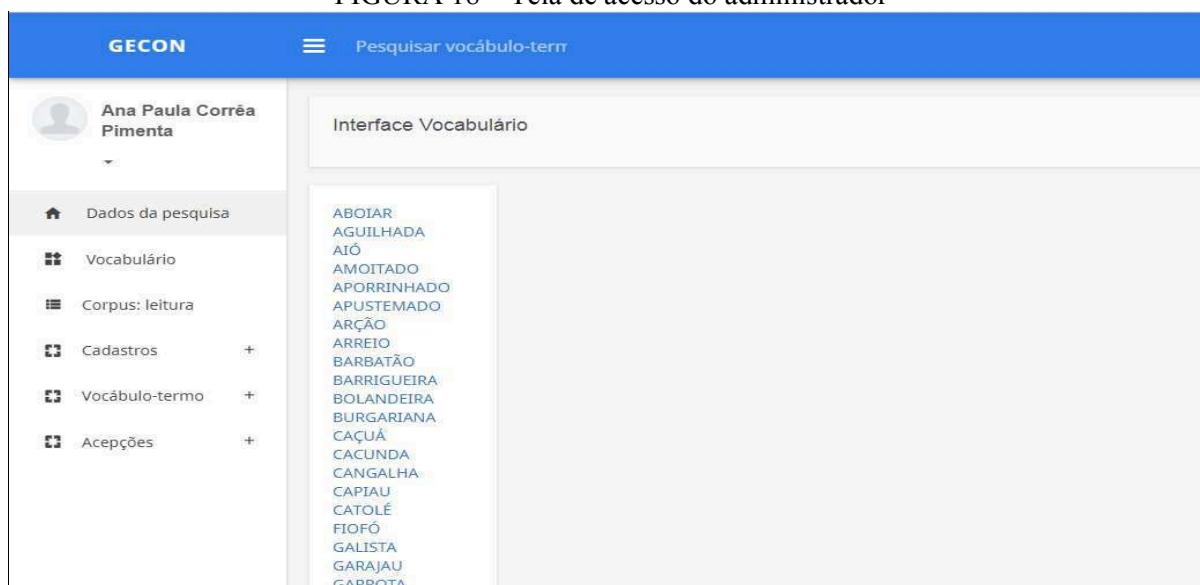
<sup>31</sup> Como já mencionamos, dado o escopo da pesquisa e por se tratar de uma proposta, não consultamos o público-alvo a respeito do modelo sugerido ou do tipo de definição que ele julgaria ideal para esse repertório. A nossa intenção é continuar o trabalho depois da defesa. Assim, teremos condições de realizar consultas com o público-alvo e, também, o cadastro de todos os vocábulos-terms encontrados no nosso *corpus* de estudo.

sua acessibilidade e aplicabilidade, decidimos, então, planejar o acesso aos dados por meio de duas interfaces: a interface **Vocabulário**, para o auxílio dos consulentes na apreensão dos conceitos; e a interface **Leitura**, trazendo o *corpus* de estudo com todas as obras e os respectivos verbetes ativados por meio de janelas *pop-up*, para o auxílio de alunos e professores nas atividades de leitura e interpretação do texto literário.

Essas duas interfaces, além de servirem como modelo para o desenvolvimento de futuras ferramentas de visualização do produto para diferentes usuários, oferecem facilidades e possibilidades de consulta, seja na rapidez do acesso, seja na forma como os dados são apresentados, por essa razão, optamos por esse formato.

Para a habilitação desses dois modos de exibição, começamos pela construção das telas para a inserção dos dados. Devido à complexidade do trabalho, buscamos o auxílio de um técnico de informática<sup>32</sup> preparado para desenhar o *layout* conforme pensamos. Assim, construímos a primeira tela de acesso do administrador, conforme mostra a FIGURA 18 :

FIGURA 18 – Tela de acesso do administrador



Fonte: <https://www.ileel.ufu.br/gecon/login>

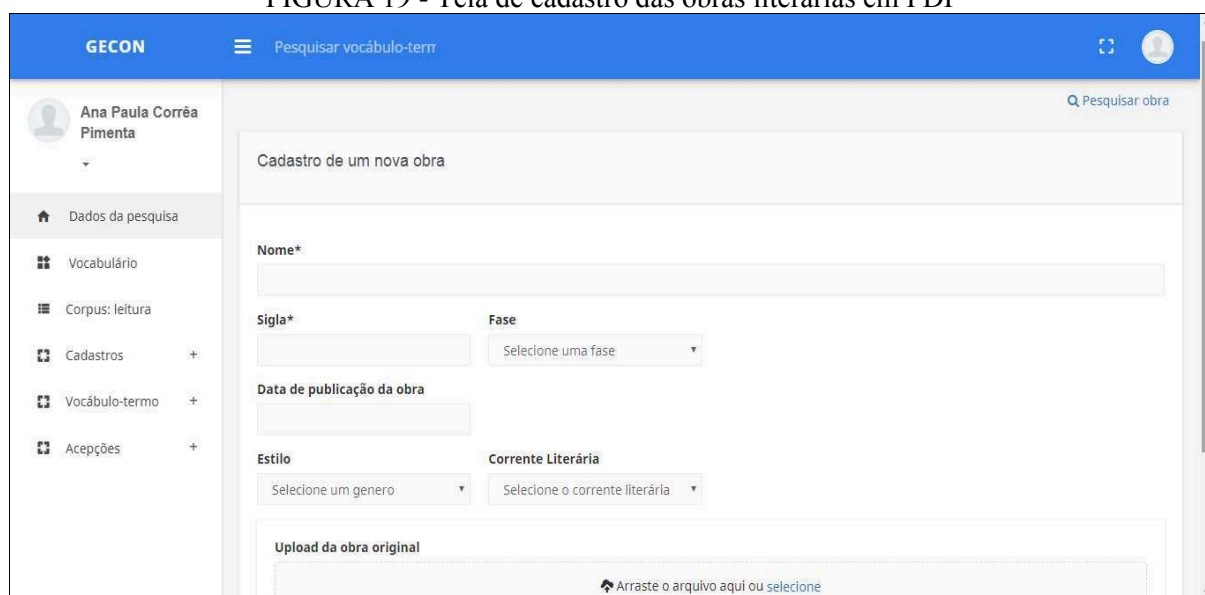
Como podemos ver, há várias opções para cadastramento na tela inicial do administrador, a saber: Dados da pesquisa, Vocabulário, Corpus-leitura, Cadastros, Vocábulo-termo e Acepções. Uma vez elaborada, começamos a inserir os dados de cada item do menu.

<sup>32</sup> Contratamos o profissional Heitor Carvalho de Almeida Neto, técnico de informática e funcionário do ILEEL, para nos auxiliar durante toda essa etapa. A sua ajuda foi de extrema importância para a concretização deste trabalho. Por isso, nada menos que os nossos sinceros agradecimentos.

Em **Dados da pesquisa**, cadastramos o trabalho completo, a linha de pesquisa, o autor e as fichas etnoterminográficas, em arquivos separados.

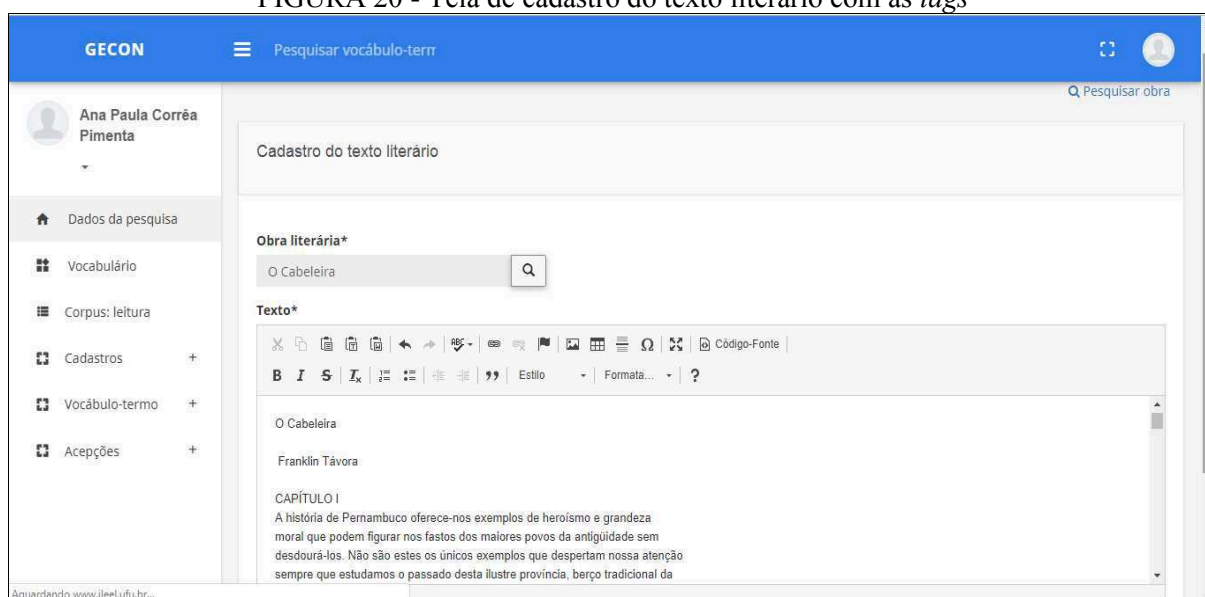
No subitem “Obras”, dentro do item **Cadastros**, inserimos todas as obras em PDF (FIGURA 19), e no subitem “Texto literário”, inserimos os textos já com as *tags* dos vocábulos-termos para leitura do programa e posterior habilitação das janelas *pop-up* (FIGURAS 20). Para exemplificar, exibimos as figuras a seguir:

FIGURA 19 - Tela de cadastro das obras literárias em PDF



Fonte: <https://www.ileel.ufu.br/gecon/login>

FIGURA 20 - Tela de cadastro do texto literário com as *tags*



Fonte: <https://www.ileel.ufu.br/gecon/login>

Após o carregamento de todas as obras, realizamos o cadastro dos vocábulos-termos para a habilitação da interface Vocabulário. Para cada vocábulo-termo, preenchemos os cinco campos da seguinte tela:

FIGURA 21 – Tela de cadastro dos vocábulos-termos

GECON Pesquisar vocábulo-termo

Ana Paula Corrêa Pimenta

Dados da pesquisa

Vocabulário

Corpus: leitura

Cadastros

Vocabulo-termo

Acepções

Cadastro de um novo vocábulo-termo

Trabalho

A REPRESENTAÇÃO DO LÉXICO SERTANISTA

Vocabulo-termo\*

ARREIO

Dicionário\*

(H/B/P/S)

Frequência\*

94,00

Informação gramatical\*

s.m.

Salvar

Gecon © 2019

Fonte: <https://www.ileel.ufu.br/gecon/login>

Da mesma forma, realizamos o cadastro das acepções:

FIGURA 22 – Tela de cadastro das acepções

GECON Pesquisar vocábulo-termo

Ana Paula Corrêa Pimenta

Dados da pesquisa

Vocabulário

Corpus: leitura

Cadastros

Vocabulo-termo

Acepções

Cadastro de uma nova acepção

Tag name\*

arreio

Obra literária\*

O Sertanejo

Vocabulo-termo\*

ARREIO

Definição

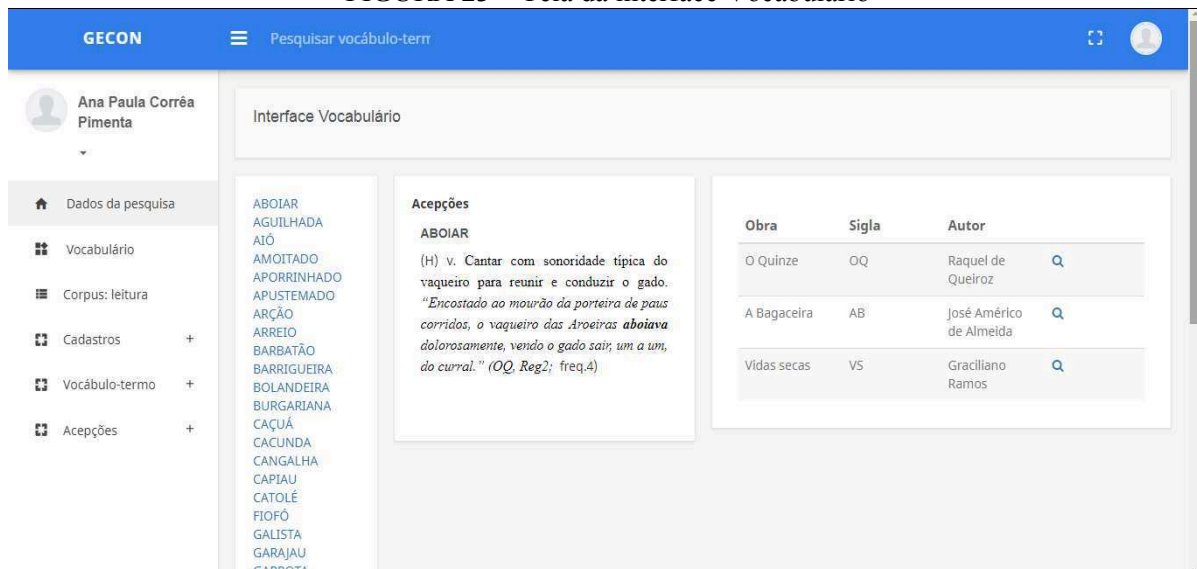
Conjunto de equipamentos aplicados nos animais de sela que os sujeita à carga e aos comandos do homem (arrieiro). "Elias saltou fora dos arreios; o cavalo levantou-se imediatamente [...]" (OG, Regl)

Fonte: <https://www.ileel.ufu.br/gecon/login>



Assim, uma vez cadastrados todos os vocábulos-termos e suas acepções, chegamos à seguinte tela da interface Vocabulário:

FIGURA 23 – Tela da interface Vocabulário



Fonte: <https://www.ileel.ufu.br/gecon/login>

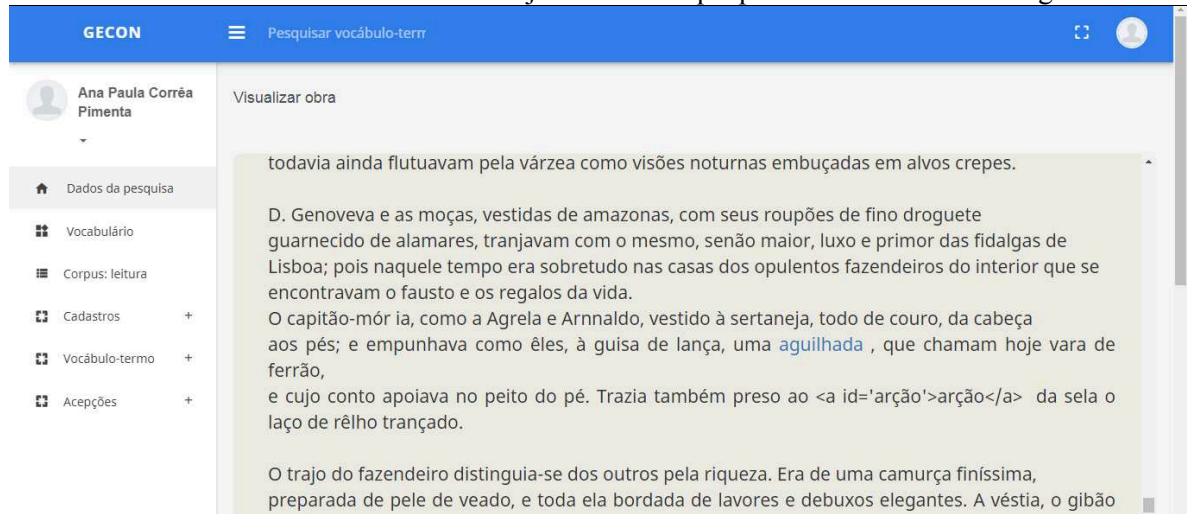
Podemos observar na figura acima que, além do verbete, na mesma janela do lado direito, aparecem os textos em que o vocábulo-termo “abojar” ocorre. Nesse ambiente, criamos um *link* que permite ao consulente ir diretamente para a interface Leitura, possibilitando o acesso dos textos literários com os vocábulos-termos e suas definições em janelas *pop-up*. A seguir, alguns exemplos:

FIGURA 24 – Tela da interface Vocabulário com os textos literários onde o vocábulo-termo aparece



Fonte: <https://www.ileel.ufu.br/gecon/login>

FIGURA 25 – Texto literário “O Sertanejo” com destaque para o vocábulo-termo “aguilhada”



Fonte: <https://www.ileel.ufu.br/gecon/login>

FIGURA 26 – Janela *pop-up* do vocábulo-termo “aguilhada”



Fonte: <https://www.ileel.ufu.br/gecon/login>

Na interface Leitura, no item **Corpus: leitura**, inserimos todas as obras literárias por fases do Regionalismo, informando a autoria, a sigla e a data de publicação. Nessa interface, como já foi dito, disponibilizamos todos os textos em PDF e, também, em HTML com os vocábulos-termos e suas respectivas definições em janelas *pop-up*. Assim, o consulente tem a opção de ler o texto literário no arquivo PDF sem as *pop-up* ou ler diretamente no sistema, visualizando as definições. Para exemplificar, exibimos as figuras abaixo:

FIGURA 27 – Tela da interface Leitura com destaque nas obras do Regionalismo Pitoresco

The screenshot shows the GECON interface. At the top, there's a blue header with 'GECON' and a search bar 'Pesquisar vocábulo-term'. On the left, a sidebar shows the user 'Ana Paula Corrêa Pimenta' and a menu with options like 'Dados da pesquisa', 'Vocabulário', 'Corpus: leitura', and 'Cadastros'. The main area is titled 'Listagem de obras' and displays a table of works.

#	Nome	Autor	Fase	Sigla	Data da publicação
	O Garimpeiro	Bernardo Guimarães	REGIONALISMO PITORESCO (fase 1)	OG	1872
	Inocência	Visconde de Taunay	REGIONALISMO PITORESCO (fase 1)	IN	1872
	O Sertanejo	José de Alencar	REGIONALISMO PITORESCO (fase 1)	OS	1875
	O Cabeleira	Franklin Távora	REGIONALISMO PITORESCO (fase 1)	OC	1876
	Pelo Sertão	Afonso Arinos	REGIONALISMO PITORESCO (fase 1)	PS	1898

Below the table, there's a pagination control showing '1'.

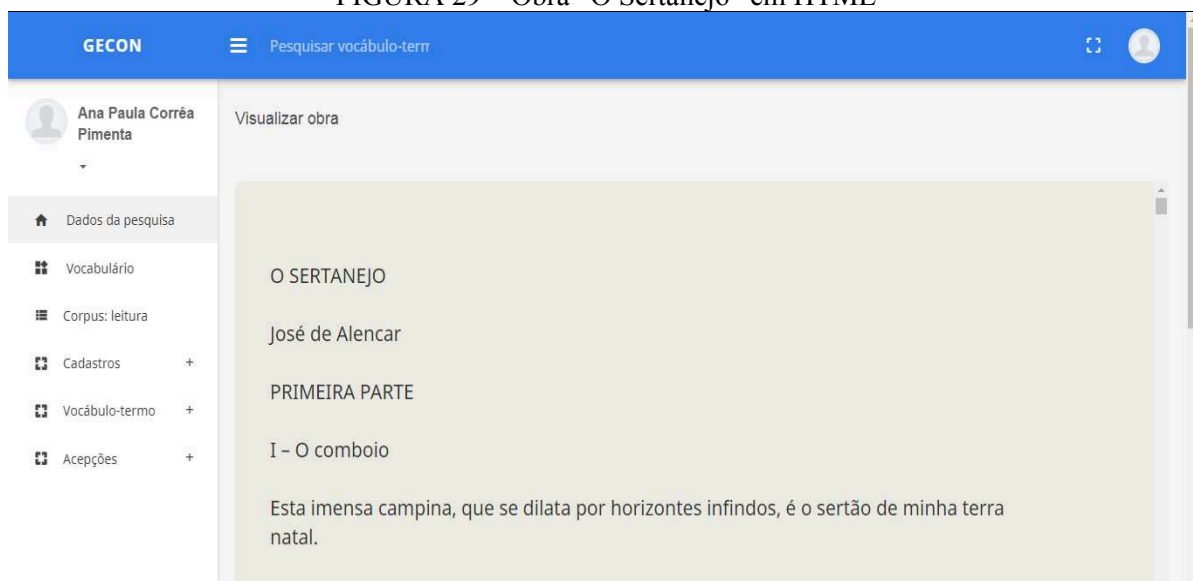
Fonte: <https://www.ileel.ufu.br/gecon>

FIGURA 28 – Obra “O Sertanejo” no arquivo PDF

The screenshot shows a PDF viewer displaying the title page of 'O Sertanejo' by José de Alencar. The title is centered at the top, followed by the author's name. Below that, the text 'PRIMEIRA PARTE' is centered. The main body of the page contains the beginning of the story, starting with 'I – O comboio'. The text is justified and includes several paragraphs. On the right side of the page, there are three circular icons for navigation: a magnifying glass, a plus sign, and a minus sign.

Fonte: <https://www.ileel.ufu.br/gecon>

FIGURA 29 – Obra “O Sertanejo” em HTML



Fonte: <https://www.ileel.ufu.br/gecon>

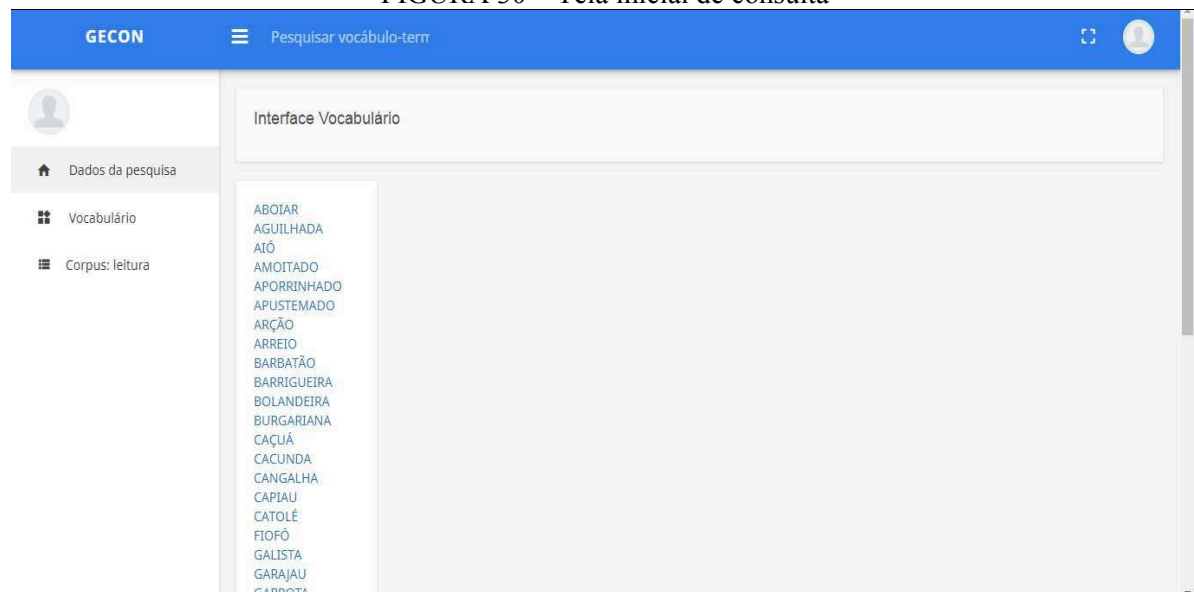
A próxima seção traz um resumo de como a página é visualizada pelo consulente e quais os modos e os procedimentos para realizar a consulta.

#### 4.4.3 Apresentação da página e o modo de consulta

Ao acessar a página o consulente se depara com uma tela inicial de consulta (FIGURA 30). Essa tela é semelhante à tela de cadastro do administrador, porém sem as possibilidades de inserção de novos verbetes ou cadastramento de novas categorias. Do lado esquerdo da tela aparece um menu com o seguintes ícones: Dados da pesquisa, Vocabulário e Corpus: leitura<sup>33</sup>.

<sup>33</sup> O endereço provisório de acesso ao site é < <http://www.ileel.ufu.br/gecon> >.

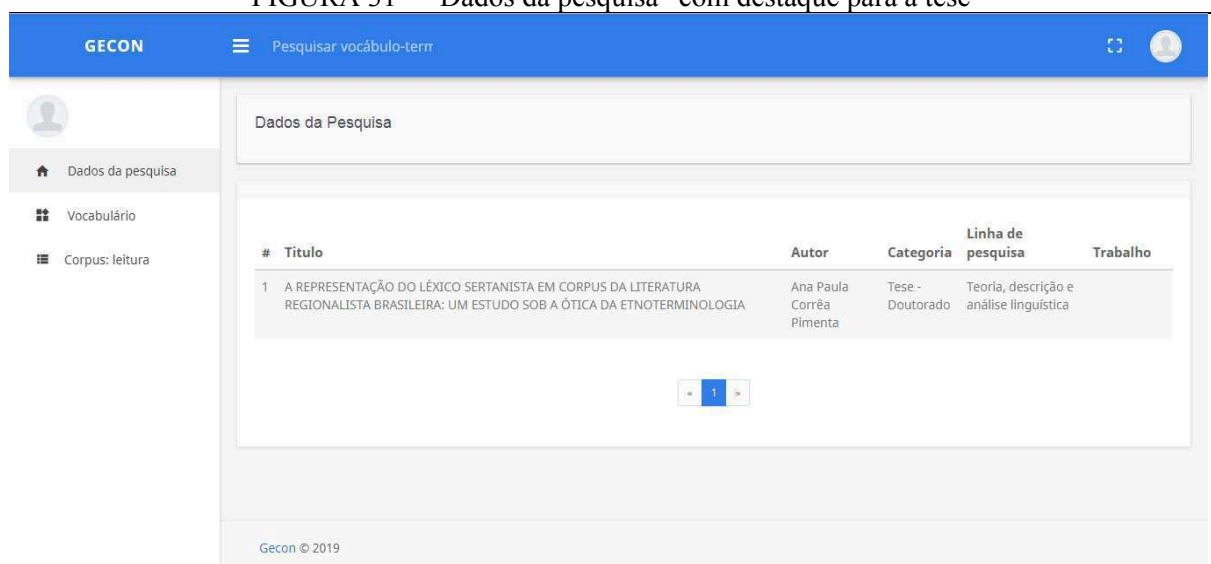
FIGURA 30 – Tela inicial de consulta



Fonte: <https://www.ileel.ufu.br/gecon>

Ao clicar em **Dados da pesquisa** o consulente visualiza uma nova página onde estão disponibilizados o trabalho completo (tese) e as fichas etnoterminográficas, como mostra a FIGURA 31:

FIGURA 31 – “Dados da pesquisa” com destaque para a tese



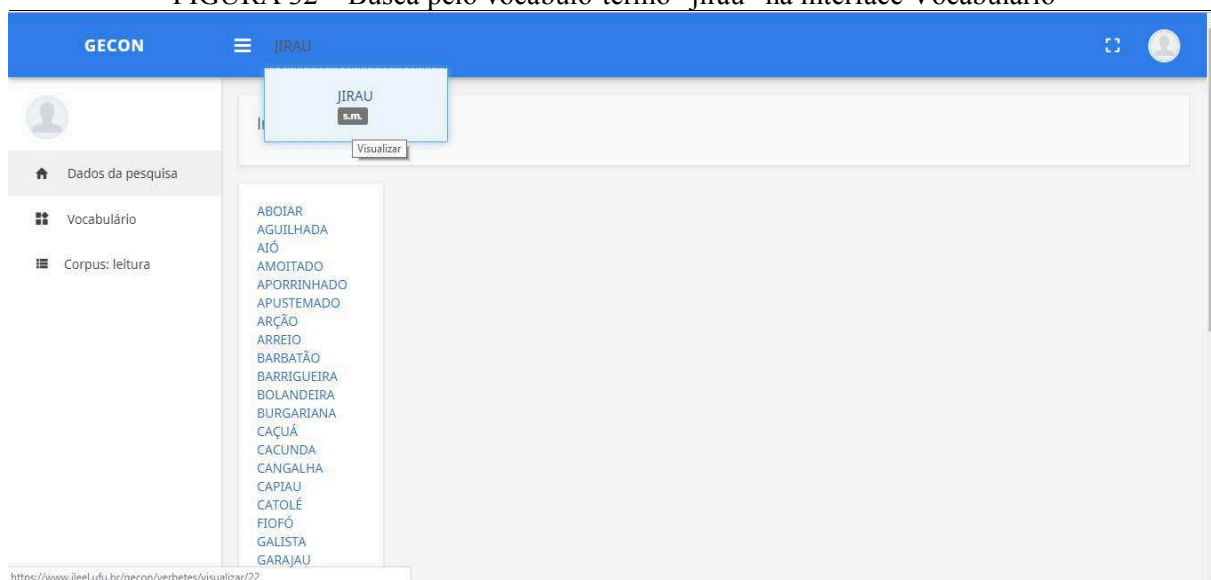
Fonte: <https://www.ileel.ufu.br/gecon>

Nos ícones abaixo o consulente tem duas opções de consulta: 1) a interface **Vocabulário** onde ele visualiza, de um lado da página, a lista com todos os vocábulos-termos

e, de outro, seus respectivos verbetes e, também, os textos literários em que ocorrem; 2) e a interface **Leitura** (Corpus: leitura) que lhe permite visualizar todos os textos literários em PDF e, também, em formato HTML com os vocábulos-termos e suas definições em janelas *pop-up*, possibilitando a leitura do texto e a sua compreensão.

Para começar a consulta pela interface Vocabulário basta o consulente digitar o vocábulo-termo desejado no item de busca ou clicar em um dos que aparecem na lista do lado esquerdo da tela para ter acesso a sua definição (FIGURA 32). Ao clicar ou digitar uma palavra aparecerá do lado direito o verbete completo: a classe gramatical, o gênero (PI), a definição, com todas as acepções conforme o contexto linguístico dos discursos-ocorrência (PD), as abonações (PP) e a frequência total e, do lado direito, as obras literárias onde o vocábulo-termo aparece (FIGURA 33). Clicando na lupa do lado direito da obra o consulente tem acesso ao texto completo e as respectivas acepções em janelas *pop-up* (FIGURAS 34, 35, 36). As figuras, a seguir, exemplificam esse processo.

FIGURA 32 – Busca pelo vocábulo-termo “jirau” na interface Vocabulário



Fonte: <https://www.ileel.ufu.br/gecon>

FIGURA 33 – Verbetes do vocábulo-termo “jirau” e os textos literários em que ocorre

**GECON** **JIRAU**

Interface Vocabulário

**Dados da pesquisa**

- Vocabulário
- Corpus: leitura

**ABOJAR**  
**AGUILHADA**  
**AIÓ**  
**AMOITADO**  
**APORRINHADO**  
**APUSTEMADO**  
**ARÇÃO**  
**ARREIO**  
**BARBATÃO**  
**BARRIGUEIRA**  
**BOLANDEIRA**  
**BURGARIANA**  
**CAÇUÁ**  
**CACUNDA**  
**CANGALHA**  
**CAPIAU**  
**CATOLÉ**  
**FIOFÔ**  
**GALISTA**  
**GARAJAU**  
**GARROTA**  
**JIRAU**

**Aceções**

**JIRAU**

(H) s.m. 1. Espécie de mesa forrada com varas para colocar utensílios de cozinha depois de lavados ou objetos domésticos. “Num **jirau**, umas panelas velhas com craveiros brotando e bogaris pelas biqueiras florindo” (ME, Reg2). 2. Armação de madeira sobre a qual se edificam as casas para evitar a água e a umidade. “A velha casa de taipa negrejava ao sol o telhado de **jirau**” (OQ, Reg2). 3. Espécie de cama feita com varas e sobre esteios fixados no chão. “Entrou no seu rancho pobre de sapé, sobre cuja esteira de **jirau** perrengueava a mulher, numa recálida de resguardo de parto” (TB, Reg2; freq.30)

Obra	Sigla	Autor
Menino de Engenho	ME	José Lins do Rego
O Quinze	OQ	Raquel de Queiroz
Tropas e boiadas	TB	Hugo de Carvalho Ramos
Os Sertões	OS	Euclides da Cunha
Luzia-Homem	LH	Domingos Olímpio
Vidas secas	VS	Graciliano Ramos

Fonte: <https://www.ileel.ufu.br/gecon>

FIGURA 34 – Busca pela obra “O Quinze” para leitura e visualização das definições

**GECON** **Pesquisar vocábulo-termo**

Interface Vocabulário

**Ana Paula Corrêa Pimenta**

**Dados da pesquisa**

- Vocabulário
- Corpus: leitura
- Cadastros
- Vocábulo-termo
- Aceções

**ABOJAR**  
**AGUILHADA**  
**AIÓ**  
**AMOITADO**  
**APORRINHADO**  
**APUSTEMADO**  
**ARÇÃO**  
**ARREIO**  
**BARBATÃO**  
**BARRIGUEIRA**  
**BOLANDEIRA**  
**BURGARIANA**  
**CAÇUÁ**  
**CACUNDA**  
**CANGALHA**  
**CAPIAU**  
**CATOLÉ**  
**FIOFÔ**  
**GALISTA**  
**GARAJAU**  
**GARROTA**  
**JIRAU**

**Aceções**

**JIRAU**

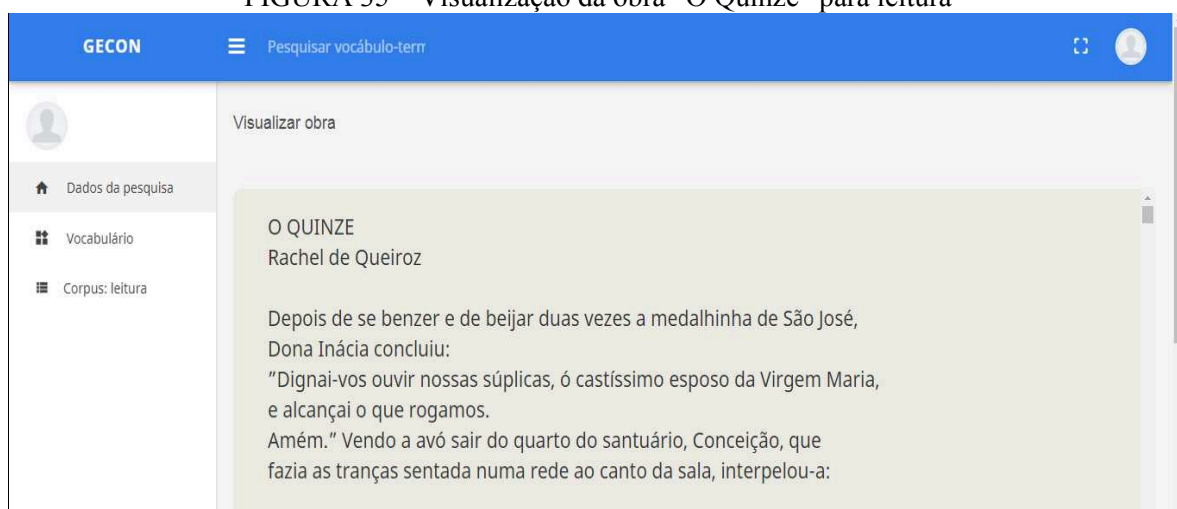
(H) s.m. 1. Espécie de mesa forrada com varas para colocar utensílios de cozinha depois de lavados ou objetos domésticos. “Num **jirau**, umas panelas velhas com craveiros brotando e bogaris pelas biqueiras florindo” (ME, Reg2). 2. Armação de madeira sobre a qual se edificam as casas para evitar a água e a umidade. “A velha casa de taipa negrejava ao sol o telhado de **jirau**” (OQ, Reg2). 3. Espécie de cama feita com varas e sobre esteios fixados no chão. “Entrou no seu rancho pobre de sapé, sobre cuja esteira de **jirau** perrengueava a mulher, numa recálida de resguardo de parto” (TB, Reg2; freq.30)

Obra	Sigla	Autor
Menino de Engenho	ME	José Lins do Rego
O Quinze	OQ	Raquel de Queiroz
Tropas e boiadas	TB	Hugo de Carvalho Ramos
Os Sertões	OS	Euclides da Cunha
Luzia-Homem	LH	Domingos Olímpio
Vidas secas	VS	Graciliano Ramos

Fonte: <https://www.ileel.ufu.br/gecon>



FIGURA 35 – Visualização da obra “O Quinze” para leitura



Fonte: <https://www.ileel.ufu.br/gecon>

FIGURA 36 – Visualização da definição de “jirau” na obra “O Quinze” em janela *pop-up*



Fonte: <https://www.ileel.ufu.br/gecon>

Na interface Leitura, o consulente pode ir diretamente aos textos clicando em **Corpus: leitura** e selecionando a fase do Regionalismo que deseja (FIGURA 37). Para ler a obra e visualizar as *pop-up*, basta o consulente clicar na lupa que aparece do lado direito (FIGURA 38), e para ler no arquivo PDF, basta clicar no balãozinho (FIGURA 39).



FIGURA 37 – Tela da interface Leitura

GECON Pesquisar vocábulo-term

Listagem de obras

#	Nome	Autor	Fase	Sigla	Data da publicação	
1	O Garimpeiro	Bernardo Guimarães	REGIONALISMO PITORESCO (fase 1)	OG	1872	<a href="#">Q</a> <a href="#">U</a>
2	Inocência	Visconde de Taunay	REGIONALISMO PITORESCO (fase 1)	IN	1872	<a href="#">Q</a> <a href="#">U</a>
3	O Sertanejo	José de Alencar	REGIONALISMO PITORESCO (fase 1)	OS	1875	<a href="#">Q</a> <a href="#">U</a>
4	O Cabeleira	Franklin Távora	REGIONALISMO PITORESCO (fase 1)	OC	1876	<a href="#">Q</a> <a href="#">U</a>
5	Pelo Sertão	Afonso Arinos	REGIONALISMO PITORESCO (fase 1)	PS	1898	<a href="#">Q</a> <a href="#">U</a>

1

<https://www.ileel.ufu.br/gecon/obras/fase/FASE1>

Fonte: <https://www.ileel.ufu.br/gecon>

FIGURA 38 – Busca pela obra “O Garimpeiro” para leitura em HTML com as *pop-up*

GECON Pesquisar vocábulo-term

Listagem de obras

#	Nome	Autor	Fase	Sigla	Data da publicação	
1	O Garimpeiro	Bernardo Guimarães	REGIONALISMO PITORESCO (fase 1)	OG	1872	<a href="#">Q</a> <a href="#">U</a> <a href="#">Ler obra</a>
2	Inocência	Visconde de Taunay	REGIONALISMO PITORESCO (fase 1)	IN	1872	<a href="#">Q</a> <a href="#">U</a>
3	O Sertanejo	José de Alencar	REGIONALISMO PITORESCO (fase 1)	OS	1875	<a href="#">Q</a> <a href="#">U</a>
4	O Cabeleira	Franklin Távora	REGIONALISMO PITORESCO (fase 1)	OC	1876	<a href="#">Q</a> <a href="#">U</a>
5	Pelo Sertão	Afonso Arinos	REGIONALISMO PITORESCO (fase 1)	PS	1898	<a href="#">Q</a> <a href="#">U</a>

1

<https://www.ileel.ufu.br/gecon/obras/visualizar/1>

Fonte: <https://www.ileel.ufu.br/gecon>

FIGURA 39 – Busca pela obra “O Garimpeiro” para leitura em PDF

The screenshot shows the GECON interface with a blue header bar containing the logo and the text 'Pesquisar vocábulo-termo'. On the left, there is a sidebar with navigation options: 'Dados da pesquisa', 'Vocabulário', and 'Corpus: leitura'. The 'Corpus: leitura' section is expanded, showing a list of corpora: 'REGIONALISMO PITORESCO (fase 1)', 'REGIONALISMO CRÍTICO (fase 2)', and 'SUPER-REGIONALISMO (fase 3)'. The main area displays a table titled 'Listagem de obras' with the following data:

#	Nome	Autor	Fase	Sigla	Data da publicação	
	O Garimpeiro	Bernardo Guimarães	REGIONALISMO PITORESCO (fase 1)	OG	1872	<a href="#">Q</a> <a href="#">Download</a>
	Inocência	Visconde de Taunay	REGIONALISMO PITORESCO (fase 1)	IN	1872	<a href="#">Q</a> <a href="#">Download</a>
	O Sertanejo	José de Alencar	REGIONALISMO PITORESCO (fase 1)	OS	1875	<a href="#">Q</a> <a href="#">Download</a>
	O Cabeleira	Franklin Távora	REGIONALISMO PITORESCO (fase 1)	OC	1876	<a href="#">Q</a> <a href="#">Download</a>
	Pelo Sertão	Afonso Arinos	REGIONALISMO PITORESCO (fase 1)	PS	1898	<a href="#">Q</a> <a href="#">Download</a>

At the bottom of the table, there is a pagination control showing '1'.

Fonte: <https://www.ileel.ufu.br/gecon>

Desse modo, na interface Leitura o consulente tem duas opções: 1) ler o texto e, a mesmo tempo, visualizar as definições dos vocábulos-terms destacados; 2) ler a obra em PDF sem as *pop-up*.

Como se pode observar, o acesso é rápido e simples, o que facilita a consulta dos que têm interesse tanto pela pesquisa ou produto etnoterminológico (vocabulário) quanto pela leitura dos textos. Embora seja apenas uma proposta de modelo de vocabulário, acreditamos na sua validade e expansão. Em uma aula de Literatura, por exemplo, essas duas interfaces poderão ser trabalhadas por professores em atividades de leitura e interpretação do texto literário, assim como poderão ser ampliadas para atender objetivos diversos.

Assim, o modelo proposto sugere possíveis desdobramentos, como: a) a alimentação do sistema, com a inclusão de todos os dados (vocábulos-terms) encontrados a partir do nosso *corpus* de estudo (etapa posterior à defesa); b) a sua ampliação e expansão, como a inserção de novos resultados obtidos a partir de outras obras literárias regionalistas de temática sertanista; a sugestão de outros tipos de definições para diferentes consulentes, por exemplo, definições adequadas aos alunos de cada etapa de ensino (Ensino Fundamental, Médio e Superior); o cadastro de novos trabalhos com essa ou outra temática; a criação de novas ferramentas para aplicação em diversas atividades de leitura e interpretação do texto literário.

Uma vez concretizado, a grande vantagem deste produto, além das contribuições para os estudos etnoterminológicos, é de servir como auxílio a alunos e professores nas atividades

de leitura e compreensão de obras consagradas da nossa Literatura e, também, como subsídio para trabalhos futuros, podendo ser ampliado, replicado ou reproduzido, seja por intermédio de um pesquisador ou através de um trabalho colaborativo entre pesquisadores, professores e alunos.

## 5 APRESENTAÇÃO, ANÁLISE E DISCUSSÃO DOS RESULTADOS

Este capítulo apresenta a síntese dos resultados obtidos a partir do nosso *corpus* de estudo e demonstra suas principais características por meio do preenchimento de fichas etnoterminográficas, de análises quantitativas, comparativas e contrastivas.

No QUADRO 11, apresentamos todos os vocábulos-termos extraídos das palavras-chave com suas respectivas fases, frequência e dicionarização. Convém esclarecer, que utilizamos o sinal ≠ para indicar aqueles vocábulos-termos que possuem acepções diferentes das que estão registradas nos dicionários informados.

QUADRO 11: Síntese dos resultados obtidos

Vocábulo-termo	Fases do Regionalismo	Frequência total no <i>corpus</i>	Dicionarização
1- abano	2	6	n/e (≠ H/B/P/S)
2- abancar	2	4	H
3- abarrear	2	3	H/P/S
4- abastança	2	7	H/B/P/S
5- abelhudo	3	4	H/B/P/S
6- abobra	3	4	H
7- aboiar	2	4	H (≠ P/S)
8- aboio	3	7	H
9- aboletar	2	7	H (≠ B/P/S)
10- abuso	2	3	n/e (≠ H/B/P/S)
11- abrasado	1	11	H/S
12- acabrunhado	2	7	H/B/S
13- acachoar	2	5	H (≠ S)
14- acauã	3	18	H
15- aceirar	2	3	B/P/S (≠ H)
16- aceiro	1, 2	25	B/P/S (≠ H)
17- acervo	2	12	n/e (≠ H/B/P/S)
18- acocorar	1, 2	14	H/B/P/S
19- acocorado	2	18	H/S
20- acuado	2	11	H/B/S
21- açude	2	23	H/B/P/S
22- adomação	3	4	n/d
23- afadigado	3	3	B/S
24- aferrar	2	4	n/e (≠ H/B/P/S)
25- aguada	2, 3	17	H (≠ S)
26- aguardenteiro	3	9	H/S
27- aguilhada	1, 2	7	H/B/P/S
28- aió	2	10	H
29- alarma	2	27	H (≠ S)
30- alastrado	2	3	H (≠ B/S)

31- alamão	1	15	n/d
32- alazão	1, 2	17	H/B/P/S
33- alçapão	1, 2	10	H/B/P/S
34- aleijão	2	5	H/P (≠B)
35- aleive	2	4	H/B/P/S
36- Alforge (*alforje)	1, 2	14	H/B/P
37- algara	2	4	H (≠S)
38- algar	2	7	H /B/P/S
39- algibeira	3	14	H/B/P/S
40- almanjarra	2	6	H /B/P/S
41- Alpargata (*alpercata, alparcata)	2, 3	56	H/S (≠B/P)
42- aluado	3	7	H/B/S
43- aluir	3	4	H/P/S
44- alviças	1	4	B/P
45- alvoroço (*alvoroto)	1, 2, 3	18	H/B/P/S
46- alvoroçado	2	4	H/B/S
47- amainar	3	4	H/B/P/S
48- amiudar	2	9	n/e (≠H/B/P/S)
49- amoitado	1, 3	13	H
50- amolentar	2	3	n/e (≠H/B/P/S)
51- amontado (*amuntado)	3	21	n/e (≠H/S)
52- amontar	3	6	n/e (≠H/P/S)
53- amuado	2	5	H/B/S
54- amunhecar	3	6	H
55- anca	1, 2, 3	39	H/B/P/S
56- ancorote	2	3	n/e (≠H/B/P/S)
57- andadura	2	9	H/B/P/S
58- anelado	2	5	n/e (≠H/S)
59- anequim	2	7	H
60- anexim	1	4	H/B/P/S
61-angico	1, 2, 3	56	H
62- angu	2	9	H
63- Aniceto	1	6	n/d
64- antolhar	2	4	H/B/P/S
65- anu	1, 2, 3	12	H
66- apartear	3	3	H
67- apeado	2, 3	13	S
68- apear	1, 2, 3	70	H/B/P/S
69- aperado	2	3	H
70- aperrear	2	4	H/B/P/S
71- aperreio	3	4	H
72- apessoado	2	4	H/B/P/S
73- apeteçido	2	4	S
74- apetrechado	3	5	n/d
75- aporrinhado	3	8	H
76- aprumo	2, 3	13	H
77- apostemado	3	3	n/d

78- araçá	2	4	H
79- aragem	1, 2, 3	24	H
80- araral	3	3	n/d
81- arção	1, 2, 3	28	H/B/P/S
82- argolão	1	9	B/S ( $\neq$ H)
83- aroeira	2	10	H/B/P/S
84- arranchamento	2	6	H
85- arraia	3	8	H ( $\neq$ B/P/S)
86- arranchar	2	3	H/B/P/S
87- arranjado	1	5	H
88- arre	3	35	H/B/S
89- arreado	2, 3	18	n/d
90- arrear	2, 3	9	H ( $\neq$ B/S)
91- arredar	1	13	H/B/P/S
92- arreio	1, 2, 3	94	H/B/P/S
93- arreliado	2, 3	13	H
94- arrenegado	3	5	H ( $\neq$ B/S)
95- arriar	2, 3	11	H ( $\neq$ B/P/S)
96- arriba	3	9	H/B/P/S
97- arribação	2	18	H/B/P/S
98- arribada	2, 3	9	H ( $\neq$ B/P)
99- arrieiro	1, 2, 3	60	H/B/P/S
100- arrimar	1, 2	10	H/B/P/S
101- arrepiar	1	3	n/e ( $\neq$ P)
102- arrojo	1, 2	16	H ( $\neq$ P/S)
103- arrostar	2	4	H/B/P/S
104- arrumadeira	3	11	H
105- acetinado	1	3	P
106- assomo	1, 2	21	H/B/S
107- assuntar	2, 3	29	H
108- ateimado	2	4	S
109- aturdido	1	3	H/S
110- avariado	3	6	H/P/S
111- avexado	3	4	H/B/S
112- avio	2	3	H
113- azado	1	3	H/B/P/S
114- azulego	2	3	H
115- azoretado	3	4	n/d
116- babugem	2	5	H ( $\neq$ B/P/S)
117- bacamarte	1, 2	44	H/B/P/S
118- bacurau	3	4	H
119- bafagem	2	4	H/B/P/S
120- bagaceira	2	28	H/S
121- bago	3	5	n/e ( $\neq$ H/B/P/S)
122- bagual	2	24	H
123- bagualada	2	6	H
124- baio	1, 2, 3	33	H/B/P/S
125- baixeiro	2, 3	14	H

126- balaio	3	10	H /B/P/S
127- balandrau	2	3	H (≠B/P/S)
128- bamburral (*bambual)	1, 2, 3	13	H/P/S
129- bandalheira	3	4	H
130- bandoleira	2, 3	10	B/P/S
131- banzar	1	4	H/B/S
132- banzeiro	2	5	H/B/S
133- baraúna	2	5	H
134- barbatão	1, 2	28	H
135- barbicacho	2	6	H /B/P/S
136- barranca	2	28	H
137- barranco	1, 3	55	H/B/P/S
138- barreiro	2	14	H/B/P/S
139- barrica	3	15	H/B/P/S
140- barrigueira	2, 3	11	H
141- barrocal	2	3	H/P/S
142- bátega	2	5	H (≠S)
143- baticum	2	4	H
144- beberagem	3	5	H (≠B/P/S)
145- beco	2, 3	34	H (≠B/ P/S)
146- bedéquer	2	3	n/d
147- beíçal	3	5	n/d
148- bendomeu	3	3	n/d
149- berne	2	5	H (≠B/P)
150- berzabum (*barzabu)	2	8	H
151- besta	1, 2, 3	146	H/B/P/S
152- bexigoso (*bexiguento)	3	31	H/B/P/S
153- biboca	2	8	H
154- bichão	3	11	H
155- bichará	2	5	H
156- bicheira	3	6	H
157- bilro	1, 2	11	H
158- binga	1, 2, 3	21	H
159- biqueira	2	6	H (≠P/S)
160- Birro	3	5	n/e (≠H/B/P/S)
161- bobéia	3	12	H
162- bocapio	3	3	n/d
163- bofe	2	6	H/B/P/S
164- bogari	2, 3	7	H/S
165- boiadeiro	2, 3	56	H
166- boitató	3	6	H
167- bolandeira	2	24	H/P/S
168- boleadeira	2	5	H
169- boqueirão	2, 3	147	H/B/P/S
170- borrar	3	3	H (≠B/P/S)
171- botar	1, 2, 3	235	B/S (≠H/P)
172- bote	1, 2, 3	65	H (≠B/P/S)
173- brabeza	3	16	H

174- braça	1, 3	71	H/B/P/S
175- breca	3	6	n/e (≠H/B)
176- brejeiro	2	15	H/S
177- brejo	3	17	H/B/S
178- brenha	1, 3	30	H/B/P/S
179- Bró	3	4	H
180- bruaca	2	5	H
181- buçal	2	3	H
182- buçalete	2	6	H
183- bucho	2	8	H/B/P/S
184- bueiro	2	7	H (≠B/S)
185- bugre	3	57	H
186- bulir	1, 2, 3	44	H (≠B/P/S)
187- burgariana	2	3	n/d
188- buriti	1, 3	77	H
189- buritizal	1, 3	27	H
190- burrame	3	3	n/d
191- busão	2	3	n/d
192- búzio (*buzo)	2, 3	15	H (≠B/P/S)
193- caapoão (*capoão)	1	6	n/d
194- caatinga	2	245	H
195- cabaça	2, 3	36	H/B/P/S
196- cabeceira	1, 3	75	H (≠B/P/S)
197- cabra	1, 2, 3	218	H (≠B/P/S)
198- cabrita	3	5	H (≠P/S)
199- cabresto	1, 2	35	H
200- cabriola	2	3	n/e (≠H/ B/P/S)
201- cabriolar	2	4	n/e (≠H/P/S)
202- cabriolé	2	5	H
203- cabrocha	2	11	H
204- cabroeira	2	8	H
205- cabrunquento	3	4	n/d
206- cabiúna	2	11	H
207- caburé	2, 3	10	H
208- cacareco	2	5	H
209- cachola	1	4	H/P/S (≠B)
210- cacimba	1, 2, 3	56	H/P/S
211- caçuá	1, 2	22	H
212- cacunda	2, 3	10	H
213- cafua	3	6	n/e (≠H/B/P/S)
214-cafuné	3	6	H/P/S
215- caipora	1, 2, 3	15	H
216- calaboca	3	3	n/d
217- calangro	3	5	H
218- calhada	3	3	n/d
219- caldeirão	2	7	H (≠B/P/S)
220- camarada	1, 2	158	H/B/P/S
221- camarinha	2	10	H (≠B/P/S)



222- cambada	2, 3	37	H (≠B/P/S)
223- cambaio	2	44	H/B/P/S
224- cambão	2	3	H
225- cambiteiro	2	11	H
226- caminheira	2	3	n/e (≠H)
227- camisu	2	9	H
228- campear	1, 2, 3	27	H/P/S (≠B)
229- campeio	3	3	H
230- campeirada	3	14	H
231- campeiro	1, 3	24	H (≠S)
232- camumbembe	2	3	H
233- canastra	1, 2	29	H/B/P/S
234- cancela	2	12	H/B (≠P)
235- cancha	2	11	H
236- candeeiro	2	7	H/B/P/S
237- canga	2	7	H/B/P/S
238- cangaceiro	2	113	H
239- cangaço	2, 3	18	H (≠P/S)
240- cangalha	1, 2, 3	52	H/B/P/S
241- cangapé	2	3	H
242- cangote	2, 3	25	H
243- canhoneio	2	19	H
244- canivete	2	6	H/B/P/S
245- cantil	2	6	H (≠B/P/S)
246- capado	3	8	H
247- capanga	3	30	H
248- capão	1, 2, 3	102	H (≠B/P/S)
249- capataz	1, 2, 3	115	H/P (≠B/S)
250- capengar	2	4	H
251- capiau	3	10	H
252- capineiro	2	3	H
253- capiongo	2	4	H
254- capoeira	1	11	H (≠B/P/S)
255- capoeirão	3	7	H (≠P/S)
256- capote	2	8	H/P/S
257- caramanchão	3	13	H/B/P/S
258- caramurú	1	7	H/B/S
259- carapina	3	16	H
260- carapinagem	3	3	n/d
261- cardão	1	10	H
262- cargueiro	1, 2	35	H
263- caritó	2	7	H
264- carnaúba (*carnaubeira)	1, 2	25	H
265- carnegão	3	8	H/P/S
266- caroá	2	6	H
267- carochinha	3	4	n/e (≠H)
268- carpina	2	6	H
269- carrascal	3	9	H/B/P/S

270- carreira	1, 2	105	H/B/P/S
271- carreiro	2	21	H/B/P/S
272- carretão	2	10	H (≠B/P/S)
273- carroção	3	9	H
274- carujo	3	3	n/d
275- casuarina	3	23	H
276- catinga	2, 3	43	H (≠B/P/S)
277- catingueira	2	16	H
278- catingueiro	2	10	H
279- catolé	1	3	H
280- catolezeiro	1	3	n/d
281- catre	1, 3	17	H/B/P/S
282- catrumano	3	28	H
283- cavalariano	3	24	H
284- cavalgadura	1, 2	20	H/B/P/S
285- cavalhada	1, 2, 3	65	H/B/P/S
286- cercadura	2	7	n/e (≠H/B/P/S)
287- ceroula	2, 3	8	H/B/P/S
288- cerradão	2	5	H
289- chakra	3	3	H
290- chamuscado	2	4	H/B/S
291- chapada	2, 3	93	H
292- chapadão	3	82	H
293- chasque	2	8	H
294- chasquear	2	6	H/B/P/S
295- chegante	3	6	n/d
296- chilena	2	11	n/d
297- chimite	3	3	n/d
298- chiqueiro	2, 3	51	H/B/P/S
299- chiru	2	34	H
300- chita	2	17	H (≠B/P/S)
301- chitado	3	4	H
302- choça	2	7	H/B/P/S
303- chocadeira	3	4	n/e (≠H/S)
304- chocalho	2, 3	28	H/B/P/S
305- chouto	2	4	H/P (≠B)
306- chupitar	2	3	n/e (≠H)
307- cilha	1, 2	8	H/B/S
308- cincha	2	4	H
309- cipó	1, 2	31	H/B/P/S
310- cipoal	2	4	H/S
311- Clavina (*clavinote)	1, 2	30	H/B/P/S
312- clavinoteiro	2	3	H
313- cocá	2	3	n/d
314- cocar	2, 3	11	n/e (≠H/B)
315- cochicholo	2	4	H (≠S)
316- cocho	2, 3	26	H (≠S)
317- cocó	2	4	n/e (≠H)

318- cócoras	2	18	H
319- cocorote	2	3	H
320- cocoruto	2	6	H
321- codorneiro	3	6	n/d
322- coimeiro	2	3	H/P/S
323- coité	1	10	H
324- coiteiro	2	6	H
325- coivara	2	7	H
326- comblém	3	3	n/d
327- comboieiros	2	4	H/S
328- comboio	1, 2	68	H/B (≠P/S)
329- comitiva	2	18	H (≠B/P/S)
330- congada	3	8	H
331- coradouro	1	4	H
332- corcel	1	4	H
333- corcovo	1	3	H/B/P/S
334- coriboca (curiboca)	1, 2	20	n/e (≠H)
335- coroca	2	4	H
336- coronilha	2	3	H (≠B/P/S)
337- corredio	2	6	B/S (≠H/P)
338- correição	3	7	H (≠B/P/S)
339- corricho	3	6	n/d
340- coscorão	3	3	H/S
341- croá	2	3	H
342- criaturo	3	4	n/d
343- cueiro	2	4	H/B
344- cuerudo	2	3	H
345- cuia	2, 3	43	H/B/S
346- Cuité	2	4	H/S
347- culatreiro	2	10	n/d
348- culo	2	6	n/e (≠H)
349- cunhão	3	3	n/d
350- cupim	2, 3	19	H
351- curral	1, 2, 3	286	H/B/P/S
352- cusco	2	7	H (≠B)
353- cutuca	2	7	H
354- cutucar	2, 3	6	H
355- danado	1, 2, 3	95	H/B (≠P/S)
356- debuxar	1	6	n/e (≠H/B/P/S)
357- demorão	3	3	n/d
358- derrear	2, 3	28	H/B/P/S
359- desacocorar	3	3	n/d
360- desamontar	3	3	n/d
361- desandar	2, 3	15	H (≠B/P/S)
362- desapare	3	3	H
363- desarrear	3	5	H
364- desarvorado	3	6	H (≠S)
365- desbarate	2	5	H (≠B/P/S)

366- descambar	2	14	H/B/P/S
367- descarecer	3	4	n/d
368- desembestar	2	3	H/B/P/S
369- desembuchar	3	6	H/S (≠B/P)
370- desensofrido	3	8	n/d
371- desgraceira	2	6	H
372- desinfluído	2	3	n/d
373- despenho	2	3	H (≠B/P/S)
374- destampatório	3	7	H (≠P/S)
375- destratar	3	4	H/S (≠B/P)
376- devesa	3	11	H/S
377- direita	2	7	n/e (≠H/B/P/S)
378- direitura	1, 3	10	P/S(≠H/B)
379- dobro	2	5	H (≠B/P/S)
380-dormideira	3	4	H/B/P/S
381- douradilha	3	6	H
382- drede	3	4	n/d
383- drongo	3	3	n/d
384- ecô	3	4	H
385- eirado	3	6	n/e (≠H/B/P/S)
386- oito	2, 3	342	H/S (≠B/P)
387- embarafustar	2	3	H
388- embastido	3	5	H
389- embira	2, 3	18	H/P/S
390- embolado	2	4	n/e (≠S)
391- embonecrar	3	3	H
392- emborcar	3	3	H (≠B/P/S)
393- embornal	2	5	H/ P/S
394- embromação	2	3	H
395- embuçado	1	6	n/e (≠H/B/P/S)
396- empacador	2	3	H
397- empalamado	1	5	H/B/P/S
398- empanzinado	2	7	P
399- empestiado	3	3	n/d
400- encabulamento	3	5	H
401- encalacrado	2	4	H
402- encanado	3	9	H ((≠B/S)
403- encourado	2	9	H/B/S
404- enfarruscado	3	4	n/e (≠S)
405- enganchado	3	4	n/d
406- engenhoca	1	16	H
407- enrabichado	2	3	H
408- enxó	2	5	H/B/P/S
409- enxurro	2	13	H/B/P/S
410- esbagaçar	2	4	H
411- esbregue	3	3	H
412- escabreado	3	3	H
413- escampo	2	3	H

414- escampado	1	3	H/S
415- escanchado	2	8	n/e (≠H/S)
416- escangalhar	2	3	H
417- escolado	3	6	H
418- esmaiado	1	3	n/e (≠H/S)
419- esmolambado	2	3	n/e (≠H)
420- espição	3	4	n/d
421- espíar	1, 2, 3	217	H/B/P/S
422- espigado	2	4	B/S (≠H, <sub>1</sub> /P)
423- espigão	2, 3	21	H/B/P/S
424- espinhaço	2	18	H/B/P/S
425- espinheiral	2	4	H
426- espinheiro	2	11	H/B/P/S
427- espoleta	3	7	H (≠P/S)
428- espora	1, 2, 3	117	H/B/P/S
429- esporar	3	8	H
430- espôrro	3	3	H
431- espreguiçadeira	2	6	H
432- estafeta	2, 3	26	H/B/P/S
433- estafetamento	2	3	H
434- esteira	2	31	n/e (≠H/B/P/S)
435- esteireiro	1	17	H/B/P/S
436- esticador	3	4	H
437- estirão	2, 3	11	B/P/S (≠H)
438- estradinho	1	3	n/d
439- estradeira	2	3	H
440-estranja	1	3	H
441- estrebuchó	3	3	n/d
442- estrebaria	2	9	H/B/P/S
443- estreleiro	3	13	H/B/P/S
444- estribeira	3	8	n/e (≠H/B/P/S)
445- estribilho	2	8	n/e (≠H/B/P/S)
446- estribo	2, 3	29	H/B/P/S
447- estropiado	2	14	H/B
448- estrumado	3	4	S
449- estúrdio	3	16	H (≠P/S)
450- excomungado	2, 3	21	H (≠B/S)
451- faceira	2	7	H/B/P/S
452- falripas	2	5	H/B/P/S
453- faroleiro	2	16	H
454- farrapo	2	29	H (≠B/P/S)
455- farreagem	3	7	n/d
456- fartão	1	3	n/e (≠H)
457- fasto	3	3	n/e (≠H/B/P/S)
458- fedido	3	4	H
459- ferra	2	7	H (≠P/S)
460- ferrada	2, 3	9	n/e (≠H/B/P/S)
461- ferragem	3	7	H/B/P/S

462- ferrão	1, 2	29	H/P/S (≠B)
463- festo	2	5	H (≠B/P/S)
464- fiar	1	9	H/B/P/S
465- fifô	2, 3	12	H
466- finório	3	4	H
467- fiofô	3	4	H
468- flete	2	12	H
469- floreado	3	6	H (≠B/S)
470- fogareiro	3	4	B/P/S (≠H)
471- foice	2, 3	23	H/B/P
472- fojo	1	7	H/B/P/S
473- fouveiro	2	3	H/B/P/S
474- franqueira	1, 2	12	H (≠B)
475- franzino	2	15	B/P/S (≠H)
476- frontear	2	4	H
477- fueiro	2	4	H/P/S
478- fundão	1, 2	12	n/e (≠H/B)
479- fundilho	3	13	H/B/P/S
480- fusco-fusco	2	13	H (≠B/P/S)
481- fuso	1	6	H/B/P/S
482- fuzuê	2	4	H
483- gabola	2	3	H/P/S
484- gacheiro	1	3	H
485- gadaria	2	5	n/d
486- gadelhudo	2	5	H/B
487- gaiaumum	3	3	n/d
488- gaitada	3	6	H (≠P/S)
489- galão	2, 3	16	H/B/P/S
490- galharada	2	3	H
491- galhofista	3	3	n/d
492- galista	3	12	H
493- gambá	3	8	H
494- gameleira	1, 2, 3	31	H
495- gangorra	2	6	H
496- garajau	3	9	H
497- garrafada	3	5	H
498- garrancheira	2	3	n/d
499- garrancho	2	13	H (≠B/P/S)
500- garrão	2	6	H
501- garrota	3	3	H
502- garrote	1, 2, 3	21	H (≠B/P/S)
503- garrucha	1, 2, 3	67	H (≠B/P/S)
504- garupeira	2	3	H
505- gastura	3	10	H
506- gateado	3	3	H (≠S)
507- gauchada	2	8	H
508- gavar	3	3	H
509- gentama	2	6	H

510- geralista	3	4	H
511- gibão	1, 2, 3	44	H/B/P/S
512- ginete	1	7	H/B/P/S
513- gogó	3	3	n/e (≠H)
514- gongolô	3	3	H
515- grameal	3	4	H
516- graúna	1	4	H
517- gravatá	2, 3	16	H/S
518- grenha	2	8	H (≠B/P/S)
519- grimpa	2	11	H (≠B/P/S)
520- grota	1, 2, 3	46	H
521- grotão	2	17	H
522- grupiara	1	4	H
523- guabiroba	2	4	H
524- guaiaca	2, 3	15	H
525- guampo	1, 3	9	H
526- guampudo	2	3	H
527- guasca	2	4	H
528- guriri	3	3	H
529- gurizote	2	5	H
530- gurungumba	3	7	H
531- icoiar	3	3	n/d
532- icozeiro	2	3	H
533- imburana	2	4	H
534- ingazeira	1	8	H
535- inhambu	3	5	H
536- inhatium	2	3	n/d
537- inteiriço	2, 3	12	n/e (≠H)
538- inventação	3	5	n/d
539- inventoria	3	3	n/d
540- ipiaça	3	3	n/d
541- ipueira	2	16	H
542- jaburu	3	4	H
543- jacarandá	1	12	H/B/P/S
544- jacuba	3	6	H
545- jagunçada	3	8	H
546- jagunço	2, 3	347	H
547- jantarola	2	5	H
548- jaracatiá (jarracatiá)	1	8	H
549- jaraguá	2	10	H
550- jararaca	2	6	H
551- jararacuçu	2	3	H
552- jarrete	2	3	H/B/P/S
553- jatobá	2	11	H
554- jebe-jebe	3	4	n/d
555- jeca	2	22	H
556- jerivá	2	5	H
557- jeriza	3	3	H

558- jia	3	36	H
559- jirau	2	30	H
560- juá	2	14	H/P/S
561- juazeiro	2	96	H/P/S
562- jugulado	2	5	n/d
563- jurema	2	5	H
564- juriti	1	3	H
565- jururu	2	3	H
566- lacraia	3	4	H
567- ladainha	2, 3	20	H (≠B/P/S)
568- lagartão	3	3	n/d
569- lagoão	2	6	H
570- lajem	2	8	H
571- lamedo	2	3	H
572- lamparina	2, 3	34	H
573- lampião	3	21	H
574- lançante	1, 3	8	H (≠S)
575- lasco	3	4	n/d
576- latada	2	21	H (≠B/P/S)
577- latagão	1	4	H/B
578- látigo	2	5	H/P/S
579- lazarento	2, 3	6	H (≠P/S)
580- lazarina	2	8	H
581- lazeira	2	4	H/B/P/S
582- léria (lereia)	2, 3	7	H
583- leso	2	4	H (≠B/P/S)
584- ligal	2	4	H
585- lingueta	3	4	n/e (≠H/B/P/S)
586- lobuno	2	3	H
587- lombeira	2	6	H
588- lombilho	2	9	H
589- lonquear	2	4	H
590- losna	2	5	H/B/P/S
591- lunanco	2	3	H
592- lubila	3	4	n/d
593- macaca	2	7	H
594- macambira	2, 3	34	H
595- maçaneta	3	5	n/e (≠H/B/S)
596- macaúba	3	4	H
597- macega	2, 3	30	H
598- macuco	3	11	H
599- madapolão	2	3	H
600- magrém	2	3	H
601- malhada	2	8	H/B/P/S
602- malhadouro	2	3	H
603- malmequer	2	3	H/B/P/S
604- manantial	2	16	H
605- mancarrão	2	8	H



606- mandacaru	2	29	H
607- mandraca	3	3	H
608- maneado	2	4	H
609- maneia	3	5	H
610- maneiro	3	9	n/e (≠H/B/P/S)
611- mangação	2	7	H
612- manguá	3	4	H
613- mangue	3	13	H/B/P/S
614- manotear	2	3	H
615- manta	2, 3	28	H/B/P/S
616- manteúdo	3	14	H (≠P/S)
617- maracanã	1	3	H
618- maravalhas	1	3	B/P/S
619- marimbu	3	5	H
620- marizeira	1, 2	13	H
621- marola	3	10	H
622- marosca	2	4	H
623- marrã	2	5	H/B/S
624- marruá	3	3	H
625- massapé	3	5	H (≠B/P/S)
626- mateiro	3	7	H (≠B/P/S)
627- matula	2, 3	15	H (≠B/P/S)
628- matungo	2	11	H
629- matutar	2	6	H
630- matuto	1, 2	81	H
631- maximé	3	7	n/d
632- mecê	1	45	H
633- meeiro	2	9	H (≠B/S)
634- meganha	3	26	H
635- melado	2	12	H/B/P/S
636- membeca	3	3	H
637- mimosal	3	5	n/d
638- mimosura	3	3	n/d
639- minhoto	1	6	n/e (≠H/B/P/S)
640- mocambeiro	1	5	H
641- mocho	2	5	H/B/P/S
642- mofa	1	7	B/P/S
643- mofino	3	6	H (≠B/P/S)
644- moirão (mourão)	1, 2, 3	84	H/B/P/S
645- molambo	2	18	H
646- monjolo	2	32	H
647- montearia	1	15	S
648- morocho	2	6	H
649- moringa	3	13	H
650- morrudo	2	3	H
651- mulada	2	4	H
652- mulita	2	3	H
653- mulungu	2	7	H

654- mundéu	2	9	H
655- mundiça	2	4	H
656- munheca	2, 3	18	n/e (≠H/B/P/S)
657- murça	3	4	n/e (≠H)
658- murici	3	9	H
659- murzelo	3	3	H/B/P/S
670- mutuca	3	10	H
671- muxoxo	2	9	H
672- nagã	3	3	H
673- nagoa	2	13	n/d
674- nhandu	2	4	H
675- nhã	3	4	n/d
676- nodoso	2	4	n/e (≠H/B/P/S)
677- oco (ôco)	2, 3	41	H (≠B/P/S)
678- ogre	2	4	H
679- oitão	1, 2, 3	24	H
680- oiteiro	2	8	H
681- oiticia	1, 2	126	H
682- ouricuri	2	5	H
683- pabulagem	2	10	H
684- paca	2	14	H/B/P/S
685- pagem	1	16	n/e (≠B/P)
686- paiol	1, 2, 3	78	H (≠B/P/S)
687- pala	2	21	H (≠B/P/S)
688- paleta	2	4	H (≠P/S)
689- paliçada	2	6	H/B/P/S
690- palmatória	2	7	H (≠B/P/S)
691- panasco	1, 2	7	H/P/S
692- papocar	3	3	H
693- papoco	3	3	H
694- papudo	3	4	H (≠B/P/S)
695- pardavasco	3	17	H
696- parelha	2	7	H/B/P/S
697- parantalha	2	3	H/P
698- parentagem	3	5	n/d
699- parnaíba	1, 2	15	H
700- parolagem	3	7	H (≠S)
701- passarinho	3	9	H/B/P/S
702- pastar	3	7	H/B/P/S
703- pastaria	3	12	H
704- pasto	1, 2, 3	169	H/B/P/S
705- patação	2	6	H/B/P/S
706- patativa	1	3	H
707- patrona	2	8	H/B/P/S
708- pautear	3	3	H
709- pavoã	3	3	H
710- Pear	1	3	H/B/P/S
711- peão	2, 3	59	H/B/P/S

712- pechada	2	3	H
713- pécidê	3	4	n/d
714- pedrez (pedrês)	1, 2	17	H/B/P/S
715- pegadio	2	5	H
716- pegureiro	2	3	H/P/S
717- peia	1, 2, 3	25	H/B/P/S
718- peitilho	3	4	n/e (≠H/P/S)
719- peitoral	3	8	H/B/P/S
720- peitoril	1	9	H/B/P/S
721- pelego	2, 3	21	H
722- penaroso	2	3	H
723- peonada	2, 3	20	H
724- perambular	3	3	H
725- perneira	2	16	H/S (≠B/P)
726- peroba	2	14	H
727- perra	2	5	n/e (≠H/B/P/S)
728- perrengue	2	6	H
729- peste	2, 3	98	H (≠B/P/S)
730- pesteador	3	4	H
731- picaço	2	19	H
732- picar	3	4	H
733- piçarra	2	4	n/e (≠H/B/P/S)
734- picuá (piquá)	2	4	H
735- picumã	2	24	n/e (≠H)
736- piguancha	2	7	H
737- pilão	2, 3	47	H/B/P/S
738- pindaíba	3	6	H
739- pindaibal	3	3	H
740- pindaibeira	3	4	n/d
741- pinguela	3	5	H/P/S
742- pinote	3	8	H/B/P/S
743- piquete	2	18	H (≠B/P/S)
744- pirão	2, 3	22	H
745- pita	1	9	H/B/P/S
746- pito	2	9	H/S
747- pitombeira	1	4	H/B/P/S
748- piúca	2	3	H
749- pixuá	2	3	H
750- planura	2	11	H/B/P/S
751- poldro	1, 2	12	H/B/P/S
752- poleiro	1, 2, 3	24	H (≠B/P/S)
753- polvadeira	2	4	H
754- polvarinho	2	4	H/P/S
755- ponche	2	7	H (≠P/S)
756- porteira	2, 3	68	H (≠B/S)
757- portinhola	3	4	n/e (≠H/B/P/S)
758- postema	3	5	H (≠B/P/S)
759- potrada	2	4	H

760- potrilho	2	4	H
761- preá	1, 2	26	H/P/S
762- pretume	3	3	H
763- pua	2	5	H/B/P/S
764- pustema	3	5	n/d
765- quartau	1, 2	6	H
766- quebranto	3	18	H/B/P/S
767- quebreira	2	3	H
768- queixada	1, 2	9	H (≠B/P/S)
769- quenga	2	6	H
770- quengo	2	4	H
771- quiba	3	4	H
772- quicé	2, 3	9	H
773- quipá	2	5	H
774- quixabeira	2	5	H
775- rabeira	3	10	H
776- rabicho	1	12	H/B/P/S
777- rabo-de-macaco	3	129	H
778- rabo-de-tatu	2	44	H
779- raçudo	3	13	H
780- ramada	2	7	H/B/P/S
789- ramaria	2, 3	7	H
790- rancharia	2	5	H
791- rancho	1, 2, 3	166	H/P/S
792- rapador	2	3	H
793- rapadura	2	30	H (≠B/P/S)
794- rasgão	2	7	H
795- rebenque	2	17	H
796- rebento	2	10	H
797- rebojo	2	6	H
798- recanteado	3	3	n/d
799- rechã	2	7	n/e (≠H)
800- recosto	2	4	H/B/P/S
801- reculutamento	2	3	n/d
802- rédea	1, 2, 3	132	H/B/P/S
803- redemunho	3	10	H
804- redomão	2	5	H
805- refugar	2, 3	9	H (≠B/P/S)
806- regato	2	9	H/B/P/S
807- rego	3	27	H/B/P/S
808- reiúna	2	8	H
809- reiúno	2	9	H
810- relho	1, 2	25	H (≠B/P/S)
811- repasto	2	10	H
812- respônsio	2	8	n/d
813- restinga	2	13	H (≠B/P/S)
814- retado	3	8	H
815- retinto	3	22	H

816- retirante	2	21	H
817- retireiro	3	5	H
818- retiro	1, 3	61	H ( $\neq$ B/P/S)
819- retouçar	2	3	H/B
820- revigorativo	3	3	n/d
821- riba	2, 3	51	n/e ( $\neq$ H)
822- ribalta	3	14	n/e ( $\neq$ H)
823- riçada	2	3	n/d
824- rincão	2	12	H/B/P/S
825- rinha	3	17	H
826- risalhada	3	4	n/d
827- riscar	3	6	H ( $\neq$ B/P/S)
828- robalo	3	7	n/e ( $\neq$ H/B/P/S)
829- roca	1	3	H/B/P/S
830- roça	2	84	H/P/S ( $\neq$ B)
831- roçado	2	17	H ( $\neq$ B/S)
832- roçar	3	12	H/B/P/S
833- rocim	2	3	H/B/P/S
834- roda (de fiar)	3	61	H/B/P/S
835- rodeio	2	12	H/B ( $\neq$ P/S)
836- rodilha	2, 3	16	H/B/P/S
837- rola	1, 2	19	H/B/P/S
838- rompância	3	3	n/d
839-ronceiro	2	10	H/B/P/S
840- roncolho	3	6	H ( $\neq$ P/S)
841- ronqueira	2	7	H/B/P/S
842- roseta	2, 3	24	H/B/P/S
843- rosilho	1, 2	17	H/B/P/S
844- roxinha	3	4	H
845- ruano	2	10	H
846- ruço	1, 3	18	H/B/P/S
847- sabugo	3	7	H/B/P/S
848- saci	2	10	H
849- safardana	3	5	H
850- salpintar	2	7	n/e ( $\neq$ H)
851- samburá	3	5	H
852- sanga	2	21	H
853- sapé	2	14	H/S
854- sapezeiro	2	4	H
855- sapiquá (*sapicuá)	2	3	H
856- saracura	1	6	H
857- sarará	3	10	H
858- sebaça	3	3	H
859- sedenho	1, 2, 3	21	H ( $\neq$ B/P/S)
860- sela	1, 2, 3	164	H/B/P/S
861- selar	3	3	H/B/P/S
862- selim	2	6	H
863- sem-vergonhista	3	4	n/d

864- sendeiro	2	4	H/B/P/S
865- seriema	2	5	H
866- serrota	2	3	H
867- sertanejo	1, 2	584	H/B/S
868- sertão	1, 2, 3	1.343	H/B/P/S
869- severoso	3	4	n/d
870- seções	1, 2	17	H (≠B/P/S)
871- silhão	2	3	H/B/P/S
872- Sobrechinha (*sobrecincha)	2	3	H
873- socavão	1, 2	9	H/S
874- sombreiro	2, 3	8	H/B/P/S
875- sorro	2	4	H
876- sortista	3	3	H
877- sorubim	1	4	n/d
878- suaçupara	3	4	H
879- suador	1	4	H
880- suadouro	3	5	H/S (≠B)
881- subaco	3	3	n/d
882- suçuarana	1, 2	19	H
883- sungar	3	3	H
884- surrão	1, 2, 3	31	H/B/P/S
885- surucucu	3	52	H
886- tabaréu	2	7	H (≠P/S)
887- tabocal	1	3	H
888- taboca	1	5	H
889- tabuleiro	2	36	H/P/S
890- tacha	2	5	H (≠B/P/S)
891- taipa	2	20	H/B/P/S
892- tala	2	4	H (≠B/P/S)
893- talabarte	3	3	H/B/P/S
894- talcoxisto	2	4	n/d
895- talim	2	3	H/B/P/S
896- taluda	2	3	n/e (≠H)
897- tamborete	2, 3	18	H/B/P/S
898- tanajura	3	4	H
899- tapera	1, 2	61	H/B/P/S
900- tapioca	3	5	H/S
901- tapuio	2	4	H
902- taramela (*tramela)	2, 3	13	H/B/P/S
903- teiú	3	8	H
904- terém	2	4	H
905- teteré-teté	3	7	n/d
906- tílburí	3	7	H
907- tinguí	1	4	H/S
908- tipóia	2	5	H
909- tiração	3	4	H
910- tiracolo (*tiracol)	1, 3	14	H/B/P/S

911- tirante	2	8	H/B/P/S
912- tiririca	3	3	H
913- toada	2	19	n/e (≠H/B/P/S)
914- tocaia	2	34	H
915- tocaiar	3	5	H
916- tocaieiro	3	5	H
917- toleima	3	14	H/P/S
918- tolhiça	2	3	n/d
919- tombador	1, 2	7	n/e (≠H/B/P/S)
920- tordilho	2, 3	13	H/B/P/S
921- toreana	2	4	H
922- torrinha	3	6	n/e (≠H/P/S)
923- torva	2	5	n/e (≠S)
924- trabuco	2, 3	25	H (≠B/P/S)
925- tralha	3	11	H (≠P/S)
926- trambolho	2	6	H/B/P/S
927- tranqueira	2	16	H/B/P/S
928- tratador	3	7	H (≠P/S)
929- tremedeira	3	8	H
930- trempe	2, 3	33	H/B/P/S
931- trepadeira	3	7	H/B/P/S
932- trilheiro	3	9	n/d
933- troa	3	3	n/d
934- troço	2	6	H/B/P/S
935- tronco	2	54	n/e (≠H/B/P/S)
936- troncudo	3	3	H
937- tronqueira	3	4	H
938- tropa	1, 2, 3	264	H (≠B/P/S)
939- tropear	2	11	H (≠S)
940- tropeiro	1, 2, 3	55	H (≠S)
941- tropilha	2	6	H
942- trotão	2	4	H/B/P/S
943- troteada	2	3	H
944- tulha	3	7	H/B/P/S
945- turco	2	10	H/B/P/S
946- umburana	2	15	H
947- umbuzeiro	2	6	H
948- urupê	2	10	H
949- ururau	3	30	H
950- urutu	3	33	H
951- uturje	3	4	n/d
952- vacada	3	6	H/B/P/S
953- vadear	3	4	H/B/P/S
954- vadia	3	11	n/e (≠H)
955- vancê	2	104	H
956- vaqueano	1, 2	18	H
957- vaqueirama	2	4	H
958- vaqueiro	1, 2, 3	223	H/B/P/S

959- vaquejada	1, 2	23	H
960- vaqueta	2	6	H/P/S
961- varado	1, 2, 3	14	H (≠S)
962- varapau	2	3	H/P/S
963- vareio	3	3	H
964- varja	3	3	H
965- vassouredo	2	3	n/d
966- vassourinha	3	7	H (≠B/S)
967- vastar	3	4	H
968- vau	3	53	H/P/S
969- vaza	2	58	H (≠B/S)
970- venta	2, 3	44	H/B/P/S
971- verdal	3	4	n/d
972- verdume	3	5	n/d
973- vereda	1, 2, 3	1033	H/B/P/S
974- verruma	2	4	H/B/P/S
975- vezada	3	3	n/d
976- vime	3	5	H/B/P/S
977- vinhático	3	6	H/B/S
978- viração	2, 3	18	H (≠B/P/S)
979- virado	2, 3	18	H (≠B/S)
980- vizindário	2	4	H
981- volteado	3	4	n/d
982- vozeio	3	5	H
983- vuvu	3	3	H
984- xexéu	2	6	H
985- xucro	2	4	H
986- zaino	2	7	H/B/P/S
987- Zebu	3	8	H (≠B)

Fonte: Elaboração da autora

### 5.1 Análise quantitativa: número de vocábulos-terms dicionarizados, não dicionarizados e com acepções diferentes dos dicionários

Dos 987 vocábulos-terms apresentados, verificamos que 827 são dicionarizados, 91 não constam em nenhum dos dicionários consultados e 69 possuem acepções divergentes daquelas registradas nos dicionários. Para melhor ilustrar esses resultados, elaboramos gráficos correspondentes a cada um desses dados encontrados.

O GRÁFICO 1 mostra o total de vocábulos-terms dicionarizados, não dicionarizados e aqueles com acepções diferentes.



GRÁFICO 1 – Percentual de vocábulos-termos dicionarizados, não dicionarizados e com acepções diferentes



Fonte: Elaboração da autora

Nesse gráfico, os 987 vocábulos-termos estão dispostos em três áreas: a área azul, representando os vocábulos-termos dicionarizados, a área laranja, representando os vocábulos-termos não dicionarizados, e a área amarela, representando os vocábulos-termos com acepções diferentes das que estão registradas nos dicionários. Assim, temos: 827 itens dicionarizados, que equivalem a 83,79% do total, 91 itens não dicionarizados (9,22%) e 69 itens com acepções diferentes (6,99%).

Dentre os 827 vocábulos-termos dicionarizados, verificamos que 787 constam no dicionário Houaiss, 225 no dicionário de Bluteau, 233 no Silva Pinto e 274 no Moraes Silva. O GRÁFICO 2, a seguir, exhibe, em números absolutos, a quantidade de vocábulos-termos presentes em cada dicionário.

GRÁFICO 2 – Número de vocábulos-terms presentes em cada dicionário



Fonte: Elaboração da autora

No GRÁFICO 2, a barra amarela representa o dicionário Houaiss, com 787 vocábulos-terms, correspondendo a 95, 16% do total de itens dicionarizados; a verde, o dicionário de Moraes Silva, com 274 vocábulos-terms ou 33, 13% dos itens dicionarizados; a azul, o dicionário de Silva Pinto, com 233 ou 28, 17%; e a laranja, o dicionário de Bluteau, com 225 ou 27,26% dos 827 vocábulos-terms dicionarizados.

Podemos perceber que há mais vocábulos-terms registrados no dicionário *Houaiss* do que nos demais. Isto se deve ao fato de que *Houaiss*, além de ser uma obra de referência nos estudos lexicográficos de língua portuguesa com um vasto repertório lexical, é uma obra que traz um grande número de regionalismos, expressões e modismos usados em várias regiões do Brasil.

## 5.2 Quanto ao número de ocorrências e fases do Regionalismo

Em relação à frequência, verificamos que: há um número significativo de vocábulos-terms com frequência 3 (a menor frequência computada), mais especificamente 178 itens; há um grande número de vocábulos-terms com frequências entre 4 e 90, totalizando 782 itens; e um pequeno número de vocábulos-terms com frequência acima de 100, exatamente 27 itens.

O vocábulo-termo de maior ocorrência é “sertão”, palavra esta que caracteriza a temática do *corpus* e possui conceitos específicos do universo de discurso etnoliterário a que pertence. Isto implica dizer que não é por acaso que este vocábulo-termo é mais recorrente que os demais, contando com 1.343 ocorrências computadas nas três fases do Regionalismo ou *subcorpora*. Além de ser uma palavra-chave do universo de discurso sertanista, ela possui diferentes acepções construídas no contexto dos discursos-ocorrência e que são inerentes a cada fase do Regionalismo, isto é: ora “sertão” faz referência a um lugar agradável, de terras extensas e produtivas, povoado de homens bons, fortes e generosos (característica do romantismo que marca o Regionalismo pitoresco); ora faz referência a um lugar inóspito, precário, perigoso, marcado pelo flagelo das secas, cujo atraso, carência e exploração vitimam os sertanejos, transformando-os, muitas vezes, em pessoas rudes e cruéis (com vimos nas obras do Regionalismo crítico); ora faz referência a um “sertão mundo”, “sertão sem lugar” como caracteriza o próprio Guimarães Rosa em *Grande Sertão Veredas*, ou seja, como um lugar vasto e sem limites (fronteiras) marcado por problemas econômicos, políticos e sociais que estão presentes não somente em uma região em particular, mas em diversas regiões do país (daí o termo universalismo).

Em relação aos vocábulos-termos computados em cada fase, verificamos que 174 aparecem na primeira, 584 na segunda, 447 na terceira e 50 aparecem nas três fases. Como o 2º *subcorpus* (que corresponde à segunda fase) é composto por 10 textos e os demais *subcorpus* (que correspondem à primeira e terceira fases) possuem 5 textos cada, é natural que o número de vocábulos-termos presentes nele seja maior. No entanto, se compararmos a quantidade de itens presentes na segunda fase com o número de itens que aparecem na terceira, percebemos que há uma diferença mínima, o que revela a alta diversidade lexical do 3º *subcorpus*.

A próxima seção apresenta as 40 fichas etnoterminográficas dos vocábulos-termos selecionados<sup>34</sup> para esta pesquisa com seus respectivos verbetes. Esclarecemos que o número de fichas foi estabelecido com base no tempo disponível para a conclusão da pesquisa e na quantidade que julgamos suficiente para ilustrar a nossa proposta de Vocabulário.

---

<sup>34</sup> Vale reiterar, que buscamos selecionar os vocábulos-termos que apresentam características diversas: substantivo, adjetivo e verbo; baixa, média e alta frequência; presente ora em uma fase, ora em duas, ora nas três fases; dicionarizado e não dicionarizado; com uma ou mais de uma acepção; e com aspectos semânticos variados.

### 5.3 Apresentação das fichas etnoterminográficas e dos verbetes

Apresentamos, a seguir, as fichas etnoterminográficas e os verbetes dos vocábulos-  
termos selecionados. São 40 fichas dispostas em ordem alfabética e transcritas conforme o  
modelo descrito na seção 4.2.6.

Ficha 1 – <b>Aboiar</b> (verbo) _____ freq. 4
<i>“Encostado ao mourão da porteira de paus corridos, o vaqueiro das Aroeiras <b>aboiava</b> dolorosamente, vendo o gado sair, um a um, do curral.” (OQ, linha 1)</i>
Registro em dicionários:
1. Houaiss: v. 1. Regionalismo: Brasil. Conduzir (gado), entoando canto plangente ou soltando brados fortes e compassados.
2. Bluteau: n/d
3. Silva Pinto: n/e
4. Moraes Silva: n/e
Nota: Encontramos a palavra nos dicionários de Silva Pinto e Moraes Silva, mas com uma acepção diferente do contexto de uso.

Verbete:

**ABOIAR** (H) v. Cantar com sonoridade típica do vaqueiro para reunir e conduzir o gado.  
*“Encostado ao mourão da porteira de paus corridos, o vaqueiro das Aroeiras **aboiava** dolorosamente, vendo o gado sair, um a um, do curral.” (OQ, Reg2 [AB, VS]; freq.4)*

Ficha 2 – <b>Aguilhada</b> (substantivo) _____ freq. 7
<i>“[...]Je galopando sempre, através de todos os obstáculos, sopesando à destra sem a perder nunca, sem a deixar no inextricável dos cipoais, a longa <b>aguilhada</b> de ponta de ferro encastada em couro, que por si só constituiria, noutras mãos, sérios obstáculos à travessia” (OS, linha 3)</i>
Registro em dicionários:
1. Houaiss: s.f. 1 . m.q. aguilhão ('vara')
2. Bluteau: Vara, que tem hum ferro agudo na ponta, com que o boeyro pica os boys.
3. Silva Pinto: s.f. Vara com ponta de ferro aguda para picar os bois.

4. Moraes Silva: s. f. Vara com púa, ou ferrão para picar os bois.

Verbetes:

**AGUILHADA** (H/B/P/S) s. f. Vara longa com uma ponta de ferro destinada a picar os bois para incitá-los ao trabalho ou a entrar no curral. “[...] e galopando sempre, através de todos os obstáculos, sopesando à destra sem a perder nunca, sem a deixar no inextricável dos cipoais, a longa **aguilhada** de ponta de ferro encaستoada em couro, que por si só constituiria, noutras mãos, sérios obstáculos à travessia” (OS, Reg2; [Reg1: OS]; freq. 7)

Ficha 3 – **Aiô** (substantivo) \_\_\_\_\_ freq. 10

“Levava no **aiô** um frasco de creolina, e se houvesse achado o animal, teria feito o curativo ordinário” (VS, linha 2)

Registro em dicionários:

1. Houaiss: s. m. Regionalismo: Brasil. bolsa de caça trançada com fibras de caroá.
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**AIÓ** (H) s. m. Espécie de bolsa de caça feita de fibras de uma planta chamada caroá. “Levava no **aiô** um frasco de creolina, e se houvesse achado o animal, teria feito o curativo ordinário” (VS, Reg2, freq. 10)

Ficha 4 – **Amoitado** (adjetivo) \_\_\_\_\_ freq. 13

“Antes de confrontarmos com a fazenda da Tapera, eu fiz sô moço entrar num capãozinho de mato e ficar ali **amoitado** (PS, linha 1)

Registro em dicionários:

1. Houaiss: adj. Regionalismo: Brasil. Que se encontra agachado; acachapado; oculto
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**AMOITADO** (H) adj. Posição pela qual alguém se encontra escondido no mato, agachado atrás de uma moita “*Antes de confrontarmos com a fazenda da Tapera, eu fiz só moço entrar num capãozinho de mato e ficar ali **amoitado***” (PS, Reg1; [Reg3: GS, CB]; freq. 13)

Ficha 5 – **Aporrinhado** (adjetivo) \_\_\_\_\_ freq. 8

“*Mas ao saber de um padecimento, lá montava seu cavaleiro de costela de fora e saía, sempre resmungão, sempre **aporrinhado**, em auxílio do aflito.*” (CL, linha 8)

Registro em dicionários:

1. Houaiss: adj. bastante aborrecido; apoquentado
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**APORRINHADO** (H) adj. Estado pelo qual alguém se encontra aborrecido, mal-humorado, por alguma contrariedade. “*Mas ao saber de um padecimento, lá montava seu cavaleiro de costela de fora e saía, sempre resmungão, sempre **aporrinhado**, em auxílio do aflito.*” (CL, Reg3, freq. 8)

Ficha 6 – **Apustemado** (adjetivo) \_\_\_\_\_ freq. 3

“*Vosmecê sabe, esse **apustemado** é de Muribeca. Povo de Muribeca não presta, tudo tabaréu, lá não tem nada, não sabe vosmecê*” (SG, linha 2)

Registro em dicionários:

1. Houaiss: n/d
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**APUSTEMADO** (n/d) adj. Característica pela qual se nomeia alguém como imprestável, sem caráter, desgraçado. “*Vosmecê sabe, esse **apustemado** é de Muribeca. Povo de Muribeca não presta, tudo tabaréu, lá não tem nada, não sabe vosmecê*” (SG, Reg3, freq. 3)

Ficha 7– **Arção** (substantivo) \_\_\_\_\_ freq. 28

“*Aquele boi que ele tinha ao **arção** da sela, era o seu triunfo como vaqueiro, pois quando ele o apresentasse, todos o proclamariam o primeiro campeador, e sua fama correria o sertão*” (OS, linha 1)

Registro em dicionários:

1. Houaiss: s. m. Regionalismo: Brasil. armação da sela de montaria, de madeira revestida de couro, formada por uma arcada na dianteira e outra na traseira.
2. Bluteau: Arsão da sella do cavallo.
3. Silva Pinto: s. f. A parte elevada da sella por diante.
4. Moraes Silva: s. m. Arção da sella; a parte elevada por diante, e por detraz.

Verbetes:

**ARÇÃO** (H/B/P/S) s. m. Peça do arreio que dá arqueamento dianteiro e traseiro à sela de montaria. “*Aquele boi que ele tinha ao **arção** da sela, era o seu triunfo como vaqueiro, pois quando ele o apresentasse, todos o proclamariam o primeiro campeador, e sua fama correria o sertão*” (OS, Reg1; [Reg2: OS, TB, AB, OQ; Reg3: GS, CB]; freq. 28

Ficha 8– **Arreio** (substantivo) \_\_\_\_\_ freq. 94

“*Elias saltou fora dos **arreios**; o cavalo levantou-se imediatamente [...]*” (OG, linha 8)

Registro em dicionários:

1. Houaiss: s. m. Regionalismo: Brasil. **3.** Conjunto de peças com que se prepara a cavalgada para montaria; arreamento (mais us. no pl.). **4.** m.q. **sela**
2. Bluteau: Adereços ordinários do cavallo, V. G. arriata, cabeçadas, suftinentes, frontal, cingola, redeas, panno da silha, rabicho.
3. Silva Pinto: s.m. Adereço ordinário do cavallo.
4. Moraes Silva: s. m. Peças que adereção as bestas de serviço, carga, carruages.

Verbete:

**ARREIO** (H/B/P/S) s. m. Conjunto de equipamentos aplicados nos animais de sela que os sujeita à carga e aos comandos do homem (arrieiro). “*Elias saltou fora dos **arreios**; o cavalo levantou-se imediatamente [...]*” (OG, Reg1[IN, OS, PS]; [Reg2: CG, TB, UR, OQ, ME, VS, SV; Reg3: GS, CL, CB, SG]; freq. 94

Ficha 9 – **Barbatão** (substantivo) \_\_\_\_\_ freq. 28

“*Seu cavalo cardão rompeu o mato a galope, como quem estava acostumado a campear o **barbatão** no mais espesso bamburral [...]*” (OS, linha 15)

Registro em dicionários:

1. Houaiss: adj e s.m. Que ou o que foi criado no mato, tornando-se bravio (diz-se de rês)
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbete:

**BARBATÃO** (H) s. m. Animal bravo, criado solto no campo, sem marca ou ferrete. “*Seu cavalo cardão rompeu o mato a galope, como quem estava acostumado a campear o **barbatão** no mais espesso bamburral [...]*” (OS, Reg1; [Reg2: LH, AB]; freq. 28)

Ficha 10– **Barrigueira** (substantivo) \_\_\_\_\_ freq. 11

“*O caipira afrouxou a **barrigueira** do pedrês, sacudiu de riba a sela, e as caçambas de pau entrechocando-se uma na outra, foi guardá-la ao paiol*” (TB, linha 1)

Registro em dicionários:

1. Houaiss: s. f. Regionalismo: Brasil. **1.** Regionalismo: Brasil. Peça de arreio que prende a sela ao cavalo, que consiste numa espécie de cinta que envolve a barriga do animal; cilha
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbete:

**BARRIGUEIRA** (H) s. f. Peça do arreio que consiste em uma espécie de correia ou cinta que prende a sela ao corpo do animal. “*O caipira afrouxou a **barrigueira** do pedrês, sacudiu de*



*riba a sela, e as caçambas de pau entrechocando-se uma na outra, foi guardá-la ao paiol” (TB, Reg2; [Reg3: CL, CB]; freq. 11)*

Ficha 11– **Bolandeira** (substantivo) \_\_\_\_\_ freq. 24

*“Seu Tomás fugira também, com a seca, a **bolandeira** estava parada” (VS, linha 1)*

Registro em dicionários:

1. Houaiss: s. f. Regionalismo: Brasil. **1.** Regionalismo: Brasil. nos engenhos de açúcar, grande roda dentada que gira sobre a moenda movimentando as mós.
2. Bluteau: n/d
3. Silva Pinto: s. f. Roda do engenho de açúcar.
4. Moraes Silva: s. f. Roda do engenho de assucar, pegada no eixo do meyo, movida pelo rodete.

Verbetes:

**BOLANDEIRA** (H/P/S) s. f. Roda grande e dentada, própria do engenho de açúcar, utilizada para moenda. *“Seu Tomás fugira também, com a seca, a **bolandeira** estava parada” (VS, Reg2, freq. 24)*

Ficha 12 – **Burgariana** (substantivo) \_\_\_\_\_ freq. 3

*“Ficara no armazém, estranha mercadoria entre a carne-sêca, a chita, a **burgariana**, a cachaça e o feijão” (SV, linha 1)*

Registro em dicionários:

1. Houaiss: n/d
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**BURGARIANA** (n/d) s. f. Espécie de tecido fino e de pouco valor, usado pelos sertanejos das regiões áridas e quentes do sertão. *“Ficara no armazém, estranha mercadoria entre a carne-sêca, a chita, a **burgariana**, a cachaça e o feijão” (SV, Reg2, freq.3).*

Ficha 13 – **Caçuá** (substantivo) \_\_\_\_\_ freq. 22

*“[...] foi atirando os sacos aqui, os **caçuás** acolá, a cangalha além, e desembestou por fim, pela margem afora, em violenta fuga” (OC, linha 5)*

Registro em dicionários:

1. Houaiss: s. m. Regionalismo: Brasil. Cesto grande e comprido de vime, cipó ou bambu, sem tampa e com alças para prender às cangalhas no transporte de gêneros diversos em animais de carga.

2. Bluteau: n/d

3. Silva Pinto: n/d

4. Moraes Silva: n/d

Verbetes:

**CAÇUÁ** (H) s. m. Cesto grande, feito de cipó de vime ou bambu, utilizado para transporte de gêneros diversos, comumente colocado no dorso dos animais de carga. *“[...] foi atirando os sacos aqui, os **caçuás** acolá, a cangalha além, e desembestou por fim, pela margem afora, em violenta fuga” (OC, Reg1; [Reg2: OS, ME, SV]; freq. 22)*

Ficha 14 – **Cacunda** (substantivo) \_\_\_\_\_ freq. 4

*“-Estou no fim. Estômago, fígado, uma dor aqui no peito que responde na **cacunda**” (UR, linha 6)*

Registro em dicionários:

1. Houaiss: s.f. 1. Regionalismo: Brasil. m.q corcunda.

2. Bluteau: n/d

3. Silva Pinto: n/d

4. Moraes Silva: n/d

Ficha 14.1 – **Cacunda** (substantivo) \_\_\_\_\_ freq. 2

*“Meu pai, me perdoe que minha **cacunda** está cansada de levar tanto pecado, de carregar tanta desgraça!” (SV, linha 2)*

Registro em dicionários:

1. Houaiss: s.f. 2. Consciência.

2. Bluteau: n/d

3. Silva Pinto: n/d

4. Moraes Silva: n/d

Verbetes:

**CACUNDA** (H) s. f. **1.** Parte do corpo referente às costas. “- *Estou no fim. Estômago, fígado, uma dor aqui no peito que responde na **cacunda***” (UR, Reg2[AB]. **2.** Por extensão, faz referência à consciência. “*Meu pai, me perdoe que minha **cacunda** está cansada de levar tanto pecado, de carregar tanta desgraça!*” (SV, Reg2; freq. 6)

Ficha 15 – **Cangalha** (substantivo) \_\_\_\_\_ freq. 52

*“Levantou-se o ofertante com toda a boa vontade e às apalpadelas começou a procurar a cama do patrão, o que só conseguiu depois de ter esbarrado na mesa e numa **cangalhas** velhas atiradas a um canto da sala” (IN, linha 3)*

Registro em dicionários:

1. Houaiss: s. f. 1 Regionalismo: Brasil. Artefato de madeira ou ferro, ger. Acolchoado, que se apõe ao lombo das cavalgaduras para pendurar carga e ambos os lados.

2. Bluteau: Armadilha de paos q formao como hua grade larga, para sustentar as quartas, que os Aguadeiros carregaõ nas bestas.

3. Silva Pinto: s.f. Duas armadilhas com feição de canastras, que se pões sobre a albarda das bestas pendentes huma de cada lado para cargas.

4. Moraes Silva: s.f. pl. Duas como canastras de grade de páo que se accommodáo no seladouro das bestas, pendendo uma de cada lado para certas cargas.

Verbetes:

**CANGALHA** (H/B/P/S) s. f. Armação de madeira que se coloca sobre o dorso dos animais para transporte de carga. “*Levantou-se o ofertante com toda a boa vontade e às apalpadelas começou a procurar a cama do patrão, o que só conseguiu depois de ter esbarrado na mesa e numa **cangalhas** velhas atiradas a um canto da sala” (IN, Reg1[OS, OC, PS]; [Reg 2: LH, CG, TB, UR, AB, OQ, ME, SV; Reg 3: GS, CL]; freq. 52 ).*

Ficha 16 – **Capiau** (substantivo) \_\_\_\_\_ freq. 8

*“Rezo que ele falou aquilo, aquele **capiau** peludo, renasceu minha alegria” (GS, linha 1)*

Registro em dicionários:

1. Houaiss: adj. e s. m. Regionalismo: Brasil. Regionalismo: Bahia, Minas Gerais. m.q. **caipira** ('roceiro').
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**CAPIAU** (H) s. m. Indivíduo característico da roça, do sertão, com hábitos e modos rudes, de pouca instrução ou convivência urbana. *“Rezo que ele falou aquilo, aquele **capiau** peludo, renasceu minha alegria” (GS, Reg3, freq. 10)*

Ficha 17 – **Catolé** (substantivo) \_\_\_\_\_ freq. 3  
*“Flor, vendo as pinhas de coquinhos amarelos, cobiçou-os, e pediu os **catolés**” (OS, linha 3)*

Registro em dicionários:

1. Houaiss: s. m. M. q. Catulé. Fruto das palmeiras; guariroba; Indaiá-do-campo.
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**CATOLÉ** (H) s. m. Fruto da guariroba, comestível, de formato pequeno e cor verde-amarelado, casca dura e polpa fibrosa. *“Flor, vendo as pinhas de coquinhos amarelos, cobiçou-os, e pediu os **catolés**” (OS, Reg1, freq. 3).*

Ficha 18 – **Fiofó** (substantivo) \_\_\_\_\_ freq. 4  
*“Para não ser censurado: no **fiofó** do cavalo fez o dinheiro guardado” (AC, linha 3)*

Registro em dicionários:

1. Houaiss: s. m. Regionalismo: Brasil. Uso: informal. ânus
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**FIOFÓ** (H) s. m. Parte do corpo humano ou de animal, em referência informal ao ânus. “Para não ser censurado: no **fiofó** do cavalo fez o dinheiro guardado” (AC, Reg3 [SG]; freq. 4)

Ficha 19 – **Galista** (adjetivo e substantivo) \_\_\_\_\_ freq. 12

“[...] soube que o major, com o desgosto sofrido, esvaziou todas as gaiolas de criação e deu por mal acabada a sua carreira de **galista**” (CL, linha 8)

Registro em dicionários:

1. Houaiss: adj. e s. m. Regionalismo: Brasil. **1.** diz-se de ou criador e treinador de galos de briga. **2.** que ou aquele que é apreciador de briga de galos
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**GALISTA** (H) adj. e s. m. Homem praticante de briga de galo, que cria ou treina galos de briga. “[...] soube que o major, com o desgosto sofrido, esvaziou todas as gaiolas de criação e deu por mal acabada a sua carreira de **galista**” (CL, Reg3, freq. 12)

Ficha 20 – **Garajau** (substantivo) \_\_\_\_\_ freq. 9

“Mas um **garajau**, os caçuás, um **garajau** cheio de barro, um **garajau** cheio de boi, ai mãe, a laranjeira murcha, os pés de árvore, tudo morto, sacudindo pela rodagem velha de Laranjeira” (SG, linha 8)

Registro em dicionários:

1. Houaiss: s. m. **1.** Regionalismo: Brasil. cesto de cipó resistente, oblongo, sobre uma base quadrada, aberto ao alto, destinado à condução de aves ao mercado; us. aos pares, pendentes das extremidades de um calão levado ao ombro. **2.** Regionalismo: Nordeste do Brasil. aparelho em que se conduz louça de barro, a cavalo ou a pé. **3.** Regionalismo: Rio Grande do Norte. atado de varas, com passadeiras de cipó ou de palha de carnaúba, onde são acondicionadas camadas sucessivas de peixe seco.
2. Bluteau: n/d

3. Silva Pinto: n/d

4. Moraes Silva: n/d

Verbetes:

**GARAJAU** (H) s. m. **1.** Espécie de cesto de cipó fechado, no qual os roceiros conduzem galinhas e outras aves ao mercado. **2.** Petrecho, geralmente feito de paus, embiras ou cipós, usado para o transporte de louças de barro, a cavalo ou a pé. **3.** Cesto trançado de taquara ou cipó usado no transporte de cargas, sobretudo preso ao lombo de animais. “*Mas um **garajau**, os caçuás, um **garajau** cheio de barro, um **garajau** cheio de boi, ai mãe, a laranjeira murcha, os pés de árvore, tudo morto, sacudindo pela rodagem velha de Laranjeira*” (SG, Reg3, freq.9)

Ficha 21 – **Garrota** (substantivo) \_\_\_\_\_ freq. 3

“*Uma vez corremos atrás de uma **garrota**, das seis da manhã até as seis da tarde, sem parar nem um momento, eu a cavalo, ele a pé*” (AC, linha 3)

Registro em dicionários:

1. Houaiss: s.f. novilha até dois anos de idade; bezerra.

2. Bluteau: n/d

3. Silva Pinto: n/d

4. Moraes Silva: n/d

Nota: Encontramos a palavra *garrote* nos dicionários de Bluteau, Silva Pinto e Moraes Silva, mas com uma aceção diferente do contexto de uso.

Verbetes:

**GARROTA** (H) s. f. Novilha, bezerra de até dois anos de idade. “*Uma vez corremos atrás de uma **garrota**, das seis da manhã até as seis da tarde, sem parar nem um momento, eu a cavalo, ele a pé*” (AC, Reg3, freq. 3)

Ficha 22 – **Jirau** (substantivo) \_\_\_\_\_ freq. 19

“*Num **jirau**, umas panelas velhas com craveiros brotando e bogaris pelas biqueiras florindo*” (ME, linha 12)

Registro em dicionários:

1. Houaiss: s. m. Regionalismo: Brasil. **1.** Armação de madeira semelhante a estrado ou palanque, que pode ser us. como cama, depósito de utensílios domésticos, secador de frutas ou, quando posta em cima de um fogão, como fumeiro de carne, toucinho, peixe etc.

2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Ficha 22.1 – **Jirau** (substantivo) \_\_\_\_\_ freq. 6

*“A velha casa de taipa negrejava ao sol o telhado de **jirau**” (OQ, linha 2)*

Registro em dicionários:

1. Houaiss: s. m. Regionalismo: Brasil. 2. Armação de madeira sobre a qual se constrói uma casa de modo a evitar a água e a umidade
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Ficha 22.2 – **Jirau** (substantivo) \_\_\_\_\_ freq. 5

*“Entrou no seu rancho pobre de sapé, sobre cuja esteira de **jirau** perrengueava a mulher, numa recaída de resguardo de parto” (TB, linha 3)*

Registro em dicionários:

1. Houaiss: s. m. Regionalismo: Brasil. 1. Armação de madeira semelhante a estrado ou palanque, que pode ser us. como cama.
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**JIRAU** (H) s.m. 1. Espécie de mesa forrada com varas, utilizada como apoio para utensílios de cozinha, depois de lavados, ou objetos domésticos. *“Num **jirau**, umas panelas velhas com craveiros brotando e bogaris pelas biqueiras florindo” (ME, Reg2 [OS, LH, TB, VS]).* 2. Armação de madeira sobre a qual se edificam as casas para evitar a água e a umidade. *“A velha casa de taipa negrejava ao sol o telhado de **jirau**” (OQ, Reg2 [OS]).* 3. Espécie de

cama feita com varas de madeira e sobre esteios fixados no chão. “*Entrou no seu rancho pobre de sapé, sobre cuja esteira de **jirau** perrengueava a mulher, numa recaída de resguardo de parto*” (TB, Reg2 [OS, LH, TB, VS]; freq. 30)

Ficha 23 – **Jururu** (adjetivo) \_\_\_\_\_ freq. 3

“*Não! que nesse caso fechava mesmo o tempo, acostava-se **jururu** à rapariga e fazia trabalhar o ferro, costurando o mais intrometido [...]*” (UR, linha 2)

Registro em dicionários:

1. Houaiss: adj. Regionalismo: Brasil. 1. Que perdeu a alegria; triste, melancólico.
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**JURURU** (H) adj. Estado pelo qual alguém se encontra triste, abatido. “*Não! que nesse caso fechava mesmo o tempo, acostava-se **jururu** à rapariga e fazia trabalhar o ferro, costurando o mais intrometido [...]*” (UR, Reg2 [TB, AB]; freq.3).

Ficha 24 – **Macambira** (substantivo) \_\_\_\_\_ freq. 34

“*Veladas por touceiras inextricáveis de **macambiras** ou lascas de pedra, não se revelavam a distancia*” (OS, linha 3)/ “*Se tivessem tido que agüentar o rojão de João Grilo, passando fome e comendo **macambira** na seca, garanto que tinham mais coragem*” (AC, linha 1)

Registro em dicionários:

1. Houaiss: s. f. **1.** planta terrestre ou epífita (*Bromelia laciniosa*), da fam. das bromeliáceas, nativa do Brasil, encontrada nas caatingas do Nordeste, de folhas verdes com linhas róseas, armadas de espinhos curvos, us. para extração de fibras ou como ração.
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**MACAMBIRA** (H) s. f. Espécie de planta de folhas rígidas e espinhosas encontrada nas regiões secas nordestinas, usada para extração de fibras ou ração, cujas folhas servem também



como alimento dos sertanejos na seca. “Se tivessem tido que agüentar o rojão de João Grilo, passando fome e comendo **macambira** na seca, garanto que tinham mais coragem” (AC, Reg3 [SG]; [Reg2: OS, VS]; freq. 34)

Ficha 25 – **Maravalhas** (substantivo) \_\_\_\_\_ freq. 3

“Ao longo do caminho, de um e outro lado, alvejavam, entre as **maravalhas** dos ramos queimados pelo sol”/ “[...]Joana fez em pedaços a tábua, e entupiu com pedras e **maravalhas** o buraco que com aquela armava ciladas aos inofensivos filhos do deserto” (OS, linha 2 e 3)

Registro em dicionários:

1. Houaiss: n/d
2. Bluteau: Aparas delgada que se tira da madeira com garlopa; ramos.
3. Silva Pinto: s.f. plur. Aparas delgadas que os carpinteiros tirão a madeira.
4. Moraes Silva: s. m. pl. Umas como fitas que os carpinteiros tirão da madeira. Accender fogo com maravalhas; coisa que faz fogo de labareda.

Verbete:

**MARAVALHAS** (B/P/S) s. f. Galhos finos e secos de plantas ou lascas de madeira, usados para acender fogo ou forrar o piso de galinheiros. “Joana fez em pedaços a tábua, e entupiu com pedras e **maravalhas** o buraco que com aquela armava ciladas (OS, Reg1 [OC]; freq. 3).

Ficha 26– **Murzelo** (adjetivo e substantivo) \_\_\_\_\_ freq. 3

“Escolhi um, animal vistoso, celheado, acastanhado **murzelo**, que bem me pareceu; e dei em erro, porque ele era meio sendeiro e historiento”/ “Alguém disse que o cavalo grande, **murzelo**-mancho, devia de ficar sendo dele mesmo” (GS, linha 1 e 3).

Registro em dicionários:

1. Houaiss: adj. e s.m. Diz-se de ou cavalo negro, da cor da amora-preta.
2. Bluteau: Mursêlo. Cor de cavallo, semelhante à da amora.
3. Silva Pinto: adj. Diz-se do cavalo, que tem côr de amora ou castanho escuro.
4. Moraes Silva: Murséllo. Cavallo murséllo; còr de amora preta.

Verbete:

**MURZELO** (H/B/P/S) adj. e s. m. Característica pela qual se nomeia o cavalo por sua pelagem de cor preta, amora-preta ou castanho escuro. “Escolhi um, animal vistoso, celheado,

acastanhado **murzelo**, que bem me pareceu; e dei em erro, porque ele era meio sendeiro e historiento [...]”(GS, Reg3, freq. 3).

Ficha 27 – **Paiol** (substantivo) \_\_\_\_\_ freq. 78

“[...] as monótonas modulações de umas chulas e modinhas, cantadas ao som da viola de três cordas pelos camaradas de Cirino, acomodados no rancho junto ao **paiol** de milho” (IN, linha 1)

Registro em dicionários:

1. Houaiss: s. m. **3**. Regionalismo: Brasil. armazém para depósito de produtos agrícolas em geral. **4**. Regionalismo: Minas Gerais, São Paulo. m.q. **celeiro** ('tulha')

2. Bluteau: n/e

3. Silva Pinto: n/e

4. Moraes Silva: n/e

Nota: Encontramos a palavra nos dicionários de Bluteau, Silva Pinto e Moraes Silva, mas com uma acepção diferente do contexto de uso.

Verbetes:

**PAIOL** (H) s. m. Armazém para depósito de milho, cereais e outros gêneros da lavoura. “[...] as monótonas modulações de umas chulas e modinhas, cantadas ao som da viola de três cordas pelos camaradas de Cirino, acomodados no rancho junto ao **paiol** de milho” (IN, Reg1 [PS]; [Reg2: CG, TB; Reg3: GS, CL, CB]; freq.78)

Ficha 28 – **Panasco** (substantivo) \_\_\_\_\_ freq. 7

“Ao tropel dos animais surdiam das touceiras de **panasco** os novilhos e garrotes mansos, que deitavam a correr pelo campo;”(OS, linha 1)

Registro em dicionários:

1. Houaiss: s.m 1. Erva que serve de alimento ao gado.

2. Bluteau: n/d

3. Silva Pinto: s.m. Espécie de herva para pasto.

4. Moraes Silva: s.m. Espécie de herva de pasto.

Verbetes:

**PANASCO** (H/P/S) s. m. Espécie de vegetal, utilizada como alimento para o gado. “Ao tropel dos animais surdiam das touceiras de **panasco** os novilhos e garrotes mansos, que deitavam a correr pelo campo” (OS, Reg1; [Reg2: OS, AB, OQ]; freq. 7).

Ficha 29 – **Patativa** (substantivo) \_\_\_\_\_ freq. 3

“No meio da orquestra concertada pelos cantos dos sabiás, das graúnas e das **patativas**, retiniam os clamores das maracanãs, os estríduos das arapongas, e os gritos dos tiés e das araras” (OS, linha 2)

Registro em dicionários:

1. Houaiss: s. f. 1. Regionalismo: Brasil. Ave passeriforme (*Sporophila plumbea*), da fam. dos emberizídeos, encontrada nas regiões meridionais e setentrionais do Brasil e em países adjacentes; com cerca de 10,5 cm de comprimento, o macho é cinza-azulado com mento, estria malar, abdome e espéculo brancos, enquanto a fêmea e o imaturo são pardacentos [Espécie canora muito procurada no mercado de aves de gaiola].

2. Bluteau: n/d

3. Silva Pinto: n/d

4. Moraes Silva: n/d

Verbetes:

**PATATIVA** (H) s. f. Espécie de pássaro de cor cinzenta e canto apreciado. “No meio da orquestra concertada pelos cantos dos sabiás, das graúnas e das **patativas**, retiniam os clamores das maracanãs, os estríduos das arapongas, e os gritos dos tiés e das araras” (OS, Reg1 [OG, OC]; freq. 3)

Ficha 30 – **Pesteado** (adjetivo) \_\_\_\_\_ freq. 4

“[...] era de beiços, mostrando a língua à grossa, colada no assoalho da boca, mas como se fosse uma língua demasiada demais, que ali dentro não pudesse caber; em bezerro **pesteado**, às vezes, se vê assim” (GS, linha 2)

Registro em dicionários:

1. Houaiss: adj. Regionalismo: Brasil. **1.** que se encontra atacado de peste. **2.** doente, enfermo

2. Bluteau: n/d

3. Silva Pinto: n/d

4. Moraes Silva: n/d

Verbete:

**PESTEADO** (H) adj. Característica daquele que se encontra doente ou vítima de alguma peste. “[...] *era de beiços, mostrando a língua à grossa, colada no assoalho da boca, mas como se fosse uma língua demasiada demais, que ali dentro não pudesse caber; em bezerro **pesteado**, às vezes, se vê assim*” (GS, Reg3 [CB]; freq. 4)

Ficha 31 – **Pindaibeira** (substantivo) \_\_\_\_\_ freq. 4

*“Fazendo picada com as mãos por entre o ramalhal e cipoama da **pindaibeira**, José de Arimateia entrou pelo capão até a água da nascente.”* (CB, linha 3)

Registro em dicionários:

1. Houaiss: n/d
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbete:

**PINDAIBEIRA** (n/d) s. f. Árvore frutífera esguia dos brejos, das cabeceiras de nascentes de água, cuja casca fornece fio branco para fazer cordas e os frutos são comestíveis. *“Fazendo picada com as mãos por entre o ramalhal e cipoama da **pindaibeira**, José de Arimateia entrou pelo capão até a água da nascente”*(CB, Reg3, freq. 4)

Ficha 32 – **Potrilho** (substantivo) \_\_\_\_\_ freq. 4

*“Nasceu o **potrilho**, lindo e gordo, filho de égua boa leiteira, crioula de campo de lei”* (CG, linha 4)

Registro em dicionários:

1. Houaiss: s. m 1. Regionalismo: Sul do Brasil. m.q. potranco: potro de menos de dois anos.
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**POTRILHO** (H) s. m. Cavalo macho novo, com menos de dois anos; diminutivo de potro. *“Nasceu o **potrilho**, lindo e gordo, filho de égua boa leiteira, crioula de campo de lei”* (CG, Reg2, freq. 4)

Ficha 33 – **Quartau** (substantivo) \_\_\_\_\_ freq. 6

*“[...]ele montado num possante **quartau** pedrez, eu à garupa”* (LH, linha 3)

Registro em dicionários:

1. Houaiss: 3. Regionalismo: Nordeste Brasil. 4. Cavalo manso, castrado. cavalo quadrado e corpulento, mas de pequeno porte; quartau.
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**QUARTAU** (H) s. m. Cavalo manso, robusto, bom de sela. *“[...]ele montado num possante **quartau** pedrez, eu à garupa”* (LH, Reg2 [AB,OQ]; [Reg1: OC]; freq. 6).

Ficha 34– **Rancho** (substantivo) \_\_\_\_\_ freq. 106

*“Entrou no seu **rancho** pobre de sapé, sobre cuja esteira de jirau perrengueava a mulher, numa recaída de resguardo de parto”* (TB, linha 24) / *“[...]Mas espiava as cabaças para bóia de anzol, sempre dependuradas na parede do **rancho**”* (GS, linha 12)

Registro em dicionários:

1. Houaiss: s. m. 9. Habitação pobre; choça, choupana. 10. Fazenda de criação.
2. Bluteau: n/e
3. Silva Pinto: n/e
4. Moraes Silva: n/e

Nota: Encontramos a palavra nos dicionários de Silva Pinto e Moraes Silva, mas com uma acepção diferente do contexto de uso. No dicionário de Bluteau, a palavra aparece no resultado de busca, mas não aparece no dicionário.

Ficha 34.1– **Rancho** (substantivo) \_\_\_\_\_ freq. 51

*“[...]os **ranchos** improvisados e cobertos de capim; as barracas e os carros de bois, outras barracas ambulantes, com seu toldo de couro, agrupados em desordem pelas Campinas e vargedos vizinhos, abrigavam uma multidão de famílias sertanejas (OG, linha 12)*

Registro em dicionários:

1. Houaiss: s. m. 9.1. Cabana us. como abrigo temporário ou para descanso de trabalhadores.
2. Bluteau: n/e
3. Silva Pinto: s.m. Tenda ou barraca móvel nos caminhos.
4. Moraes Silva: s.m. Casa, ou tenda móvel, que se faz pelos caminhos.

Nota: Não encontramos a palavra no dicionário de Bluteau, apesar de aparecer no resultado de busca.

Ficha 34.2 – **Rancho** (substantivo) \_\_\_\_\_ freq. 9

*“Quando o rancho dos meninos chegava à casa, apareceu-lhe Arnaldo, com uma coleção de ninhos de anuns” (OS, linha 4)*

Registro em dicionários:

1. Houaiss: s. m. 1. Grupo de pessoas reunidas para determinado fim, esp. em marcha ou jornada.
2. Bluteau: n/e
3. Silva Pinto: n/e
4. Moraes Silva: n/e

Nota: Encontramos a palavra nos dicionários de Silva Pinto e Moraes Silva, mas com uma acepção diferente do contexto de uso.

Verbetes:

**RANCHO** (H/P/S) s. m. **1.** Casa pobre da roça, do sertão. *“Entrou no seu **rancho** pobre de sapé, sobre cuja esteira de jirau perrengueava a mulher, numa recaída de resguardo de parto” (TB, Reg2 [OS, LH, CG, UR, AB, OQ, ME]; [Reg1: OG, IN, OS, OC, PS; Reg3: GS, CB]).* **2.** Construção rústica e provisória, geralmente feita de capim, ramos ou palha, usada como abrigo ou descanso de trabalhadores sertanejos (especialmente tropeiros ou boiadeiros). *“[...]os **ranchos** improvisados e cobertos de capim; as barracas e os carros de bois, outras barracas ambulantes, com seu toldo de couro, agrupados em desordem pelas Campinas e vargedos vizinhos, abrigavam uma multidão de famílias sertanejas (OG, Reg1[IN, OS, PS]; [Reg2: OS, LH, OQ, SV; Reg3: GS, CB]).* **3.** Bando de pessoas reunidas para um fim qualquer.

*“Quando o rancho dos meninos chegava à casa, apareceu-lhe Arnaldo, com uma coleção de ninhos de anuns” (OS, Reg1; [Reg2: LH, UR; Reg3: CB]; freq. 166)*

Ficha 35 – **Rosilho** (adjetivo e substantivo) \_\_\_\_\_ freq. 17

*“[...] atravessando a vila ao lado da sua amada, montado no próprio **rosilho** em que tantas brilhaturas fizera nas cavalcadas” (OG, linha 3)*

Registro em dicionários:

1. Houaiss: adj. ou s. m. 1. que ou o que tem o pelo avermelhado entremeado de branco, o que dá um aspecto de cor rosada (diz-se de cavalo).
2. Bluteau: Cavallo rofilho. *Vid.* Rutilho.
3. Silva Pinto: Russilho.
4. Moraes Silva: Rusilbe.

Verbetes:

**ROSILHO** (H//B/P/S) adj. e s. m. Nome dado ao cavalo, derivado do tom de pelagem avermelhada e branca, dando aspecto de cor rosada. *“[...] atravessando a vila ao lado da sua amada, montado no próprio **rosilho** em que tantas brilhaturas fizera nas cavalcadas” (OG, Reg1; [Reg2: TB, UR]; freq. 17)*

Ficha 36 – **Suaçuapara** (substantivo) \_\_\_\_\_ freq. 4

*“**Suaçuapara** corria da gente, com a cabeça empinada quase nas costas, protegendo para não prender nas árvores sua galhadura dele” (GS, linha 1)*

Registro em dicionários:

1. Houaiss: s. m. Regionalismo: Brasil. m.q. cervo-do-pantanal (*Blastocerus dichotomus*); veado-campeiro (*Ozotoceros bezoarticus*)
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**SUAÇUAPARA** (H) s. m. Espécie de veado-campeiro com cerca de 1 m de comprimento, de pelagem marrom, galhada com três pontas e cerca de 30 cm de altura. *“**Suaçuapara** corria da*

gente, com a cabeça empinada quase nas costas, protegendo para não prender nas árvores sua galhadura dele” (GS, Reg3, freq. 4)

Ficha 37 – **Tiririca** (substantivo) \_\_\_\_\_ freq. 1

“Mandei abrir o casarão, vassourar salas e varandas, tosar as **tiriricas**, limpar a chácara das ervas daninhas e carurus”(CL, linha 2)

Registro em dicionários:

1. Houaiss: s.f. Regionalismo: Brasil. 1. Planta de até 40 cm (*Cyperus rotundus*), com inflorescências umbeliformes, de distribuição cosmopolita, considerada uma das piores ervas daninhas, us. como estimulante, febrífuga, diurética e adstringente, e outrora em perfumaria; junça-aromática, pé-de-galinha, tiririca-comum.

2. Bluteau: n/d

3. Silva Pinto: n/d

4. Moraes Silva: n/d

Ficha 37.1 – **Tiririca** (substantivo) \_\_\_\_\_ freq. 2

“Agora o que se construi é ele contra mim e até Amaro dá sua parte e depois fica com suas **tiriricas**, mirando o rabo da menina com cara de cachorro que quebrou o prato/ Lá era melhor, pelo menos tinha os bois e as jias para a gente ficar falando mal e Amaro ficava cortando as **tiriricas** deles,”(SG, linha 1 e 3)

Registro em dicionários:

1. Houaiss: n/e

2. Bluteau: n/d

3. Silva Pinto: n/d

4. Moraes Silva: n/d

Nota: Encontramos no dicionário *Houaiss* apenas a definição descrita na ficha 1

Verbetes:

**TIRIRICA** (H) s. f. **1.** Espécie de erva daninha; praga. “Mandei abrir o casarão, vassourar salas e varandas, tosar as **tiriricas**, limpar a chácara das ervas daninhas e carurus” (CL, Reg3). **2.** Expressão usada para se referir ao órgão genital humano ou de animal. “Lá era melhor, pelo menos tinha os bois e as jias para a gente ficar falando mal e Amaro ficava cortando as **tiriricas** deles” (SG, Reg3; freq.3).

Ficha 38 – **Tronqueira** (substantivo) \_\_\_\_\_ freq. 4



*“Tem uma errada, mas é só no corredor; a porteira certa é a **tronqueirinha** de arame liso — o senhor pega por ela, arrodeia a cabeceira de um resfriado de buriti, e vai sair nos currais da fazendinha” (CB, linha 2)*

Registro em dicionários:

1. Houaiss: s. f. Tronqueira. Regionalismo: Sul do Brasil. cada uma das estacas grossas que sustentam as varas de porteira ou cancela; mourão, tronqueiro.
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**TRONQUEIRA** (H) s. f. Estaca de madeira que sustenta a cerca de arame ou varas da porteira. *“Tem uma errada, mas é só no corredor; a porteira certa é a **tronqueirinha** de arame liso — o senhor pega por ela, arrodeia a cabeceira de um resfriado de buriti, e vai sair nos currais da fazendinha” (CB, Reg3, freq. 4)*

Ficha 39 – **Ururau** (substantivo) \_\_\_\_\_ freq. 30

*“Nesse entrementes, sei lá saído de onde, apareceu aquele gogó de espinho para dizer que **ururau** nunca foi jacaré de correr na frente de qualquer comedor de farinha” (CL, linha 27)*

Registro em dicionários:

1. Houaiss: s. m. Regionalismo: Brasil. m.q. jacaré-de-papo-amarelo (Caiman latirostris)
2. Bluteau: n/d
3. Silva Pinto: n/d
4. Moraes Silva: n/d

Verbetes:

**URURAU** (H) s. m. Espécie de jacaré de papo amarelo que, segundo uma lenda, seria a transformação de um rapaz por causa de uma paixão proibida. *“Nesse entrementes, sei lá saído de onde, apareceu aquele gogó de espinho para dizer que **ururau** nunca foi jacaré de correr na frente de qualquer comedor de farinha” (CL, Reg3, freq. 30)*

Ficha 40 – **Xexéu** (substantivo) \_\_\_\_\_ freq. 6

*“O xexéu de minha terra que me ensinou a cantar... Antes me tirasse o canto e me ensinasse a voar...” (AB, linha 3)*

Registro em dicionários:

1. Houaiss: s. m. Regionalismo: Brasil. m.q. **japiim** (*Cacicus cela*). Ave passeriforme (*Cacicus cela*), da fam. dos emberizídeos, florestal e comum, de ampla distribuição na região tropical americana, atingindo em sua dispersão este-meridional o litoral sul da Baía; de coloração geral negra com baixo dorso, coberteiras das asas e base da cauda amarelo-brilhante; baguá, xexéu [Famosa pelo canto variadíssimo e pela habilidade de produzir imitações dos sons de outras aves e mamíferos de seu *habitat*.

2. Bluteau: n/d

3. Silva Pinto: n/d

4. Moraes Silva: n/d

Verbetes:

**XEXÉU** (H) s. m. Espécie de ave passeriforme de coloração negra e amarela, famosa por seu canto e pela capacidade de imitar sons de outras aves e mamíferos de seu ambiente. *“O xexéu de minha terra que me ensinou a cantar... Antes me tirasse o canto e me ensinasse a voar...” (AB, Reg2, freq.6)*

#### 5.4 Análise dos vocábulos-termos e discussão dos resultados

Inicialmente, convém reiterar que o Regionalismo, além de ser uma tendência literária que contribuiu decisivamente para a constituição da literatura brasileira, foi um projeto nacionalista de valorização e expressão dos aspectos físicos e humanos do Brasil, isto é, foi uma forma de identificar e apresentar o “tipicamente brasileiro”: seja por meio de um *Regionalismo Pitoresco*, de idealização do homem do sertão e da sua terra, seja através de um Regionalismo Crítico, de denúncia e crítica dos problemas regionais, ou mesmo por meio de um Super-regionalismo, da transcendência do regional para o universal. Fato é que o homem do sertão se tornou tema central do Regionalismo brasileiro como representação humana autêntica do Brasil da época.

Como vimos, a cada fase, a figura do sertanejo e de sua terra foi sendo moldada e representada de acordo com as aspirações da época e, assim, o Regionalismo foi adquirindo novas roupagens até se consolidar dentro da conjuntura literária. Então, não por acaso,

podemos ver várias nuances do sertanejo e de seu mundo nas descrições trazidas pelas obras literárias.

Prova disso, encontramos vários vocábulos-terms que evidenciam realidades dos sertões do Brasil (aspectos geográficos, históricos, políticos e sociais) e também aspectos linguísticos e culturais do sertanejo, permitindo-nos conhecer o homem do sertão, suas características físicas, comportamentais e, especialmente linguísticas. Como é o caso de **aboiar, aió, aguilhada, amoitado, aporrinhado, apustemado, arção, arreio, barbatão, barrigueira, bolandeira, burgariana, caçuá, cacunda, cangalha, capiau, catolé, fiofó, galista, garajau, garrota, jirau, jururu, macambira, maravalha, murzelo, paiol, panasco, patativa, pestead, pindaibeira, potrilho, quartau, rancho, rosilho, suaçuapara, tiririca, tronqueira, ururau e xexéu.**

Observamos que essas unidades lexicais, de modo geral, têm sememas específicos do universo sertanejo, apresentando traços semântico-conceptuais característicos de uma identidade cultural. Isto porque, ao apresentar um mundo ficcional exclusivo do universo sertanista, o escritor constrói seu próprio contexto referencial a partir de dados do mundo real e de uma “linguagem especial” que engendra significados próprios do universo de discurso a que pertencem e do grupo étnico que conceptualizou seus signos-símbolos.

Desse modo, essas unidades retratam: a) características humanas - **capiau, galista**; b) características comportamentais – **aboiar, amoitado, aporrinhado, apustemado, pestead, jururu**; c) características de animais – **barbatão, garrota, murzelo, quartau, potrilho, rosilho, suaçuapara**; d) elementos da natureza – **catolé, macambira, maravalhas, panasco, patativa, pindaibeira, tiririca, xexéu**; e) objetos – **aió, aguilhada, arção, arreio, barrigueira, bolandeira, caçuá, cangalha, burgariana, garajau, jirau, paiol, tronqueira**, f) partes do corpo – **cacunda, fiofó, tiririca**; g) tipos de habitação – **rancho**; h) mitos e lendas – **ururau**.

Elas têm, portanto, características específicas: “de um lado, são vocábulos metasemióticos, de outro, são quase-terms técnicos, pois pertencem a uma linguagem especial/especializada” (BARBOSA, 2005, p. 105). É nesse sentido que entendemos que as essas unidades “[...] têm sememas muito especializados, construídos com semas específicos do universo de discurso em causa, provenientes das narrativas, cristalizados, tornando-se verdadeiros símbolos dos temas envolvidos” (BARBOSA, 2007, p. 434).

Dito de outro modo, são palavras de caracterização temática que possuem especificidades semânticas advindas de uma visão humana particular de uma realidade

antropocultural: a realidade dos sertanejos das regiões interioranas do Brasil. Logo, é no contexto de seus discursos-ocorrência e no contexto sociocultural do grupo linguístico em questão, que entendemos o sentido dessas unidades, conforme nos explica Barbosa (2005, p. 105): “é preciso estar familiarizado com as histórias, conhecer o pensamento e o sistema de valores da cultura em questão, para poder compreendê-las bem”.

Nessas condições, examinando os itens selecionados para esta análise, verificamos que 3 (três) deles não são dicionarizados (**apustemado**, **burgariana** e **pindaibeira**) e 27 aparecem apenas no dicionário *Houaiss*, o que indica que tivemos que recorrer ao contexto linguístico de uso (observando várias linhas de concordância) para elaborar uma definição exata de cada vocábulo-termo (fato que possibilitou uma compreensão mais precisa das acepções). Na verdade, todas as definições foram elaboradas de acordo com o contexto linguístico dos discursos-ocorrência, mas, indubitavelmente, a apreensão do conceito dos itens dicionarizados resultou ser um elemento facilitador no processo.

Acerca dessas unidades lexicais não dicionarizadas, convém destacar, que o vocábulo-termo **apustemado**, em seu contexto linguístico, não deriva de “apostemar” - que está registrada no *Houaiss* como “4. tornar-se irado, zangado”. Em *Sargento Getúlio*, a palavra está engendrada no texto com o sentido de “desgraçado, sem caráter, imprestável”, como podemos ver nos seguintes excertos: “[...] porque não foi assim que Vencecavalo agarrou as balas com os dentes e cuspiu elas no chão e disse: com essas balas, **apustemado**, vosmecê me tirou uma lasca do dente queiro de cima do lado direito” / “Vosmecê sabe, esse **apustemado** é de Muribeca. Povo de Muribeca não presta!”/[...] fugir pra Paulo Afonso, pra Paulo Afonso, ia nos infernos, viu, cão da pustema **apustemado**, lhe faço uma desgraça, pirobo, senvergonho, pirobão, sacano, xibungo, bexiguento, chuparino do cão da gota do estupor balaio, mija na vareta, tem ginásio, tem ginásio! nunca vi ginásio fazer caráter”.

Por outro lado, **pindaibeira**, no contexto de *Chapadão do Bugre*, deriva de pindaíba – “arbusto (*Guatteria vilosissima*) nativo do Brasil, de ramos flexíveis, cuja casca fornece fio branco us. em cordoaria, folhas lanceoladas grandes e frutos comestíveis; araticum-peludo, embira-de-caçador, embira-pindaíba, embiratanha, pixiricum” (HOUAISS, 2009). Para exemplificar, citamos: “Nem que tivesse de rodear toda a **pindaibeira**, seguir apalpando e farejando palmo a palmo”/“Fazendo picada com as mãos por entre o ramalhal e cipoama da **pindaibeira**, José de Arimateia entrou pelo capão até a água da nascente”/ “Seu Persilva apeou o burro e amarrou-o num ramo da **pindaibeira**”.

Já o vocábulo-termo **burgariana** não tem relação com nenhuma outra palavra, o que indica que tivemos que recuperar o seu sentido somente no contexto de *Seara Vermelha*. Os fragmentos: “Seu corpinho raquítico treme sob o vestido de **burgariana**”/ “Ficara no armazém, estranha mercadoria entre a carne-sêca, a chita, a **burgariana**, a cachaça e o feijão”/ “Na caatinga não faz frio, se fizesse os sertanejos teriam todos morrido porque vestem farrapos de roupas, calças de mescla azul, camisa de **burgariana**”, revelam o significado desta palavra e a sua representatividade no universo sertanejo: trata-se de uma espécie de tecido fino e de pouco valor usado pelos sertanejos nas regiões áridas e quentes do sertão.

Convém, ainda, destacar, que os vocábulos-terminos **aboiar** e **paiol** estão registrados nos dicionários de Bluteau, Silva Pinto e Moraes Silva, mas com acepções diferentes do contexto de uso dos discursos-ocorrência. A palavra **aboiar** aparece em Silva Pinto e Moraes Silva com o sentido de boiar, colocar boias; já **paiol** (dicionarizada em B/P/S) aparece nos três dicionários como “lugar do navio onde se guarda a pólvora”. Isso revela que tais vocábulos-terminos foram engendrados nos discursos-ocorrências das obras de forma a apresentarem conceitos específicos em referência ao universo sertanejo.

Sobre o vocábulo-termo **tiririca**, observamos que, das duas acepções encontradas nos textos (1. Espécie de erva daninha; 2. Órgão genital), apenas a primeira aparece no dicionário *Houaiss*, o que evidencia a sua particularidade semântica no universo do discurso em questão. Por outro lado, acerca do vocábulo-termo **rancho**, verificamos que, das três acepções encontradas nos textos (1. Casa pobre da roça; 2. Construção rústica e provisória usada como abrigo ou descanso de trabalhadores; 3. Bando de pessoas reunidas para um fim qualquer), todas aparecem no dicionário *Houaiss* e somente a segunda aparece em Silva Pinto e Moraes Silva. No dicionário de Bluteau, apesar de a palavra aparecer no resultado de busca, não conseguimos encontrá-la no texto.

Os demais vocábulos-terminos estão dicionarizados, ora em todos, ora em um ou outro, como: **agulhada** (H/B/P/S); **arção** (H/B/P/S); **arreio** (H/B/P/S); **cangalha** (H/B/P/S); **murzelo** (H/B/P/S); **rosilho** (H/B/P/S); **bolandeira** (H/P/S); **maravalhas** (B/P/S); **panasco** (H/P/S). Em síntese, temos: 36 itens dicionarizados no *Houaiss*, 7 no Bluteau, 10 no Silva Pinto e também no Moraes Silva; e 3 itens não dicionarizados.

Quanto à frequência, verificamos que 8 vocábulos-terminos têm frequência 3 (a menor frequência computada), 30 possuem frequências que variam entre 4 a 90 e apenas 2 possuem frequência acima de 90 (**arreio**- freq. 94 e **rancho** –freq.166). Vale aqui reiterar que a frequência não está relacionada apenas a questões estilísticas, situacionais ou circunstanciais,

mas também a questões de significância semântica no universo de discurso em causa. Assim, percebemos que tanto **arreio** como **rancho**, além de serem unidades de caracterização temática, possuem uma alta significância e representatividade semântica no universo sertanejo, por isso são tão recorrentes.

Em relação aos vocábulos-terms computados em cada fase, constatamos que 15 aparecem na primeira, 22 na segunda, 23 na terceira, e apenas 5 aparecem nas três fases. Como o 2º *subcorpus* (que corresponde a segunda fase) é composto por 10 textos e os outros dois *subcorpora* estão compostos por 5 textos cada um, é natural que o número de vocábulos-terms presentes no 2º *subcorpus* seja maior. No entanto, se compararmos a quantidade de itens presentes na segunda fase com o número de itens que aparecem na terceira, percebemos que existe uma diferença mínima, o que revela uma alta variação lexical do 3º *subcorpus*.

Em suma, todas essas unidades lexicais adquirem estatuto de vocábulo-termo no discurso literário regionalista sertanista pelo fato de: a) fazerem parte de um sistema conceptual estruturado dentro de uma temática específica: temática sertanista; b) atuarem na composição de um mundo ficcional semioticamente construído pela força modelizante da linguagem literária; c) possuírem intertextualidade e interdiscursividade intra e interuniverso de discurso com discursos etnoliterários; d) atualizarem um sistema de valores, um sistema de crenças, enfim, a axiologia de um grupo específico: o sertanejo dos sertões brasileiros; e) designarem conceitos formados com semas do universo de discurso em que são usadas; f) referirem-se às particularidades de um mundo ficcional quanto à sua função simbólica; g) atuarem no plano da memória, de maneira que é possível resgatar nas narrativas elementos da cultura e da linguagem do sertanejo ainda presentes em nossa sociedade.

Por isso, é imprescindível interpretar, com toda a clareza, de um lado, os traços semântico-conceptuais das unidades lexicais que integram um discurso etnoliterário e, de outro, as relações de significação que estabelecem no interior de um texto, sempre lembrando que cada vocábulo-termo constitui um registro linguístico, resultado das produções culturais humanas, que atrelados a uma identidade cultural corroboram para a construção de sentidos de um universo literário particular.

## CONSIDERAÇÕES FINAIS

O diálogo promovido nesta pesquisa entre Linguística e Literatura pela via etnoterminológica não somente demonstrou que é possível encontrar elementos de natureza especializada em um texto literário, como também evidenciou usos etnoterminológicos no discurso literário regionalista sertanista, corroborando para o reconhecimento dessas unidades como vocábulos-terminos.

Enquanto integrantes de uma linguagem literária, especialmente organizada, que modeliza a concepção de um mundo semioticamente construído, vimos que os vocábulos-terminos apresentam características lexicais e semântico-conceptuais específicas do universo de discurso a que pertencem e são reveladoras de uma identidade cultural. De um modo geral, essas unidades retratam realidades dos sertões do Brasil, como: o espaço geográfico (**gerais, veredas, estirão, capão**); elementos da fauna (**barbatão, suaçuapara, murzelo**); elementos da flora (**catolé, macambira, panasco, pindaibeira**); ocupações (**boiadeiro, tropeiro, cangaceiro**); objetos (**coité, aió, arreio**); crenças e costumes (**ururau, caipora, berzabum**) e, também, o tipo humano sertanejo, suas características físicas (**aleijão, apessoado**), psicológicas (**aluado, danado**), comportamentais (**amoitado, amuado**), linguísticas (**arre, abobra, aniceto**) e socioculturais (**caburé, capiau, bugre, jeca, curiboca**).

Em uma apreciação mais ampla, vimos que os escritores regionalistas, na tentativa de resgatar o que era tipicamente brasileiro, buscaram na literatura representar a nossa identidade nacional: o espaço físico e geográfico do Brasil da época e o tipo humano brasileiro, com suas características físicas, comportamentais, culturais e, especialmente, linguísticas. Então, não é por acaso que o sertanejo e sua terra se tornaram tema central das obras regionalistas e, também, não é por acaso que encontramos unidades lexicais portadoras de especificidades semânticas advindas de uma conceptualização étnica do sertanejo.

São vocábulos que se fazem termos à medida que designam conceitos específicos referentes ao universo sócio-linguístico-cultural do sertanejo, que erigem da própria conceptualização linguística desse grupo concebido em seu meio. E é nessa particularidade semântica, nesse limite, que a fronteira existente entre o vocábulo e o termo se torna densa, perceptível.

O **sertanejo, capiau, catrumano, tabaréu, geralista**, que mora em **ranchos, taperas ou choças**, situadas no alto das **chapadas**, nas **veredas** ou **brejos**, não é um tipo humano meramente ficcional, um ser fantástico criado pela imaginação humana, mas um ser real que

vive, sente e percebe o mundo físico e humano do sertão. É o **vaqueiro** que **aboia** a **boiada**, o **tropeiro** ou **arrieiro** que conduz a **tropa**, o **comboieiro** que guia o **comboio** de **bestas**, o **campeiro** que trabalha no campo, o **cangaceiro**, o **jagunço** e o **capanga** que atuam como guarda-costas, o **esteireiro** que fabrica **esteiras**, o **brejeiro** que trabalha nos engenhos, etc.

Em um cenário figurativo, vislumbramos: o **sertanejo boiadeiro**, em sua lida diária, ora **acabrunhado**, ora **alvorçado**, ora **arreliado**, ora **amuado**, levantando-se bem cedo do seu **jirau** de palha. Ele veste a camisa de **burgariana** e a calça de **chita**, calça suas **botinas** ou **alpercatas**, coloca em sua **algibeira** o **pito** e a **binga**, pega a **alforje** e a **bruaca**. Em direção ao **bamburral** vai atrás do **cardão carujo**. No **piquete**, coloca sobre ele o **arreio**, aperta a **barrigueira** e com a **agulhada** na mão vai **troteando** em busca da **boiada**. No **capoão**, em meio as **maravalhas**, **macambiras**, **pindaibeiras**, **tiriricas** e **jaraguás**, o **camumbembe** **aboia** o **gado** com seu canto rítmico e sonoro, **picando** com a **espora** o **quartau** e **ferroando** o **espinhaço** da **vacada**. E assim, com a ajuda de seu **cusco**, toca a **gadaria** rumo ao **curral**. De um lado **aparta** os **garrotes**, de outro a **vaca leiteira**, e ainda, os **bois pesteados** para curar a **bicheira**. **Apeando-se** do **baio**, segue a **labuta**, tira leite, cura o **marruco**, marca o **bezerro** e vacina a **camurça**. Esta é a lida do **boiadeiro**, **caboclo** forte e **chibante** que luta diariamente por sua sobrevivência nesse **sertão** inconstante.

O aspecto referencial desses vocábulos-terms demonstra que eles representam particularidades culturais em mundos ficcionais construídos a partir de dados do mundo real que não só fazem parte da vida do sertanejo das regiões interioranas do Brasil, como da nossa vida, da nossa cultura e memória enquanto brasileiros. Assim, devido ao seu valor linguístico, artístico e cultural e no afã de conhecê-las melhor e de torná-las conhecidas no contexto do discurso literário regionalista é que propomos um Vocabulário dessas unidades.

Enquanto contribuição, acreditamos que esta proposta sirva como respaldo a novas propostas etnoterminográficas, abrindo caminho para trabalhos e pesquisas futuras relacionadas ao tema e, ainda, colabore para o reconhecimento de outros fenômenos linguísticos relacionados ao vocábulo-termo. Por isso, esperamos que ela possa ser replicada e que, a partir dela, surjam novas pesquisas que levem em conta outros textos literários e outros universos de discurso.

Acreditamos ter cumprido com os objetivos propostos neste estudo e, também, confirmado a nossa hipótese inicial sobre a possibilidade de encontrarmos elementos de natureza especializada (vocábulos-terms) em obras da Literatura Regionalista sertanista. Como pode ser verificado em nossa consolidação teórico-metodológica, os resultados desta



investigação foram obtidos por meio de observações empíricas de usos das unidades lexicais em contextos presentes no *corpus* de estudo (abordagem direcionada pelo *corpus*), ou seja, foi a partir dos dados que propomos as articulações teóricas. Logo, não foram os dados que se adequaram às teorias, foram as teorias que se adequaram aos dados.

Desse modo, além de confirmar a nossa hipótese sobre o construto lexical do *corpus*, conseguimos cumprir com o objetivo de propor um modelo de Vocabulário para o acesso *online*. Assim, referendamos o posicionamento de Barbosa (2009a, p. 1) de que “as unidades lexicais do universo de discurso etnoliterário têm um estatuto próprio e exclusivo: possuem simultaneamente duas funções, vocábulo e termo”, e que “é preciso estar familiarizado com as histórias, conhecer o pensamento e o sistema de valores da cultura em questão, para poder compreendê-las bem” (BARBOSA, 2006, p 50).

Entendemos que as unidades lexicais presentes nos textos literários regionalistas sertanistas são dotadas de uma configuração especial, híbrida e fluída, tanto pelo fato de serem multifuncionais, vocábulos-termos, quanto pela sua heterogeneidade semântica própria do universo sertanejo. Portanto, um Vocabulário dessas unidades é pertinente, não apenas por ser um produto inédito, mas por auxiliar no entendimento de obras clássicas da nossa Literatura que fazem parte do nosso universo histórico, político e sociocultural.

Contudo, não podemos deixar de reconhecer as limitações de nossa investigação. Uma delas refere-se ao público-alvo. Como já foi mencionado, dado o escopo da pesquisa e por se tratar de uma proposta, não consultamos o público-alvo a respeito do modelo sugerido ou do tipo de definição que ele julgaria ideal para esse repertório. Assim, julgamos com base em Barbosa (2001) e Barros (2004), que a definição por compreensão seria a mais adequada ao nosso tipo de repertório.

Por outro lado, devido à pouca notabilidade de alguns escritores, não levamos em conta outras obras regionalistas de temática sertanista e, além disso, também não selecionamos outras obras dos escritores escolhidos, o que pressupõe possíveis desdobramentos de nosso trabalho. Pesquisas que utilizem como *corpus* outras obras literárias regionalistas e que levem em conta outros usos etnoterminológicos nessas manifestações são importantes para o entendimento dos vocábulos-termos, haja vista a necessidade de ampliação de estudos na área.

Como não analisamos ou descrevemos o processo de conceptualização dos vocábulos-termos, entendemos que há mais uma possibilidade de abordagem pela via etnoterminológica, o que revela que não esgotamos as possibilidades de investigação dentro do escopo da

pesquisa. Observar, por exemplo, o processo de construção do conceito - o *modus operandi* conceptual dessas unidades lexicais portadoras de especificidades semânticas ou o processo de semiotização (conversão do conceito em grandeza-signo) são, ainda, outros pontos interessantes de investigação. Por isso, acreditamos no desdobramento desta pesquisa, a fim de que novos estudos possam ser realizados.

Apesar de reconhecer as limitações de nossa investigação, cremos ter contribuído para um melhor entendimento e reconhecimento dos vocábulos-termos no universo de discurso etnoliterário e, sobretudo, acreditamos ter contribuído para o enriquecimento e expansão dos estudos na área, colaborando para o surgimento de novas possibilidades de investigação no âmbito dos estudos etnoterminológicos.

## REFERÊNCIAS

- ALMEIDA, G. M. de B. A teoria comunicativa da terminologia e a sua prática. *Alfa*, São Paulo, v. 50, n. 2, p. 85-101, 2006.
- ALMEIDA, J. M. G. de. *A tradição regionalista no romance brasileiro: 1857-1945*. 2. ed. Rio de Janeiro: Achiamé, 1981.
- ANDRADE, M. M. Lexicologia, terminologia: definições, finalidades, conceitos operacionais. In: OLIVEIRA, A. M. P. P. de; ISQUERDO, A. N. (Orgs.) *As ciências do léxico: lexicologia, lexicografia, terminologia*. Campo Grande: Ed. da UFMS, 1998. p. 89-98.
- ARAGÃO, M. do. S. S. de. A Socioterminologia e Etnoterminologia das plantas medicinais do Nordeste. *Acta Semiótica et Lingvistica*, UFPB/ SBPL: v.15, nº1, ano 34, p. 34 – 49, 2010. Disponível em: <<http://periodicos.ufpb.br/index.php/actas/article/viewFile/14645/8296>>. Acesso em: 8 fev. 2018.
- ARAÚJO, G. P. de. *O conhecimento etnobotânico dos Kalunga: uma relação entre língua e meio ambiente*. 218f. Tese (Doutorado em Linguística)- Departamento de Linguística, Português e Línguas clássicas, Universidade de Brasília, Brasília, 2014.
- ARINOS, Afonso. Pelo Sertão. In: ARINOS, Afonso. *Obras Completas*. Rio de Janeiro: 1968.
- ATHAYDE, T. de. Parte introdutória. In: ALMEIDA, José Américo de. *A bagaceira*. 37. ed. Rio de Janeiro: José Olympio, 2004.
- BARBOSA, M. A. Lexicologia, lexicografia, terminologia, terminografia: objeto, métodos, campos de atuação e de cooperação. *Anais do XXXIX Seminário do GEL*. Franca, Unifran, 1991, p. 182-189.
- BARBOSA, M. A. Dicionário de língua, vocabulários técnico-científicos, glossários: estatuto semântico-sintático das unidades-padrão. *Anais do XLI Seminário do GEL*. Ribeirão Preto, 1994, p. 289-294.
- BARBOSA, M. A. Dicionário, vocabulário, glossário: concepções. In: ALVES, I. M. (Org.). *A constituição da normalização terminológica no Brasil*. São Paulo: FFLCH/CITRAT, 2001, p. 23-45. Disponível em: <<http://citrat.fflch.usp.br/sites/citrat.fflch.usp.br/files/u10/Cad.%20Terminologia%201.pdf>>. Acesso em: 22 mar. 2018.
- BARBOSA, M. A. Estrutura e formação do conceito nas línguas especializadas: tratamento terminológico e lexicográfico. *Rev. Brasileira de Linguística Aplicada*, Belo Horizonte, v. 4, n. 1, p. 55-86, 2004. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1984-63982004000100006](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1984-63982004000100006)>. Acesso em: 12 mar. 2018. <https://doi.org/10.1590/S1984-63982004000100006>

BARBOSA, M. A. Terminologia e Lexicologia: plurissignificação e tratamento transdisciplinar das unidades lexicais nos discursos etno-literários. *Revista de Letras*, Ceará, v. 1/2, n. 27, p. 103-107, 2005.

BARBOSA, M. A. Para uma Etno-Terminologia: recortes epistemológicos. *Cienc. Cult.*, São Paulo, v. 58, n. 2, jun. 2006a. Disponível em: <[http://cienciaecultura.bvs.br/scielo.php?script=sci\\_arttext&pid=S000967252006000200018&lng=en&nrm=iso](http://cienciaecultura.bvs.br/scielo.php?script=sci_arttext&pid=S000967252006000200018&lng=en&nrm=iso)>. Acesso em: 03 fev. 2018.

BARBOSA, M. A. Terminologia aplicada: teorias, práticas e desenvolvimento técnico-científico. *Anais da 58ª Reunião Anual da SBPC* - Florianópolis, SC - Julho/2006b. Disponível em: <[http://www.sbpnet.org.br/livro/58ra/atividades/TEXTOS/texto\\_263.html](http://www.sbpnet.org.br/livro/58ra/atividades/TEXTOS/texto_263.html)>. Acesso em: 5 fev. 2018.

BARBOSA, M. A. Etno-terminologia e Terminologia Aplicada: objeto de estudo, campo de atuação. In: ISQUERDO, A. N.; ALVES, I. M. (Orgs.). *As ciências do léxico: lexicologia, lexicografia, terminologia*. Campo Grande: UFMS; São Paulo: Humanitas, 2007. p. 433-445.

BARBOSA, M. A. Cultura popular amazônica em Etno-Terminologia. *Anais da 61ª Reunião Anual da SBPC*- Manaus, AM. Julho/2009a.

BARBOSA, M. A. *Terminologia Aplicada: percursos interdisciplinares*. Polifonia, Cuiabá, n.17, p. 29-44, 2009b.

BARROS, L. A. *Curso básico de Terminologia*. São Paulo: Editora da Universidade de São Paulo, 2004.

BECHARA, Evanildo. Bernardo Élis: apresentação. In: ÉLIS, Bernardo. *Seleção* (organização de Gilberto Mendonça Teles) 3.ed. Rio de Janeiro: J. Olympio, 1991.

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

BERBER SARDINHA, T. *Metáfora*. São Paulo: Parábola Editorial, 2007.

BERBER SARDINHA, T. *Pesquisa em Linguística de Corpus com WordSmith Tools*. Campinas: Mercado de Letras, 2009.

BERBER SARDINHA, T. MCI, um identificador de candidatos a metáfora em corpora. In: SHEPHERD, T. M. G.; BERBER SARDINHA, T.; PINTO, M. V. *Caminhos da linguística de corpus*. Campinas: Mercado de Letras, 2012. p. 87-105.

BERBER SARDINHA, T.; ALMEIDA, G. M. de B. A Linguística de *Corpus* no Brasil. In: TAGNIN, S. E. O; VALE, O. A. *Avanços da Linguística de Corpus no Brasil*. São Paulo: Humanitas, 2008. p. 17-40.

BEVILACQUA, C. R. Por que e para que a linguística de corpus na terminologia. In: TAGNIN, S.; BEVILACQUA, C. (Orgs.). *Corpora na Terminologia*. São Paulo: Hub Editorial, 2013. p. 11-27.

BIBER, D.; CONRAD, S.; REPPEN, R. *Corpus linguistics: investigating language structure and use*. Cambridge: Cambridge University Press, 1998. <https://doi.org/10.1017/CBO9780511804489>

BIDERMAN, M. T. C. *Teoria Linguística* 2. ed. São Paulo: Martins Fontes, 2001a.

BIDERMAN, M. T. C. As Ciências do Léxico. In: OLIVEIRA, A. M. P. P.; ISQUERDO, A. N. *As Ciências do Léxico: Lexicologia. Lexicografia. Terminologia*. Campo Grande/MS: Ed. UFMS, 2001b. p. 13-22.

BILAC, O. Resposta do Sr. Olavo Bilac. In: ACADEMIA BRASILEIRA DE LETRAS. *Discursos Acadêmicos (1897-1906)*. Rio de Janeiro: Civilização Brasileira, 1934. 1 v.

BLUTEAU, R. *Vocabulario Portuguez & Latino*. Coimbra: Collegio das Artes da Companhia de Jesus, 1712-1728. Disponível em: <<http://dicionarios.bbm.usp.br/en/dicionario/>>. Acesso em: 2 fev. 2017

BOWKER, L.; PEARSON, J. *Working with specialized language: a practical guide to using corpora*. London/New York: Routledge, 2002. <https://doi.org/10.4324/9780203469255>

CABRÉ, M. T. *Terminology: theory, methods, and applications*. Philadelphia/PA: John Benjamins, 1999a. <https://doi.org/10.1075/tlrp.1>

CABRÉ, M. T. Hacia una teoría comunicativa de la terminología: aspectos metodológicos. In: CABRÉ, M. T. *La Terminología: representación y comunicación*. Barcelona: IULA, 1999b, p.129-150.

CANDIDO, A. Literatura e subdesenvolvimento. In: CANDIDO, A. *A educação pela noite & outros ensaios*. São Paulo: Ática, 1989, p. 140-162. Disponível em: <<https://joaocamillopenna.files.wordpress.com/2014/03/antonio-candido-a-educacao-pela-noite.pdf>>. Acesso em: 2 ago. 2018.

CANDIDO, A. *Formação da literatura brasileira: momentos decisivos*. 9. ed. Belo Horizonte: Itatiaia Ltda, 2000. 2 v.

CANDIDO, A. *Textos de Intervenção*. Org. Vinicius Dantas. São Paulo: Duas Cidades/Ed. 34, 2002.

CANDIDO, A. A literatura e a formação do homem. In: CANDIDO, A. *Textos de intervenção*. São Paulo: Duas Cidades/34, 2002a. p. 77-92. Disponível em: <<https://books.google.com.br>>. Acesso em: 2 ago. 2018.

CANDIDO, A. O homem dos avessos. In: CANDIDO, A. *Tese e Antítese: ensaios*. 4. ed. São Paulo: T. A. Queiroz, 2002b. p.121-139. Disponível em: <<https://filosoficabiblioteca.files.wordpress.com/2017/10/antonio-candido-o-homem-dos-avessos-estudo-de-grande-sertc3a3o-veredas-de-guimarc3a3es-rosa-in-tese-e-antc3adtese.pdf>>. Acesso em: 2 ago 2018.

CANEVER, F. Efeitos de frequência no uso do infinitivo flexionado em português brasileiro. In: SHEPHERD, T. M. G.; BERBER SARDINHA, T.; PINTO, M. V. (Orgs.). *Caminhos da Linguística de Corpus*. Campinas, SP: Mercado das Letras, 2012, p. 405-426.

CARDOSO, A. L. M. *A transcodificação de textos científicos em textos etnoliterários, o cordel: o desenvolvimento da cognição com reflexão crítica*. 175f. Tese (Doutorado em Semiótica e Linguística Geral) - Departamento de Linguística, Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2011.

CARDOSO, S. A. F. *Termosteo: a elaboração de vocabulários monolíngues de Termos da teologia em um estudo conduzido por corpus*. 2017.315f. Tese (Doutorado em Estudos Linguísticos) - Programa de Pós-graduação em Estudos Linguísticos, Universidade Federal de Uberlândia, Uberlândia, 2017.

CARNEIRO, R. M. O. *Discurso literário de fantasia infantojuvenil: proposta de descrição terminológica direcionada por corpus*. 2016. 281f. Dissertação (Mestrado em Estudos Linguísticos)- Programa de Pós-Graduação em Estudos Linguísticos, Universidade Federal de Uberlândia, Uberlândia, 2016.

CARVALHO, F. M. de. *O dicionário do folclore brasileiro: um estudo de caso da etnoterminologia e tradução etnográfica*. 252f. Dissertação (Mestrado em Estudos da Tradução)- Departamento de Línguas estrangeiras e tradução, Programa de Pós-graduação em Estudos da Tradução, Universidade de Brasília (UNB), Brasília, 2013.

CLAS, A. A pesquisa terminológica e a formulação de parâmetros em função das necessidades dos usuários. In: ISQUERDO, A. N.; KRIEGER, M. G. (Org.). *As ciências do léxico: lexicologia, lexicografia, terminologia*. v. II. Campo Grande: UFMS, 2004. p. 223-238.

COSTA, N. M. P. *Estudo etnoterminológico preliminar do sistema de cura e cuidados do povo Mundurukú (Tupi)*. 150f. Dissertação (Mestrado em Linguística) – Departamento de Linguística, Português e Línguas clássicas, Universidade de Brasília, Brasília, 2013.

COSTA, N. M. P.; GOMES, D. M. A Etnoterminologia da língua Mundurukú- Tupi e as contribuições da Ecolinguística. *Cadernos de Linguagem e Sociedade*, Brasília: v. 14, nº1, p. 252 - 274, 2013. Disponível em: <<http://periodicos.unb.br/index.php/les/article/view/9176/6890>>. Acesso em: 8 fev. 2018. <https://doi.org/10.26512/les.v14i1.22249>

COUTINHO, A. *A literatura no Brasil*. 6.ed. São Paulo: Global, 2001.

DIKI-KIDIRI, M. Un enfoque cultural de la terminologia. *Debate Terminológico*, n. 5, 2009.

DUTRA, D. P.; MELLO, H. (Org.). *Anais do X Encontro de Linguística de Corpus: aspectos metodológicos dos estudos de corpora*. Belo Horizonte: Faculdade de Letras da UFMG, 2012.

FERREGUETTI, K; PAGANO, A; FIGUEREDO, G. Significados existenciais no português brasileiro: um estudo contrastivo em textos traduzidos e não traduzidos. In: DUTRA, D. P.; MELLO, H. (Org.). *Anais do X Encontro de Linguística de Corpus: aspectos metodológicos dos estudos de corpora*. Belo Horizonte: Faculdade de Letras da UFMG, 2012, p. 280-293.

FINATTO, M. J. B. Estudos sobre linguagens e textos científicos e técnicos: o que é uma terminologia textual? In: BATTISTI, E. ; COLLISCHONN, G. (Org.). *Língua e Linguagem: perspectivas de investigação*. 1ed. Pelotas - RS: EDUCAT, 2011. p. 153-172.

FRANKENBERG-GARCIA, A. Compilação e uso de corpora paralelos. In: TAGNIN, S. E. O.; VALE, O. A. (Orgs.). *Avanços da Linguística de Corpus no Brasil*. São Paulo: Humanitas, 2008. p. 117-136.

FROMM, G. *Proposta para um modelo de glossário de informática para tradutores*. 2002. 82f. Dissertação (Mestrado em Semiótica e Linguística Geral) – Departamento de Linguística, Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2002.

FROMM, G. *VoTec: a construção de vocabulários eletrônicos para aprendizes de tradução*. 2007. 215f. Tese (Doutorado em Estudos Linguísticos e Literários em Inglês) - Departamento de Linguística, Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2007.

GALVÃO, W. N. *Saco de gatos: ensaios críticos*. 2. ed. São Paulo: Duas Cidades, 1976.

GAUDIN, F. Socioterminologia: um itinerário bem-sucedido. In: ISQUERDO, A. N.; DAL CORNO, G. O. M. (Org.). *As ciências do léxico: lexicologia, lexicografia, terminologia*. Campo Grande: UFMS, 2014. p. 293-309.

GONÇALVES, L. B. Linguística de *corpus* e análise literária: o que revelam as palavras-chave. In: TAGNIN, S. E. O.; VALE, O. A. (Orgs.). *Avanços da Linguística de Corpus no Brasil*. São Paulo: Humanitas, 2008, p. 387-405.

HALLIDAY, M. A. K. Language as system and language as instance: the corpus as a theoretical construct. In: SVARTVIK, J. (Org.). *Directions in corpus linguistics: proceedings of Nobel Symposium 82*. Berlin; Nova York: Mouton de Gruyter, 1992. p. 61-76.

HOUAISS, A. *Dicionário eletrônico Houaiss da língua portuguesa*. Versão 3.0, 2009.

ISO – INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. *Norma 1087-1: Terminology work – Vocabulary*. Genebra, 2000.

ISO – INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. *Norma 704: Terminology – Vocabulary*. Genebra, 2009.

KRIEGER, M. da G. Terminologia: uma entrevista com Maria da Graça Krieger. *ReVEL*, v. 9, n. 17, p. 443-452, 2011.

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

LATORRE, V. R. D. As denominações no léxico de Grande Sertão: Veredas sob uma perspectiva Etnoterminológica. *Acta Semiótica et Lingvistica*, UFPB/ SBPL: v.15, nº1, ano

34, p. 82 – 92, 2010. Disponível em: <<https://www.usp.br/bibliografia/obra.php?cod=23703&s=grosa>>. Acesso em: 8 fev. 2018.

LATORRE, V. R. D. *Uma abordagem etnoterminológica de Grande Sertão: Veredas*. 2011. 156f. Dissertação (Mestrado em Semiótica e Linguística Geral). Departamento de Linguística, Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2011.

LATORRE, V. R. D. A dialética entre os extremos: da Terminologia à Etnoterminologia. *Caderno Seminal Digital*, Rio de Janeiro: v. 19, nº 19, p. 70 – 94, jan/jun. 2013. Disponível em: <[http://www.usp.br/gmhp/publ/CadSeminal\\_N%2019.pdf](http://www.usp.br/gmhp/publ/CadSeminal_N%2019.pdf)>. Acesso em: 4 fev. 2018. <https://doi.org/10.12957/cadsem.2013.12062>

LATORRE, V. R. D. A Etnoterminologia no âmbito dos estudos da tradução. *Acta Semiótica et Lingvistica*, UFPB/ SBPL: v.21, nº1, p. 86 – 95, 2016. Disponível em: <<http://www.usp.br/bibliografia/obra.php?cod=23694&s=grosa>>. Acesso em 8 fev. 2018.

LEECH, G. Corpora and theories of linguistics performance. In: SVARTVIK, Jan (Ed.). *Directions in corpus linguistics: proceedings of Nobel Symposium 82*. Berlin: Mouton de Gruyter, 1992. p.105-122.

LEITÃO JÚNIOR, A. M. As imagens do sertão na Literatura Nacional: o projeto da modernização na formação territorial brasileira a partir dos Romances Regionalistas da Geração de 1930. *Terra Brasilis- Revista da Rede Brasileira de História da Geografia e Geografia Histórica* São Paulo: v.1, p. 1-20, nov. 2012. Disponível em: <<http://terrabilis.revues.org/468>>. Acesso em: 15 jul. 2018. <https://doi.org/10.4000/terrabilis.468>

LYONS, J. Linguagem e Cultura, In: LYONS, J. *Lingua(gem) e Linguística: uma introdução*. Tradução de Marilda Winkler Averbug e Clarisse Sieckenius de Souza. Rio de Janeiro: LTC, [1987] 2009, p. 223-244.

NOMURA, M. Conceitos linguísticos de linguagem literária. *Alfa*, São Paulo, v. 40, n. 1, p. 189-204, 1996.

NOMURA, M. *Linguagem funcional e literatura: presença do cotidiano no texto literário*. São Paulo: Annablume, 1993.

NOVODVORSKI, A; FINATTO, M. J. B. Linguística de Corpus no Brasil: uma aventura mais do que adequada. *Letras & Letras*, Uberlândia v. 30, n. 2, p. 7-16, 2014. Disponível em: <<http://www.seer.ufu.br/index.php/letraseletras/article/view/28516>>. Acesso em: 8 maio 2018. <https://doi.org/10.14393/LL60-v30n2a2014-1>

ORENHA-OTTAIANO, A. Algumas contribuições advindas da compilação de *corpora* especializados via Web e WebBootCat para a tradução, terminologia e fraseologia. In: TAGNIN, S. E. O.; VALE, O. A. (Orgs.). *Avanços da Linguística de Corpus no Brasil*. São Paulo: Humanitas, 2008, p. 137-165.



PAIS, C. T.; BARBOSA, M. A. Da análise de aspectos semânticos e lexicais dos discursos etno-literários: a proposição de uma etno-terminologia. *Matraga*, Rio de Janeiro, v.11, n. 16, p. 79-100, jan./dez. 2004.

PARODI, G. *Lingüística de Corpus: de la teoría a la empiria*. Madrid: Iberoamericana/Vervuert, 2010. <https://doi.org/10.31819/9783865278715>

PIMENTA, A. P. C. *A Representação do léxico rural em Ermos e Gerais de Bernardo Élis*. 2013. 206f. Dissertação (Mestrado em Estudos da Linguagem) – Programa de Pós-Graduação em Estudos da Linguagem, Universidade Federal de Goiás, Catalão, 2013.

PINTO, L. M. S. *Diccionario da Lingua Brasileira*. Ouro Preto: Typographia de Siva, 1832. Disponível em: <<http://dicionarios.bbm.usp.br/en/dicionario/>>. Acesso em: 2 fev. 2017

PIRES, G. da S.; SALOMÃO, M. M. M. Contribuições metodológicas para o desenvolvimento da plataforma FrameNet Brasil: a descrição de algumas unidades lexicais dos frames Fechamento e Movimento \_corporal. In: DUTRA, D. P.; MELLO, H. (Org.). *Anais do X Encontro de Linguística de Corpus: aspectos metodológicos dos estudos de corpora*. Belo Horizonte: Faculdade de Letras da UFMG, 2012, p. 172-195.

POTTIER, B. *Théorie et analyse en linguistique*. 2 éd. Paris: Hachette, 1991.

PROENÇA, M. C. Introdução. In: ALMEIDA, José Américo de. *A bagaceira*. 37. ed. Rio de Janeiro: José Olympio, 2004.

RASO, T; MELLO, H. (Orgs). *C- ORAL-BRASIL I: Corpus* de referência do português brasileiro falado informal. Belo Horizonte: Editora UFMG, 2012.

SAGER, J. C. *A Practical course in Terminology Processing*. Amsterdam; Philadelphia: J. Benjamins Publishing Company, 1990. <https://doi.org/10.1075/z.44>

SCOTT, M. *WordSmith Tools*. Version 6. Liverpool: Lexical Analysis Software, 2012.

SHEPHERD, T. M. G.; BERBER SARDINHA, T.; PINTO, M. V. (Orgs.). *Caminhos da Linguística de Corpus*. Campinas, SP: Mercado das Letras, 2012.

SILVA, A, M. *Diccionario da Lingua Portuguesa*. Lisboa: Officina de Simão Thaddeo Ferreira, 1789. Disponível em: <<http://dicionarios.bbm.usp.br/en/dicionario/>>. Acesso em: 2 fev. 2017.

SILVA, B. de C. T. *O espaço e o imaginário popular nos contos de Afonso Arinos*. 2008. 103f. Dissertação (Mestrado em Teoria Literária) – Programa de Pós-graduação em Letras, Universidade Federal de Uberlândia, Uberlândia- MG, 2008.

SILVA, M. M. A. da; CHAGURI, J. de P. A etnoterminologia da culinária baiana na obra Dona Flor e seus dois maridos: análise dos aspectos do discurso etnoliterário na versão para o inglês. *Dialogia*, São Paulo, v. 9, n. 1, p. 37-46, 2010. <https://doi.org/10.5585/dialogia.v9i1.1649>

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

STEGER, H. O que é linguagem literária? *Fragmentos*, Florianópolis, n. 3, jan./dez, p. 101-140, 1987.

SVARTVIK, J. Corpus linguistics comes of age. In: SVARTVIK, J. (Org.). *Directions in corpus linguistics: proceedings of Nobel Symposium 82*. Berlin; Nova York: Mouton de Gruyter, 1992. p. 7-13. <https://doi.org/10.1515/9783110867275>

TAGNIN, S. E. O. Linguística de corpus e fraseologia: uma feita para a outra. In: ALVAREZ, M. L. O.; UNTERNBAUMEN, E. H. (Orgs.) *Uma (re)visão da teoria e da pesquisa fraseológicas*. Campinas: Pontes Editores, 2011. p. 277-302

TAGNIN, S. E. O. Glossário de Linguística de *Corpus*. In: VIANA, V.; TAGNIN, S. E. O. (Orgs.). *Corpora no ensino de línguas estrangeiras*. São Paulo: Hub, 2011a. p. 357-361.

TAGNIN, S. E. O.; VALE, O. A. *Avanços da Linguística de Corpus no Brasil*. São Paulo: Humanitas, 2008.

TAGNIN, S. E. O.; BEVILACQUA, C. (Org.). *Corpora na Terminologia*. São Paulo: Hub Editorial, 2013.

TEMMERMAN, R. Questioning the univocity ideal. The difference between socio-cognitive Terminology and traditional Terminology. *Hermes, Journal of Linguistics*, n. 18, p. 51-90, 1997. <https://doi.org/10.7146/hjlc.v10i18.25412>

TEMMERMAN, R. *Towards New Ways of Terminology Description: the sociocognitive approach*. Amsterdã/Filadélfia: John Benjamins, 2000. <https://doi.org/10.1075/tlrp.3>

TOGNINI-BONELLI, E. *Corpus Linguistics at Work*. Amsterdam: John Benjamins, 2001. <https://doi.org/10.1075/scl.6>

TUFANO, D. *Estudos de Literatura brasileira*. 4.ed. São Paulo: Moderna, 1988.

VIANA, V. Linguística de corpus: conceitos, técnicas e análises. In: VIANA, V. ; TAGNIN, S. E. O. (Orgs.). *Corpora no ensino de línguas estrangeiras*. São Paulo: Hub Editorial, 2010. p. 25-95.

VIANA, V.; TAGNIN, S. E. O. (Org.). *Corpora no ensino de línguas estrangeiras*. São Paulo: Hub Editorial, 2010.

VIANA, V.; TAGNIN, S. E. O. *Corpora no ensino de línguas estrangeiras*. São Paulo: Hub Editorial, 2011.

VIANA, V.; TAGNIN, S. E. O. (Org.). *Corpora na tradução*. São Paulo: Hub Editorial, 2015.

URUPÊS. *Guia do estudante*. Disponível em: <<https://guiadoestudante.abril.com.br/estudo/urupes-resumo-e-analise-da-obra-de-monteiro-lobato/>>. Acesso em: 20 jul. 2018.

URURAU. Disponível em: <<https://pt.wikipedia.org/wiki/Ururau>>. Acesso em: 23 jul. 2018.

## 1.1 Referências das obras que compõem o *corpus* de estudo

ALENCAR, José de. *O sertanejo*. São Paulo: Melhoramentos, 1995. Disponível em: <<http://www.bibvirt.futuro.usp.br>>. Acesso em: 3 mar 2017.

ALMEIDA, José Américo de. *A bagaceira*. 13. ed. Rio de Janeiro: José Olympio, 1974.

AMADO, Jorge. *Seara Vermelha*. São Paulo: Companhia das Letras, 2009. Disponível em: <<https://licoedecasa.files.wordpress.com/2012/07/jorge-amado-seara-vermelha.pdf>>. Acesso em: 3 abr. 2017.

ARINOS, Afonso. *Pelo sertão*. Rio de Janeiro: F. Briguiet & CIA, 1947. Disponível em: <<http://bibdig.biblioteca.unesp.br/handle/10/25989>>. Acesso em: 4 abr. 2017.

CARVALHO, José Cândido de. *O coronel e o lobisomem*. São Paulo: Companhia das Letras, 2014. Disponível em: <<https://pt.scribd.com/document/366764352/Jose-Candido-de-Carvalho-O-coronel-e-o-lobisomem-pdf>>. Acesso em: 5 jul. 2018.

CUNHA, Euclides da. *Os sertões*. São Paulo: Três, 1984 (Biblioteca virtual do estudante brasileiro). Disponível em: <<http://www.bibvirt.futuro.usp.br>>. Acesso em: 3 abr. 2017.

GUIMARÃES, Bernardo. *O Garimpeiro*. São Paulo: Ática, 1997. Disponível em: <<http://www.bibvirt.futuro.usp.br>>. Acesso em: 5 mar. 2017.

LOBATO, Monteiro. *Urupês*. 37. ed. São Paulo: Brasiliense, 1994. Disponível em: <<https://ufprbrasilaluis.files.wordpress.com/2015/02/bocatorta-e-urupc3aas.pdf>>. Acesso em: 4 abr. 2017.

LOPES NETO, João Simões. *Contos gauchescos*. 9. ed. Porto Alegre: Globo, 1976. Disponível em: <<http://www.bibvirt.futuro.usp.br>>. Acesso em: 3 abr. 2017.

OLÍMPIO, Domingos. *Luzia-Homem*. São Paulo: Ática, 1983. (Universidade da Amazônia) Disponível em: <<https://www.nead.unama.br>>. Acesso em: 4 abr. 2018.

PALMÉRIO, Mário. *Chapadão do Bugre*. 12.ed. Rio de Janeiro: José Olympio, 2006. Disponível em: <<https://docero.com.br/doc/ns81c>>. Acesso em: 3 maio 2017.

QUEIROZ, Rachel de. *O Quinze*. 93. ed. Rio de Janeiro: J. Olympio, 2012. Disponível em: <[https://vivelatinoamerica.files.wordpress.com/2016/03/o\\_quinze\\_obra\\_rachel\\_de\\_queiroz.pdf](https://vivelatinoamerica.files.wordpress.com/2016/03/o_quinze_obra_rachel_de_queiroz.pdf)>. Acesso em: 8 mar. 2018.

RAMOS, Graciliano. *Vidas secas*. 45. ed. São Paulo: Record, 1982. Disponível em: <<https://dynamicon.com.br/wp-content/uploads/2017/02/Vidas-secas-de-Graciliano-Ramos.pdf>>. Acesso em: 3 abr. 2016.

RAMOS, Hugo de Carvalho. *Tropas e boiadas*. Goiânia (GO): Instituto Centro-Brasileiro de Cultura, 1917. Disponível em: <<http://www.dominiopublico.gov.br/download/texto/cb000001.pdf>>. Acesso em: 3 abr. 2017.

REGO, José Lins do. *Menino de engenho*. 80. ed. Rio de Janeiro: José Olympio, 2001. Disponível em: <<https://www.skoob.com.br/livro/pdf/menino-de-engenho/livro:970/edicao:1278>>. Acesso em: 3 abr. 2017.

RIBEIRO, João Ubaldo. *Sargento Getúlio*. 5. ed. São Paulo: Ponto de Leitura, 2010. Disponível em: <<https://www.skoob.com.br/livro/pdf/sargento-getulio/livro:3408/edicao:173278>>. Acesso em: 5 jul. 2018.

ROSA, João Guimarães. *Grande Sertão: Veredas*. Rio de Janeiro: Editora Nova Aguilar, 1994. Disponível em: <<http://www.blam.com.br>>. Acesso em: 3 abr. 2017.

SUASSUNA, Ariano. *Auto da compadecida*. 36. ed. Rio de Janeiro: Nova Fronteira, 2014. Disponível em: <[https://moodle.ufsc.br/pluginfile.php/1905669/mod\\_resource/content/1/Auto%20da%20Compadecida.pdf](https://moodle.ufsc.br/pluginfile.php/1905669/mod_resource/content/1/Auto%20da%20Compadecida.pdf)>. Acesso em: 5 jul. 2018.

TAUNAY, Alfredo Maria Adriano d'Escragnoles. *Inocência*. 25. ed. São Paulo: Ática, 1998. Disponível em: <<http://www.dominiopublico.gov.br/download/texto/bn000002.pdf>>. Acesso em: 3 abr. 2017.

TÁVORA, Franklin. *O cabeloira*. 8. ed. São Paulo: Ática, 1973. Disponível em: <<https://www.nead.unama.br>>. Acesso em 3 abr. 2017.