



**UNIVERSIDADE FEDERAL DO PARÁ  
INSTITUTO DE LETRAS E COMUNICAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS  
MESTRADO EM LINGUÍSTICA**

**Davi Pereira de Souza**

**FRASEOLOGISMOS NO DISCURSO POLÍTICO  
BRASILEIRO: uma proposta de glossário**

**Volume I**

**Belém**

**2018**

**DAVI PEREIRA DE SOUZA**

**FRASEOLOGISMOS NO DISCURSO POLÍTICO BRASILEIRO:  
uma proposta de glossário**

Dissertação apresentada ao Programa de Mestrado em Letras, da Universidade Federal do Pará, como requisito parcial para obtenção do título de Mestre em Linguística.

Orientador: Prof. Dr. Abdelhak Razky  
Coorientador: Prof.<sup>a</sup>. Dr.<sup>a</sup>. Maria Luísa Ortiz Alvarez

**Belém**

**2018**

**Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD  
Sistema de Bibliotecas da Universidade Federal do Pará  
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo autor**

---

S719f Souza, Davi Pereira de  
FRASEOLOGISMOS NO DISCURSO POLÍTICO BRASILEIRO: uma  
proposta de glossário / Davi Pereira de Souza. — 2018.  
263 f.

Orientador(a): Prof. Dr. Abdelhak Razky  
Coorientação: Prof<sup>a</sup>. Dra. Maria Luiza Ortiz Alvarez  
Dissertação (Mestrado) - Programa de Pós-Graduação em Letras, Instituto de  
Letras e Comunicação, Universidade Federal do Pará, Belém, 2018.

1. Fraseologia. 2. Glossário fraseológico. 3.  
Linguística de corpus. 4. Discurso político. I. Título.

CDD 410

---

---

**DAVI PEREIRA DE SOUZA**

**FRASEOLOGISMOS NO DISCURSO POLÍTICO BRASILEIRO: uma  
proposta de glossário**

Dissertação apresentada ao Programa de Mestrado em Letras, da Universidade Federal do Pará, como requisito parcial para obtenção do título de Mestre em Linguística.

**Banca examinadora**

---

Prof. Dr. Abdelhak Razky – orientador  
UFPA/UNB

---

Prof.<sup>a</sup>. Dr.<sup>a</sup>. Maria Luísa Ortiz Alvarez – Coorientadora  
UNB

---

Prof.<sup>a</sup>. Dr.<sup>a</sup>. Marcela Moura Torres Paim – Membro externo  
UFBA

---

Prof.<sup>a</sup>. Dr.<sup>a</sup>. Regina Célia Fernandes Cruz – Membro interno  
UFPA

---

Prof. Dr. Alcides Fernandes de Lima – Suplente  
UFPA

**Belém**

**2018**

À minha mãe, exemplo de mulher batalhadora e sonhadora!

## AGRADECIMENTOS

Se não fossem pelas limitações impostas pela tradição acadêmica, esta seção de agradecimentos teria certamente um volume muito extenso, pois, mesmo que a cultura atual de produção do conhecimento tenda a ser egoísta e solitária, não deveríamos nos esquecer dos mais simples gestos e atitudes das pessoas, do carinho, apoio e amor dispensados a quem, muitas vezes, prefere o isolamento social para poder estudar, sacrificando momentos que talvez nunca voltem a repetir-se.

Nesse sentido, devo aqui primeiramente registrar meus sinceros e mais profundos agradecimentos a Deus, ser pessoal cuja benevolência em favor de seus filhos jamais poderá ser plenamente retribuída. Mas vale a pena o esforço, o qual deve transcender esta singela menção honrosa neste trabalho. Ao meu pai eterno, agradeço pela saúde, pela paz que excede todo o entendimento, pelas bênçãos incontáveis e não merecidas e, também, pelas lutas e dificuldades nas quais a fraqueza revelou-se forte, o deserto tornou-se manancial e aquilo que parecia impossível pôde, enfim, ser alcançado.

Sou imensamente grato à minha família, especialmente mãe e irmãos; sem o apoio de vocês, não teria conseguido chegar até aqui! Muito obrigado pelos sacrifícios, por sonharem comigo e oferecerem o suporte necessário quando precisei. Mãe, desculpe-me pela luz acesa durante tantas noites de estudo, sei que a senhora não dormia enquanto eu não me deitava; desculpe-me, ainda, pela minha ausência constante e por parecer estar distante, mesmo compartilhando o mesmo teto. Sei que a senhora me compreendeu, mas não deixa de ser injusto para ambos. Por isso, sou eternamente grato à senhora!

Agradeço também ao meu orientador, prof. Dr. Abdelhak Razky, pelas observações, sugestões e críticas ao trabalho. Sou igualmente grato à minha coorientadora, professora Dr<sup>a</sup>. Maria Luísa Ortiz Alvarez, pelas orientações, acompanhamento, conselhos e incentivos; sua parceria nesta pesquisa foi fundamental, sua paciência e palavras de conforto jamais serão esquecidas por mim! Aproveito, também, para agradecer aos professores que participaram da banca de qualificação e da de defesa, respectivamente: Dr<sup>a</sup>. Regina Cruz, Dr<sup>a</sup>. Marcela Paim e Dr. Alcides Fernandes. Obrigado pelas valiosas contribuições!

Faço um agradecimento especial aos colaboradores da pesquisa, que validaram os fraseologismos sem registro nas obras de referência consultadas. Obrigado pela gentileza e por aceitarem compartilhar seus conhecimentos linguísticos, reconhecendo e definindo alguns fraseologismos.

Gostaria também de agradecer aos amigos e amigas que, a seu modo, contribuíram direta e indiretamente com este trabalho, oferecendo apoio psicológico, financeiro e até abrigo. Especialmente, sou grato a Carlene Salvador, amiga que o GeoLinTerm me deu, a qual se tornou como uma irmã; obrigado, Carlene, por tudo, pelas sugestões e orientações para a pesquisa fraseológica, mas principalmente pelas risadas e por ter sido uma de minhas confidentes durante todo esse árduo processo. Também sou grato a Rejane Garcez, pela companhia sempre agradável e por formar, junto comigo e com a Carlene, o grupo dos “gordinhos da fraseologia” (risos). Também não poderia esquecer de mencionar o meu amigo Romário Sanches, que me aceitou em sua residência durante o primeiro ano do mestrado e, ainda, o meu amigo Thiago Azevedo, cujo carinho e gratidão ficam aqui reconhecidos. Por extensão, agradeço a todas as pessoas, amigos, coordenadores e voluntários do/no Projeto GeoLinTerm. Tenho muito orgulho de ser um geolintermiano, e devo isso ao professor Alcides Lima.

Agradeço aos amigos Joelson Sousa, Raquel Rocha, Francisca Imaculada e Larisse Silva; vocês são inesquecíveis! Sou igualmente agradecido à equipe de voluntários do Projeto Logos e aos irmãos em Cristo da igreja evangélica Assembleia de Deus, Congregação Jaderlândia 09! Obrigado por compreenderem minhas ausências e me ajudarem em oração! Que Deus os abençoe em dobro!

Ao CNPq, agradeço, enfim, pela concessão da bolsa de mestrado, sem a qual talvez a presente dissertação não teria sido finalizada.

## RESUMO

Esta pesquisa, de abordagem quali-quantitativa, tem como objetivo geral produzir um glossário, em versão impressa e eletrônica, de fraseologismos utilizados no discurso político brasileiro. Por sua vez, os objetivos específicos consistem em: descrever os fraseologismos que caracterizam o discurso político brasileiro; identificar padrões de combinatórias sintagmáticas recorrentes no *corpus* e; verificar possíveis variantes fraseológicas. Os fraseologismos, ou unidades fraseológicas, são combinações sintagmáticas recorrentes (MEJRI, 1997, 2012), caracterizadas, dentre outros aspectos, pela sua polilexicalidade, fixidez, frequência, congruência e idiomaticidade. Para tanto, adotou-se uma metodologia orientada pelos pressupostos gerais da Linguística de *Corpus* (BERBER SARDINHA, 2004) e sua relação com a Fraseologia (TAGNIN, 2005, 2011). A pesquisa foi dividida em cinco etapas principais, a saber: i) revisão da bibliografia sobre a área em foco, particularmente as pesquisas fraseológicas desenvolvidas no Brasil e na França; ii) constituição e tratamento do *corpus*; (iii) seleção do *corpus* de referência; (iv) procedimentos de análise dos resultados e; v) elaboração do glossário fraseológico. Os 570 textos que compõem o *corpus* são provenientes de *blogs ou websites* de 4 (quatro) colunistas que assinam matérias sobre política nas revistas *Istoé*, *Época*, *Carta Capital* e no jornal *Folha de São Paulo*, tendo sido escolhido um colunista por periódico. Os textos foram publicados entre janeiro de 2014 e dezembro de 2016. Foram utilizados os *softwares Words Smith Tools* (SCOTT, 2008), que realiza busca semiautomática em grandes *corpora* textuais, e o *Lexique pro* – versão 3.6 (SIL, 2004-2012), para preenchimento da ficha fraseológica de cada verbete, resultando posteriormente na organização do glossário, adotando-se microestrutura formada por *entrada, categoria gramatical, definição, contexto, variante fraseológica, remissiva e notas*. Quanto ao referencial teórico, o trabalho ancorou-se na abordagem francesa da Fraseologia, sobretudo na perspectiva de Salah Mejri (1997, 1998, 1999, 2002, 2005, 2010, 2011, 2012, 2018). O glossário produzido contém 438 entradas, lematizadas e organizadas alfabeticamente pela primeira unidade lexical da sequência. Os resultados demonstram a predominância de fraseologismos da língua geral em detrimento de unidades fraseológicas específicas do discurso político, o que está relacionado ao fato de o *corpus* não ser especializado, uma vez que os colunistas não são tecnicamente cientistas políticos, mas jornalistas e comentaristas que lidam com assuntos da área e se direcionam para um público geral de leitores, em grande parte formado também por não especialistas. Além disso, a política, sendo de natureza interdisciplinar, produz um discurso que se constitui no cruzamento de outros domínios, como o direito, as ciências sociais, a linguística, entre outros (DORNA, 1995; CHARAUDEAU, 2006). De todo modo, os fraseologismos desempenham um papel peculiar nesse domínio, servindo para produzir diferentes efeitos de sentido, particularmente, os de caráter irônico e ambíguo, presentes nas relações estabelecidas pelos interlocutores inseridos nas tensões ideológicas e político-partidárias que se acirram em momentos de crise política e econômica.

**Palavras-chave:** Fraseologia. Glossário fraseológico. Linguística de *Corpus*. Discurso político.

## ABSTRACT

This qualitative-quantitative research aims to produce a glossary, in print and electronic, of phraseologisms used in the Brazilian political discourse. In turn, the specific objectives consist in: describing the phraseologies that characterize Brazilian political discourse; identify patterns of recurrent syntagmatic combinatorics in the corpus; check possible phraseological variants and. Phraseologisms, or phraseological units, are recurrent syntagmatic combinations (MEJRI, 1997, 2012), characterized, among other aspects, by its polilexicality, fixity, frequency, congruence and idiomaticity. In order to do so, a methodology guided by the general assumptions of Corpus Linguistics (BERBER SARDINHA, 2004) and its relationship with Phraseology (TAGNIN, 2005, 2011). The research was divided into five main stages, namely: i) review of the bibliography on the area in focus, particularly the phraseological researches developed in Brazil and France; ii) constitution and treatment of the corpus; (iii) selection of the reference corpus; (iv) procedures for analyzing results and; v) elaboration of the phraseology glossary. The 570 texts that make up the corpus come from blogs or websites of 4 (four) columnists who sign matters about politics in the periodicals *Istoé*, *Época*, *Carta Capital* and in the newspaper *Folha de São Paulo*, being chosen a columnist by periodical. The texts were published between January 2014 and December 2016. We used the software Words Smith Tools (SCOTT, 2008), which performs semiautomatic search in large corpora textual, and Lexique pro - version 3.6 (SIL, 2004-2012), to fill the phraseological form of each entry, resulting in the organization of the glossary, organization of the glossary, adopting microstructure formed by entry, grammatical category, definition, context, phraseological variant, reference and notes. As for the theoretical reference, the work was anchored in the French approach to Phraseology, especially in the perspective of Salah Mejri (1997, 1998, 1999, 2002, 2005, 2010, 2011, 2012, 2018). The glossary produced contains 438 entries, lemmatized and organized alphabetically by the first lexical unit of the sequence. The results demonstrate the predominance of phraseologisms of the general language to the detriment of specific phraseological units of political discourse, which is related to the fact that the corpus is not specialized, since the columnists are not technically political scientists, but journalists and commentators who deal with subjects of the area and are directed towards a general public of readers, largely formed also by non-specialists. In addition, politics, being of an interdisciplinary nature, produces a discourse that is constituted by the crossing of other domains, such as law, social sciences, linguistics, among others (DORNA, 1995, CHARAUDEAU, 2006). In any case, the phraseologisms play a peculiar role in this domain, serving to produce different effects of sense, particularly those of ironic and ambiguous character, present in the relations established by the interlocutors inserted in the ideological and political-partisan tensions that arise in moments of political and economic crisis.

**Key-words:** Phraseology. Glossary of terms. Linguistics of *Corpus*. Political discourse.

## LISTA DE ILUSTRAÇÕES

<b>Ilustração 1</b> – Texto em <i>Word</i> antes da limpeza .....	96
<b>Ilustração 2</b> – Exemplo de texto em formato <i>TXT</i> após o processo de limpeza.....	97
<b>Ilustração 3</b> – Tela inicial do programa <i>WordSmith Tools 6.0</i> .....	100
<b>Ilustração 4</b> – Janela de trabalho do <i>WordList</i> , contendo um exemplo do <i>corpus</i> de estudo.....	100
<b>Ilustração 5</b> – <i>WordList</i> elaborada após inserção da <i>stoplist</i> .....	101
<b>Ilustração 6</b> – Tela do <i>Keywords</i> indicando a seleção de duas <i>wordlists</i> para a composição de palavras-chave.....	102
<b>Ilustração 7</b> – <i>KeyWords</i> com as 30 primeiras palavras-chave.....	103
<b>Ilustração 8</b> – Concordâncias para a palavra <i>Dilma</i> .....	104
<b>Ilustração 9</b> – Concordância para <i>Dilma</i> , ordenada pela primeira palavra à esquerda.....	107
<b>Ilustração 10</b> – Visualização do fraseologismo <i>gol contra</i> no seu contexto real de uso, exibida na plataforma do <i>Concord</i> .....	109
<b>Ilustração 11</b> – Verbetes de <i>política externa</i> no modo editar.....	117
<b>Ilustração 12</b> – Verbetes de <i>política exterior</i> no modo editar.....	118
<b>Ilustração 13</b> – Visualização em nuvem dos componentes lexicais dos fraseologismos que remetem à política.....	120
<b>Ilustração 14</b> – Consulta para <i>bode expiatório</i> na versão eletrônica do dicionário Aurélio.....	122
<b>Ilustração 15</b> – Nomes relativos a partes do corpo que compõem fraseologismos.....	127
<b>Ilustração 16</b> – Nomes de objetos e acessórios que formam fraseologismos..	129

**Ilustração 17** – Visualização em nuvem de palavras do reino animal que se mostraram produtivas em relação aos fraseologismos coletados no *corpus*...130

## LISTA DE GRÁFICOS

**Gráfico 1** – Estrutura sintagmática dos fraseologismos.....124

**Gráfico 2** – Distribuição dos fraseologismos por agrupamentos temáticos .....126

**Gráfico 3** – Relação entre a estrutura polilexical e as ocorrências no *corpus*...132

## LISTA DE QUADROS E ESQUEMAS

**Quadro 1** – Organização dos paradigmas terminológicos relacionados aos resultados do *phraséologie* e do *figement*.....62

**Quadro 2** – Caracterização de Barbosa (2001) para dicionário de língua, vocabulário técnico-científico/especializado e glossário.....75

**Quadro 3** – Projetos ligados à fraseologia cadastrados no diretório do CNPq.90

**Quadro 4** – Relação dos colunistas, periódico e quantidade de textos selecionados.....94

**Quadro 5** – Dimensões do *corpus* de estudo.....97

**Quadro 6** – Dimensões do *corpus* de referência.....98

**Quadro 7** – Listagem dos 30 primeiros agrupamentos com a palavra *Dilma*.105

**Quadro 8** – perfil dos colaboradores do **grupo A** que validaram os fraseologismos de uso geral e de outros domínios diferentes do político.....110

<b>Quadro 9</b> – Perfil dos colaboradores do <b>grupo B</b> que validaram os fraseologismos do discurso político.....	111
<b>Esquema 1</b> – Microestrutura do glossário.....	113
<b>Quadro 10</b> – Etiquetas para o verbete principal.....	117
<b>Quadro 11</b> – Análise de fraseologismos verbais quanto à fixidez sintática.....	133
<b>Quadro 12</b> – Análise de fraseologismos verbais quanto à fixidez paradigmática.....	135
<b>Quadro 13</b> – Análise de sequências nominais quanto ao cruzamento das noções de fixidez e congruência.....	139

## LISTA DE ABREVIATURAS E SIGLAS

ALiB	Atlas Linguístico do Brasil
BFQS	Projeto Bélgica-França-Quebec-Suíça
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CELL	Complexidade, estudos em Linguística Aplicada e Linguagem
CITRAT	Centro Interdepartamental de Tradução e Terminologia da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo.
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
COMET	<i>Corpus</i> Multilíngue para Ensino e Tradução
EURALEX	European Association for Lexicography
EUROPHAS	European Society of Phraseology
GECon	Grupo de Estudos Contrastivos
GeoLinTerm	Geossociolinguística e Socioterminologia
ISO	International Standardization Organization
LADL	Laboratoire d'Automatique Documentaire et Linguistique
LC	Linguística de <i>Corpus</i>
LDI	Lexiques Dictionnaires Informatique
LexFras	Lexicologia, Lexicografia, Terminologia e Fraseologia
LLI	Laboratoire de Linguistique Informatique
PLN	Processamento de Língua Natural
PPGL	Programa de Pós-Graduação em Letras
PT	Partido dos Trabalhadores
SIL	<i>Summer Institute of Linguistics</i>
TERMISUL	Projeto Terminológico Cone Sul
TRALEX	Ensino – Tradução, Léxico e Ensino de Línguas Estrangeiras
UEAP	Universidade Estadual do Amapá
UEFS	Universidade Estadual de Feira de Santana
UEG	Universidade Estadual de Goiás
UFBA	Universidade Federal da Bahia
UFGD	Universidade Federal da Grande Dourados
UFMS	Universidade Federal do Mato Grosso do Sul
UFPA	Universidade Federal do Pará
UFRGS	Universidade Federal do Rio Grande do Sul
UFRN	Universidade Federal do Rio Grande do Norte
UFU	Universidade Federal de Uberlândia
UNB	Universidade de Brasília
UNESP	Universidade Estadual Paulista Júlio de Mesquita Filho
URSS	União das Repúblicas Socialistas Soviéticas
USP	Universidade de São Paulo
VALEXTRA	Projeto Variação Lexical, Teorias, Recursos e Aplicações

## SUMÁRIO

### VOLUME I

<b>1 INTRODUÇÃO</b> .....	<b>14</b>
<b>CAPÍTULO 2 – A POLÍTICA</b> .....	<b>19</b>
2.1 CONCEPÇÕES DE POLÍTICA .....	19
2.2 O DISCURSO POLÍTICO .....	23
<b>CAPÍTULO 3 – FUNDAMENTAÇÃO TEÓRICA</b> .....	<b>32</b>
3.1 FRASEOLOGIA: QUESTÕES GERAIS .....	32
3.2 A FRASEOLOGIA E SEUS DESDOBRAMENTOS .....	33
3.3 BREVE HISTÓRICO DOS ESTUDOS FRASEOLÓGICOS .....	39
3.4 ESTUDOS FRASEOLÓGICOS NO BRASIL: BREVE PANORAMA.....	41
3.5 A TRADIÇÃO FRANCESA DE ESTUDOS FRASEOLÓGICOS.....	46
3.5.1 <b>As pesquisas em torno do conceito de <i>figement</i> (cristalização)</b> .....	47
3.5.2 <b>A proposta de Salah Mejri</b> .....	65
3.5.2.1 <i>Critérios para a delimitação de fraseologismos</i> .....	72
3.6 CONCEPÇÃO DE DICIONÁRIO, VOCABULÁRIO E GLOSSÁRIO .....	80
3.7 LINGUÍSTICA DE <i>CORPUS</i> .....	92
3.7.1 <b>O que é Linguística de <i>Corpus</i>?</b> .....	93
3.7.2 <b>Tipos de <i>corpus</i></b> .....	94
<b>CAPÍTULO 4 – METODOLOGIA</b> .....	<b>96</b>
4.1 NATUREZA DA PESQUISA.....	96
4.2 DESCRIÇÃO DAS ETAPAS DE PESQUISA.....	96
4.2.1 <b>Recensão da literatura especializada</b> .....	96
4.2.2 <b>Constituição e delimitação do <i>corpus</i> de estudo</b> .....	101
4.2.2.1 <i>Levantamento dos textos escritos</i> .....	101
4.2.2.2 <i>Limpeza e armazenamento dos textos</i> .....	102
4.2.3 <b>Seleção do <i>corpus</i> de referência</b> .....	106
4.2.4 <b>Procedimentos de análise</b> .....	107
4.2.4.1 <i>Elaboração das listas de frequência e de palavras-chave</i> .....	107
4.2.4.2 <i>Levantamento de candidatos a fraseologismos</i> .....	111
4.2.4.3 <i>Validação dos fraseologismos identificados no corpus</i> .....	116
4.2.5 <b>Elaboração do glossário fraseológico</b> .....	119

4.2.5.1 <i>Organização macroestrutural do glossário</i> .....	119
4.2.5.2 <i>Organização microestrutural do glossário</i> .....	119
4.2.5.3 <i>O uso do Lexique Pro na organização do glossário</i> .....	123
<b>CAPÍTULO 5 – DISCUSSÃO E ANÁLISE DOS RESULTADOS</b> .....	<b>127</b>
5.1 DESCRIÇÃO GERAL DOS RESULTADOS .....	127
5.1.1 <b>Quanto ao registro nos dicionários de língua portuguesa</b> .....	129
5.1.2 <b>Padrões formais de combinação</b> .....	131
5.1.3 <b>Distribuição dos fraseologismos em agrupamentos temáticos</b> .....	133
5.2 CARACTERIZAÇÃO DAS UNIDADES FRASEOLÓGICAS .....	139
5.2.1 <b>Quanto à polilexicalidade</b> .....	139
5.2.2 <b>Quanto à fixidez</b> .....	140
5.2.3 <b>Quanto à congruência</b> .....	146
5.3 <i>A COBRA VAI FUMAR NA POLÍTICA OU AINDA É POSSÍVEL DAR A VOLTA POR CIMA? ASPECTOS DO USO DE FRASEOLOGISMOS NO DISCURSO POLÍTICO</i> .....	149
5.3.1 <b>A capacidade expressiva e de concisão do fraseologismo</b> .....	150
5.3.2 <b>Os efeitos de ironia implicados no uso de fraseologismos</b> .....	152
5.3.3 <b>Efeitos de ambiguidade implicados no uso de fraseologismos</b> .....	154
<b>CONSIDERAÇÕES FINAIS</b> .....	<b>158</b>
<b>REFERÊNCIAS</b> .....	<b>161</b>

## VOLUME II

<b>6 APRESENTAÇÃO DO GLOSSÁRIO</b> .....	165
<b>ÍNDICE</b> .....	241
<b>REFERÊNCIAS DO GLOSSÁRIO</b> .....	249

## 1 INTRODUÇÃO

Dois fatos inerentes à espécie humana, que se cruzam teoricamente desde as origens mais remotas de organização social, são certamente a linguagem e a política, sendo a primeira responsável pela estruturação da segunda (DORNA, 1995). De fato, pode-se considerar, em sentido amplo, que todo ato de linguagem é também um ato político na medida em que busca, em última análise, agir sobre o outro, para influenciá-lo. Assim, tal como a política, a linguagem presta-se a negociações, a tensões ideológicas, visando a algum tipo de poder. E, em nome do poder, tanto uma quanto a outra podem tornar-se prejudiciais à sociedade, provocando guerras, crises e descontentamento. Não é fortuito que atualmente, no Brasil, os rótulos *coxinha* e *mortadela*, dentre tantos outros que viralizam com o suporte das mídias virtuais, figuram como instrumentos linguísticos que demarcam posicionamentos, interesses e afiliações político-partidárias, servindo-se, muitas vezes, a algum tipo de desconstrução da biografia de alguém, criando assim um clima hostil, no qual a simples menção a A ou a B pode desencadear discussões e até pancadaria.

Além dessas palavras de cunho marcadamente pejorativo, outras unidades lexicais, de maior extensão sintagmática, descrevem, nomeiam ou denunciam atos, práticas, processos e relações estabelecidos no âmbito da política. Desse modo, sequências como *crime de colarinho branco*, *balcão de negócios*, *caixa dois* e *lavagem de dinheiro* fazem referência a uma parte corrosiva, viral e repugnante da política que se instalou em praticamente todos os segmentos e repartições sociais, indo do ambiente familiar a mais alta cúpula de magistrados. A sensação, principalmente, no Brasil, é de uma terra *sem lei* (e justiça!), onde os cidadãos parecem estar constantemente *no olho do furacão*.

Ao interessar-se por essas unidades lexicais cristalizadas, específicas ou não ao discurso político, mas utilizadas nesse domínio, esta dissertação, vinculada ao projeto de pesquisa Geossociolinguística e Socioterminologia – GeoLinTerm (RAZKY; LIMA; OLIVEIRA, 2010), constitui a primeira pesquisa a nível de mestrado no âmbito do referido projeto na área da Fraseologia, seguindo-se à tese pioneira intitulada *Fraseologismos do futebol brasileiro: proposta de um dicionário eletrônico* (SALVADOR, 2017) e à tese de Garcez (2018), recentemente defendidas no Programa de Pós-graduação em Letras

(PPGL) da Universidade Federal do Pará (UFPA). Pode-se dizer, portanto, que o presente estudo inaugura, juntamente com as teses já citadas, um novo eixo de investigação no contexto do GeoLinTerm e no quadro mais geral das pesquisas acadêmicas do PPGL.

A investigação ancora-se nos aportes teóricos da Fraseologia francesa e na abordagem metodológica da Linguística de *Corpus*, tendo como objetivo final produzir um glossário, em versão impressa e eletrônica, de fraseologismos utilizados no discurso político brasileiro, a partir de um *corpus* escrito constituído de textos que tratam acerca de política, extraídos de *blogs*/colunas pessoais de quatro colunistas que assinam matéria sobre o tema nas revistas *Época*, *Istoé*, *Carta Capital* e no jornal *Folha de São Paulo*. E, como objetivos específicos, foram delineados: a) *descrever os fraseologismos que caracterizam o discurso político brasileiro*; b) *identificar padrões de combinações sintagmáticas recorrentes no corpus*; c) *verificar e descrever possíveis variantes fraseológicas*.

Para tratar do tema selecionado, partiu-se de três questões gerais que configuram a problemática da pesquisa: (i) *Quais unidades fraseológicas caracterizam a linguagem da política no Brasil?*; (ii) *Em que medida os grandes acontecimentos no meio político do país influenciam o uso ou a criação de fraseologismos?*; e (iii) *Em qual direção os fraseologismos identificados no corpus se movimentam: da língua geral ao domínio especializado da política, ou vice-versa?*

Diante dessas questões, foram propostas inicialmente algumas hipóteses. Em relação à primeira questão, acreditava-se que, dentre as unidades fraseológicas, deveriam ser mais recorrentes no *corpus* expressões idiomáticas, ditos populares e colocações. Essa hipótese estava associada à ideia de que a política constitui um domínio que, apesar de seu caráter especializado, mantém interface com outras áreas do conhecimento humano estreitamente relacionadas ao cotidiano, o que poderia explicar um maior número de fraseologismos que apresentam mais opacidade do ponto de vista semântico, como as expressões idiomáticas. Já com respeito à segunda questão, assumia-se que determinados fatos políticos, de grande alcance nacional, como os recentes casos de corrupção na Petrobrás e o *impeachment* da ex-presidente Dilma Rousseff, certamente favoreceriam a criação e intensificação de fraseologismos. Por fim, em se tratando da terceira questão colocada acima, defendia-se que

possivelmente as unidades fraseológicas identificadas seriam em sua maioria provenientes do uso da língua comum, em direção ao domínio especializado da política, pelo fato de a política ser uma área que atravessa outros domínios, sobretudo o cotidiano, além de os jornalistas/comentaristas selecionados se dirigirem para um público amplo de interessados no assunto, formado majoritariamente por cidadãos comuns, sem formação específica em política.

A justificativa para tal dissertação reside em três motivos principais. O primeiro, de caráter mais geral, traduz-se na importância de se investigar o campo da política em seu aspecto linguístico, mas também ideológico e cultural, na medida em que as fraseologias podem revelar concepções de mundo e ideologias presentes numa dada sociedade como a brasileira. O segundo motivo diz respeito à existência de poucos estudos fraseológicos no campo da política brasileira<sup>1</sup>. Até o momento, sabe-se de apenas dois trabalhos que analisam fraseologismos utilizados no discurso político, como o artigo *La fraseología en el discurso político y económico de los medios de comunicación*, de Carmen Navarro (2002), e o livro *A CPI das palavras: origem de palavras e expressões da linguagem política*, do professor Ari Riboldi (2008). Esta obra de Riboldi, apesar de sua relevância, não constitui tecnicamente um repertório lexical específico de fraseologismos, já que o autor também trata de unidades monolexicais. Com exceção destes dois trabalhos, não se tem conhecimento de que haja glossário ou dicionário fraseológico do discurso político<sup>2</sup>; portanto, neste aspecto, esta pesquisa apresenta-se como inédita. Por fim, a terceira justificativa para este estudo deve-se à necessidade de haver mais pesquisas no Brasil sobre o complexo fenômeno fraseológico cuja delimitação e proposição de critérios de classificação precisam ser mais abrangentes e coerentes com as distintas tipologias de unidades fraseológicas conhecidas. Com efeito, somente

---

<sup>1</sup> Há no domínio da política trabalhos terminológicos, no campo dos neologismos e na análise do discurso, como: *A terminologia política no período pré-eleitoral* (ALVES, 1983); *Aspectos conceituais, terminológicos e políticos das denominações: discurso científico e discurso político* (BARBOSA, 2005); *Uma terminologia do discurso neoliberal?* (CAMPOS, 2007); e *Língua, discurso e política* (FIORIN, 2009). Em fraseologia, os únicos de que se tem conhecimento até o momento são *La fraseología en el discurso político y económico de los medios de comunicación* (NAVARRO, 2002) e os textos de Ortiz Alvarez, *A denominação fraseológica no humor e na política* (2005a) e *A dinâmica e potencialidade da denominação fraseológica no discurso político e humorístico* (2005b). Porém, estes trabalhos não propõem a produção de glossário fraseológico com unidades oriundas da política.

<sup>2</sup> Há um dicionário da política disponível na internet (<https://www.politize.com.br>), com 270 conceitos da área, mas não se constitui trabalho fraseológico.

a pesquisa descritiva, num primeiro momento, poderá oferecer subsídios empíricos para possíveis e necessárias reformulações no campo da Fraseologia, minimizando assim a opacidade terminológica e os inconvenientes ainda existentes na prática lexicográfica e na fraseografia de modo geral.

Em se tratando da metodologia, a pesquisa foi dividida em cinco etapas principais, a saber: (i) revisão da bibliografia sobre a área em foco, particularmente as pesquisas fraseológicas desenvolvidas no Brasil e na França; (ii) constituição e tratamento do *corpus*; (iii) seleção do *corpus* de referência, no capítulo 3; (iv) análise dos resultados; e v) elaboração do glossário com as unidades fraseológicas identificadas no âmbito político brasileiro. A descrição das etapas será detalhada posteriormente, no capítulo de metodologia. Os textos que servem como fonte de consulta para o levantamento e reconhecimento das fraseologias são provenientes de *blogs* ou *websites* de 4 (quatro) colunistas que assinam matérias sobre política nas revistas *Istoé*, *Época*, *Carta Capital* e no jornal *Folha de São Paulo*, sendo um colunista por periódico.

Para otimizar a tarefa de tratamento dos dados, foi utilizado o *software WordsSmith Tools* (SCOTT, 2008) que realiza busca semiautomática em grandes *corpora* textuais. Já a terceira etapa culminou com a produção de um glossário fraseológico de unidades utilizadas no discurso da política brasileira. Para tanto, foi usado o programa *Lexique pro* – versão 3.6 (SIL, 2004-2012), para preenchimento da ficha fraseológica de cada verbete, resultando posteriormente em uma versão impressa e outra eletrônica do glossário.

Do ponto de vista teórico, o trabalho adota principalmente os pressupostos da corrente francesa da Fraseologia, baseando-se nos estudos desenvolvidos por Maurice Gross (1982, 1993), Gaston Gross (1988, 1996), Béatrice Lamiroy (2003, 2005, 2008) e sobretudo Salah Mejri (1997, 1998, 1999, 2002, 2005, 2008, 2012, 2018). Deste modo, considera-se unidade fraseológica ou fraseologismo a combinatória sintagmática recorrente (MEJRI, 2012) que apresenta, em diferentes graus, propriedades como polilexicalidade, fixidez, congruência, opacidade, frequência de uso, idiomaticidade, entre outras.

A dissertação está dividida em dois volumes, sendo o volume I composto por cinco capítulos. No capítulo 1, apresentam-se os objetivos, a justificativa e breve sinalização metodológica da pesquisa. No capítulo 2, discutem-se os

conceitos de política, desde sua origem clássica, até sua concepção moderna, abordando-se, também, o discurso político e suas características.

No capítulo 3, apresentam-se e discutem-se os principais conceitos e fundamentos de base da teoria fraseológica, realizando-se uma breve atualização dos trabalhos brasileiros na área, destacando-se, em seguida, a vertente francesa de estudos fraseológicos. Definem-se, também, as concepções de dicionário, glossário e vocabulário, além dos conceitos operatórios oriundos da Linguística de *Corpus*.

O capítulo 4 apresenta a metodologia utilizada na pesquisa. Descrevem-se as etapas, os recursos e critérios utilizados desde a constituição do *corpus*, passando pela análise e elaboração do glossário.

O capítulo 5 trata dos resultados da pesquisa. Realiza-se primeiramente uma descrição geral dos fraseologismos, com destaque para a análise linguística e ênfase nos efeitos de sentido implicados no uso das sequências cristalizadas no discurso político.

Encerram o volume I as considerações finais, nas quais se retomam os objetivos delineados e alcançados, apontando-se algumas dificuldades encontradas e possíveis contribuições da pesquisa.

Já o volume II constitui o glossário fraseológico produzido. Primeiramente, realiza-se uma apresentação do repertório lexical, explicando-se alguns aspectos da macro e da microestrutura. Por fim, apresentam-se as referências utilizadas especificamente no glossário.

## CAPÍTULO 2 – A POLÍTICA

Neste capítulo, aborda-se a política, conceituando-a a partir das visões clássica e moderna. Discute-se, conseqüentemente, a noção de poder político, contrastando os diferentes pontos de vista a respeito do assunto. Por fim, trata-se do discurso político, enfatizando-o como elemento fundamental e estruturante do campo político.

### 2.1 CONCEPÇÕES DE POLÍTICA

Etimologicamente, o termo política provém do adjetivo grego *politikós*, que deriva de *pólis*<sup>3</sup>, relativo à “cidade autônoma e soberana, cujo quadro institucional é caracterizado por uma ou várias magistraturas, por um conselho e por uma assembléia de cidadãos” (BOBBIO, MATTEUCCI, PASQUINO, 1998, p. 949). Em seu sentido clássico, tributado em geral aos filósofos gregos Platão e sobretudo Aristóteles, política diz respeito à arte ou ciência do Governo. Trata-se, pois, de uma **forma de saber** a respeito de atividades humanas que se referem de algum modo ao Estado (BOBBIO, MATTEUCCI, PASQUINO, 1998, p. 954). Modernamente, o termo passou a designar a **atividade ou o conjunto de atividades** relacionadas ao Estado.

Max Weber (2011 [1967, 1968]) enfatiza que o conceito de política é abrangente, podendo incluir qualquer espécie de atividade diretiva e autônoma. Neste sentido, pode-se falar de política escolar, política de sindicato adotada durante uma determinada greve, política de descontos etc. Ao levar em consideração esse caráter mais fluido da definição, José Luís Fiorin (2009, p. 148), por exemplo, indica que: “fala-se em política do corpo, politizar a sexualidade, política de cotas, política de ação afirmativa etc.” Na mesma direção, Wolfgang Leo Maar (2006), embora reconheça certa unanimidade com relação à política institucional, afirma que, na verdade, o que existem são políticas, ou pelo menos diferentes propostas políticas na sociedade.

Entretanto, o sociólogo alemão, acrescentando a noção de poder à definição moderna de política, entende por esse termo: “o conjunto de esforços

---

<sup>3</sup> Aristóteles (2001) considera a *pólis* como uma sociedade política, sendo a mais importante das associações humanas, cuja vantagem consiste em favorecer o bem mais alto dentre todos.

feitos com vistas a participar do poder ou a influenciar a divisão do poder, seja entre Estados, seja no interior de um único Estado” (WEBER, 2011 [1967, 1968], p. 67). De fato, desde Aristóteles, já se discutia o aspecto do poder político. Logo, a noção de poder não pode ser dissociada do conceito de política, pois esta, sendo concebida como forma de *práxis* humana, está intimamente relacionada à noção de poder (BOBBIO, MATTEUCCI, PASQUINO, 1998).

Assim, o poder constitui um traço definidor da política. Por esse motivo, Maar (2006, p. 8-9) explica que:

Entre o voto e a força das armas está uma gama variada de formas de ação desenvolvidas historicamente visando resolver conflitos de interesses, configurando assim a atividade política em sua questão fundamental: sua relação com o poder.

De fato, o poder está na base da significação do termo político, em qualquer um de seus usos, seja na linguagem comum, seja na linguagem dos especialistas e profissionais. Segundo João Ubaldo Ribeiro (1998), em ambos os casos, o termo se refere ao exercício de alguma forma de poder e às inúmeras consequências desse exercício.

O poder político, por sua vez, se estabelece pela posse dos instrumentos pelos quais se exerce a força física, por meio das armas de todo tipo e potência. Caracteriza-se como uma das formas de poder do homem sobre o próprio homem (BOBBIO, MATTEUCCI, PASQUINO, 1998). Os autores, no entanto, esclarecem que:

Embora a possibilidade de recorrer à força seja o elemento que distingue o poder político das outras formas de poder, isso não significa que ele se resolva no uso da força; tal uso é uma condição necessária, mas não suficiente para a existência do poder político. Não é qualquer grupo social, em condições de usar a força, mesmo com certa continuidade (uma associação de delinqüência, uma chusma de piratas, um grupo subversivo, etc), que exerce um poder político. O que caracteriza o poder político é a **exclusividade** do uso da força em relação à totalidade dos grupos que atuam num determinado contexto social, exclusividade que é o resultado de um processo que se desenvolve em toda a sociedade organizada, no sentido da monopolização da posse e uso dos meios com que se pode exercer a coação física (BOBBIO, MATTEUCCI, PASQUINO, 1998, p. 956, grifo nosso).

Nesse contexto, o Estado se destaca como o “instrumento máximo de institucionalização de todo o poder político” (BONAVIDES, 2000, p. 150). A citação de Weber (2011), a seguir, é explicativa dessa condição de primazia do Estado em relação ao uso do poder político, pois, para o autor, o poder configura uma dominação legítima, baseada no uso da força e da violência. Segundo o autor:

Tal como todos os agrupamentos que historicamente o precederam, o Estado consiste em uma relação de dominação do homem sobre o homem, fundada no instrumento da violência legítima (isto é, da violência considerada como legítima). O Estado só pode existir, portanto, sob condição de que os homens dominados se submetam à autoridade continuamente reivindicada pelos dominadores (WEBER, 2011, p. 67-68).

Tal condição do Estado se deve pelo fato de ele ser detentor majoritário dos meios pelos quais se imputa a força física. O que lhe dá sustentação é o acordo estabelecido entre os cidadãos, seja de forma voluntária e “consciente” (nas democracias, por exemplo), seja de forma autoritária, no casos dos regimes ditatoriais. Nos regimes democráticos, é o voto dos cidadãos o principal instrumento que confere legitimidade ao Estado, a qual está fundamentada na noção de consenso (cf. ARENDT, 1972). A origem desse acordo em prol do Estado está relacionada ao fato de o ser humano, como animal gregário que é, caracterizar-se fundamentalmente por ser comunitário, buscando, no coletivo, satisfazer suas necessidades básicas de sobrevivência (DUSSEL, 2006). Neste sentido, a política é “uma atividade que organiza e promove a produção, reprodução e aumento da vida de seus membros” (DUSSEL, 2006, p. 24)<sup>4</sup>.

Semelhantemente à noção de campo proposta por Pierre Bourdieu, Dussel (2006) trata do campo político. Para o autor argentino, integra o campo político tudo o que se denomina por político, isto é, as ações, instituições, princípios etc., os quais se organizam em sistemas e subsistemas dentro dos quais se situa o sujeito. Segundo Dussel (2006, p. 16), “Todo campo político é um âmbito atravessado por forças, por sujeitos singulares com vontade e com

---

<sup>4</sup> “la política es una actividad que organiza y promueve la producción, reproducción y aumento de la vida de sus miembros” (DUSSEL, 2006, p. 24).

certo poder”.<sup>5</sup> Assim como um campo pode ser atravessado por outros campos e sistemas, o sujeito que nele está inserido é um sujeito intersubjetivo, relacionado a estruturas de poder ou a instituições no âmbito das quais atua como agente que se define em relação aos demais.

A política, embora seja atividade delimitada em um campo, está constantemente em contato com outros campos e sistemas, como a economia, o direito, o esporte etc. Conseqüentemente, pode haver entre eles cruzamento, como exemplifica o autor no trecho a seguir:

Assim como os campos se cruzam (o campo econômico pode cruzar ao político), os sistemas de cada campo podem por sua vez cruzar-se entre si (o sistema capitalista pode cruzar-se com o sistema liberal ou com um sistema pós-colonial de elites formadas na dependência política). A burguesia, com a Revolução inglesa do século XVII, criou um sistema político parlamentar que lhe permitiu desenvolver o sistema capitalista econômico até alcançar a revolução industrial (sistema tecnológico subsumido materialmente dentro do sistema capitalista) (DUSSEL, 2006, p. 18)<sup>6</sup>.

De fato, a política constitui-se na interdisciplinaridade. Assim, de acordo com Patrick Charaudeau (2006), o fenômeno político possui natureza complexa, resultando de um conjunto de fatos que, embora pertencentes a ordens distintas, se cruzam simultaneamente. Trata-se de fatos políticos, fatos sociais, fatos jurídicos e fatos morais, os quais são respectivamente objetos das diferentes disciplinas: ciências políticas, sociologia, direito, filosofia política. Por essa razão, o autor considera que, diferentemente do que alguns poderiam pretender, não há um domínio exclusivo para a análise do fenômeno político, ressaltando seu caráter interdisciplinar.

Conseqüentemente, esse diálogo com outras áreas também se manifesta no nível do discurso político, o qual será tratado na seção a seguir.

---

<sup>5</sup> “Todo campo político es un ámbito atravesado por fuerzas, por sujetos singulares con voluntad, y con cierto poder” (DUSSEL, 2006, p. 24).

<sup>6</sup> “Así como los campos se cruzan (el campo económico puede cruzar al político), los sistemas de cada campo pueden a su vez cruzarse entre ellos (el sistema capitalista puede cruzarse con el sistema liberal o con un sistema poscolonial de elites formadas en la dependencia política). La burguesía, con la Revolución inglesa del siglo XVII, creo un sistema político parlamentario que le permitio desarrollar el sistema capitalista económico hasta alcanzar la revolución industrial (sistema tecnológico subsumido materialmente dentro del sistema capitalista)” (DUSSEL, 2006, p. 18).

## 2.2 O DISCURSO POLÍTICO

Em virtude de o fenômeno político está situado no cruzamento de outras dimensões, o universo da linguagem política não constitui um sistema fechado em si, isto é, estritamente independente e específico, comunicando-se naturalmente com os universos contíguos, tais como a economia, a sociologia, o direito etc. (BOBBIO, MATTEUCCI, PASQUINO, 1998). Essa característica do discurso político também é apontada e discutida por outros autores, como Alexandre Dorna (1995) e Charaudeau (2006). Dorna (1995), por exemplo, cita a linguística e até a teologia como disciplinas com as quais o fenômeno político se intercrusa, além das ciências políticas, da psicologia e da sociologia.

Além disso, a linguagem política é ambígua, tendo a maioria dos termos significados diversos. Segundo Bobbio, Matteucci e Pasquino (1998), isso se deve a dois fatores. O primeiro diz respeito ao fato de muitos termos, majoritariamente de origem grega, terem passado por transformações ao longo da história, como *democracia*, *aristocracia*, *política* etc. O segundo fator está relacionado à ausência de uma ciência política capaz de precisar de modo universal o conceito dos termos mais frequentes e basilares do campo político.

Os referidos autores também apontam que a linguagem política se alimenta de muitas palavras da língua comum. Inversamente, muitos termos técnicos (tirania, ditadura, democracia), jorjados por teóricos e especialistas do campo, comumente migram para o uso comum, integrando-se à linguagem do cotidiano, sofrendo, portanto, variações e alterações de sentido. Explicam ainda que, não raro, mesmo os termos mais técnicos, também ensejam interpretações diferentes, pois designam fatos históricos complexos e se vinculam a doutrinas muitas vezes controvertidas.

De acordo com Bobbio, Matteucci e Pasquino (1998), a linguagem política não é isenta de ideologia; pelo contrário, consideram que não há termo político ideologicamente neutro. Por isso, ao uso dos termos e fraseologismos no discurso político subjazem posicionamentos e juízos de valor. Com efeito, “O discurso político não é menos ideológico do que no passado, mas é tão habitado pelas ideologias predominantes do momento” (BOLTANSKI, CHIAPELLO, 1999

apud BART, 2010, p. 82)<sup>7</sup>. Um caso ilustrativo seria o uso da palavra *golpe*, no contexto recente da política brasileira, como se vê abaixo, em dois trechos extraídos do *corpus* da pesquisa.

O **golpe** que por ora afasta Dilma Rousseff figura apenas no começo de uma pauta mais complexa e extensa, muito além da confirmação do impeachment. É também do conhecimento do mundo mineral que a mira da casa-grande está alçada na direção de Lula e do PT, a repetir a obsessão de Catão, o Censor, em relação a Cartago, delenda esse, destruída há de ser, repetia incansavelmente (RCMC16M7e). (grifo nosso).

Para Mino Carta, jornalista fundador e chefe de redação da revista Carta Capital, nota-se claramente que o processo de *impeachment* que afastou a ex-presidente Dilma Rousseff configurou, na verdade, um golpe. Posição divergente é assumida pela colunista Ruth de Aquino, da revista *Época*, conforme se vê em:

Na semana que vem, será morto outro garoto, outra mulher, outro policial. Porque está tudo errado. Policiais são afastados, depois voltam. Mães e avós, com bebês e crianças, fazem fila de madrugada, lutando por vaga na creche ou escola. Os sem-teto e desempregados lotam albergues para alimentar a família. Os traficantes dominam áreas carentes num país em que se rouba de tudo, de merenda a remédio, verba de cultura e obras, contracheque, Fundo de Garantia. Bilhões de reais precisam voltar aos cofres públicos para dar paz e dignidade aos brasileiros. Esse é o verdadeiro **golpe** que viola a Constituição e saqueia nossos sonhos (RERA16M7a). (grifo nosso).

Ao considerar como verdadeiro golpe as mazelas e injustiças que assolam o Brasil, a colunista, na verdade, nega implicitamente a ideia de que o *impeachment* da ex-presidente Dilma constituiu um golpe de Estado. Tais exemplos ilustram o caráter ideológico do discurso político.

Ao discutir as características do discurso político, Dorna (1995) denuncia que poucos autores refletem explicitamente sobre o *status* desse gênero de discurso. De acordo com Dorna (1995), há quem considere o discurso político apenas um caso do discurso de influência, fato que o submeteria às regras e princípios de uma teoria da comunicação. Por outro lado, há teóricos que defendem a existência de funções particulares desempenhadas pelo discurso

---

<sup>7</sup> “Le discours politique n’est pas moins idéologique que par le passé, mais il est comme habité par les idéologies dominantes du moment” (BOLTANSKI, CHIAPELLO, 1999, apud BART, 2010, p. 82).

político. As principais características apontadas por Dorna (1995) podem ser assim sintetizadas: a) o discurso político é a “busca da aprovação de uma determinada forma de organização social”<sup>8</sup> (MORRIS, 1946, *apud* DORNA, 1995, p. 132); b) constitui o veículo das ideologias (REBOUL, 1980 *apud* DORNA, 1995); c) figura como instrumento de ação, podendo servir como ferramenta de prescrição e de valorização; d) produz um efeito de aproximação e de adesão; e) visa a mudança ou a manutenção do *statu-quo* da ordem existente e; f) propicia um lugar de interação entre os membros de uma sociedade.

Em se tratando das funções atribuídas ao discurso político, Dorna (1995) destaca: a) *uma função estruturante* – o discurso é considerado o “cimento” do sistema político, isto é, a condição fundamental para a existência da política (ALMOND & POWELL, 1966 *apud* DORNA, 1995); b) *uma função decisional ou decisiva* – uma vez que o principal mecanismo da política é a decisão (DEUTSCH, 1963 *apud* DORNA, 1995), o discurso desempenha um papel central nesse processo, em que o poder sempre necessita persuadir, convencer, produzir e obter informação; c) *uma função pedagógica* – como consequência de outras funções, a função pedagógica deriva justamente do fato de o discurso político ser estruturante, coeso, visando a adesão e facilitando a manutenção ou reprodução do sistema político. Segundo o autor, “o poder político, a política como um todo é feita de uma boa dose de manipulação, mesmo quando é pedagógica”<sup>9</sup> (DORNA, 1995, p. 133) e; d) *uma função terapêutica* – o discurso político tende a fornecer uma “coerência simbólica a toda massa ideacional que constitui a organização do ego” (ANSART, 1976 *apud* DORNA, 1995, 133). Isso significa que esse tipo de discurso inclina-se a agradar as expectativas criadas no âmbito da própria ideologia, evitando-se o ressurgimento de dúvidas, a fim de manter uma racionalização constante para todas as relações vividas. Com efeito, “o discurso político tende a evitar a angústia pela renovação das certezas assumidas” (ANSART, 1976 *apud* DORNA, 1995, 133).

---

<sup>8</sup> “l'enjeu du discours politique est la recherche de l'approbation d'une certaine forme d'organisation sociale” (MORRIS, 1946, *apud* DORNA, 1995, p. 132).

<sup>9</sup> “Le pouvoir politique, la politique tout court est faite d'une bonne dose de manipulation, même quand elle est pédagogique” (DORNA, 1995, p. 133).

Convém, ainda, focar o discurso político relacionando-o à linguagem, à ação e ao poder, visto que esse tipo de discurso visa exatamente a agir sobre o outro, levando-o a aderir posicionamentos. Nesse sentido, revelam-se muito pertinentes as análises e reflexões desenvolvidas por Patrick Charaudeau (2006), o qual apresenta uma síntese dos trabalhos de Max Weber, Hannah Arendt e Jürgen Habermas a respeito da noção de poder político.

De acordo com Charaudeau (2006), Weber concebe o poder político sob a forma de uma dominação legitimada pela violência, já que as relações entre os humanos se baseiam nas relações estabelecidas entre a figura do dominante e a dos dominados, sendo o Estado, no campo político, o agente que impõe sua autoridade pela força e pela violência legitimada sob aparente legalidade. Diferentemente, para a filósofa Hannah Arendt, o poder político estaria fundamentado no consentimento e vontade dos homens de ser e viver em comunidade. “É esse <estar junto> que funda o fato político no qual poder e ação se definem reciprocamente: todo poder é poder de agir conjuntamente” (CHARAUDEAU, 2006, p. 256). Sob esse viés, o poder não estaria subordinado à opressão e à violência, mas à liberdade dos cidadãos.

Jürgen Habermas, por sua vez, teria conseguido reconciliar os dois pontos de vista mencionados, ao distinguir um poder comunicacional e um poder administrativo. No primeiro caso, o detentor do poder é o próprio povo, longe da esfera de dominação do Estado, situado, pois, no espaço público, no qual os cidadãos interagem e debatem suas opiniões pela via argumentativa, gerando assim a opinião pública. Já o poder administrativo seria reservado às relações de dominação. Compõem este campo os dispositivos que organizam a ação social, os mecanismos de regulação por lei, as sanções aplicadas a quem possa violar a vontade de agir. Dessa maneira, “se institui um sistema político que tende a se defender contra toda tentativa de desestabilização e, para fazê-lo, exclui, seleciona, busca ser eficaz e, portanto, impõe” (CHARAUDEAU, 2006, p. 256). Coadunando com o ponto de vista de Habermas, Charaudeau (2006) sintetiza que o poder comunicacional e o poder administrativo articulam-se em torno de um mesmo aspecto, ou seja, a legitimidade, ressaltando-se, porém, que:

poderíamos dizer que o poder comunicacional é o que torna possível a construção de um espaço político ao colocar em cena a questão da legitimidade, e que o poder administrativo, ao se fundar sobre essa

legitimidade, ao tirar partido de uma vontade popular, implementa um dispositivo de realização concreta do poder ao se impor às próprias pessoas que o fundaram (CHARAUDEAU, 2006, p. 257).

Ao filiar-se a essa concepção de poder político, dividido em dois tipos de poder, Charaudeau (2006) defende uma concepção de linguagem política proveniente da interação de dois eixos que determinam dois tipos de atividade social, a saber: (i) *a do dizer político*, no qual se situa o debate de ideias, no campo da opinião pública e; (ii) *a do fazer político*, restrito ao espaço privilegiado para a tomada de decisão e realização de atos. Para o autor, esses dois eixos se legitimam reciprocamente, mas diferenciam-se pelo estabelecimento de relações de força e de um jogo de dominação particular. Deste modo, quanto ao campo do *dizer político*, há domínio da linguagem, que se manifesta através de uma arena discursiva na qual são permitidas diferentes estratégias, como manipulação, proselitismo, ameaça, promessa, tendo como objetivo a imposição de uma opinião. Por outro lado, no campo do *fazer político*, o que predomina é a ação como o espaço no qual atua uma instância política considerada soberana e uma instância cidadã, que frequentemente pode solicitar resultados, acompanhando-os como uma espécie de fiscais das ações realizadas. Neste caso, objetiva-se uma dominação baseada na regulamentação, na sanção e na reivindicação (CHARAUDEAU, 2006).

Em decorrência desses dois campos, o autor distingue, então, dois tipos de atividade discursiva. O primeiro volta-se para as ideias e seu valor de verdade, constituindo o espaço de fabricação das ideologias; é o político. O segundo tipo focaliza os atores e sua força de ação, constituindo o espaço de fabricação das relações de força; é a política. Esses dois tipos de atividade discursiva fundamentam o pressuposto da linguagem-ação no campo político.

Para demonstrar como esse pressuposto se aplica na análise do discurso político, e quais são as estratégias discursivas utilizadas, o autor esclarece primeiramente que o sujeito político pode encontrar-se em duas situações distintas de enunciação: *fora da governança*, quando se busca obter o poder participando-se de um pleito na condição de candidato, e *dentro da governança*, quando o sujeito já participa do sistema, ocupando uma pasta ou exercendo outra atividade dentro do Governo. Disso decorrem diferentes estratégias discursivas (palavra de promessa, palavra de decisão, palavra de justificação,

palavra de dissimulação), uma vez que o sujeito se encontra em posições distintas de legitimidade.

A palavra de promessa ou de advertência deve levar em conta simultaneamente três critérios inter-relacionados: – o discurso deve pretender ser ao mesmo tempo idealista e realista, definindo uma idealidade social e sistemas de valores, mas explicitando, também, os meios para que sejam alcançados os objetivos; – o sujeito que enuncia deve possuir credibilidade ante a instância cidadã, o que o leva a elaborar uma imagem de si de convicção e de suporte de identificação e; – a fim de conseguir adesão ao projeto, o sujeito político busca o convencimento de seu público, apelando, para isso, para a razão ou para a emoção, em diferentes meios de divulgação (na mídia, em comícios, em declarações escritas, panfletos, cartazes, redes sociais etc.

Segundo o autor, *fora da governança*, a palavra de promessa ou de advertência não encontra tantas dificuldades, desde que assuma a força do profetismo. Não se pode dizer o mesmo para a situação de enunciação *dentro da governança*. Neste contexto, a palavra de promessa encontra-se relacionada à ação, sob constante crítica do adversário e do cidadão. Em função disso, surgem três tipos de estratégias discursivas: palavra de decisão, de justificação e de dissimulação.

Quanto à palavra de decisão, o autor a caracteriza como sendo essencialmente uma palavra de ação, que se fundamenta sobre uma posição de legitimidade. No tocante ao campo político, ela indica três coisas: (i) diz que há uma desordem social, julgada inaceitável, que pode ser uma situação, fato ou acontecimento; (ii) anuncia que é preciso tomar uma providência para solucionar essa anormalidade, a fim de reinseri-la em uma nova ordem ou em um novo quadro jurídico e; (iii) revela, no próprio ato de sua enunciação, a medida a ser aplicada. Deste modo, seriam exemplos de atos realizados por uma palavra decisional a decisão de intervir ou não em um conflito, de orientar a política econômica em uma ou outra direção, de editar leis etc.

O autor menciona um exemplo específico, no qual se encontram a evidência de uma desordem social, a necessidade de uma nova ordem e a realização de um conjunto de atos pela própria declaração. Trata-se da declaração no rádio proferida pelo General De Gaulle, em maio de 1968, em seu retorno de Baden Baden: “Nas circunstâncias atuais, eu não renunciarei, nem

substituirei meu Primeiro Ministro (...). Eu dissolvo, hoje mesmo, a Assembleia Nacional” (CHARAUDEAU, 2006, p. 260).

Com respeito à palavra de justificação, o autor explica que, até mesmo as tomadas de decisão ou anúncio de ação feitos por autoridades precisam ser relegitimados em virtude das interrogações ou questionamentos realizados com frequência pelos concorrentes políticos ou pelos movimentos sociais. O discurso de justificação incide, pois, sobre a ação a fim de dar/lembrar sua razão de ser. Não significa enunciar uma confissão nem culpa, “Trata-se de passar uma eventual posição de a uma posição de benfeitor responsável por seus atos” (CHARAUDEAU, 2006, p. 261). Um exemplo prototípico desse tipo de palavra seriam os discursos que predominam nos relatórios proferidos por porta-vozes do Governo ao final das reuniões ministeriais.

Já a palavra de dissimulação, intrínseca ao discurso político, estabelece um jogo de máscaras envolvendo palavra, pensamento e ação. Diferentemente do que se costuma imaginar, o ator político jamais utiliza uma frase fortuitamente. Ele consegue prever “as críticas de seus adversários, os efeitos perversos da informação midiática e os movimentos sociais que deve tentar antecipar e neutralizar” (CHARAUDEAU, 2006, p. 261).

Intimamente relacionada à palavra de dissimulação está a questão da mentira no discurso político. Segundo o autor, é de conhecimento de todo homem político a impossibilidade de dizer tudo, a todo instante, da maneira como pensa ou imagina. A razão disso está no fato de que as palavras não devem impedir a ação, o que leva ao surgimento de várias estratégias.

Dentre os recursos utilizados, há a estratégia do vago, a qual se combina com a palavra de promessa. Por exemplo, diante de uma promessa feita por um candidato à presidência da República, o sujeito pode até demonstrar engajamento pela causa, mas de modo vago, para ganhar tempo ou apostar que a promessa seja esquecida. De acordo com o autor, a estratégia consiste em permanecer no vago sem que isso implique perda da credibilidade. Charadeau (2006) ressalta que o sujeito político não poder falhar nesse ponto.

Outra estratégia é a do silêncio. Diante de questões que poderiam levantar polêmicas e reações as mais diversas e complexas possíveis, como as que violam direitos humanos e ferem as relações diplomáticas entre países, muitas vezes se prefere não dizer ou anunciar nada a respeito. Entende-se,

nesses casos, que o anúncio das ações a serem conduzidas poderia incentivar numerosas manifestações violentas que poriam em risco a realização da própria ação.

Há, também, a estratégia de denegação. O homem político, envolvido em graves processos jurídicos, nega sua participação ou a de uma outra pessoa (seu cúmplice) no crime sobre o qual está sendo acusado. Charaudeau (2006) explica que, na hipótese de ter realmente ocorrido envolvimento no processo, negar é o mesmo que mentir, seja negando a realidade dos fatos, seja dando um falso testemunho. Segundo o autor, o mais importante é que não seja possível provar a implicação das pessoas nos casos. A política brasileira exhibe fartos exemplos desse tipo de estratégia, como os fatos mais recentes relacionados à prisão de Eduardo Cunha, ao *impeachment* da ex-presidente Dilma Rousseff, algumas delações premiadas etc.

Para Charaudeau (2006), há uma versão mais nobre desse tipo de estratégia. Trata-se do *blefe*: “deixar entender que se sabe alguma coisa quando não se sabe nada e assumir o risco de ser levado a provar” (CHARAUDEAU, 2006, p. 263). O autor cita como exemplo a situação ocorrida em 1974, durante o debate presidencial entre Giscard e Mitterrand. Na ocasião, Giscard ameaçou retirar de sua pasta as provas das alegações que fazia ao seu adversário político, mas só havia páginas vazias.

Por fim, o autor trata da estratégia da razão suprema. Não se emite nenhuma declaração, usa-se um discurso falso ou se faz acreditar em nome de uma <razão de estado>. Justifica-se a mentira pública, nesta situação, porque objetiva-se manter um bem supremo ou uma base identitária do povo, apesar da opinião pública ou os cidadãos individualmente não concordarem com isso. Com isso, tem-se a impressão de que não existem realmente mentiras, uma vez que “se trata de um discurso que não remete a uma vontade de enganar o outro, mas, ao contrário, de um desejo de torná-lo cúmplice de um imaginário que todo mundo teria interesse em conservar” (CHARAUDEAU, 2006, p. 263).

Do conjunto dessas estratégias, o autor acredita que apenas a de denegação seja claramente condenável, porque ela incide no laço de confiança estabelecido entre o cidadão e seus representantes. Os demais casos podem ser discutíveis, como o foram ao longo da história. Para Maquiavel, por exemplo, citado por Charaudeau (2006), o príncipe deve ser um “grande simulador e

dissimulador”. Assim também, algumas questões não precisam chegar ao conhecimento do povo, que “sente mais do que pensa” (TOCQUEVILLE *apud* CHARAUDEAU, 2006, p. 264). Conclui, então, que:

Poderíamos mesmo dizer com algum cinismo que o homem político não tem de dizer a verdade, mas parecer dizer a verdade: o discurso político se interpõe entre a instância política e a instância cidadã (sic.) criando entre os dois um jogo de espelhos (CHARAUDEAU, 2006, p. 264).

Não se pode negar a influência sofista nessa maneira de perceber, “com algum cinismo”, a relação do homem político com a linguagem e a verdade. Como é sábio, para os sofistas, fazer parecer verdadeiro importava mais que a própria verdade do enunciado. Ao levar em conta esse princípio, Charaudeau (2008) considera o discurso político como um lugar de “verdade presa”, no qual o que importa não é somente a verdade da declaração lançada publicamente, mas sua força de veracidade. Assim, muito mais que um objetivo de demonstração, o discurso político, sendo sempre um discurso dirigido, pretende incitar a pensar e a agir, sendo os próprios receptores desse discurso, os cidadãos, também integrantes desse processo, na condição de opinião pública.

Charaudeau (2006) conclui, enfim, que, no discurso político, ocorre uma fusão da verdade das aparências, encenada pelo discurso, e uma verdade das ações, decorrente das decisões. Funda-se assim “um <verossímil>, sem o qual não haveria ação possível no espaço público. Talvez esteja aí um dos fundamentos da palavra política” (CHARAUDEAU, 2006, p. 268).

Portanto, a política, como campo de atuação que põe em cena a linguagem, a ação e o poder, constitui um complexo fenômeno da sociedade humana, cobrindo praticamente todos os aspectos da vida social, influenciando e sendo influenciada por outros campos e sistemas existentes. Entretanto, o que lhe dá sustentação material e sócio-histórica, permitindo visualizar suas contradições e ideologias, é o discurso, visto alhures como “cimento”, estruturante de todo o sistema. Daí a importância do estudo do discurso político como forma de compreender os muitos aspectos em que a política se apresenta, como o linguístico e, em especial, sob o ponto de vista fraseológico, tal como se propõe nesta dissertação.

## CAPÍTULO 3 – FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, são apresentados e discutidos os conceitos de base que fundamentam teoricamente o trabalho. Aborda-se também, brevemente, o problema terminológico que tem caracterizado a área da Fraseologia, evidenciando-se as convergências consideradas relevantes para o estabelecimento e refinamento do campo de estudo em questão.

### 3.1 FRASEOLOGIA: QUESTÕES GERAIS

A observação de Saussure (2006 [1916]) de que, em matéria de linguagem, o ponto de vista cria o objeto, nunca foi apenas uma frase de efeito. De fato, o mestre genebrino, nos primórdios da linguística, já vislumbrava uma característica marcante dos estudos da linguagem, a saber: os múltiplos pontos de vista (fonético, fonológico, morfológico, morfossintático, sintático, semântico, pragmático, lexical, discursivo, sociolinguístico, entre outros) a que está sujeita a língua enquanto objeto constantemente modificado pelas diversas lentes teóricas do analista.

Essa tônica de Saussure parece reverberar com mais intensidade no campo de estudos comumente denominado de Fraseologia, haja vista a profusão teórica e terminológica que caracteriza a área. Esse fato tem despertado diferentes posicionamentos a respeito da delimitação e do alcance da área, como se verá mais adiante.

De modo geral, a bibliografia sobre fraseologia exhibe uma falta de consenso dos estudiosos acerca da terminologia a ser adotada, da delimitação do campo de estudo, bem como dos critérios ou propriedades pelos quais se deve identificar uma unidade fraseológica. Por esse motivo, Corpas Pastor (1996), Oliveira Silva (2011) e Tagnin, (2011), entre outros, pontuam que o termo *fraseologia* não se isenta de controvérsias.

Em decorrência disso, Monteiro-Plantin (2014, p. 23) salienta que:

os limites da Fraseologia são difíceis de estabelecer, principalmente por não haver consenso entre os linguistas quanto à delimitação das unidades que podem ser objeto de investigação, e tampouco em relação à categorização dessas unidades.

Alinha-se a esse posicionamento a constatação de Carballo (1997-1998), que caracteriza a Fraseologia como um intricado universo, em função das múltiplas abordagens e da diversidade denominativa que a área apresenta. Carballo (1997-1998, p. 67-68) acredita que tal cenário poderia, inclusive, impossibilitar o próprio campo de investigação de cumprir, por muitas vezes, alguns objetivos fundamentais para a disciplina, como: “Delimitar suas fronteiras, esclarecer os tipos de unidades que a integram, e dar definições precisas livres de ambiguidades”. Deve-se ressaltar que a referida publicação da autora está situada no final da década de 1990.

Por sua vez, Ortiz Alvarez (2000, p. 82) defende que “esta heterogeneidade existente reflete as vias de busca, os diversos critérios sobre o objeto da fraseologia e as tarefas e funções inerentes a essa disciplina”. Acredita-se, de um lado, que essa diversidade pode ser interpretada como prejudicial à definição do referido campo como disciplina científica. Por outro lado, pode ser indicativa da natureza complexa do objeto, que se situa num cruzamento (ou na interface) entre os níveis lexical, sintático, semântico e pragmático.

### 3.2 A FRASEOLOGIA E SEUS DESDOBRAMENTOS

De início, convém destacar o caráter polissêmico do termo *fraseologia* (cf. ORTIZ ALVAREZ, 2000; BARBOSA, 2012; MONTEIRO-PLANTIN, 2014). Há, de modo geral, pelo menos três acepções conhecidas para o referido termo<sup>10</sup>: a) unidade polilexical caracterizada, dentre outros critérios, por diferentes graus de fixidez sintática, restrições semânticas e frequência de uso; 2) conjunto formado por essas unidades pertencentes ao uso geral de uma língua ou a de um domínio especializado (fraseologia do português, fraseologia do futebol, fraseologia do meio ambiente); 3) campo de estudo ou disciplina que tem tais unidades como principal objeto de investigação. Para cada uma dessas acepções, é possível

---

<sup>10</sup> Costuma-se grafar a inicial maiúscula para referir à disciplina *Fraseologia*, distinguindo-se da grafia em letra inicial minúscula, relativa à *fraseologia* como unidade polilexical ou conjunto de sequências polilexicais características de uma língua ou de uma especialidade.

encontrar divergências denominativas e conceituais que resultam, em geral, de diferentes orientações teóricas e metodológicas.

Com respeito à primeira acepção, Charles Bally (1951), considerado o “pai” da Fraseologia, ao tratar sobre a terceira possibilidade de se delimitar os fatos da expressão<sup>11</sup>, inclui sob o termo genérico *phraséologie* (fraseologia) as *locuções compostas*. Para ele, essas locuções constituem-se unidades lexicológicas e se caracterizam por serem formadas por várias palavras que, na combinação, perdem parte de seu significado usual para adquirirem um outro significado para a unidade resultante. O autor exemplifica o fenômeno com a expressão adverbial *tout de suite* (imediatamente), em que não se manteve o significado de nenhum dos componentes da expressão.

Maurice Gross (1982), ao se referir a essas unidades polilexicais, utiliza os termos *phrase, formes* ou *expressions figées* (expressões fixas ou cristalizadas), os qualificadores *proverbiales, idiomatiques, composées, méthapores* e *clichés*. Segundo o autor, embora constituam fenômeno de envergadura, as expressões fixas foram por muito tempo ignoradas pelas teorias predominantes da linguística, pela falta de estudos sistemáticos na área.

Por sua vez, Alberto Zuluaga (1980, p. 16) utiliza o termo unidade fraseológica para denominar “toda combinação fixa de palavras”<sup>12</sup>. Sob tal designação, o autor latino-americano reúne combinações linguísticas formadas por duas ou mais palavras em que haja fixidez, ilustrando com saudações do tipo: *boa noite, feliz Páscoa* etc.

Já Ortiz Alvarez (2000, p. 73) define a fraseologia como a “combinação de elementos linguísticos de uma determinada língua, relacionados semântica e sintaticamente, que não pertencem a uma categoria gramatical específica e cujo significado é dado pelo conjunto de seus elementos”. Para a autora, essas combinatórias são unidades semânticas que possuem traços categoriais próprios, distinguindo-se das palavras e das combinações livres, podendo-se, então, falar de “um sistema fraseológico da língua” (ORTIZ ALVAREZ, 2000, p. 75).

---

<sup>11</sup> Para o autor, a terceira possibilidade de se delimitar os fatos da expressão ocorre quando “l'unité psychologique excède les limites de l'unité graphique et s'étend stir plusieurs mots” (BALLY, 1951, p. 5).

<sup>12</sup> “[...] toda combinación fija de palabras” (ZULUAGA, 1980, p. 16).

Monteiro-Plantin (2014) trata de unidades fraseológicas, entendendo-as como sequências linguísticas que integram o objeto de estudo da Fraseologia. Sob esta denominação, a autora inclui as sentenças proverbiais, expressões idiomáticas (EI), pragmatemas e fórmulas situacionais, colocações, locuções fixas, frases feitas, clichês e chavões.

Ao definir as unidades fraseológicas, Corpas Pastor (1996) as considera unidades lexicais que constituem grupos de palavras, formadas por no mínimo duas palavras gráficas, podendo chegar ao nível das orações compostas<sup>13</sup>.

Nesse contexto, Ortiz Alvarez (2012a) afirma haver muitos hiperônimos para fazer referência à fraseologia, tais como: *unidade fraseológica*, *expressão pluriverbal*, *unidade pluriverbal lexicalizada*, *expressão fixa*, *fraseolexema*, *frasema*, *fraseologismo*, *combinatória lexical*. Porém, segundo a autora, os termos mais recorrentes são *unidade fraseológica* e *fraseologismo*. O primeiro, inclusive, é o termo mais utilizado na Espanha, como assinala Pamies (2014). Essa profusão denominativa caracteriza, de acordo com Biderman (2005), a primeira questão, de natureza terminológica, que se apresenta no exame da problemática teórica das unidades complexas do léxico.

Ao criticar a profusão terminológica, a fim de se evitarem confusões teóricas, Salah Mejri (2012) propõe uma ordenação do campo terminológico da área, a partir da oposição entre o fenômeno fraseológico e o processo por meio do qual ele se exprime. Define, assim, a “*phraséologie*” (fraseologia) como **fenômeno linguístico**, comum a todas as línguas vivas, que se manifesta por meio das associações sintagmáticas recorrentes. Atua nesse fenômeno o **processo** de “*figement*” (cristalização, congelamento), do qual resultam os fraseologismos, que apresentam diferentes graus de fixidez, polilexicalidade, congruência e idiomatidade. Citam-se, como exemplos, “bater as botas”, “sair de fininho”, “quebrar um galho”<sup>14</sup>, entre outros, cujos sentidos não provêm da mera somatória do significado individual dos componentes da estrutura complexa, mas trata-se de um sentido da unidade como um todo, em geral, de caráter idiomático, mas não apenas.

---

<sup>13</sup> “[...] las unidades fraseológicas (UFS) – objeto de estudio de la fraseología – son unidades léxicas formadas por más de dos palabras gráficas en su límite inferior, cuyo límite superior se sitúa en el nivel de la oración compuesta” (CORPAS PASTOR, 1996, p. 20).

<sup>14</sup> Respectivamente: *morrer*, *fugir discretamente de uma situação* e *dar uma ajuda a alguém em momento de necessidade*.

Com relação às propriedades dos fraseologismos, alguns autores comumente mencionam a idiomaticidade, a complexidade da estrutura morfológica, a fixidez, a frequência de uso, entre outras. Salah Mejri (2012), por exemplo, adota a noção de *continuum*, isto é, de escalaridade, para determinadas propriedades do fraseologismo, como a fixidez e a congruência.

Assim, conforme Mejri (2012), a fixidez constitui um parâmetro para explicar o fenômeno fraseológico e descrever o mecanismo de cristalização por meio do qual as solidariedades sintagmáticas se apropriam das regras da combinatória sintagmática, no plano sintático e no semântico. No primeiro caso, o autor explica que não é possível, por exemplo, no fraseologismo “perder a mão”, modificar o determinante ou acrescentar um adjetivo à palavra “mão”. Fazendo essas alterações, a unidade se desfaria, produzindo estruturas inexistentes como: \*perder uma mão, \*perder as mãos, \*perder uma mão pequena.

Já a fixidez no plano semântico manifesta-se por meio da propriedade em que determinadas sequências são fixadas de uma só vez, com seu respectivo sentido, como “engolir água” e “ter um gato na garganta”. Para o autor, essa fixidez também é de natureza paradigmática. Exemplifica-se essa noção a partir de uma expressão idiomática conhecida no Brasil: “bater as botas”. Não dá para comutar, por exemplo, “botas” com “sapatos” ou “calçados”, sob pena de criar uma forma que não existe como fraseologismo no uso da língua portuguesa.

Outro critério apresentado por Mejri (2012) diz respeito à noção de congruência. Para ele, trata-se de um “processo de adaptação das unidades lexicais pelo qual elas se integram naturalmente na combinatória” (MEJRI, 2009, 79 apud MEJRI, 2012, p. 143). O autor esclarece que o referido processo atua tanto no nível morfológico e sintático quanto semântico. Em termos de fraseologismos, a congruência seria a adequação da estrutura sintagmática às regras de formação das sequências fixas. Inversamente, seria incongruente toda sequência de palavras que violasse essas regras. O autor cita como exemplo de incongruência sintática os casos \*Ele pegou um/os touros por um/os (só, dois...) chifre, quando a formação sintagmática e o uso fixam apenas a forma “Ele pegou o touro pelos chifres”.

Em relação à segunda acepção mencionada, entende-se que a fraseologia é um fenômeno linguístico, de caráter marcadamente lexical. Sendo

assim, pode originar-se tanto do uso geral da língua ou das línguas especializadas (cf. ARAGÃO, 2016; KRIEGER & FINATTO, 2004).

De acordo com Aragão (2016), a depender do enfoque dado, pode-se falar de fraseologia da língua geral e de fraseologia especializada, as quais, por sua vez, se distinguem da paremiologia, cujo objeto de estudo são especificamente os provérbios. A fraseologia especializada difere-se da geral pelo seu conteúdo temático específico a uma área ou domínio especializado. Neste sentido, a fraseologia aproxima-se da unidade terminológica por compartilhar da função de representar e transmitir conhecimento especializado. Bevilacqua (2004) propõe ainda outra característica distintiva para a fraseologia especializada: entre seus componentes, deve haver necessariamente um ou mais termos que representam o núcleo terminológico da estrutura, como captar *energia*, consumo de *energia*, próprios do âmbito da energia solar.

Para Ortiz Alvarez (2012), a fraseologia ou terminologia especializada, é recolhida de textos e de bancos de dados especializados, orais ou escritos, constando em dicionários de especialidade. Já a fraseologia comum, ou fraseologia popular, abrange as demais unidades fraseológicas de uso geral, tais como as expressões idiomáticas, locuções, colocações etc. (ORTIZ ALVAREZ, 2012).

Krieger & Finatto (2004) corroboram essa distinção, destacando o papel do contexto comunicacional para que se possa tratar de fraseologias da língua geral ou de fraseologias especializadas. Esta última, inclusive, tornou-se um dos objetos de estudo da Terminologia, ao lado do termo e da definição (KRIEGER & FINATTO, 2004).

Em sua terceira acepção, referente ao campo do conhecimento designado pelo termo Fraseologia, há pelo menos dois posicionamentos que divergem quanto ao *status* de autonomia científica da área. Para alguns autores, como Bally (1951), Corpas Pastor (1996), Xatara & Parreira (2011), Barbosa (2012) e Aragão (2016), a Fraseologia seria um ramo da linguística, figurando como uma subdisciplina da lexicologia, integrando parte das ciências que se ocupam do léxico.

De acordo com Dobrovolskij (2012), a vinculação da Fraseologia à linguística remete ao trabalho de Charles Bally (1909), *Traité de stilistique française*, no qual é utilizado pela primeira vez o termo *phraséologie* com o

sentido especificado alhures. No entanto, conforme pontua o autor, as ideias do linguista suíço não foram aceitas e desenvolvidas de imediato pelos estudiosos da época. Seus estudos e princípios norteadores alcançaram êxito e reconhecimento apenas na década de 1940, com os trabalhos dos linguistas russos, em especial os de Victor Vladimirovic Vinogradov. Por esse motivo, Ortiz Alvarez (2000) afirma que o ano de 1940<sup>15</sup> representa uma data histórica para o estabelecimento da Fraseologia como disciplina linguística.

Por outro lado, autores como Polivánov (1931 *apud* ORTIZ ALVAREZ, 2000), Tristá (1988) e Monteiro-Plantin (2014), advogam para a Fraseologia uma condição de autonomia científica, independente da Lexicologia, com método e objeto de estudo específicos. Nesse sentido, vale a pena refletir a respeito da observação de Mejri (2008, p. 01), segundo o qual pelo menos duas condições são exigidas para que se defina epistemologicamente uma ciência ou disciplina, a saber: “a delimitação de um objeto de estudo específico e a construção metodológica”<sup>16</sup>. Com efeito, será que o campo fraseológico já conseguiu responder a tais demandas epistemológicas?

Apesar das contradições e problemas terminológicos ainda existentes no campo dos estudos fraseológicos, não se pode negar o crescimento significativo que essa área vem experimentando nos últimos anos. A própria Corpas Pastor, umas das maiores pesquisadoras atuais da Fraseologia, mudando seu ponto de vista em relação ao que defendeu em 1996, no seu *Manual de Fraseología Española*, assim se expressa em recente entrevista:

A Fraseologia é considerada tradicionalmente um ramo da Linguística, concretamente uma subdisciplina dentro da Lexicologia. Também se estudou a Fraseologia a partir da Etnolinguística, especialmente no que se refere à Paremiologia (os provérbios são considerados um repositório de sabedoria popular, cf. o folclore popular). Mas desde finais da década de 1990 e muito especialmente desde o início do século XXI, a Fraseologia experimentou um auge a tal ponto que já pode se considerar uma disciplina autônoma e independente, que desenvolveu um aparato teórico próprio e interdisciplinar (CORPAS PASTOR, 2017, p. 262).

---

<sup>15</sup> Já Corpas Pastor, na apresentação do seu *Manual de Fraseología Española* (1996), considera a década de 1950 como o marco temporal do surgimento dessa disciplina científica, destacando, nesse contexto, a obra pioneira de Julio Casares, no cenário hispânico.

<sup>16</sup> “Epistémologiquement, une science ou une discipline se définit par au moins deux exigences : la délimitation d’un objet d’étude spécifique et la construction méthodologique” (MEJRI, 2008, p. 01).

Para cada uma dessas acepções, é possível encontrar divergências teóricas e denominativas que talvez sejam melhor compreendidas, colocando-se os estudos fraseológicos em perspectiva histórica, como proposto na seção a seguir.

### 3.3 BREVE HISTÓRICO DOS ESTUDOS FRASEOLÓGICOS

Embora o campo da Fraseologia venha apresentando um considerável desenvolvimento nos últimos anos, o fenômeno sobre o qual se investiga foi negligenciado por muitos anos no âmbito da linguística, tendo sido considerado como objeto marginal e de menor importância na agenda científica do século XX, principalmente. Segundo Corpas Pastor (1996), o estabelecimento da Fraseologia como disciplina científica só veio a ocorrer na década de 1950, no contexto da antiga União das Repúblicas Socialistas Soviéticas (doravante URSS), a partir dos trabalhos do linguista russo V. V. Vinogradov.

Entretanto, antes mesmo de ganhar *status* de objeto científico no âmbito de uma disciplina específica, o fenômeno fraseológico já despertava o interesse de muitos estudiosos da linguagem, que passaram, então, a ser considerados seus precursores: Ferdinand de Saussure, Charles Bally, Emile Benveniste, Bernard Pottier e os linguistas russos Polivanov e V. V. Vinogradov. Cada um destes autores, embora assumindo pontos de vista diferentes e termos distintos, podem ser chamados de pioneiros ao tratar do fenômeno da cristalização lexical e das combinatórias sintagmáticas.

Ao abordar as relações sintagmáticas e paradigmáticas, Saussure (2006) reconhece o mecanismo geral de associação dos signos, do qual resulta o que ele chama de sintagma, isto é, unidade formada por dois ou mais elementos consecutivos. Para o autor, o termo sintagma abrange tanto palavras quanto grupos de palavras e, portanto, as combinatórias complexas do léxico. Ao levar em consideração a dicotomia *língua x fala*, o mestre genebrino enfatiza que pertence à língua um número significativo de expressões em que o uso inibe qualquer modificação; trata-se das *frases feitas*.

Discípulo de Saussure, Charles Bally é reconhecido na literatura da área como o principal responsável pelo início de uma série de estudos a respeito da fraseologia, tratando-a como um fenômeno lexicológico no seu famoso *Traité de*

*stilistique française*, publicado originalmente em 1909<sup>17</sup>. Bally reconhece que, no estudo da delimitação dos fatos da expressão, a fraseologia constitui a terceira possibilidade de demarcar a unidade lexicológica, pois insere-se exatamente no caso em que “l'unité psychologique excède les limites de l'unité graphique et s'étend sur plusieurs mots” (BALLY, 1951, p. 5). Propõe, então, índices externos e internos para seu reconhecimento, baseando-se na relação da forma linguística dessas unidades e no pensamento, bem como na maneira com a qual o falante compreende o uso das combinatórias (cf. BALLY, 1909 [1951]).

Apesar de seu estudo pioneiro, as ideias do linguista suíço Charles Bally não foram reconhecidas de imediato à época em que foram publicadas. Foram, então, os russos que, ao resgatarem os textos do pai da Estilística, deram o passo fundamental para instituir a Fraseologia como uma disciplina científica na metade do século XX, no rol dos estudos linguísticos que se vinham realizando no contexto da então URSS. Corroboram essa constatação histórica inúmeros pesquisadores da área, dentre os quais Corpas Pastor (1996), Ortiz Alvarez (2000) e Dobrovolskij (2012).

Dentre os fraseólogos russos, destacam-se Vinogradov e Polivanov. O primeiro, representante da escola soviética, estudou a fraseologia com base em três aspectos fundamentais: “a) as propriedades internas da unidade fraseológica; b) o papel da unidade fraseológica dentro do contexto de uso; c) as relações manifestadas por essas unidades com os outros subsistemas, principalmente com o sistema lexical e sintático” (ORTIZ ALVAREZ & UNTERNBÄUMEN, 2011). Polivanov, por sua vez, estava interessado em estudar as expressões fixas e seus significados individuais.

Assim, esses primeiros trabalhos no campo da Fraseologia contribuíram para delimitar seu objeto de estudo, lançando as bases teóricas que posteriormente seriam aperfeiçoadas por outros estudos. Depois de surgir oficialmente na URSS, a Fraseologia passou a outros países e continentes, desenvolvendo-se especialmente na Alemanha, Espanha, França, Estados Unidos até chegar aos países latino-americanos, como Cuba e Brasil. Testemunham em favor desse crescimento vertiginoso da área o surgimento de associações, projetos e congressos importantes, como a EUROPHAS

---

<sup>17</sup> Segundo Ortiz Alvarez & Unternbäumen (2011), o termo fraseologia foi utilizado por Bally pela primeira vez no Ocidente em 1909.

(European Society of Phraseology), em 1999, a EURALEX (European Association for Lexicography), em 1983, a Associação Brasileira de Fraseologia, fundada em 2011, em Brasília, durante o I Congresso brasileiro de fraseologia, ocorrido naquele ano, e o Projeto Variação Lexical, Teorias, Recursos e Aplicações – VALEXTRA (CAPES/COFECUB nº 838/150), projeto celebrado entre a universidade francesa Paris 13 e a Universidade Federal da Bahia (UFBA) que visa à produção de um dicionário de fraseologismos brasileiros presentes nos dados do projeto Atlas Linguístico do Brasil – ALiB.

### 3.4 ESTUDOS FRASEOLÓGICOS NO BRASIL: BREVE PANORAMA

No Brasil, os primeiros trabalhos fraseológicos remontam ao início do século XX, com a compilação de adágios, ditos e provérbios populares, sendo pioneiro o estudo de João Ribeiro, *Frases feitas: estudo conjectural de locuções, ditados e provérbios*, publicado em 1908. Contudo, ainda não figuravam como estudos sistemáticos, visto que geralmente não eram baseados em metodologia rigorosa de coleta de dados e análise. Mas houve trabalhos desse período que se destacaram pelo pioneirismo na descrição das unidades repertoriadas. Nesse contexto, surgem, por exemplo, os trabalhos do pesquisador e folclorista Luís da Câmara Cascudo, do linguista e dialetólogo Antenor Nascentes, dentre outros.

Todavia, de acordo com Ortiz Alvarez, a década de 1990 representou uma época marcante para os estudos em fraseologia no Brasil. Para a autora, surge, nesse período, uma produção volumosa e sistemática de estudos descritivos sobre o fenômeno fraseológico, sendo em sua maioria caracterizada por pesquisas contrastivas entre o português e outras línguas como o inglês, espanhol, francês, alemão etc.

O levantamento bibliográfico exposto a seguir baseia-se principalmente em três trabalhos que também objetivam elencar as pesquisas fraseológicas realizadas no país. Trata-se do texto *Estudos Fraseológicos no Brasil: estado da arte*, de Ortiz Alvarez, publicado em 2012, do artigo *Les études phraséologiques dans le contexte brésilien : un panorama*, de Cláudia Xatara & Ortiz Alvarez (2012), e do capítulo *Estudos Fraseológicos no Brasil – Santo de Casa também Faz Milagre*, que integra o livro-manual *Fraseologia: era uma vez um Patinho*

*Feito no Ensino de Língua Materna* (MONTEIRO-PLANTIN, 2014). Os trabalhos inventariados estão organizados cronologicamente, conforme se vê abaixo.

- 1908. **João Ribeiro** – *Frases feitas: estudo conjectural de locuções, ditados e provérbios.*
- 1961. **João Perez.** *Provérbios brasileiros.*
- [s.d] **Cid Franco.** *Dicionário de expressões populares brasileiras.*
- 1974. **Ático Vilas Boas da Mota.** *Provérbios em Goiás.*
- 1977. **Luiz da Câmara Cascudo.** *Locuções tradicionais do Brasil.* / **Raimundo Magalhães Júnior.** *Dicionário de provérbios, locuções e ditos curiosos.*
- 1978. **Maria Tereza Camargo Biderman.** Publicação do livro *Teoria Linguística: linguística quantitativa e computacional.*
- 1979. **Anne-Marie Loffler-Lauriant; Lúcia Pinheiro Lobato; Marc Tukia.** Publicação do artigo *Pour une étude contrastive des lexies complexes: cas particulier des lexies à chiffres en français, portugais et finnois.* *Cahiers de Lexicologie*, v. 34, n. 1, p. 61-86.
- 1981. **Márcio Pugliesi.** *Dicionário de expressões idiomáticas.*
- 1982. **Oswaldo Serpa.** *Dicionário de expressões idiomáticas inglês-português, português-inglês.* / **Leonardo Mota.** *Adagiário Brasileiro.*
- 1985. **Martha Steinberg.** *1001 provérbios em contraste.*
- 1986. **Antenor Nascentes.** *Tesouro da Fraseologia brasileira.*
- 1989. **Sidney Camargo & Martha Steinberg.** *Dicionário de expressões idiomáticas metafóricas inglês-português.* / **Stella Ester Ortweiller Tagnin.** Publicação do livro *Expressões idiomáticas e convencionais.* / **Flamínia Manzano Moreira Lodovici.** Defesa da dissertação intitulada *Elementos constitutivos dos idiomatismos no português do Brasil.*
- 1994. **Cláudia Maria Xatara.** Defesa da dissertação *As expressões idiomáticas de matriz comparativa.*
- 1996. **Eliane Roncolato.** Dissertação *Estudo contrastivo das expressões idiomáticas do português e do espanhol.*
- 1998. **Cláudia Maria Xatara.** Tese de doutorado *A tradução para o português de expressões idiomáticas em francês.*

- 1999. **Marcos Vinícius Fornicola**. *Expressões idiomáticas da língua francesa e respectivas formas equivalentes em língua portuguesa: tratamento léxico-semântico*.
- 2000. **Alexandra Caramori**. *É o bicho: é bestiale*, dicionário de expressões idiomáticas no domínio dos animais com equivalências em italiano e respectivas listas temáticas. / **Maria Luisa Ortiz Alvarez**. Tese de doutorado *Expressões idiomáticas do português do Brasil e do espanhol de Cuba: estudo contrastivo e implicações para o ensino do português como língua estrangeira*.
- 2001. **Eliane Roncolato**. Tese de doutorado *Expressões idiomáticas do português do Brasil e do espanhol da Colômbia: análise, classificação e equivalências*. / **Oto Araújo Vale**. *Expressões cristalizadas do português do Brasil: uma proposta de tipologia*.
- 2002. **Maria Luisa Ortiz Alvarez**. Artigo intitulado *O mundo da fraseografia*. / **Paula Christina Falcão Pastore**. Dissertação *A tradução para o português de expressões idiomáticas em inglês como nomes de animais*.
- 2004. **Susana Quinteros de Creus**. Organização de número temático dedicado à Fraseologia, da revista *Letras de Hoje*, v. 39, n. 1, com a participação de pesquisadores brasileiros e estrangeiros. / **Tatiana Helena Rios**. Dissertação *Idiomatismos com nomes de partes do corpo humano*. / **Cláudia Maria Xatara**. Publicação do capítulo *As unidades fraseológicas e terminológicas em dicionários bilíngues gerais*. In: ISQUERDO, Aparecida Negri; KRIGER, Maria da Graça. (Orgs.). *As ciências do léxico*.
- 2005. **Stella Ester Ortweiller Tagnin**. Reedição do livro *O jeito que a gente diz – expressões idiomáticas e convencionais, incluindo-se as colocações e um capítulo sobre Linguística de Corpus*.
- 2007. **Maria Luisa Ortiz Alvarez**. Artigos *As expressões idiomáticas nas aulas de ele: um bicho de sete cabeças?*, *Tabus linguísticos e expressões cristalizadas* / **Flamínia Manzano Moreira Lodovici**. Tese de doutorado *O idiomatismo como lugar de reflexão sobre o funcionamento da língua*. / **Maria Eugênia Olímpio de Oliveira Silva**. Publicação do livro *Fraseografía Teórica y Práctica*.
- 2008. **Maria Luisa Ortiz Alvarez**. Artigos *Cada macaco no seu galho/cada oveja con su pareja: a metáfora animal nas expressões idiomáticas*, *Cada macaco no seu galho: um estudo dos padrões de uso dos provérbios brasileiros e seus equivalentes em espanhol e russo*, *A fraseologia do meio ambiente: análise conceitual e terminológica*. / **Beatriz Fanciane Camacho**. *Estudo comparativo de expressões idiomáticas do português do Brasil e de Portugal e do francês da França e do Canadá*. / **Simone Rosa Nunes Reis**. *Uma comparação do tratamento de expressões idiomáticas em quatro dicionários bilíngues francês/português*

- e português/francês. / **Cláudia Maria Xatara**. Artigo *Expressions idiomatiques en interface bilingue électronique*. / **Cláudia Maria Xatara & Wanda Lacerda Oliveira**. Publicação do dicionário de provérbios, idiomatismos e palavrões em uso francês-português/português-francês. / **Cláudia Maria Xatara & Wanda Lacerda Oliveira**. Publicação do dicionário de provérbios, idiomatismos e palavrões em uso francês-português/português-francês (2 ed.). / **Cláudia Maria Xatara & Thais Succi Marini**. Artigo *Revisitando o conceito de provérbio*. / **René Gottlieb Strehler**. Artigo *Considerações sobre a fraseologia partindo de Portugiesische Redewendungen*.
- 2009. **Maria Luísa Ortiz Alvarez**. Artigo *A motivação metafórica nas expressões idiomáticas como parte do patrimônio cultural dos povos latino-americanos*. **Huélinton Cassiano Riva**. Tese de doutorado *Dicionário onomasiológico de expressões idiomáticas*. / **Paula Christina Falcão Pastore**. Tese de doutorado *A simbologia dos animais em expressões idiomáticas inglês-português: uma proposta lexicográfica*. / **René Gottlieb Strehler**. Artigo *Fraseologismos e cultura*. / **Maria Eugênia Olímpio de Oliveira Silva**. Artigo *Discurso metalexigráfico sobre los diccionarios bilingües: la equivalencia y la equivalencia fraseológica*.
  - 2010. **Maria Luisa Ortiz Alvarez & Percilia Santos**. *Aspectos culturais relevantes no ensino de português para falantes de espanhol: as expressões idiomáticas e a carga cultural compartilhada*. / **Percilia Santos & Maria Luisa Ortiz Alvarez** (orgs.). Organização da coletânea *Língua e cultura no contexto de português língua estrangeira*. / **Cleci Regina Bevilacqua**. Artigo *Caracterización de los corpus textuales y la identificación de Combinatorias Léxicas Especializadas en la lengua portuguesa y sus equivalentes en lengua española*. / **Cleci Regina Bevilacqua** et al. Artigo *Grupo Termisul: do projeto Acervo ao estudo de combinatórias léxicas*. / **Elizabete Aparecida Marques**. Artigo *Metáforas orientacionais: um exame cognitivocontrastivo de fraseologismos somáticos do português e do espanhol e suas implicações para o ensino da língua portuguesa a hispano-falantes*. / **Cláudia Zavaglia**. Artigo *Quem tem boca vai a Roma : as cores dos provérbios*.
  - 2011. **Maria Luisa Ortiz Alvarez & Enrique Huelva Unternbäumen**. Publicação da coletânea *Uma (re)visão da teoria e da pesquisa fraseológicas*. / **Elizabete Aparecida Marques**. Artigo *La mano como base metonímica de locuciones verbales y adverbiales: un estudio comparativo entre el español y el portugués*. / **Cláudia Maria Xatara**. Artigos *La traduction phraséologique, Les avantages de la perspective onomasiologique pour la phraséographie*. / **Maria Eugênia Olímpio de Oliveira Silva**. Artigo *Enfoque onomasiológico y fraseografía: cuestiones teórico-prácticas*. / **Rosemeire Selma Monteiro-Plantin**. Artigo *La phraséologie au Brésil : un peu de ce qu'on y fait*.
  - 2012. **Suzete Silva Nascimento**. Publicação da coletânea *Fraseologia & CIA*. / Lançamento da revista *Frasema*, o primeiro periódico científico

brasileiro dedicado à Fraseologia, pela Universidade Federal do Ceará, sob a direção da professora Rosemeire Selma Monteiro-Plantin. / **Maria Luisa Ortiz Alvarez**. Organização das coletâneas *Tendências atuais na pesquisa descritiva e aplicada em fraseologia e paremiologia*, volumes 1 e 2.

- 2014. **Claudia Zavaglia**. Organização do número 2 (*Fraseologia e Paremiologia*), volume 8, da revista Domínios de Lingu@gem.

Apesar de relativamente extensa, a lista acima poderá certamente ser atualizada e acrescida de novos estudos que estão sendo feitos no Brasil, sobretudo no âmbito dos programas de pós-graduação em Letras das diferentes universidades brasileiras.

Observação importante feita por Ortiz Alvarez & Unternbäumen (2011) acerca do levantamento bibliográfico que apresentam, integrado na lista compilada acima, diz respeito ao fato de ter havido maior profusão de trabalhos desenvolvidos na área da Fraseologia a partir da década de 1990. Para esses autores, a partir desse período, ocorre aumento do número de pesquisas na área, com a defesa de dissertações e teses, além da proliferação de artigos publicados em revistas e em livros. Tal observação é compartilhada por Xatara & Ortiz Alvarez (2012).

Dentre os pesquisadores brasileiros que orientaram/desenvolveram e/ou ainda orientam/desenvolvem estudos no âmbito da Fraseologia, Monteiro-Plantin (2014) cita: Maria Aparecida Barbosa, Maria do Socorro Silva de Aragão, Maria Tereza Camargo Biderman, Maria Helena de Moura Neves, Margarida Basílio, Ieda Maria Alves, Enilde Leite de Jesus Faulstich, Stela Ester Ortweiller Tagnin, Maria Luísa Ortiz Alvarez e Cláudia Maria Xatara. Há, todavia, muitos outros nomes que se destacam nessa área de pesquisa, como Oto Araújo Vale, Cleci Regina Bevilacqua, Elizabete Aparecida Marques, Maria Eugênia Olímpio de Oliveira Silva, a própria Rosemeire Selma Monteiro-Plantin, Antônio Luciano Pontes, entre tantos outros.

Não se pode negar, pois, o crescimento dos estudos fraseológicos no Brasil. Da mesma maneira, não se pode negar a influência e contribuição da tradição de estudos fraseológicos de outros países no processo de consolidação da área no contexto brasileiro. Em particular, destaca-se a tradição francesa sobre o fenômeno *figement*, como se verá na seção seguinte.

### 3.5 A TRADIÇÃO FRANCESA DE ESTUDOS FRASEOLÓGICOS

Curiosamente, as bases para o estudo da fraseologia foram dadas por um linguista francês, Charles Bally, ainda no início do século XX. Mas, apesar desse pioneirismo francês, a Fraseologia como disciplina científica surgiu oficialmente no contexto da década de 1950, no bojo dos estudos linguísticos desenvolvidos na então URSS. Como se viu anteriormente, foram os russos que, ao retomarem as lições de Bally, efetivamente começaram a realizar estudos sistemáticos sobre as unidades fraseológicas cujos resultados permitiram paulatinamente o aprofundamento das análises e a expansão da área para outros países e continentes.

Essa dinâmica e percurso dos estudos fraseológicos fundamentam-se em razões históricas e políticas que os delinearam, mas o aprofundamento dessas questões escapam ao escopo da presente Dissertação.

Cabe, no entanto, em virtude da escolha pelo referencial teórico adotado, salientar a importância das pesquisas efetuadas pelos franceses no que se refere ao campo fraseológico. Assim como os autores soviéticos, alemães e os de língua inglesa, espanhola e portuguesa, os franceses representam, no cenário científico internacional relativo à Fraseologia, um movimento de vanguarda que tem produzido fecundas investigações na área em tela, com destaque para os centros de pesquisa, como o *Laboratoire d'Automatique Documentaire et Linguistique* (LADL), iniciado e coordenado por Maurice Gross, o *Laboratoire de Linguistique Informatique* (LLI), sob a coordenação do prof. Salah Mejri, e o projeto liderado por Igor Mel'čuk, para citar alguns.

Uma das contribuições da vertente francesa da fraseologia, que tem se alinhado à abordagem léxico-gramática (cf. CORPAS PASTOR, 2017), tendo como principais representantes Maurice Gross e Gaston Gross, diz respeito ao postulado de que as expressões fixas são regidas, de modo geral, pelas mesmas regras sintáticas que operam nas ditas sentenças livres. Esta hipótese, formulada e confirmada a partir das evidências empíricas fornecidas pelos grandes bancos de dados da língua francesa, contrapõe-se à ideia pseudocientífica do caráter anômalo dos fraseologismos. Com efeito, o trabalho com o processamento automático da língua, sobretudo com as unidades

fraseológicas, tem propiciado resultados e reflexões que direcionam a fraseologia para o centro da discussão linguística.

Outra importante contribuição dos estudos fraseológicos franceses para a área especificamente e também para a linguística, de modo geral, se refere à proposta teórica de Salah Mejri, para a terceira articulação da linguagem. A divisão bipartida formulada por André Martinet seria insuficiente para dar conta de explicar a origem e o funcionamento sintático, semântico e pragmático dos fraseologismos, uma vez que tais unidades se estruturam tanto a partir de elementos da segunda articulação quanto da primeira, figurando, pois, nesse quadro teórico, como estruturas estranhas.

Além disso, a própria noção problemática de palavra, tradicionalmente vista apenas pelo aspecto da monolexicalidade, ainda não foi resolvida teoricamente, especialmente por que os fraseologismos são unidades polilexicais mas funcionam similarmente a palavras simples, motivo pelo qual o conceito de palavra precisa ser refundado ou superado. Para tanto, a terceira articulação da linguagem, conforme defendida por Mejri (1997, 2012, 2018), representaria um passo fundamental, visto que permitiria analisar a palavra tanto em seu aspecto monolexical quanto polilexical, acomodando assim os fraseologismos no quadro analítico da linguística, de acordo com as demandas empíricas levantadas pelos fatos observados. Essa concepção sobre a linguagem foi possível graças aos estudos acerca do processo de *figement*, que opera nas línguas vivas produzindo inúmeras unidades fraseológicas, como se verá na subseção abaixo.

### **3.5.1 As pesquisas em torno do conceito de *figement* (cristalização)<sup>18</sup>**

Muitas vezes confundido com o próprio conceito de fraseologia, o *figement*, cristalização ou congelamento, constitui o processo linguístico que atua teoricamente em todos os níveis da língua, notadamente no campo sintático, semântico e lexical. Age de maneira gradual, produzindo combinações

---

<sup>18</sup> Embora não seja tão simples encontrar equivalentes fraseológicos entre as línguas, procurouse, na medida do possível, oferecer exemplos em português para as questões levantadas e analisadas pelos autores franceses. Buscou-se especialmente usar fraseologismos recolhidos no próprio glossário produzido por esta pesquisa.

de estruturas que se associam no eixo sintagmático, com implicações no eixo paradigmático, para formar unidades cada vez maiores, tendo em vista as necessidades comunicativas e expressivas dos falantes.

Ao se tornar um dos maiores linguistas franceses a pesquisar o assunto, Maurice Gross publica, em 1982, o trabalho intitulado *Une classification des phrases “figées” du français*, em que apresenta resultados e reflexões oriundas de suas pesquisas realizadas no âmbito do LADL<sup>19</sup>, a partir de um grande banco de dados da língua francesa, sob a orientação da léxico-gramática. Trata-se de um estudo que privilegia o aspecto formal das frases cristalizadas.

Pautando-se numa abordagem gerativa transformacional, com vistas ao tratamento automático do léxico, Maurice Gross reconhece que, no âmbito desse quadro, as frases cristalizadas eram consideradas as exceções às regras sintáticas formuladas geralmente a partir de um conjunto de formas. Entretanto, dispõe-se a demonstrar que, ao contrário, essas sequências não são excepcionais, nem em relação à sua sintaxe, nem ao seu léxico. Considera que o *figement* é um fenômeno de envergadura para a pesquisa em linguística, porém carente de estudos.

Seu ponto de partida é a observação intuitiva de que o sentido da combinatória não é a soma dos sentidos dos constituintes. Considera essa intuição de “*figées*” um teste operatório que pode ser aplicado tanto às frases cristalizadas quanto a outras categorias sobre as quais age o *figement*, como nas construções nominais, adverbiais e verbais. No caso das estruturas nominais, o autor toma como exemplo os nomes compostos *cul de sac* (beco sem saída) e *pommer de terre* (*batata*), ressaltando a importância do traço de união como a melhor maneira de representá-los ortograficamente (*cul-de-sac*, *pommer-de-terre*), recomendando que tais unidades deveriam figurar como entradas ordinárias nos dicionários. Apesar de reconhecer os nomes compostos como estruturas de forma e conteúdo heterogêneos, demonstra que, na distribuição nominal, não se distinguem dos nomes simples, podendo exercer todas as funções gramaticais que lhes são típicas. Assim, por exemplo, *cul-de-*

---

<sup>19</sup> O *Laboratoire d'Automatique Documentaire et Linguistique* (LADL), coordenado anteriormente por Maurice Gross, na universidade Paris VII, é um projeto responsável pela constituição de um banco de dados linguísticos do francês hexagonal, que visa à classificação sintática de frases simples, sob uma abordagem formal no âmbito da léxico-gramática.

*sac e pommer-de-terre* equivalem respectiva e distribucionalmente a *impasse e batata*.

Nas estruturas com advérbio, como nos exemplos “o político foi pego <com a mão na massa>, aliás, com a mão na cueca, cheia de dólares” e “Entre seus militantes não faltam operários e intelectuais e mesmo ex-comunistas <de carteirinha>” (RCMC16M12a), o autor observa a impossibilidade de separar a preposição do sintagma nominal que lhe segue, sob pena de mudar o sentido da unidade fixa. Igualmente, nota essa restrição em algumas locuções prepositivas correspondentes a nomes simples, como *insu* (inconscientemente) e <à l’insu de> (sem conhecimento de).

Quanto às estruturas com verbo, Maurice Gross (1982) demonstra que existem determinados verbos compostos que se classificam geralmente da mesma maneira que os verbos simples, como: “*Ida a envoyé (promener son fils = son fils promener)*”, “*Ce travail fait suer Max*”, “*Ida a laissé tomber Max*” e “*Max est allé trouver Luc*”.

Para exemplificar o conjunto de questões a que se debruça, o autor apresenta uma análise para a sequência *casser sa pipe*<sup>20</sup>, equivalente a *morrer*, como na frase “*Max a cassé sa pipe*”. Para o autor, o sentido de “morrer” não resulta da soma dos significados individuais de “cassé” (quebrar) e “pipe” (cachimbo). Ressalta que a posição de sujeito da oração é ocupada por um nome variável, mas ao verbo e ao objeto não cabem modificações, sem que se desfaça a unidade de sentido. Não se pode, por exemplo, substituir “cassé” e “pipe” por formas sinônimas mantendo-se o mesmo sentido da sequência. Em português, isso seria equivalente ao fraseologismo <bater as botas> (morrer), cujas comutações com as palavras “chocar/colidir” e “sapato/chuteira” quebrariam a sequência cristalizada.

Assim, usando o asterisco (\*) para representar estruturas que, embora gramaticalmente aceitas, não mantêm o sentido idiomático da sequência original, demonstra a impossibilidade de substituição pelas sequências “\*Max a (brisé + rompu) sa pipe” (*Max quebrou/rompeu seu cachimbo*) – e “\*Max a cassé (son brûle-gueule + sa bouffarde + son fume-cigarette, etc.)” (*Max quebrou seu cigarro*). Nota ainda a impossibilidade de se modificar o determinante, como em

---

<sup>20</sup> No caso do português, seria algo como o fraseologismo *bater as botas*, a que se atribui o sentido de morrer.

“Max a cassé sa **propre** pipe” (*Max quebrou seu próprio cachimbo*) e “\*Max a cassé (**une + la + cette**) pipe” (*Max quebrou um, o, este cachimbo*), sendo o possessivo obrigatoriamente correferente ao sujeito: “\*Max a cassé leur(s) pipe(s)” (*Max quebrou seu(s) cachimbo(s)*), “\***Luc et Max** ont cassé leur pipe” (*Lucas e Max quebraram seu cachimbo*). Evidencia também que não é possível introduzir modificador para “pipe”, como em: “\*Max a cassé sa **vieille** pipe” (*Max quebrou seu cachimbo velho*), “\*Max a cassé sa pipe **bien remplie** (*Max quebrou seu cachimbo bem ocupado*)”. Além disso, observa que o adjetivo possessivo apresenta uma fonte diferente do que costuma ter, ou seja: um complemento nominal preposicionado (de N) ou uma relativa, produzindo estruturas sem o sentido idiomático característico do fraseologismo, como: “\*(La maladie + il) a cassé la pipe de Max” – *a doença dele quebrou o cachimbo de Max* – e “\*Max a cassé la pipe que le Seigneur lui a donnée” – *Max quebrou o cachimbo que o senhor lhe deu*.

Com isso, o autor mostra que as expressões fixas, tal como a que utiliza para ilustrar os problemas da análise, rejeitam as modificações comumente aceitas para qualquer outra frase. O bloqueio a essas alterações caracteriza a sequência como cristalizada.

Entretanto, ao aprofundar sua análise, reconhece que a própria sequência *cassé sa pipe* admite certas modificações estruturais, como a concordância de número e pessoa entre o possessivo e o núcleo do sujeito; a mudança do modo verbal “Max **va** casser sa pipe” (*Max vai quebrar seu cachimbo*), “Je crains que Max n'ait cassé sa pipe” (*Eu temo que Max tenha quebrado seu cachimbo*) e “Max aurait cassé sa pipe” (*Max teria quebrado seu cachimbo*); ausência de restrições para a inserção de advérbios de tempo “Max va casser sa pipe **dans peu de temps**” (*Max vai quebrar seu cachimbo em pouco tempo*); e inserções entre o verbo e o complemento: “Nous casserons **tous notre** pipe un jour” (*Nós vamos todos quebrar nosso cachimbo um dia*), “Tu casseras **aussi** ta pipe” (*Você também quebrará seu cachimbo*). Sendo assim, o autor chega à conclusão de que essa sequência figura, na verdade, como uma estrutura verbal ou predicativa, em que *sa pipe* constitui complemento verbal. Ressalta, contudo, algumas outras restrições, como a não passivação “\**La pipe est cassée par Max*” (*O cachimbo é quebrado por Max*), “*La pipe de Max est cassée*” (*O cachimbo do*

*Max está quebrado*) e a não relativização “\**La pipe que Max a cassée fascine Luc*” (*O cachimbo que Max quebrou fascina Lucas*).

Com esse trabalho, Maurice Gross apresenta inegáveis contribuições para o melhor conhecimento das unidades fraseológicas, especialmente do francês. Define e diferencia sequências livres das fixas, salientando, porém, a dificuldade para delimitar esses dois tipos. Segundo ele, as primeiras são “frases em que o sujeito e o complemento têm distribuição livre, isto é, apenas restrições de um ponto de vista semântico” (GROSS, 1982, p. 159), ao passo que, nas frases cristalizadas, um ou vários constituintes possuem restrição lexical. Em virtude da dificuldade para se estabelecer os traços fundamentalmente distintivos entre essas sequências, o autor propõe a noção de *continuum* entre formas fixas e livres. Dessa forma, demonstra que a variação atua conjuntamente na configuração das sequências cristalizadas.

Outro resultado significativo que convém ressaltar nesse trabalho de Maurice Gross refere-se à posição sintática em que o *figement* atua com mais força, pelo menos no que tange a frases simples. Em um estudo realizado a partir de 8.000 frases, observou menos de 600 sujeitos fixos, mais de 1.300 objetos diretos livres e mais de 1700 objetos indiretos livres. Em contrapartida, ocorreram mais de 1000 complementos de nomes fixos. Em resumo, concluiu que os complementos parecem ser mais favoráveis ao *figement* que a posição de sujeito.

Ex-orientando de Maurice Gross e depois colega de pesquisa no LADL, Gaston Gross também se destaca pelos estudos e publicações que vem realizando acerca do *figement*, notadamente a partir da década de 1990, quando sua principal obra vem a lume em 1996, a saber: *Les expressions figées en français: noms composés et autres locutions*. Nessa obra, o autor expõe os problemas gerais que surgem no tratamento, sobretudo automático, do *figement*, oferecendo uma análise detalhada dos fraseologismos, com base em dados recolhidos do *corpus* estabelecido para o *Laboratoire de Linguistique Informatique*, coordenado por ele na universidade Paris XIII. (cf. MEJRI, 1998). Mejri (2018) considera que essa obra de Gaston Gross representa uma boa síntese relativa à questão estudada.

Segundo Gaston Gross (1996), o *figement* constitui uma propriedade das línguas naturais, sendo um dos seus fenômenos mais importantes (CLAS &

GROSS, 1998). Todavia, durante muito tempo, foi ignorado pelos gramáticos e pela linguística, tendo sido considerado objeto de estudo marginal comparando-se a outros que gozavam de maior prestígio no quadro geral dos estudos da linguagem. Além da percepção coletiva e simplista de que as palavras compostas são aquelas que possuem o traço de união, o autor ressalta também o aspecto flutuante e heterogêneo da terminologia adotada, como fator que obscurecia o entendimento do fenômeno, em face das frequentes incompatibilidades das definições estabelecidas.

Ao discutir a noção de *figement*, Gross considera o uso desse termo apropriado porque permite explicar, ao mesmo tempo, fenômenos distintos que se inter-relacionam, abrangendo, portanto, não só a sintaxe, ou apenas a semântica, mas elementos diversos da estrutura gramatical e lexical da língua. Explícita, no entanto, as propriedades gerais do fenômeno, as características que considera comuns, como a polilexicalidade, a opacidade semântica, o bloqueio das propriedades transformacionais, a não-atualização dos elementos, o escopo do *figement*, os graus de *figement*, o bloqueio de paradigmas sinonímicos, a não-inserção de elementos no interior das expressões fixas, a *défigement* (descristalização) e a etimologia.

Sendo uma das principais características do *figement*, a **polilexicalidade** é a propriedade que uma sequência possui de ser formada por mais de duas palavras que, em outros contextos, possuem existência autônoma, como o fraseologismo **descascar um abacaxi** (resolver uma situação embaraçosa, complicada): “[Sérgio] Quintella sabe <descascar abacaxi>. Foi ele quem ajudou o governo federal na solução de mercado para o gigantesco Projeto Jari na Amazônia” (RIRB14M12d).

Pelo critério acima, ficam de fora as palavras derivadas, pelo fato de os elementos não-lexicais não possuírem autonomia sintática. Para o autor, poder-se-ia objetar que existe uma restrição na formação de derivadas, que poderia ser considerada um caso de *figement*, como a preferência do sufixo *-eza* pela base *gentil*, para derivar *gentileza*, em detrimento de outros sufixos que também formam substantivos (*-idade*, *-eira*). Entretanto, conforme esclarece o linguista, esse tipo de preferência faz parte de uma restrição mais geral que é do âmbito da sintaxe, pois “cada predicado tem um domínio de argumentos que lhe é próprio” (GROSS, 1996, p. 10).

A segunda propriedade apontada se refere à **opacidade semântica**. Diferentemente da noção de composicionalidade, segundo a qual o sentido de uma frase resultaria simplesmente da soma dos significados individuais de seus componentes, o caráter opaco possibilita que uma determinada sequência possa assumir um sentido diferente da interpretação literal que se poderia fazer caso se conheçam os significados individuais dos constituintes da estrutura. Por exemplo: **engolir sapo**, que pode significar literalmente a ação de se alimentar de sapo, como o faz uma cobra, ou, em virtude da opacidade semântica, a mesma combinação poderia significar “tolerar coisas ou situações desagradáveis sem responder por incapacidade ou conveniência” (HOUAISS, 2009), como se pode notar em: "Há cinco anos, em agosto de 2009, depois de <engolir muito sapo>, Marina trocou o PT pelo PV para se candidatar à Presidência" (RERA14M8c). Gross (1996, p. 11) afirma que, em casos como este último, estar-se-ia diante “de uma sequência opaca ou semanticamente fixa e restrita lexicalmente”. Adverte, porém, que a opacidade se trata de um fenômeno escalar, podendo ser total, parcial ou inexistente. Para Lamiroy (2008), essa opacidade semântica integra um fenômeno semântico mais geral, que é a polissemia das palavras, comum a todas as línguas naturais.

A terceira característica discutida pelo autor é o **bloqueio de propriedades transformacionais** e está estreitamente relacionada à anterior. Conforme o linguista, verbos e complementos podem ter mudanças estruturais que dependem de sua organização interna. Deste modo, a partir de uma frase como *a criança leu este livro*, é possível chegar às seguintes transformações: passivação – *este livro foi lido pela criança*; pronominalização – *a criança o leu*; *détachement* – *o livro, a criança o leu*; extração – *é este livro que a criança leu*; e relativização – *o livro que a criança leu*.

Segundo o autor, há verbos que dispensam algumas dessas transformações, e a causa disso não é simples de perceber. O caso limite seria a sequência que não aceitaria nenhuma das transformações, constituindo uma estrutura sintaticamente fixa e semanticamente opaca, já que, para o autor, “a opacidade semântica e as restrições sintáticas andam juntas” (GROSS, 1996, 12). Para ilustrar um caso limite, cita-se, por exemplo, o fraseologismo **bater boca**, utilizado no trecho que segue: “Para o manda chuva petista [Lula], a presidente [Dilma] só tem a perder <batendo boca> com o deputado fluminense

[Eduardo Cunha]" (RIRB15M12a). Neste caso, todas as transformações estruturais vistas anteriormente são negligenciadas pela unidade em foco, a menos que perca o sentido do fraseologismo, produzindo formas como: passivação – \* a boca foi batida pela Dilma; pronominalização – \*a Dilma a bateu; *détachement* – \* esta boca, a Dilma a bateu; extração – \*é esta boca que a Dilma bateu; e relativização – \*a boca que a Dilma bateu.

Gross (1996) mostra ainda que o bloqueio de transformações pode ocorrer também nas estruturas nominais, como em um grupo formado de substantivo e adjetivo (um livro difícil), como segue: nominalização – a dificuldade deste livro; adição de advérbios intensificadores – este livro muito difícil; e predicação – este livro é difícil. Tais modificações, porém, não se aplicam numa sequência nominal cristalizada como **braço direito** "principal e eficaz auxiliar; braço forte" (HOUAISS, 2009), utilizada no excerto: "Neca Setubal, socióloga, educadora, autora de mais de dez livros, defensora do desenvolvimento sustentável e herdeira do banco Itaú, é o <braço direito> de Marina" (RERA14M8e). Em qualquer tentativa de modificações dessa natureza, o fraseologismo se desfaz como unidade significativa, como comprovam as estruturas: nominalização – \*a direita do braço; adição de advérbios intensificadores – \*braço muito direito; e predicação – \* o braço é direito.

Diante dessas questões, Gross (1996, p. 13) reconhece que de fato "a cristalização é um fenômeno que transcende isto que nós chamamos geralmente de diferentes níveis de análise linguística e que uma descrição que seria apenas sintática ou semântica reteria apenas parte dos fatos" (tradução minha).

A quarta propriedade discutida por Gross (1996) diz respeito ao que ele chamou de **não-atualização de elementos**. Tal critério permite tratar, por exemplo, das sequências compostas, visto que o autor as define justamente como estruturas em que algum de seus elementos lexicais componentes não pode ser atualizado. O conceito de locução defendido pelo autor está relacionado a essa propriedade, uma vez que, para ele, as locuções são exatamente as sequências em que seus componentes não se atualizam, como na locução preposicional **a rodo** "em grande quantidade" (SILVA, 2013, p. 1295), presente no trecho: "Para cobrar ética de governos, não podemos patrocinar para os filhos formaturas com álcool a rodo" (RERA16M12d). No âmbito dessa combinação, tanto a preposição quanto o substantivo não se atualizam, não

podem ser substituídos por equivalentes, nem aceitam inserções de elementos no meio da estrutura sintagmática.

A quinta e a sexta propriedades apontadas por Gross estão relacionadas e referem-se, respectivamente, ao **escopo do figement** e aos **diferentes graus** que ele pode apresentar. Segundo o autor, a situação não exhibe tanta dificuldade quando a cristalização afeta completamente as sequências, como nos casos em que não é possível modificar nenhum elemento da combinatória. Por exemplo: um provérbio (**à noite todos os gatos são pardos**, uma sequência verbal (**ter olho maior que a barriga** –), uma substantiva (**calça-curta** – marido dominado pela mulher; homem frouxo, covarde (AURÉLIO, 2004), uma sequência adjetiva (**de ponta** – avançado), uma adverbial (**a ferro e fogo** – "de todas as formas possíveis, incluindo a violência" MICHAELIS, versão on-line) e uma locução prepositiva (**a toque de caixa** – "A toda a pressa, com açodamento" HOUAISS, 2009).

Para o autor, as sequências totalmente cristalizadas não constituem paradigmas e funcionam de modo compactado, em bloco, semelhante às formas simples, razão pela qual o linguista recomenda que pudessem constituir entradas nos dicionários.

Contrariando ao que se possa parecer a regra mais comum, Gross (1996) observa que os casos de cristalização total não são os mais frequentes, pois, em geral, nas sequências dadas apenas um dos componentes é cristalizado. Por esse motivo, adverte que: "uma boa descrição linguística deve refletir com precisão, para uma determinada sequência, o que é fixo e o que não é, especialmente em uma perspectiva de processamento automático" (GROSS, 1996, p. 16).

A menor frequência de sequências totalmente cristalizadas ocorre exatamente porque a cristalização se dá em graus, variando num *continuum*. Com efeito, "as variantes são mais frequentes que a cristalização total" (GROSS, 1996, p. 16). Assim, uma sequência nominal como **vinho tinto** não é completamente cristalizada, pois o adjetivo pode ser substituído por branco, *rosé* ou cinza. Conforme o autor, esses adjetivos, na posição de epíteto, não podem funcionar como predicativos e também não estão no seu sentido habitual, pois referem-se, na verdade, a diferentes tipos de vinho. Observa, então, que a possibilidade de comutação está relacionada, de certo modo, a determinadas

posições na estrutura sintagmática da combinatória. Conclui, pois, que as línguas apresentam diferentes graus de cristalização, configurando um *continuum* entre sequências livres e outras plenamente restritas.

Outra propriedade discutida por Gross (1996) é o **bloqueio de paradigmas sinonímicos**. Para explicar essa característica, o autor retoma a dicotomia saussureana dos eixos sintagmático e paradigmático e esclarece que, no âmbito de uma sequência livre, as palavras, ao se inserirem em paradigmas, podem ser comutadas no plano sintagmático por outras sinônimas, conforme a natureza dos predicados. Entretanto, no caso das sequências cristalizadas, essas possibilidades de substituição por sinônimo são bloqueadas. Assim, em um fraseologismo como **pagar o pato** ("ser responsabilizado por atos praticados por outro(s)" MICHAELIS, versão on-line), visto no trecho "Com as finanças públicas sem controle no Brasil de Lula e Dilma, quem <paga o pato> não é o PT, é você." (RERA14M4b), não se pode substituir o substantivo *pato* por outro, mesmo que seja da mesma classe de animais, como *ganso*, *cisne*, *marreco* etc. Não há, neste caso, paradigma sinonímico.

Quanto à **não inserção de elementos no interior da sequência cristalizada**, Gross observa que, nas construções livres, é comum a inserção de adjetivos, advérbios e de relativas, especialmente nos grupos nominais, mas não apenas. Deste modo, um sintagma como *um livro* admite, por exemplo, o acréscimo de adjetivo (um *bom* livro), de advérbio (um livro *muito* bom), de relativa (um livro *que é* bom). Tais possibilidades, porém, são menos frequentes em se tratando de sequências cristalizadas. Em **carta branca** ("autorização conferida a alguém para agir do modo que julgue melhor; plenos poderes" HOUAISS, 2009), usada no trecho "É porque aqui [no Brasil] a imprensa e a classe política dão <carta branca> para a polícia agir da maneira como convier na defesa dos interesses do seu partido" (JFVS16M9b), não se pode inserir outro adjetivo ao que já existe, nem advérbio (\*dar carta muito branca), nem relativa (\*uma carta que é branca).

Gross (1996) e Class & Gross (1998) salientam que essa regra não é absoluta, havendo a possibilidade de inserção de elementos<sup>21</sup> como a negação e outros advérbios. O fraseologismo **dar bola** ("Dar atenção a; ligar importância

---

<sup>21</sup> Gross (1996) afirma que o locutor pode modificar a estrutura da sequência cristalizada para provocar efeitos metalinguísticos e humorísticos.

a" (AURÉLIO, 2004), por exemplo, é comumente usado também na sua forma negativa, como no trecho a seguir: "Na gestão de Márcio Thomaz Bastos criou-se o Departamento de Recuperação de Ativos e Cooperação Jurídica Internacional no Ministério da Justiça. Mas Dilma Rousseff **não** <deu bola> ao DRCI, que anda a míngua" (RIRB15M10b).

O autor também apresenta o **défigement** (descristalização, descongelamento) como mais uma característica das sequências cristalizadas. Define-a como o rompimento da "camisa de força que caracteriza as sequências fixas" (GROSS, 1996, p. 20). Diferentemente das transgressões (consideradas falhas) às regras sintáticas que regulam as relações entre as estruturas das combinações livres, como o uso da passiva com verbo intransitivo, a descristalização ocorre exatamente a partir da violação às leis impostas pela cristalização, nas sequências cristalizadas. Entretanto, neste caso, a "quebra" das regras atende a efeitos lúdicos, como o que se verifica no trecho de fala da personagem Magda, interpretada pela atriz Marisa Orth, no programa humorístico *Sai de baixo*, da Rede Globo de televisão. Em uma das cenas, Magda, segurando seu esposo pelos braços, fala: "Como diz o ditado, quem tem ciúme vai a Roma, hein!". A descristalização ocorre justamente pelo fato de Magda substituir a palavra "boca", que integra realmente o ditado popular *Quem tem boca vai a Roma*, por "ciúme", provocando, com tal alteração, o humor.

Por fim, Gross (1996) discute a **etimologia** como um parâmetro relativo à origem do *figement*. Conforme explicação do autor, as sequências cristalizadas não são criadas livremente e reguladas pelo falante a qualquer momento. Pelo contrário, sua combinação sintagmática é imposta aos sujeitos, tendo fonte histórica, ainda que não seja muitas vezes possível acessar sua origem. De acordo com o linguista, o processo de cristalização pode originar-se tanto de fatores externos, como fatos históricos (*Quem tem boca vai a Roma*), mitológicos (*maçã da discórdia*), religiosos (*separar o joio do trigo*), ou estar vinculado a passagens literárias (*rasgar ceda*). Além disso, o estudo etimológico das sequências cristalizadas pode elucidar questões da história linguística interna, por meio de construções sintáticas e vocabulário pertencentes a estágios anteriores da língua.

Percebe-se, portanto, que o conjunto de trabalhos na esteira da léxico-gramática, representado particularmente por Maurice Gross e Gaston Gross<sup>22</sup>, deu importantes contribuições para o entendimento do processo de cristalização na língua, especificamente no caso do francês utilizado na França Hexagonal. Todavia, como é de se esperar no universo científico, outras pesquisas têm sido desenvolvidas, de modo a ampliar e aprofundar as análises, a partir de pressupostos teóricos distintos ou de metodologia diferente. Dentre os estudos, destacam-se as pesquisas de Béatrice Lamiroy, Igor Mel'čuk, Salah Mejri, entre outros.

Baseando-se nos trabalhos oriundos do LADL, mas ampliando o escopo da pesquisa para integrar quatro variedades regionais do francês, Lamiroy coordena o projeto Bélgica-França-Quebec-Suíça (BFQS), incluindo a francofonia partilhada pela Bélgica, Québec e Suíça. Os resultados desse projeto têm oferecido questões interessantes sobre o processo de cristalização, como a discussão a respeito das expressões cristalizadas de caráter regional, confirmando a maioria das hipóteses levantadas por Maurice Gross e *cia*, mas também problematizando conceitos e critérios já utilizados em descrições anteriores.

Em três dos muitos trabalhos publicados por Lamiroy – *Expressions verbales figées et variation en français: le projet BFQS* (LAMIROY, B. et al., 2003), *Le problème central du figement est le semifigement* (LAMIROY & CLEIN, 2005) e *Les expressions figées: à la recherche d'une définition* (LAMIROY, 2008) –, a autora sintetiza as principais reflexões em torno da cristalização. Chama a atenção para a expansão do interesse a respeito da cristalização ao longo dos anos, seja por lexicógrafos e gramáticos, seja por linguistas filiados a diferentes abordagens teóricas e metodológicas. No âmbito dos estudos da linguagem, a autora destaca o interesse da linguística descritiva ou teórica, das pesquisas cognitivas e psicolinguísticas e da linguística textual e contrastiva. Apesar disso, ressalta a permanência de uma certa flutuação terminológica e conceitual no domínio em pauta.

Ao realizar um balanço da bibliografia produzida na área, Lamiroy & Klein (2005) resumem sete pontos consensuais: (i) *a cristalização é um fenômeno*

---

<sup>22</sup> Mejri (2018) considera que a obra de Gaston Gross de 1996 fornece uma síntese dos trabalhos desenvolvidos no âmbito da léxico-gramática.

*fundamental do funcionamento da linguagem; (ii) possui caráter muito complexo e polifatorial, escapando a qualquer definição simples; (iii) há múltiplos critérios estabelecidos pelos especialistas, mas todos convergem para questões centrais (opacidade semântica ou não-composicionalidade do sentido, arbitrariedade das expressões, sentido figurado, convencionalidade, pré-fabricação, restrições sintáticas, bloqueio de propriedades transformacionais, bloqueio da variação paradigmática); (iv) alguns critérios são menos gerais e de caráter complementar, como traços de estágios mais antigos da língua, não-atualização de um elemento da combinatória, restrições de seleção no plano paradigmático, impossibilidade de tradução das sequências fixas para outras línguas; (v) a cristalização é de natureza gradativa; (vi) justamente esse caráter gradual representa dificuldade para a delimitação das estruturas cristalizadas, especialmente nas zonas em que a cristalização é mínima e os casos são menos protótipos; e (vii) a cristalização atua tanto na sincronia quanto na diacronia da língua.*

Apesar de elencar esses pontos consensuais, no bojo da bibliografia especializada, Lamiroy & Klein (2005) reconhecem a incapacidade dos linguistas de elaborarem um traço definatório que possa ser compartilhado por todos os casos a serem incluídos no rol das expressões cristalizadas. Segundo os autores, essa situação levou alguns pesquisadores, tais como Misri (1987), Danell (1992), Achard & Fiala (1997), a defenderem a intuição como um critério relevante para se decidir o que é cristalizado ou não em uma língua. Deste modo, sob um ponto de vista psicolinguista, “a noção de cristalização seria antes uma categoria cognitiva existente na memória dos falantes, que uma noção propriamente linguística” (LAMIROY & KLEIN, 2005, p. 137)<sup>23</sup>.

Nesse sentido, Klein & Lamiroy (2016), ao definirem as sequências cristalizadas, mantêm alguns dos critérios já utilizados, como a natureza pré-fabricada das sequências e o fato de não procederem de uma combinação sintático-semântica livre, mas as situam como **parte da competência lexical do falante nativo, em sua memória**, o que lhe permitiria completar partes dos fraseologismos que porventura estivessem faltando em determinado uso, como

---

<sup>23</sup> “la notion de figement serait plutôt une catégorie cognitive existant dans la mémoire des locuteurs, qu'une notion proprement linguistique” (LAMIROY & KLEIN, 2005, p. 137).

neste caso: *ao bom entendedor...* (espera-se que o falante nativo consiga completar a sequência com *meia palavra basta*).

A autora admite, no entanto, a dificuldade prática de se operar com esse traço cognitivo da cristalização. Embora esse aspecto da inscrição na memória do falante seja parte essencial da cristalização, a raiz do problema deriva do fato de essa noção ser dinâmica. Assim, uma sequência pode ser considerada semicristalizada por ainda estar em vias de implementação no seio de uma comunidade, ou seja, sua sedimentação na memória do falante não está completa. Para Lamiroy & Klein (2005), isso está relacionado ao que Bolinger (1976) afirmou sobre a linguagem, o fato de que esta mantém em equilíbrio a memória e a produção. Entretanto, segundo Lamiroy & Klein (2005), em se tratando da cristalização, ocorre que frequentemente estamos na fronteira desses dois aspectos da linguagem, sendo seus contornos nem sempre nítidos e demarcados. Assim, “O que é automático, memorizado ou congelado é de fato apenas ‘post factum’, tendo sido um dia engendrado como uma produção livre” (LAMIROY & KLEIN, 2005, p. 137)<sup>24</sup>.

Os autores, portanto, embora reconheçam a importância fundamental da intuição dos falantes, revelam a dificuldade de se trabalhar com essa noção. Ressaltam, inclusive, que, mesmo já dispendo de experiência com análises relativas à cristalização das frases do francês, hesitam na seleção de algumas sequências. Segundo eles, um dos principais obstáculos da equipe que integra o projeto BFQS é justamente decidir pela manutenção ou não de determinadas combinações na lista das sequências cristalizadas.

Há que se destacar a importância do projeto BFQS para as reflexões que Lamiroy têm apresentado em suas publicações a respeito da cristalização. Uma importante contribuição no que se refere à definição do *figement* na língua encontra-se em Lamiroy (2008), em que a autora sintetiza os critérios largamente adotados na literatura da área, demonstrando, de um lado, os problemas da definição e do caráter polifatorial e gradual do fenômeno e, de outro, a não-especificidade dos critérios atribuídos na análise.

Ao tratar especificamente do caráter polifatorial e gradual da cristalização, Lamiroy (2008) observa que a existência de protótipos revela que

---

<sup>24</sup> “Ce qui est automatique, mémorisé ou figé ne l'est en effet que « post factum », ayant été un jour engendré comme une production libre” (LAMIROY & KLEIN, 2005, p. 137).

os diferentes fatores da cristalização atuam de forma desigual no quadro das expressões cristalizadas. As restrições morfossintáticas, por exemplo, geralmente utilizadas para descrever as sequências, não agem em todos os casos. Comparando-se os fraseologismos **abrir mão** ("desistir de algo" MICHAELIS, versão online) e **aliviar a barra** ("ajudar a sair de uma enrascada, relaxar uma penalidade" FULGÊNCIO, 2008, p. 479), nota-se que, enquanto a forma passiva é permitida em *A barra foi aliviada pelo pai, que pagou a conta no supermercado*, a mesma transformação é interdita para *\*A mão foi aberta por Pedro*.

A autora mostra ainda que fatores semânticos, lexicais e morfossintáticos podem agir ao mesmo tempo ou separadamente numa sequência cristalizada. Por exemplo, em **abrir mão**, visto anteriormente, o sentido é opaco, não há paradigma que permita substituir o item lexical *mão* por outra palavra, além de haver restrição sintática, a exemplo da impossibilidade da passiva.

Diferentemente, no caso de **braço direito**, nem todos os fatores podem ser percebidos. O fraseologismo apresenta opacidade semântica e restrição morfossintática (não é possível a nominalização, nem a pluralização do sintagma), mas admite substituição lexical, como a substituição de *direito* por *forte*, que dá origem à variante **braço forte**, conforme registrada no excerto: "O que veio à tona com o homem-bomba e ex-<braço forte> de Dilma e Lula no Congresso [Dalcídio] é um conjunto de crimes contra a democracia, que vão muito além da corrupção passiva ou compra de votos" (RERA16M3a).

Por último, há sequências em que apenas um dos três fatores se manifesta de forma mais evidente, como **canoa furada** ("empreendimento temerário; situação enganosa, problemática ou arriscada" HOUAISS, 2009): "O mesmo Covas que ameaçou largar o PSDB caso FHC aceitasse o convite de Collor para ser seu chanceler. E não é que o homem quase embarcou na <canoa furada>?" (RCMC14M9a). Do ponto de vista morfossintático, nota-se, por exemplo, a possibilidade da predicação – *Isto está me cheirando mal, esta canoa é furada*. Quanto ao léxico, *canoa* admite a substituição por *barca*, formando um paradigma. Entretanto, é inegável a opacidade semântica do fraseologismo.

Para a autora, essa distribuição desigual dos fatores mencionados revela que de fato o problema da delimitação das sequências cristalizadas não

está nos protótipos, isto é, no limite superior do fenômeno. Pelo contrário, conforme já explicitou em Lamiroy & Klein (2005), a dificuldade maior está no grau inferior da cristalização, na semicristalização.

Quanto ao caráter escalar da cristalização, Lamiroy (2008) adverte que a explicação não se limita à distribuição desigual dos fatores usados na delimitação das estruturas. Esse aspecto também decorre do fato de o *figement* inscrever-se no tempo, tal como os demais processos de inovação lexical. Assim sendo, “as expressões, que são polilexicais por definição, requerem um determinado tempo de soldura para se constituírem como expressões cristalizadas” (LAMIROY, 2008, p. 7)<sup>25</sup>. Tal observação conduz ao entendimento de que o grau de cristalização pode também estar associado à diacronia da língua, conforme o estado em que a sequência em via de cristalização se encontra.

Outra observação da autora concerne à falta de especificidade dos critérios adotados na descrição das sequências cristalizadas. Segundo Lamiroy (2008), os fatores arrolados acima, de ordem semântica, lexical e morfossintática, não constituem condições necessárias e suficientes para uma definição mais precisa do fenômeno.

Em relação à opacidade semântica, a autora destaca o fato de esse aspecto integrar um fenômeno mais geral das línguas naturais, a saber, a polissemia das palavras. Essa observação já havia sido feita por outros linguistas, Gross (1996) e Mejr (1997).

De fato, como ressalta Lamiroy (1996), muitas locuções idiomáticas baseiam-se em metáforas e metonímias de um dos componentes da combinatória ou de todos eles em conjunto, podendo a mesma sequência possuir duplo sentido. Nos casos de homonímia, por exemplo, observam-se tanto o sentido literal quanto o sentido figurado, tal como **andar na contramão**, que pode de fato representar uma *violação às leis de trânsito* como também, de forma figurada, indicar uma *atitude contrária, diferente do que a maioria espera*:

Uma pesquisa Datafolha relata que mais de 70% dos entrevistados aprovam a violência praticada pela polícia contra Lula, na sexta 4. Onde sobraram os brasileiros que foram beneficiados durante a

---

<sup>25</sup> “les expressions, qui sont par définition polylexicales, requièrent un certain temps de soudure pour se constituer en tant qu’expressions figées” (LAMIROY, 2008, p. 7).

Presidência do ex-metalúrgico? Talvez o habitante do limbo careça de memória. Este é, porém, traço comum à maioria. Tal a chave de entendimento dos eventos dos dias de hoje, ao menos uma das razões. Permito-me <andar na contramão> (RCMC16M3d).

Segundo a autora, a depender da distância entre o sentido figurado da expressão e o sentido primário das palavras que a constituem, costuma-se considerar a sequência como cristalizada ou como uma metáfora simples. A fronteira entre os sentidos é muito tênue, não havendo muitas vezes uma razão objetiva para justificar a inclusão de determinadas expressões em um dicionário e a exclusão de outras.

Com respeito à ruptura paradigmática, Lamiroy (2008) esclarece que o critério lexical que restringe a comutação dos componentes da estrutura faz parte, na verdade, de um fenômeno mais geral, qual seja, o das solidariedades lexicais. Segundo a linguista, esse fenômeno corresponde ao “*the idiom principle*”, formulado por Sinclair (1991), segundo o qual a escolha de uma palavra exerce influência sobre a escolha das demais que as cerca. Lamiroy (2008) destaca o fato de esse mesmo princípio ser utilizado tanto para as expressões cristalizadas quanto para as colocações (FIRTH, 1957 *apud* LAMIROY, 2008), um tipo fraseológico específico, comumente distinto das expressões cristalizadas.

Segundo a autora, as colocações são definidas como associações caracterizadas pela coaparição arbitrária mas estatisticamente frequente dos elementos que as compõem. Diferem-se das expressões cristalizadas por serem geralmente binárias, formadas por uma base e um colocado, e semanticamente transparentes. Porém, há colocações mais transparentes e outras de sentido mais opaco, razão que origina uma série de dúvidas, em particular nas estruturas verbais, para se diferenciar uma simples colocação de uma sequência cristalizada.

Quanto às restrições morfossintáticas, Lamiroy (2008) identifica dois obstáculos. O primeiro refere-se ao reconhecimento de que, mesmo as frases ditas “livres”, sofrem a ação de restrições morfossintáticas, fato que tem motivado o surgimento de pesquisas que propõem os termos *colocação gramatical* e *colonstrução* (colocação + construção), levando em conta tanto as relações lexicais convencionais como também os traços gramaticais. Segundo a

autora, esse ponto de vista permitiria rever as dimensões tradicionalmente atribuídas ao processo da cristalização e à sintaxe das frases “livres”. Para ela, a existência abundante de colocações no nível do léxico somada às restrições léxico-gramaticais reduzem o quadro das frases “livres” e acentua a presença da cristalização por toda parte.

Outro ponto levantado pela autora relacionado ao critério das restrições morfossintáticas diz respeito à função discursiva da fraseologia. De acordo com a linguista, uma pesquisa recente nesse campo demonstrou que, além da função referencial, as unidades fraseológicas cumprem função organizacional no nível textual e também função interacional. Muitos fraseologismos, por exemplo, são usados como marcas discursivas ou constituem formas rotineiras conversacionais. Lamiroy (2008) ressalta que essas unidades são abundantes e originam-se frequentemente no registro oral, não figurando, muitas vezes, nos dicionários. Todavia, reivindica que tais combinações estruturais devam fazer parte do domínio fraseológico, uma vez que, apesar da transparência semântica, são afetadas por restrições lexicais e morfossintáticas.

Após realizar esses apontamentos, Lamiroy (2008) apresenta algumas propostas suas para tentar solucionar os problemas levantados, em especial, o do caráter flutuante da definição, além de analisar sugestões de outros autores que poderiam ser alternativas para o tratamento adequado das expressões cristalizadas.

Uma das alternativas propostas tem sido a adoção de um sistema de ponderação que se baseia no cálculo do grau de cristalização. Segundo esse sistema, são propostos alguns testes (geralmente formais) (GUIMIER & OUESLATI, 2006 *apud* LAMIROY, 2008) envolvendo os diferentes fatores relacionados e, com base nas respostas negativas das expressões a esses testes, dimensiona-se o grau de cristalização; quanto mais negativas, mais cristalizadas serão as sequências, e vice-versa. Bolly (2008 *apud* LAMIROY, 2008), por exemplo, tentou uma definição objetiva para o *figement* a partir da análise de 200 sequências verbais formadas com o verbo *prendre* (tomar), submetidas a um conjunto de critérios relativos à opacidade semântica, à ruptura paradigmática e às restrições morfossintáticas.

Segundo Lamiroy (2008), o sistema de ponderação se revela contraproducente por dois motivos. Primeiro porque ele se mostra inviável do

ponto de vista prático, uma vez que sua aplicação a uma quantidade muito grande de dados se constituiria materialmente irrealizável. Já o segundo motivo está relacionado à dificuldade de medir-se a dimensão semântica da expressão cristalizada.

Nesse contexto, a autora admite que, se for levada em conta a definição em sentido amplo de fraseologia, postura muito mais adotada pela tradição anglo-saxônica que pelo próprios franceses, será ilusório tentar encontrar uma definição realmente satisfatória do fenômeno, que seja totalmente imune a imprecisões. Por esse motivo, a autora recomenda que:

Devemos nos contentar com uma definição mais geral segundo a qual uma expressão cristalizada é uma unidade fraseológica composta de várias palavras, contíguas ou não, que apresentam certo grau de cristalização semântica, certo grau de cristalização lexical e certo grau de fixidez morfossintática (LAMIROY, 2005, p. 12)<sup>26</sup>.

Talvez como uma forma de consolo, a autora utiliza o advérbio *heureusement* (infelizmente) para dizer que o problema em torno da cristalização não constitui um caso isolado na linguística. Há outros conceitos basilares da análise linguística que também ainda não foram satisfatoriamente definidos, como a noção de palavra e de frase. No entanto, conforme destaca Lamiroy (2008), esse cenário não impossibilitou que os linguistas continuassem a fazer seus trabalhos continuamente. Na subseção abaixo, serão apresentadas e discutidas exatamente as reflexões de Salah Mejri para o estudo da cristalização lexical, com destaque para a sua proposta referente à terceira articulação da linguagem, o que, segundo ele, permitiria analisar o conceito de palavra em termos de sua monolexicalidade e de sua polilexicalidade.

### 3.5.2 A proposta de Salah Mejri

Salah Mejri é professor de linguística da universidade Paris 13 e diretor do laboratório *Lexiques Dictionnaires Informatique* (LDI). Desde a década de 1990, desenvolve pesquisas sobre vários fenômenos relacionados a diferentes

---

<sup>26</sup> “on doit se contenter d’une définition très générale selon laquelle une expression figée est une unité phraséologique constituée de plusieurs mots, contigus ou non, qui présentent un certain degré de figement sémantique, un certain degré de figement lexical et un certain degré de fixité morphosyntaxique” (LAMIROY, 2005, p. 12).

níveis da análise linguística, sobretudo o léxico. Sua obra é muito vasta, cobrindo a neologia lexical, os problemas da tradução, tratamento automático dos dados linguísticos, atlas linguísticos, aspectos lexicológicos, lexigráficos e terminológicos, ensino de língua etc. Todavia, no conjunto de seus trabalhos, destaca-se a investigação acerca do *figement*, principal objeto de suas pesquisas cujas contribuições para o campo da fraseologia de forma geral e da francesa, em particular, se revelam na síntese crítica que o autor realiza dos trabalhos anteriores sobre o fenômeno em questão, na tentativa de ordenamento terminológico da área, na formulação de novos conceitos teóricos e metodológicos e, principalmente, no refinamento de critérios e categorias de análise que tem feito a partir de problemáticas observadas no tratamento de dados empíricos, particularmente no contexto do reconhecimento automático das sequências cristalizadas.

Uma de suas principais obras, resultante de sua tese de doutoramento, é certamente o livro *Le figement lexical – descriptions linguistiques et structuration sémantique*, publicado em 1997, pela faculdade de letras de Manouba, Tunísia. Nessa obra, prefaciada por Robert Martin, Mejri apresenta uma síntese crítica dos trabalhos produzidos anteriormente a respeito do fenômeno fraseológico de modo geral, como os estudos de Saussure, Bally, Sechehaye, Pottier, Benveniste, Darmasteter etc., bem como das pesquisas e descrições de aspectos particulares do processo do *figement*. Além da revisão bibliográfica, acompanhada de reflexões pessoais, Mejri propõe uma abordagem sistemática e inovadora para o estudo do processo de cristalização linguística, a saber: a descrição da estrutura semântica das sequências cristalizadas. Com isso, o autor objetiva “mostrar que a cristalização não é um fato isolado, mas que está, ao contrário, no centro do funcionamento da linguagem” (MEJRI, 1997, p. 34)<sup>27</sup>. Trata-se, portanto, de uma obra basilar para os estudos fraseológicos, em especial para a vertente francesa.

Assim como muitos outros autores que o antecederam, Mejri (1997) reconhece a complexidade do fenômeno da cristalização e a quantidade ínfima de estudos exaustivos sobre tal objeto, justamente como consequência dessa complexidade. Para ele, o estudo do *figement* deveria levar em conta os

---

<sup>27</sup> “montrer que le figement n’est pas un fait isole mais qu’il est au contraire au centre même du fonctionnement de la langue” (MEJRI, 1997, p. 34).

conceitos metodológicos que a linguística moderna tem forjado para a descrição linguística e as mudanças realizadas.

Embora considere que o *figement* possa ser abordado sob distintos pontos de vista e diferentes níveis de análise, uma vez que constitui um fenômeno geral das línguas naturais, Mejri trata-o a partir de uma abordagem lexicológica, concebendo as sequências cristalizadas como unidades do léxico. Tal concepção tem implicações importantes para o domínio de estudo fraseológico e para a linguística, uma vez que põe em xeque alguns conceitos, como, por exemplo, o de palavra, arbitrariedade do signo linguístico; além de propiciar uma visão mais dinâmica e solidária dos níveis linguísticos, mormente os que se referem à morfologia, sintaxe e semântica, visto que essas unidades cristalizadas apresentam significante plural, formado de pelo menos duas palavras autônomas, restrições sintáticas e traços semânticos peculiares. Assim, em vez do estudo do *figement* limitar-se apenas a um dos aspectos envolvidos em sua origem e funcionamento, a abordagem lexicológica permite relacionar todos eles em uma visão de conjunto, sem perder de vista o *status* de unidade lexical.

Em face das características do *figement*, Mejri aponta o interesse epistemológico desse fenômeno. Para o autor, as sequências cristalizadas se situam tanto no domínio da *langue* (as frases feitas e provérbios, por exemplo), quanto no da *parole* (as gírias, por exemplo), relativizando, assim, uma das dicotomias saussurianas e pondo em evidência o caráter integrador da cristalização. Segundo Mejri (1997, p. 11), o *figement* constitui “de fato um fenômeno que reconcilia de algum modo o individual e o social, a atividade linguageira e a virtualidade da langue, etc.”<sup>28</sup>.

Ademais, de acordo com Mejri (1997, 2004, 2018), a polilexicalidade está para a sequência cristalizada assim como a polissemia está para a monolexicalidade na qual se situam as unidades lexicais simples. Para ele, enquanto, no quadro da monolexicalidade, as palavras mudam de paradigma de acordo com as significações que contraem nos diferentes contextos, no caso do *figement*, os componentes lexicais dos fraseologismos são fixados pela saturação lexical em um significante plural. O autor explica que a esse

---

<sup>28</sup> “en effet d’un phénomène qui réconcilie en quelque sorte l’individuel et le social, l’activité langagière et la virtualité de la langue, etc.” (MEJRI, 1997, p. 11).

significante subjaz uma dupla estruturação semântica, que agencia os significados literal dos componentes da sequência e o sentido global da estrutura sintagmática. Devido a esse traço de origem, o autor defende que os fraseologismos guardam uma memória lexical, cujos sentidos podem ser acessados quando se utilizam as sequências.

Diante disso, o autor enfatiza a importância do *figement* como um objeto de estudo de grande relevância para a linguística. Entretanto, nota que durante muito tempo tal fenômeno foi marginalizado no quadro das descrições, sendo muito raras as obras linguísticas que se dedicassem inteiramente ao seu estudo. Mesmo no âmbito da lexicologia, poucos estudos eram realizados, se comparados a outros processos, como a derivação, composição e truncação.

Em contrapartida, identifica que o interesse sobre o fenômeno tem crescido, o que pode ser atestado pelas realizações de eventos científicos e aumento do número de publicações referentes a esse objeto (cf. MEJRI, 1997). Observa, então, que tal crescimento está relacionado a uma tomada de consciência da importância do fenômeno para o sistema geral da língua, da necessidade de superar o quadro de pesquisas até então permanecido fragmentado, incoerente e aproximativo, bem como da urgência de novas abordagens que possam resultar em conceitos metodológicos mais objetivos, a fim de delimitar com mais precisão o objeto de estudo e conseqüentemente “abrir caminhos para contribuições que ponham fim à letargia que caracteriza os estudos sobre a cristalização” (MEJRI, 1997, p. 23)<sup>29</sup>.

No quadro dessas contribuições, o autor menciona alguns resultados de pesquisa que levam em conta as sequências cristalizadas nas análises efetuadas. Cita, por exemplo, a obra de Maurice Gross, que demonstrou a importância quantitativa e qualitativa dessas sequências ao lado das sequências consideradas livres. Especificamente sobre o aspecto quantitativo, Mejrri faz alusão à equipe de lexicometria de Saint-cloud, segundo a qual as sequências cristalizadas constituem 20% dos textos tratados.

Para Mejrri, além desses dados, outros fatos permitem argumentar em prol de uma concepção sobre o *figement* que lhe destacaria no âmbito das preocupações dos linguistas. Dentre esses fatos, o autor destaca que o *figement*

---

<sup>29</sup> “et ouvrent la voie à des contributions qui pourraient mettre fin à la léthargie caractérisant les études sur le figement” (MEJRI, 1997, p. 23).

engaja todas as dimensões do sistema linguístico, indo do fonético ao discursivo. Trata-se de um recurso pelo qual a língua pode equipar-se de todo tipo de unidades (simples e frásticas), pertencentes a todas as categorias gramaticais, sem se restringir a um tipo de discurso ou registro linguístico. Neste aspecto, Mejri (1997) aborda o fato de que a principal diferença relativa à presença de unidades cristalizadas num registro e outro será de cunho quantitativo, uma vez que, apesar da abundância de fraseologismos em registros marcadamente populares e orais, eles também são significativos no discurso jornalístico, em textos literários, na terminologia científica etc. Além disso, mostra que, embora os lexicógrafos não tratem tradicionalmente as sequências cristalizadas como entradas nos dicionários, elas figuram em grande quantidade no corpo dos verbetes da palavra considerada mais importante da sequência, como sub-entradas.

Mejri considera que uma das formas de se verificar a importância do *figement* no contexto das pesquisas está relacionada ao número de termos forjados para denominar as sequências cristalizadas. De fato, nesse campo de estudos, há uma profusão terminológica, conforme mencionado alhures. Por outro lado, Mejri (1997) observa que tal profusão denominativa se deve a duas razões principais: (i) ao caráter impreciso e flutuante do conceito de palavra e (ii) à tentativa de forjar uma terminologia mais precisa para as unidades sob análise que rompa com as antigas denominações. Quanto ao primeiro caso, o autor reconhece que a noção de palavra, apesar de muito criticada, resiste e se mantém nas análises linguísticas, sendo em relação a ela que as denominações para as sequências cristalizadas são forjadas. A fim de evitar a imprecisão decorrente da noção de palavra, a gramática tradicional propôs denominações que privilegiam a dimensão complexa das sequências cristalizadas, mas sem estabelecer suficientemente os limites conceituais entre os termos, razão pela qual coexistem: locução, expressão / frase feita, expressão idiomática, galicismo, palavra composta, entre outras formas. Já o segundo motivo, relacionado ao primeiro, tem a ver com o desejo dos linguistas, tais como Benveniste, Pottier e Martinet, de se afastarem das denominações aproximadas, procurando ser mais precisos em suas formulações, propondo assim, respectivamente, os termos *sinapse*, *lexia* e *synthème*.

Sobre esse aspecto, o autor considera que:

a questão terminológica é fundamental no tratamento da questão a ser estudada porque somos convencidos de que um fato confusamente denominado é frequentemente mal apreendido. Sobre o plano puramente epistemológico, a precisão conceitual anda de mãos dadas com o grau de precisão científica (MEJRI, 1997, p. 26)<sup>30</sup>.

Assim, diante do quadro de proliferação denominativa, Mejri (2012) propõe um esclarecimento conceitual seguido de um ordenamento terminológico. Por um lado, *phraséologie* (fraseologia) e *figement* (cristalização, congelamento) designam conceitos distintos, mas complementares. O primeiro termo, integrante da terminologia utilizada no Leste europeu (cf. Mejri, 2005), refere-se ao **fenômeno** linguístico que se manifesta nas associações sintagmáticas recorrentes. Já o *figement*, embora frequentemente também tratado como fenômeno<sup>31</sup>, é, na verdade, em comparação a *phraséologie*, o **processo** pelo qual as referidas associações sintagmáticas se combinam. Considerado um processo universal próprio das línguas vivas, o *figement* atua tanto na diacronia quanto na sincronia, de forma sistemática, em todos os níveis linguísticos, ocorrendo independentemente da vontade dos locutores (MEJRI, 2012). Afirma o autor que o referido mecanismo de combinação coloca o nível sintagmático a serviço do lexical, tornando, assim, cada sintagma um possível candidato a constituir uma unidade polilexical (MEJRI, 2012).

Ao tomar como base essa primeira distinção, o autor organiza um primeiro grupo de termos, a partir da delimitação do fenômeno, processo e resultados. Deste modo, tem-se, de um lado, o termo *phraséologie*, designativo do fenômeno já explicitado, e os resultados, produtos ou diferentes expressões desse fenômeno, a saber: *phraséologismes* (fraseologismos) e *phrasèmes* (frasemas), tendo este último uma implicação teórica específica no âmbito da teoria Sentido-Texto, de Igor Mel'cuk. De outro lado, tem-se, como já se viu, o *figement* como o processo que dá origem às estruturas cristalizadas, a partir do qual se forjam

---

<sup>30</sup> "la question terminologique est fondamentale dans le traitement de la question à étudier parce que nous sommes convaincu qu'un fait confusément dénommé est souvent mal appréhendé. Sur le plan purement épistémologique, la précision conceptuelle va de pair avec le degré de précision scientifique" (MEJRI, 1997, p. 26).

<sup>31</sup> Além de Maurice Gross (1982), Gaston Gross (1996), Lamiroy (2008) e tantos outros autores, o próprio Salah Mejri já se referiu a *figement* como sendo um fenômeno, o que, porém, não inviabiliza vê-lo tecnicamente como um processo, como propõe Mejri (2012).

termos como sequências cristalizadas, expressões cristalizadas etc. (MEJRI, 2012).

Depois de opor o fenômeno ao processo, Mejrí (2012) finalmente identifica três paradigmas terminológicos atinentes tanto aos resultados da *phraséologie* quanto aos do *figement*. O quadro abaixo sintetiza e organiza os termos relacionados.

Quadro 01 – Organização dos paradigmas terminológicos relacionados aos resultados da *phraséologie* e do *figement*.

<b>CRITÉRIOS DE AGRUPAMENTO</b>	<b>PARADIGMAS TERMINOLÓGICOS</b>
herdados da tradição gramatical e lexicográfica	locução, expressão feita, <i>tournure</i> , galicismo
marcados por abordagens teóricas específicas, como a teoria Sentido-Texto	frasema, semi-locução, pragmatema
privilegiam aspectos específicos, como:	idiomaticidade, idiotisme, expressão idiomática, idioma
- o aspecto <i>idiomático</i>	
- o aspecto <i>parêmico</i>	parêmia, provérbio, sentença, truísmo, expressão proverbial, máxima, dito, adágio etc.

Fonte: adaptado de Mejrí (2012).

As diferentes denominações estão relacionadas, portanto, a abordagens distintas. Ao discutir essa problemática terminológica já em sua obra de 1997, Mejrí demonstra a ausência de um termo que pudesse ser usado sistematicamente, cobrindo de forma neutra todas características comumente identificadas. Nota que o termo *lexia complexa*, proposto por Pottier, demonstrase profícuo em muitos casos na literatura da área, pelo fato de inserir também a dimensão frástica das unidades, mas não engloba explicitamente, segundo Mejrí, as locuções prepositivas e conjuntivas.

Por outro lado, o termo *locução*, embora apresente a vantagem de descrever as unidades inferiores à frase (*locução verbal, nominal, adjetival, preposicional, conjuntiva* etc), também revela-se limitada em seu alcance. Por esse motivo, o autor elege o termo *sequência* que, segundo ele, abarcaria todas as denominações propostas e seria suficientemente neutro em relação à

definição desses termos. Acrescentando-se o adjetivo *figé* (cristalizado) à sequência, indica-se claramente o processo linguístico que está na origem das unidades, diferenciando-se dos demais fatos do léxico que poderiam ser confundidos com a ação do *figement*. Define, assim, a sequência cristalizada como “um sintagma formado conforme a sintaxe da língua e que, uma vez reutilizado e usado, será uma sequência congelada” (MEJRI, 1997, p. 29)<sup>32</sup>. De acordo com o autor, o termo sequência cristalizada poderia, então, abranger todos os segmentos cristalizados, indo da unidade simples a estruturas superiores, integrando todos os tipos de interjeições, locuções, palavras compostas, entre outras (MEJRI, 1997).

Ao tratar dessa abundância de termos utilizados no domínio fraseológico, seja para se referir ao fenômeno, ao processo ou aos produtos e resultados da manifestação dos dois primeiros, Mejrj (2012) chama a atenção para o fato de que frequentemente essa proliferação terminológica tem sido vista negativamente. Para muitos, trata-se de uma questão mais de caráter meramente terminológico, revelando-se pouco útil para o estudo do caudal fraseológico. Todavia, como salienta o autor, se houver mudança de perspectiva, considerando a complexidade do objeto de estudo, essa variação denominativa seria, com efeito, “a prova de uma grande dinâmica teórica e descritiva” (MEJRI, 2012)<sup>33</sup>.

### 3.5.2.1 *Crítérios para a delimitação de fraseologismos*

Assim como ocorre para a denominação do fenômeno fraseológico, do processo que o origina e dos produtos ou unidades a que se chegam, há vários critérios pelos quais se pode proceder à delimitação das sequências cristalizadas. Conforme já explicitado, a obra de Gaston Gross (1996) apresenta, por exemplo, um conjunto de propriedades e testes para se identificarem os fraseologismos, tais como a polilexicalidade, a idiomaticidade, o bloqueio das regras transformacionais, ausência de paradigma sinonímico etc. Todos esses critérios são importantes para a descrição das unidades, mas podem tornar

---

<sup>32</sup> “séquence figée est un syntagme formé conformément à la syntaxe de la langue et qui, une fois réutilisé et entré dans l'usage, sera lui aussi une séquence figée” (MEJRI, 1997, p. 29).

<sup>33</sup> “la preuve d'une grande dynamique théorique et descriptive” (MEJRI, 2012, p. 140).

exaustivo o trabalho descritivo em grande escala, como já observou Lamiroy (2005). Por esse motivo, sem negar a funcionalidade desses recursos, mas, pelo contrário, integrando-os em duas ferramentas metodológicas principais, adota-se, nesta Dissertação, os elementos metodológicos assumidos por Mejri (2012) para a delimitação das unidades fraseológicas, a saber: a fixidez e a congruência.

A fixidez, conforme o autor, designa uma noção nova que serve para descrever o processo da cristalização pelo qual as associações sintagmáticas são bloqueadas pelas regras da combinatória sintagmática. A fixidez pode ocorrer nos planos sintático, semântico, manifestando-se assim tanto no eixo sintagmático como no paradigmático.

No plano sintático, por exemplo, o fraseologismo *abrir mão* não permite o acréscimo de um determinante (\*abrir uma mão), nem de adjetivo para modificar a palavra mão (\*abrir mão fria, quente, leve).

Quanto ao nível semântico, são fixos no léxico os significados globais atribuídos a *baixar a bola* e *botar a mão no fogo (por)*, respectivamente: “eliminar ou diminuir o motivo de orgulho ou vaidade de (alguém)” (MICHAELIS, versão on-line) e “não ter dúvida alguma a respeito da integridade, da competência e do caráter de (alguém)” (HOUAISS, 2009).

Da mesma forma, a fixidez alcança o eixo paradigmático. Assim, por exemplo, o fraseologismo **conversa para boi dormir** “*palavreado com intuito de enganar; falsidade*” (HOUAISS, 2009) geralmente não aceita substituições para nenhuma das palavras que compõem a unidade, a menos que a estrutura se desfaça como fraseologismo, produzindo estas formas: \***diálogo**<sup>34</sup> para boi dormir, \*conversa **a fim de** o boi dormir, \*conversa para **vaca, cavalo, bode** dormir, \*conversa para boi **descansar, tirar um cochilo**.

Com base no critério da fixidez, o autor define a sequência cristalizada como a sequência que apresenta “uma fixidez total ou parcial das regras da combinatória sintagmática, da comutatividade paradigmática e da composicionalidade semântica” (MEJRI, 2012, p. 143)<sup>35</sup>. Percebe-se, portanto,

---

<sup>34</sup> Silva (2013), na versão preliminar do dicionário brasileiro de fraseologia, registra a variante **história pra boi dormir**.

<sup>35</sup> “une fixité totale ou partielle des règles de la combinatoire syntagmatique, de la commutativité paradigmaticque et de la compositionnalité sémantique” (MEJRI, 2012, p. 143).

nessa definição, a noção de grau como um pressuposto da fixidez. De fato, o fraseologismo *botar a mão no fogo (por)*, já visto em relação à fixidez semântica, permite a substituição do verbo “botar” pelo seu concorrente “pôr” (pôr a mão no fogo por), sem desfazer o significado da unidade.

Em relação à noção de congruência, o autor a define como “um processo de adaptação de unidades lexicais pelo qual elas se integram naturalmente na combinatória” (MEJRI, 2009, p. 79 *apud* MEJRI, 2012, p. 143). Tal adaptação ocorreria tanto no nível morfológico, sintático e semântico. Segundo o linguista, essa definição de congruência objetiva elaborar uma definição de palavra, cuja noção tem sido problemática há muito tempo na linguística.

No caso dos fraseologismos, a noção de congruência é pertinente pelo seu corolário, isto é, a incongruência, pela qual se considera que uma determinada sequência seja incongruente caso não atenda às regras que estão na base da formação das sequências cristalizadas (MEJRI, 2012). Concretamente, isso significa dizer, por exemplo, que a impossibilidade de se dizer \**pelo(s) andar(es) da/de uma carruagem/carroça/automóvel* se deve ao fato de a sintaxe cristalizada impor apenas *pelo andar da carruagem* “do jeito que as coisas vão” (HOUAISS, 2009).

Para Mejri (2012), a noção de (in)congruência mostra-se relevante por possuir natureza distinta do conceito de gramaticalidade, que segundo ele, recai exclusivamente sobre a boa formação gramatical. Difere-se, ainda, da noção de aceitabilidade, de cunho normativo. Pelo contrário, a (in)congruência pode ser aplicada a todos os níveis da língua, especialmente na sintaxe, semântica e pragmática. Com esse novo elemento metodológico, Mejri (2012) estabelece que tudo o que se encaixa nas regras da combinatória é considerado congruente, enquanto o que a contraria se torna incongruente.

Desse princípio, resulta a noção da previsibilidade. Define-se como previsível o que se espera de acordo com as “possibilidades inscritas na aplicação das regras” (MEJRI, 2012, p. 143). Este é um dos motivos que explica, por exemplo, o fato de os locutores completarem as sequências cristalizadas em caso de truncação, seja por qual razão for. No *corpus* desta pesquisa, há inúmeros exemplos que ilustram esse fato, como este que segue:

**Crise  
Toma lá...**

Nas reuniões informais em Brasília, como na semana passada, onde se discute à (sic) necessidade de união da base do governo e a divisão de responsabilidades da crise com o Legislativo, se aconselha Dilma Rousseff a ter paciência com os políticos. Termina com ela parecendo concordar. Mas logo a seguir ataca a classe. Presidente da Câmara, Eduardo Cunha dá o troco com a moeda que tem: toma CPIs e aumento na manobra para acelerar a abertura de um processo de impedimento de Dilma Rousseff (RIRB15M8a).

No excerto, o fraseologismo serve como título da seção sobre assuntos relativos à crise política no Brasil. Mas a sequência está incompleta, truncada, cabendo ao leitor engajar-se na tarefa de reconhecer a outra metade da estrutura, já que o fraseologismo em sua forma plena é *toma lá dá cá*. Para tanto, o princípio da previsibilidade testemunha a competência do leitor diante das regras da combinatória.

Em casos de incongruência, isto é, quando há violação das regras do *figement*, o falante percebe e corrige. Assim, para um falante do português brasileiro, se, em vez de *cair a ficha*, por exemplo, lhe fosse dito ou escrito *\*derrubar a ficha*, pretendendo-se o mesmo significado da primeira, certamente a atitude imediata seria de corrigir a sequência, preferindo a forma *cair a ficha*, como utilizada no excerto seguinte, extraído do *corpus*: "O PT já se uniu à esquerda. É verdade, o partido demorou a perceber isso, <a ficha só caiu> no golpe" (RCMC16M9a).

Observa o autor que a tripla característica das sequências cristalizadas favorece a previsibilidade sintagmática, ou seja, o fato de serem polilexicais, bem formadas e fixas. Por esse motivo, é possível que o falante realize a mesma operação de completar ou corrigir a sequência, seja em relação a uma unidade frástica, como o provérbio, seja se tratando de colocações, especializadas ou de uso na língua geral. No discurso especializado, por exemplo, as restrições paradigmáticas atuam de forma muito intensa, o que está relacionado evidentemente à tendência das comunicações especializadas serem normalmente forjadas para serem muito precisas, pois designam conceitos distintos na estrutura de conhecimento de um determinado domínio. Na política, por exemplo, *um governo é governo de direita/de esquerda/de centro etc.*

Mejri (2012) ressalta que essas características das sequências cristalizadas (polilexicais, bem formadas e fixas) auxiliam também na identificação das incongruências obtidas quando se pretende reestruturá-las.

Dessa maneira, a partir de um fraseologismo como *dar bola a* "dar atenção a; ligar" (AURÉLIO, 2004), utilizado no trecho abaixo:

Na gestão de Márcio Thomaz Bastos criou-se o Departamento de Recuperação de Ativos e Cooperação Jurídica Internacional no Ministério da Justiça. Mas Dilma Rousseff não <deu bola ao> DRCI, que anda a míngua (RIRB15M10b).

algumas tentativas de reestruturação revelam-se incongruentes, como a passivação \*a bola não foi dada por Dilma Rousseff; a pronominalização \*Dilma Rousseff não a deu e; destacamento \*a bola, Dilma Rousseff não deu.

Da mesma forma, a incongruência também se verifica no plano pragmático, ao contrariar, por exemplo, as regras estabelecidas para a adequação entre as fórmulas e o contexto de enunciação. Assim, empregar indistintamente os pragmatemas *boa noite* e *bom dia* violaria as normas pragmáticas.

De acordo com o autor, o cruzamento das noções de fixidez e (in)congruência permite a delimitação das sequências cristalizadas e, conseqüentemente, sua diferenciação em relação às combinatórias livres. Com efeito, as incongruências podem mensurar o grau de cristalização da sequência, em seu aspecto sintático, semântico, pragmático etc.

Como referido anteriormente, a fixidez pode ocorrer tanto no eixo sintagmático quanto no eixo paradigmático. No primeiro caso, refere-se às regras da concatenação sintagmática, podendo ser uma fixidez geral ou específica. Na fixidez geral, a inserção de novos elementos na sequência cristalizada produz estruturas completamente incongruentes, desfazendo a unidade: \*a voz do povo é (menor, maior, melhor que) (igual) a voz de Deus; \*trem da (de muita) alegria; \*telhado de vidro (laminado).

Por sua vez, a fixidez específica leva em consideração o conjunto das reestruturações previstas, em geral, pela sintaxe das sequências cristalizadas, como as modificações já discutidas por Gaston Gross (1996) (MEJRI, 2012). Para uma sequência nominal de estrutura Adj. + N, como o fraseologismo *viva alma*, utilizado pelo ex-presidente Lula para defender-se das acusações que lhe foram feitas: "Não tem, neste país, uma <viva alma> mais honesta do que eu", disse Lula" (RERA16M1c), haveria as seguintes reestruturações específicas,

que seriam interdidas pela cristalização da sequência: posposição do adjetivo \*alma viva; coordenação ou agrupamento dos adjetivos \*feliz e viva alma<sup>36</sup>; modificação por advérbios \*muito, apenas viva alma; acréscimo de adjetivos pospostos ou de modificadores preposicionais \*viva alma *penada*, \*viva alma *de gente*.

Em se tratando da fixidez no plano paradigmático, o autor esclarece que ela pode ser total ou parcial. Embora menos frequente, a fixidez paradigmática total impede qualquer comutação para os itens lexicais que compõem a sequência, como nos exemplos **engolir sapo** "tolerar coisas ou situações desagradáveis sem responder por incapacidade ou conveniência" (HOUAISS, 2009) e **escreveu não leu o pau comeu** "se não cumprir conforme o combinado, a punição virá com certeza" (MICHAELIS, versão on-line). Para esses fraseologismos, qualquer tentativa de substituição lexical é interdida pela fixidez total da sequência.

Com respeito à fixidez paradigmática parcial, Mejri (2012) distingue dois casos. No primeiro caso, situam-se as variantes de uma sequência que não revogam o seu caráter cristalizado: *ter (adotar) dois pesos e duas medidas; além (demais) da conta; apostar (todas) as (suas) fichas; pegar (muito) mal*. Já a segunda situação diz respeito a variantes que contêm diferenças de nuances semânticas ou de níveis da língua, como *colocar (meter) alguma coisa na cabeça de alguém; empurrar (meter, enfiar) goela abaixo; não despertar o leão (cão) que dorme; pôr (botar) as barbas de molho*.

O autor explica ainda, assim como para a fixidez paradigmática, que a fixidez semântica age tanto na combinatória interna quanto na combinatória externa. A interna tem relação com as figuras, tropos e estereótipos implicados no uso das sequências, bem como com as transferências de domínios. Assim, por exemplo, se, para o fraseologismo **tiro no pé**, ocorresse uma substituição da palavra "pé" por outras que pertencem ao mesmo paradigma de partes do corpo, se estaria afetando a fixidez semântica interna da sequência, produzindo formas incongruentes como \*tiro no olho/na mão/na cabeça. Outro exemplo, neste caso com um estereótipo, seria a unidade **mão de vaca**, que, devido à

---

<sup>36</sup> É possível, no entanto, acrescentar o adjetivo **única** à sequência, para intensificar a ausência de pessoa. Assim, a fala de Lula poderia, por exemplo, ter sido assim: "Não tem, neste país, uma (única) viva alma mais honesta do que eu".

fixidez interna, não aceitaria a comutação da palavra *vaca* por *boi*, *novilha* ou *cabra*, ainda que pertençam à mesma família dos ruminantes, sendo inadmissível, portanto, no contexto brasileiro: \*mão de boi/novilha. Por sua vez, em relação à transferência de domínios, observa-se, quanto à expressão idiomática **abotoar o paletó**<sup>37</sup> (morrer), cuja transferência se dá do domínio da vestimenta para o contexto das relações humanas de natureza metafísica, a impossibilidade de substituição do domínio de origem por outro qualquer objetivando-se a manutenção do mesmo significado. No próprio domínio da vestimenta, a simples substituição de paletó por outra peça do vestuário seria interdita pela fixidez semântica interna, não reconhecendo formas como: \*abotoar o casaco, \*abotoar a jaqueta<sup>38</sup>.

Por outro lado, a fixidez semântica externa implica o fator pragmático das sequências. Por isso, pode ser exemplificada a partir de fatos que demonstrem a fixidez pragmática, que prevê uma relação estabelecida entre uma sequência e uma situação específica de uso. Conseqüentemente, “a adequação da sequência à situação faz parte do sentido da sequência” (MEJRI, 2012, p. 146)<sup>39</sup>. Tal pressuposto é fundamental para esclarecer muitos casos de inconveniência ou gafes cometidos pelo uso equivocado de determinados fraseologismos em contextos inadequados. Imagine-se, por exemplo, uma situação de velório, no Brasil, onde a ideia da morte é culturalmente cercada por comoção e tristeza. Um indivíduo, conhecido da família ou não, ao chegar ao local do velório, cumprimenta os parentes ou amigos mais próximos da pessoa que faleceu, dizendo algo como: *Meus parabéns! Muitas felicidades!* Certamente seria criticado ou expulso do recinto, pois violaria a fixidez pragmática, a qual pressuporia, para situações dessa natureza, fórmulas do tipo *meus pêsames!*, *meus sentimentos!*, *minhas condolências!* Portanto, cada vez que a fixidez não é respeitada, ocorre incongruência pragmática (MEJRI, 2012).

Para Mejri (2012), esses dois elementos metodológicos (fixidez e congruência) estão relacionados a dois fatos importantes, a saber: o *defigement* (descristalização, descongelamento) e a competência fraseológica. Em relação

---

<sup>37</sup> Silva (2013), em sua versão preliminar do dicionário brasileiro de fraseologia, registra, além de **paletó de madeira** (caixão) e **abotoar o paletó**, a variante **abotoar o paletó de madeira**.

<sup>38</sup> Silva (2013), porém, registra a variante **abotoar o peito**.

<sup>39</sup> “l'adéquation de la séquence à la situation fait partie du sens de la séquence” (MEJRI, 2012, p. 146).

ao primeiro, o autor lhe atribui um duplo aspecto, que decorre tanto do processo de expressão da sequência, no que tange ao seu emprego lúdico, quanto de seu emprego metalinguístico. Segundo o autor, o uso lúdico geralmente explora a dualidade semântica das sequências cristalizadas, ativando o sentido literal e o sentido global da unidade, como no exemplo abaixo, extraído do *corpus* da pesquisa:

Digna, às vezes cabisbaixa, Alexandra perdoou publicamente o ex-marido, o deputado Pedro Paulo Teixeira, que a agrediu com socos, tapas e chutes no rosto e no corpo por duas vezes, em 2008 e 2010. Uma das agressões aconteceu dentro do carro, na frente da filha de 2 anos. 'As pessoas erram, e eu vim <dar a cara a tapa>', disse Alexandra. Alexandra **deu mesmo a cara a tapa** (RERA15M11b).  
(grifo nosso)

Como se vê, no excerto acima, explora-se, no comentário da articulista da revista *Época*, o duplo sentido da sequência **dar a cara a tapa**. De fato, em “Alexandra deu mesmo a cara a tapa”, a descristalização opera justamente para produzir, de forma lúdica, o efeito de ambiguidade e ironia.

No que se refere ao emprego metalinguístico das sequências cristalizadas, Meiri (2012) explica que, neste caso, a descristalização resulta de operações realizadas por linguistas no âmbito de suas análises sobre as unidades fraseológicas. Cita, como exemplo, um comentário feito por Gaston Gross (1996), sobre o bloqueio de paradigmas sinonímicos para as sequências nominais: “– os substantivos: uma caixa preta, \*uma caixa escura; [...] um curto-circuito, \*um breve-circuito; uma chave inglesa, \*uma chave britânica”. Meiri (2012) prefere tratar essas manipulações como descristalizações, não como estruturas agramaticais, uma vez que, segundo ele, *caixa escura*, por exemplo, não configura uma sentença agramatical, sendo apenas a descristalização de **caixa-preta**.

Com respeito à competência fraseológica, Meiri enfatiza que o reconhecimento das sequências cristalizadas mostra-se completamente dependente do cruzamento das noções de fixidez e congruência, vistas como elementos resultantes das dicotomias fixidez/variação, congruência/incongruência. De acordo com o autor, os estrangeiros não nativos têm mais dificuldade justamente nesses dois aspectos, seja pela violação à fixidez das sequências cristalizadas, seja por não perceberem as

(in)congruências já assinaladas. Para o autor, as unidades mais difíceis para serem delimitadas e utilizadas pelos estrangeiros são as colocações, porque são sequências que apresentam, simultaneamente, uma grande liberdade combinatória e uma cristalização relativa, razão pela qual não se consegue delimitar claramente seus contornos e abrangência.

### 3.6 CONCEPÇÃO DE DICIONÁRIO, VOCABULÁRIO E GLOSSÁRIO

Desde épocas remotas, o homem se interessa pela palavra, em seus vários aspectos (fônico, morfológico, sintático, semântico, pragmático e discursivo). Da simples confecção de listas de palavras ou termos de textos antigos, de difícil compreensão, à produção de inúmeros repertórios lexicais, com todos os recursos tecnológicos altamente sofisticados e o conhecimento científico acumulado, o ser humano, ao longo de seu desenvolvimento histórico, sempre esteve em contato com a palavra, operando com a face mais dinâmica e fluida da língua, isto é, o léxico, ao criar numerosas formas de compilar e analisar as unidades lexicais.

Progressivamente, com o passar dos tempos, as primeiras práticas de colecionar ou reunir palavras, antes feitas de forma mais assistemática e sem tanta preocupação científica, foram sendo aperfeiçoadas e desencadearam reflexões que propiciaram o surgimento de métodos e de técnicas específicas, bem como de princípios e de pressupostos teóricos, dando origem assim à lexicografia<sup>40</sup>, lexicologia, terminologia, terminografia, fraseologia e fraseografia. Com isso, a despeito de suas especificidades teórico-metodológicas e epistemológicas, as chamadas ciências do léxico têm produzido um volume muito grande de dados linguísticos, sobretudo de caráter lexical, seja na organização de inúmeros repertórios lexicógrafos, seja no campo mais teórico da reflexão crítica sobre a produção realizada.

Entretanto, essa profícua produção de obras ou repertórios lexicográficos/terminográficos/fraseográficos tem apresentado uma árdua tarefa para os pesquisadores, que consiste em definir, com precisão, os conceitos e os termos referentes à tipologia dessas obras, principalmente no que tange a

---

<sup>40</sup> Biderman (1984) situa o início da verdadeira lexicografia nos tempos modernos.

dicionário, glossário e vocabulário. Em geral, os autores que se debruçam sobre essa problemática reconhecem nesse domínio a existência de uma profusão denominativa e de caráter conceitual, como observa Barbosa (2001, p. 25): “ainda nos tempos atuais, persiste pluralidade de denominações de um mesmo conceito de obra lexicográfica e, inversamente, pluralidade de conceitos para uma mesma denominação”. O termo vocabulário, por exemplo, pode referir-se tanto a um conjunto de vocábulos de um universo de discurso quanto a um tipo de dicionário (BARBOSA, 1995).

Considerado o protótipo das obras lexicográficas de modo geral, pela sua frequência de uso, importância linguística, social, cultural, política e valor de consumo, o dicionário exhibe um paradoxo. Trata-se de “um dos objetos culturais mais usuais e mais mal conhecidos” (REY-DEBOV, 1984, p. 63).

Em sua tipologia para as obras lexicográficas, Barbosa (1995, 2001), combinando diferentes aspectos (níveis de atualização da língua – sistema, norma e falar concreto –; etapas do percurso gerativo da codificação e decodificação; diferentes tipos de conjuntos lexicais – universo léxico, conjuntos vocabulares, conjuntos palavras-ocorrência –; variações diacrônica, diatópica, diastrática e diafásica; e sistema de macroestrutura, de microestrutura e de remissivas), distingue dicionário, por um lado, vocabulário e glossário, por outro.

Segundo a autora, o dicionário de língua costuma reunir unidades do sistema, tendo o lexema como unidade padrão, ou seja, processa unidades lexicais da língua geral (BARBOSA, 2001); ao passo que os dois últimos tratam de unidades lexicais situadas na norma e no falar concreto, respectivamente: vocábulos e palavras. Especificamente, os vocabulários técnico-científicos e especializados integram uma norma linguística e sócio-cultural, constituindo-se conjuntos-vocabulários. Já o glossário, situando-se no falar concreto, pode ser definido de maneira geral como inventário das palavras-ocorrências e de suas acepções em um texto manifestado, figurando, segundo essa perspectiva, como um anexo de um texto ou obra de um autor. Baseando-se em Dominguez (1985), Barbosa (1995) considera, então, que o dicionário é uma obra de codificação, sendo o vocabulário e o glossário repertórios de decodificação.

Em sentido estrito, um glossário seria:

a obra lexicográfica que apresentasse unidades lexicais extraídas de um único texto manifestado e definidas em suas significações específicas, correspondentes a cada palavra-ocorrência, no mais alto nível de densidade sêmica, sem reunir, em um só verbete, duas ou mais palavras-ocorrências com a mesma forma de expressão. Cada palavra-ocorrência poderia corresponder, portanto, a uma entrada (BARBOSA, 1995, p. 4).

Assim delineado, um glossário seria exclusivo a um texto ou discurso concretamente manifestado, razão pela qual a autora o caracteriza como sincrônico, sintópico, sinstrático e sinfásico. Diferentemente, ao dicionário caberiam as variações diacrônicas, diatópicas, diastráticas e diafásicas; ao vocabulário, apenas a perspectiva sincrônica, admitindo-se, porém, em alguns casos, as variações socioterminológicas. Barbosa (2001, p. 36) ressalta que os vocabulários “definem-se, contudo, por uma rigorosa perspectiva sinfásica, própria de um tecnoleto, representativa de um universo de discurso”.

Barbosa (1995) utiliza também um critério quantitativo-qualitativo para distinguir vocabulário de glossário. Para a autora, o primeiro tipo de repertório pretende ser representativo de um universo de discurso, contemplando inúmeros discursos manifestados, constituindo uma norma lexical discursiva. O glossário, por seu turno, busca ser representativo da situação lexical de um único texto, considerando sua especificidade léxico-semântica e semântico-sintática, em um contexto enunciativo exclusivo e preciso.

Em relação à estruturação semântico-sintática proposta por Pais (1984 *apud* BARBOSA, 2001), a autora explica os diferentes níveis de coerção que as unidades-padrão dos três tipos de obras lexicográficas mencionadas podem apresentar. De um lado, o lexema, ao nível do sistema, caracteriza-se por apresentar maior grau de polissemia em virtude de possuir mais semas lexicais e gramaticais em torno de seu semema polissêmico. De outro lado, o vocábulo, pertencente a uma norma, sofre restrições semântico-sintáticas do universo de discursivo que integra. A palavra-ocorrência, por sua vez, apresenta muito mais restrição pelo fato de sua significação ser específica ao texto, recebendo, porém, acréscimos de semas conceituais, no eixo sintagmático (epissemema). Essas observações conduzem à autora a concluir que “à medida que o semema de língua se restringe, no processo de atualização, aumenta o grau de previsibilidade semântico-sintática” (BARBOSA, 2001, p. 36). A síntese e

ilustração dessas relações na estruturação semântico-sintática podem ser verificadas no excerto:

Observe-se, por exemplo, o lexema voar, em português do Brasil, com seu semema polissemêmico, ao nível de sistema, e que se distribui em três significados mais especializados: o “voar” do pássaro, o “voar” da máquina, o “voar” como “sonhar”, correspondentes, assim, a três vocábulos, ou, se preferirmos, a três sememas ligados ao mesmo sobressememema. Entretanto, essa mesma unidade lexical, considerada no discurso da aviação, deve ser analisada como um vocábulo monossememático, termo técnico-científico, que constitui uma marca daquele discurso: “voar da máquina, exclusivamente”. Esse vocábulo sofre, por sua vez, restrições semânticas em diferentes contextos intra-universo de discurso e, ao mesmo tempo, enriquece-se com semas contextuais provenientes da combinatória sintagmática: vôo rasante, vôo cego etc. (BARBOSA, 2001, p. 37).

Assim, levando em consideração os aspectos discutidos, a autora apresenta sua definição para o dicionário de língua, vocabulário e glossário, conforme sintetiza o quadro abaixo.

Quadro 2 – Caracterização de Barbosa (2001) para dicionário de língua, vocabulário técnico-científico/especializado e glossário.

<b>Obras lexicográficas</b>	<b>Características</b>
dicionário de língua	recupera, armazena e compila lexemas efetivos, de frequência regular, integrantes de diferentes normas
vocabulário técnico-científico/especializado	recupera e armazena vocábulos de um universo de discurso, enquanto elementos configuradores de uma norma discursiva, ou seja, vocábulos de alta frequência e distribuição regular, restritos a uma <i>phasis</i> , que podem, eventualmente, relacionar-se a vários <i>topoi</i> e <i>strata</i>
Glossário	recupera, armazena e compila palavras-ocorrências de um <i>chronos</i> , de um <i>topos</i> , de uma <i>phasis</i> , ou, noutros termos, extraídas de um único discurso concretamente realizado.

Fonte: adaptado de Barbosa (1995, 2001).

Embora tenha apresentado uma proposta para o ordenamento terminológico e conceitual dessas três obras lexicográficas, Barbosa (2001) admite que ainda está longe um consenso conceitual e denominativo para o conjunto dos repertórios lexicográficos e terminográficos. Reconhece que, pelo

menos na prática terminográfica, seria desejável uma normalização nesse sentido, o que favoreceria a comunicação entre especialistas. Por outro lado, adverte que uma normalização excessiva na área científica pode tornar-se um fator perturbador, com implicações nocivas ao próprio papel da ciência, vista “como processo de investigação, de livre, ampla e profunda discussão de teorias, de modelos e de sua evolução, e que poderia traduzir-se, pois, em obstáculo ao avanço científico, assim também em constrangimento da liberdade acadêmica” (BARBOSA, 2001, p. 44).

Em sua *Pequena introdução à lexicografia*, Herbet Andreas Welker (2004) também trata da problemática em torno da delimitação de obras lexicográficas/terminográficas. Especificamente com respeito à distinção entre dicionário, vocabulário e glossário, retoma a tipologia de Barbosa (1995), vista acima, e acrescenta outros pontos de vista, como o de Mário Vilela (1995).

Ao analisar duas das grandes obras de referência do português, *Michaelis* e *Aurélio*, Welker identifica que o primeiro trata como sinônimos de *dicionário* os termos *léxico*, *vocabulário* e *glossário*. No tocante ao *Aurélio*, o autor observa que a primeira acepção fornecida para *vocabulário* “conjunto das palavras de uma língua” coloca o termo em relação de sinonímia com *léxico*. O aspecto mais restrito de um vocabulário, conforme salientado por Welker (2004), só é identificado nas demais acepções fornecidas pelo Aurélio para o termo em questão, como em: “conjunto das palavras em certo estágio da língua”, “conjunto das palavras especializadas em qualquer campo de conhecimento ou atividade; nomenclatura; terminologia”, “conjunto das palavras e expressões conhecidas e/ou empregadas por pessoa(s) de determinada faixa etária, social etc.”, “conjunto das palavras usadas por um autor em sua obra, ou em parte dela”.

Das proposições de Vilela (1995), retomadas por Welker (2004), convém transcrever a parte em que se distinguem os conceitos de dicionário, vocabulário e glossário, a partir de uma perspectiva de “coleção de unidades”, tal como se vê em: “o dicionário é a recolha ordenada dos vocábulos duma língua, o vocabulário é a recolha de um sector determinado duma língua e o glossário é o vocabulário difícil de um autor, de uma escola ou de uma época” (VILELA, 1995, p. 14 apud WELKER, 2004, p. 25). Percebe-se, nessas definições, que a abrangência dos itens lexicais repertoriados em relação à língua é um critério determinante, sendo o dicionário o repertório de maior abrangência, estendo-se

a toda a língua<sup>41</sup>, o vocabulário compreendendo um setor da língua (uma especialidade, por exemplo) e o glossário restringindo-se ao conjunto vocabular de um autor, época ou escola.

Apesar da contribuição que o trabalho de Welker (2004) apresenta para a lexicografia brasileira, já que se trata de uma das primeiras introduções de caráter geral nessa área, no Brasil, ele não oferece elementos satisfatórios para a clara diferenciação entre dicionário, vocabulário e glossário, não indo muito além da reprodução das propostas de Barbosa (1995) e de Vilela (1995). Deve-se ressaltar, no entanto, que, no que tange à discussão da tipologia para os dicionários, o trabalho de Welker é um dos mais amplos, analisando diferentes propostas e sugerindo critérios objetivos para se delimitar especificamente os tipos de dicionários.

Podem-se sintetizar os critérios propostos por Welker (2004) para a delimitação de dicionários da seguinte maneira: (i) primeiramente, a distinção entre obras de consulta em formato de livro e as computadorizadas; (ii) a segunda entre dicionários monolíngues e bilíngues/multilíngues, sendo estes os mais raros e; (iii) por fim a distinção entre dicionários gerais e especiais<sup>42</sup>, propondo, ainda, que apenas um tipo de dicionário seja considerado geral, cuja caracterização seria “alfabético, sincrônico, da língua contemporânea, arrolando sobretudo os lexemas da língua comum” (WELKER, 2004, p. 43), sendo os demais dicionários especiais.

Por sua vez, Guilherme Fromm, em artigo intitulado *Obras lexicográficas e terminológicas: definições* (2004), apresenta uma descrição dos muitos tipos de repertórios lexicográficos e terminológicos, com o intuito de demonstrar que o termo *dicionário* é frequentemente inadequado para designar essas obras. Para isso, o autor discute a proposta de normalização da ISO (1990), as tipologias de Barbosa (2001), de Haensch (1982) e de CITRAT/USP (ALVES, 2001).

---

<sup>41</sup> Krieger (2006, p. 144) explica que “a totalidade não significa o registro exaustivo das palavras de uma língua, desde as mais antigas até os neologismos mais recentes, até porque a lexicografia não consegue acompanhar o dinamismo lexical; mas caracteriza a abrangência do componente léxico, sem privilegiar uma temática específica”.

<sup>42</sup> Para Barbosa (2001), a distinção entre dicionários gerais e específicos parece ser mais clara, contudo, há maior complexidade entre os dicionários especiais, “que registram vocabulários de normas de universos de discurso e vocabulários de discursos-ocorrências” (BARBOSA, 2001, p. 34).

Da norma ISO 1087 (ISO, 1990 *apud* FROMM, 2004, não paginado), transcrevem-se as definições a seguir:

6.2.1 dictionary: Structured collection of lexical units with **linguistic information** about each item

6.2.1.1 terminological dictionary (admitted term: technical dictionary): Dictionary (6.2.1) containing **terminological data** (6.1.5) from one or more specific subject fields (2.2).

6.2.1.1.1 vocabulary (admitted term glossary): terminological dictionary (6.2.1.1) containing **the terminology** (5.1) of a specific field (2.2) or of related subject fields (2.2) and **based on terminology work** (8.2) (grifo nosso).

A partir dessas definições recomendadas pela ISO, percebe-se uma diferenciação mais ou menos objetiva entre o dicionário e os demais tipos de repertório (dicionário terminológico/técnico, vocabulário e glossário). O primeiro, isto é, o dicionário (de língua) distingui-se por inventariar as unidades lexicais, fornecendo informações linguísticas (fonéticas, morfológicas, sintáticas, semânticas etc.) sobre cada uma. Já os demais repertórios restringem-se às línguas de especialidade que caracterizam uma ou mais campos de conhecimento. Segundo Fromm (2004), a definição dada ao dicionário é clara, mas sua distinção em relação a dicionário terminológico/técnico, vocabulário e glossário é imprecisa. Diante disso, o autor faz a seguinte reflexão:

Afinal, o que basicamente diferenciaria um dicionário técnico de um vocabulário é que o segundo se basearia em um trabalho terminológico. As perguntas que ficam são a (sic) seguintes: um dicionário terminológico não precisaria também de um trabalho terminológico de preparação? Afinal, existe ou não diferença entre glossário e vocabulário (que aqui são considerados sinônimos) (FROMM, 2004, não paginado).

Para tentar responder a tais questões, o autor retoma a classificação linguística proposta por Barbosa (2001), que já foi discutida acima, o estudo de Haensch (1982) e algumas considerações feitas pela equipe do CITRAT/USP (ALVES, 2001), a qual traduziu para o português a referida norma ISO. Do trabalho de Haensch (1982), Fromm afirma que o lexicógrafo alemão considera sinônimos os termos *vocabulário* e *glossário*, sendo este confinado às palavras de um texto ou autor, ou ainda como designativo de um:

Repertorio de palabras, en muchos casos de términos técnicos (monolingüe o plurilingüe), que no se pretende ser exhaustivo, y en que la selección de palabras se ha hecho más o menos al azar; por ejemplo, glosario de términos ecológicos español-inglés (HAENSCH, 1982, p. 106 *apud* FROMM, 2004, não paginado).

Assim sendo, permanece a imprecisão conceitual entre vocabulário e glossário. De fato, Haensch (1982), assim como Welker (2004), priorizam a tipologia de dicionários, deixando em segundo plano a distinção entre as demais obras lexicográficas/terminográficas.

No que tange às considerações da equipe do CITRAT/USP, Fromm (2004) destaca a adoção do termo *glossário* para denominar um repertório terminológico do campo da economia. De acordo com Aubert (2001 *apud* FROMM, 2004), o referido glossário segue as diretrizes do CITRAT/USP, que, por sua vez, filia-se ao modelo franco-canadense, privilegiando a situação de uso dos termos, que se dá na interação com os usuários.

Face à apresentação e discussão das propostas tipológicas para as obras lexicográficas e terminográficas, Guilherme Fromm (2004) conclui que, de fato, ainda não há uma definição satisfatoriamente objetiva. Dito em suas próprias palavras, “Em vista de todas essas conceituações, normativas ou não, fica patente que a definição quanto ao título das obras nas áreas lexicográfica e terminológica não é unívoca” (FROMM, 2004, não paginado).

Por seu turno, Maria da Graça Krieger (2006) discute a produção lexicográfica, considerando a complexidade dessa produção, as tipologias para os dicionários linguísticos, entre outros aspectos. Enfatiza, principalmente, as diferenças teórico-metodológicas entre os dicionários gerais e os dicionários que tratam de temáticas específicas, nos quais se enquadram os dicionários técnico-científicos ou terminológicos.

Como seu enfoque é a tipologia de dicionários, a autora não apresenta uma clara distinção conceitual entre essas obras e as demais, como vocabulário e glossário. Deixa entrever, no entanto, que, pelo menos no âmbito terminográfico, dicionário terminológico e glossário são utilizados indistintamente. Assim, segundo a autora, “Essas produções que, muito comumente, também aparecem sob a forma de glossários, restringem-se a repertoriar o que é convencionalmente chamado de léxico especializado” (KRIEGER, 2006, p. 143).

De forma semelhante, Cleci Regina Bevilacqua e Maria José Bocorny Finatto, no artigo intitulado *Lexicografia e terminografia: alguns contrapontos fundamentais* (2006), discutem as diferenças e os pontos em comum entre o fazer lexicográfico e o fazer terminográfico, considerados pelas autoras como dois “estilos” de produzir obras dicionarísticas. Desse trabalho, importa para esta dissertação a distinção que as autoras fazem, em nota de rodapé, entre dicionário e glossário, no contexto do Grupo TERMISUL (Projeto Terminológico Cone Sul). Segundo as pesquisadoras, o critério que os diferencia tem a ver com o nível de abrangência na exploração do *corpus* de estudo, conforme se vê no excerto:

Adotamos junto ao Grupo TERMISUL uma distinção entre dicionário e glossário. Um glossário, na nossa concepção, implica uma abrangência menor de corpus de estudo que a de um dicionário. Um dicionário da legislação ambiental brasileira pressupõe um trabalho exaustivo de reconhecimento da terminologia. Um glossário, por sua vez, compreenderia, usualmente, um segmento dessa mesma legislação (BEVILACQUA & FINATTO, 2006, p. 48)<sup>43</sup>.

Depreende-se, pois, que a distinção feita no âmbito do TERMISUL é válida especialmente para a terminologia/terminografia. Dicionário e glossário contrastam quanto à abrangência do *corpus* estudado.

Por último, ressalta-se a tipologia apresentada por Enilde Faulstich (1995) para os repertórios lexicais (lexicográficos e terminográficos), considerando a abordagem da Socioterminologia. Para tanto, a autora retoma algumas definições propostas no *Vocabulaire Systématique de la Terminologie* (BOUTIN-QUESNEL et al., 1985), com tradução feita pela própria linguista brasileira. Do conjunto dessas definições, destacar-se-ão, nesta dissertação, apenas as que dizem respeito a dicionário, vocabulário e glossário e seus tipos principais, para ser coerente com o objetivo desta seção e com os demais autores explicitados anteriormente.

Dessa maneira, para o termo *dicionário*, são forjadas as seguintes definições:

**Dicionário** – Repertório de unidades lexicais que contém informações de natureza fonética, gramatical, conceitual, semântica, referencial.

---

<sup>43</sup> Krieger & Finatto (2004), também em nota de rodapé, coadunam dos mesmos critérios para diferenciar dicionário de glossário, no contexto terminológico e terminográfico.

**Dicionário de língua** – Dicionário que contém informações fonéticas, gramaticais, semânticas acerca das unidades lexicais de uma língua.

**Dicionário geral** – Dicionário de língua que descreve as unidades lexicais de uma língua

**Dicionário especial** – Dicionário de língua que descreve unidades lexicais selecionadas por algumas de suas características. Ex.: dicionário de sinônimos, dicionário de gíria etc.

**Dicionário terminológico** – Dicionário que apresenta a terminologia de um ou de vários domínios. (FAULSTICH, 1995, p. 5).

De início, nota-se que os três primeiros tipos de dicionários mencionados estão implicados entre si, sendo muito tênue a fronteira conceitual entre um tipo e outro. Apenas os dicionários especial e terminológico apresentam traços distintivos relativamente mais precisos com relação aos demais: o dicionário especial, pelo fato de suas unidades lexicais serem selecionadas por “determinadas” (quais?) características, e o terminológico, por apresentar especificamente a terminologia de um ou de vários domínios.

Em se tratando de *vocabulário*, tem-se as definições:

**Vocabulário** – Repertório que inventaria os termos de um domínio e que descreve os conceitos designados por estes termos por meio de definições ou de ilustrações.

**vocabulário alfabético** – Vocabulário apresentado em ordem alfabética com ou sem remissivas.

**vocabulário sistemático** – Vocabulário apresentado em ordem sistemática e geralmente acompanhado de um *index*.

**vocabulário unilíngüe** – Vocabulário que repertoria os termos de uma única língua.

**vocabulário multilíngüe** – Vocabulário que repertoria os termos acompanhados de seus equivalentes de uma ou várias línguas. (FAULSTICH, 1995, p. 6).

Nesse caso, a diferenciação conceitual entre os tipos de vocabulário mostra-se muito mais evidente. Entretanto, apesar de o vocabulário restringir-se aos termos de **um só** domínio, o que neste caso seria distintivo no tocante à parte da definição de dicionário terminológico, ainda persiste a possibilidade de confundir seu conceito com o de dicionário terminológico, uma vez que, tal como visto anteriormente, este tipo de dicionário também pode inventariar a terminologia de apenas um campo de conhecimento. Essa aproximação conceitual, certamente, motiva o fato de muitas obras terminológicas serem designadas indistintamente por dicionário terminológico/técnico/especializado ou vocabulário técnico-científico/terminológico etc.

Finalmente com relação ao termo *glossário*, Faulstich (1995), em nota de rodapé, informa que considerou insatisfatória a definição fornecida pelo *Vocabulaire Systématique de la Terminologie*, isto é, de que um glossário seria o "Repertório que define ou explica os termos antigos, raros ou desconhecidos." (BOUTIN-QUESNEL et al., 1985 *apud* FAULSTICH, 1995, p. 13). Por essa razão, a socioterminóloga propõe três acepções para o termo *glossário*:

- a) Repertório que define termos de **uma área científica ou técnica**, dispostos em ordem alfabética, podendo apresentar ou não remissivas.
- b) Repertório em que os termos, **normalmente de uma área**, são apresentados em ordem sistemática, acompanhados de informação gramatical, definição, remissivas podendo apresentar ou não contexto de ocorrência.
- c) Repertório em que os termos são apresentados em ordem alfabética ou em ordem sistemática seguidos de informação gramatical e **do contexto de ocorrência**. (FAULSTICH, 1995, p. 6, grifo nosso).

De fato, ao propor definições mais detalhadas para o termo *glossário*, Faulstich (1995) amplia a noção dada pelo *Vocabulaire*, conferindo maior sistematicidade e objetividade a esse tipo de repertório terminológico, se comparado à definição anterior, excessivamente simplificada, de que glossário "define ou explica os termos antigos, raros ou desconhecidos". Nota-se que a distinção dessas três acepções de glossário estabelece claramente uma relação com a subdivisão feita pela equipe do *Vocabulaire* para os conceitos de vocabulário e seus tipos alfabético e sistemático. Ao fazer esse paralelo, seria possível designar, por exemplo, cada uma das três acepções da seguinte maneira: à acepção a), corresponderia o termo *glossário alfabético*; à acepção b), o termo *glossário sistemático*; e à acepção c), simplesmente o termo *glossário*, o qual serviria como hiperônimo dos demais subtipos de glossário.

Assim, pode-se concluir que a proposta de Faulstich (1995) avança no sentido de inserir novos elementos, sobretudo à definição de glossário. Porém, de alguma maneira, ainda é possível observar certa imprecisão com respeito à delimitação dos conceitos de dicionário terminológico, vocabulário e glossário, pelo menos se for levado em consideração o fato de cada uma destas obras terminográficas poder repertoriar termos de um só domínio especializado<sup>44</sup>.

---

<sup>44</sup> Ao reconhecer as dificuldades de se delimitar de forma clara e exhaustiva uma classificação taxonômica das obras lexicográficas, Bugenõ Miranda (2014, p. 224) ancora sua proposta de

Essa breve revisão dos trabalhos assinalados ao longo desta seção poderia significar a impossibilidade de se poder adotar um termo e uma definição adequados para se referir ao produto lexicográfico/fraseográfico desta Dissertação, uma vez que não se configura como um trabalho de natureza estritamente terminológica, nem somente da língua geral e muito menos exaustivo com relação ao *corpus* de estudo. Não se trata de termos, nem de unidades lexicais situadas apenas ao sistema da língua ou da norma, mas de fraseologismos, os quais, como já defendeu Mejri (1997), reconcilia, de algum modo, a virtualidade do sistema às práticas linguageiras.

Outro aspecto a ser considerado para essa aparente impossibilidade referida deve-se à ausência de uma fraseografia, principalmente brasileira, que apresente um conjunto de princípios e de denominações satisfatoriamente precisos, a ponto de delimitar claramente os repertórios produzidos. Xatara & Parreira (2011) apresentam reflexões e orientações muito importantes que devem presidir a confecção do projeto fraseográfico, mas discutem apenas a produção de dicionários fraseológicos, não o distinguem, por exemplo, de um glossário fraseológico. Evidentemente, como se viu, não se trata de um problema específico da fraseografia; as próprias ciências do léxico, em particular a lexicologia e a lexicografia, a despeito das inegáveis contribuições, ainda não conseguiram fornecer uma solução efetivamente satisfatória, a fim de atender às demandas que são inerentes tanto à ciência de um modo geral quanto à própria prática lexicográfica.

Por outro lado, deve-se enfatizar que justamente grande parte dessa pluralidade terminológica, de pontos de vistas teóricos e de abordagens metodológicas revela que a ciência é exatamente o lugar do diverso, da diferença, do diálogo, sem que com isso se abra mão do rigor metodológico e dos pressupostos que caracterizam as inúmeras áreas e subáreas, respeitando-se suas especificidades. Isso significa que os novos conceitos, termos e perspectivas não precisam ser necessariamente inéditos ou serem mutuamente exclusivos, pois isso levaria ao que Lima (2010) chamou de *ciência balcanizada*, fragmentada em ilhas em conflito; pelo contrário, o conhecimento pode progredir

---

classificação em três axiomas, sendo o primeiro deles: “Axioma 1: Uma classificação exaustiva e de aplicação universal é impossível, já que se detecta o surgimento constante de novas obras lexicográficas”.

muito mais pelas possibilidades que se abrem a partir de uma visão de conjunto, de complementaridade.

Por conseguinte, tendo em vista essa concepção de complementaridade, acredita-se que o termo *glossário* possa satisfazer de forma adequada o produto lexicográfico desta Dissertação. A razão principal que justifica essa opção deve-se a características do *corpus* de estudo selecionado e aos objetivos delineados. Trata-se de um conjunto de 570 textos, produzidos por 4 colunistas que comentam sobre política e assuntos afins em seus blogs ou colunas ancoradas nos sites de *revistas* e jornal de circulação abrangentes no Brasil. Os textos foram publicados entre janeiro de 2014 e dezembro de 2016. Desse *corpus*, que não é estritamente especializado em política, só interessaram à pesquisa os fraseologismos utilizados no discurso sobre política; portanto, não houve pretensão de exaustividade, nem de totalidade em relação à língua portuguesa.

Assim, baseando-se principalmente nos critérios propostos por Krieger & Finatto (2004), Bevilacqua & Finatto (2006) e Faulstich (1995), defende-se o uso do termo *glossário* para o repertório de fraseologismos do discurso sobre política. Adota-se aqui uma definição de glossário que poderia ser assim resumida: *repertório, de natureza descritiva, de termos ou fraseologismos coletados sistematicamente, mas não exaustivamente, de um corpus de estudo preferencialmente específico a um domínio, apresentando definição das unidades e seus contextos de ocorrência, informação gramatical, podendo conter ou não sistemas de remissivas, podendo ainda ser organizado alfabeticamente ou de maneira onomasiológica, ou ainda combinando as duas possibilidades.*

No contexto dessa profícua produção de repertórios lexicais, deve-se, ressaltar a grande contribuição de outras áreas, como a da Linguística de *Corpus*, que tem modificado particularmente o modo de se trabalhar com o léxico. Por isso, na seção abaixo, abordam-se, em linhas gerais, os conceitos e termos desse ramo da linguística.

### 3.7 LINGUÍSTICA DE *CORPUS*

Nesta seção, apresentam-se os principais conceitos e termos utilizados na área da Linguística de *Corpus* (doravante LC), particularmente os que possuem relação direta com esta pesquisa.

### 3.7.1 O que é Linguística de *Corpus*?

De uma maneira geral, a Linguística de *Corpus* pode ser concebida, quanto ao seu estatuto, como uma abordagem (disciplina) ou como uma metodologia, havendo inclusive quem a considere ambas as coisas, a depender do modo que é empregada (TAGNIN, 2011). No contexto específico desta dissertação, emprega-se a LC como uma metodologia que orientou a constituição e processamento do *corpus*, tendo auxiliado, principalmente, na identificação e extração semiautomática das unidades fraseológicas.

Sardinha (2006) define a LC como um campo dedicado à criação e à análise de *corpora*, isto é, de conjuntos de textos (escritos e falados) coletados sistematicamente e armazenados em arquivos processados por computador. Para o autor, a LC protagoniza uma revolução nos estudos da linguagem, oferecendo aos analistas acesso a grandes quantidades de dados linguísticos, antes inacessíveis. No entanto, isso só foi possível graças ao desenvolvimento da Informática e conseqüentemente ao advento do computador, razão pela qual o autor considera que a LC contemporânea não existiria sem a Informática. Em função das relações que estabelece com outras áreas, como, por exemplo, a Lexicografia, Terminologia, Fraseologia, Tradução, Linguística Aplicada e estudos em Processamento de Língua Natural (PLN), a LC tem sido concebida como uma área interdisciplinar cujo desenvolvimento se iniciou na Europa, na década de 1980, tendo alcançado outras partes do mundo, como o Brasil (BERBER SARDINHA & ALMEIDA, 2008).

De acordo com Tagnin (2011, 2015), a Linguística de *Corpus* caracteriza-se por ser uma abordagem empírica baseada na observação de um volume muito grande de dados autênticos, ou seja, de textos efetivamente produzidos nos variados segmentos da sociedade, não forjados simplesmente pelo analista segundo sua intuição. Trata-se, portanto, de uma perspectiva que concebe a língua como um sistema probabilístico (TAGNIN e TEIXEIRA, 2012), em que alguns traços mostram-se mais frequentes que outros, privilegiando assim as

ocorrências prováveis e não apenas as possíveis, pois, segundo a autora, “nem tudo que é **possível** ocorrer, é **provável** que ocorra” (TAGNIN, 2011, p. 278, grifo da autora). Decorre desse pressuposto a importância da investigação linguística baseada em *corpus* ou *corpora*, definido por Berber Sardinha (2000, p. 325) como “conjuntos de dados lingüísticos textuais que foram coletados criteriosamente com o propósito de servirem para a pesquisa de uma língua ou variedade linguística”. Similarmente, Sinclair (1991, p. 171) define o *corpus* como “um corpo de linguagem natural (autêntica) que pode ser usado como base para pesquisa linguística” (*apud* BERBER SARDINHA, 2000, p. 336). Mas não existe apenas um tipo de *corpus*, podendo variar de acordo com a natureza e os propósitos da pesquisa, como se verá na próxima subseção.

### 3.7.2 Tipos de *corpus*

Berber Sardinha (2000) apresenta uma tipologia de *corpus* organizada segundo determinados critérios, relativos ao: modo (falado/escrito); ao tempo (sincrônico, diacrônico, contemporâneo, histórico); à seleção (de amostragem, monitor, dinâmico ou orgânico, estático, equilibrado); ao conteúdo (especializado, regional ou dialetal, multilíngue); à autoria (de aprendiz, de língua nativa); disposição interna (paralelo, alinhado) e; à finalidade (de estudo, de referência, de treinamento ou teste).

O autor propõe ainda outros meios pelos quais o *corpus* pode ser classificado conforme sua composição, a saber: pluralidade de autoria (os textos foram produzidos por um autor ou mais de um autor); origem de autoria (o(s) autor(es) é(são) falante(s) nativo(s) ou não-nativo(s)); meio (o texto é falado ou escrito); integralidade (os textos foram recolhidos na íntegra ou apenas fragmentos); especificade (os textos são variados ou apenas de um único tipo de texto); dialeto (as variedades linguísticas presentes no *corpus* são do tipo padrão ou regionais); equilíbrio (as variedades do *corpus* se distribuem equitativamente ou não); fechamento (há possibilidade da inclusão de conteúdos novos ou não); renovação (o conteúdo do *corpus* espelha um período definitivo de tempo ou se renova); temporalidade (retratam-se períodos históricos de tempo ou não); e plurilinguismo (são apenas textos originais ou o *corpus* também

inclue as traduções dos textos para uma ou mais língua estrangeira). Para a descrição completa de cada tipo, recomendamos a leitura de Sardinha (2000).

Em virtude da natureza do *corpus* desta pesquisa, serão abordados mais especificamente os *corpora* relacionados ao modo, ao tempo, ao conteúdo, à autoria, à finalidade e à temporalidade.

Em relação ao **modo**, o *corpus* pode ser *falado*, quando composto de trechos de fala transcritos, ou *escrito*, no qual os textos são escritos, sendo impressos ou não.

Com respeito ao critério relativo ao **tempo**, o *corpus* pode ser: *sincrônico*, compreendendo um determinado período de tempo; *diacrônico*, abrangendo vários períodos de tempo; *contemporâneo*, representando o período de tempo corrente e; *histórico*, representando um período de tempo passado.

Em se tratando do **conteúdo**, pode-se classificar o *corpus* como *especializado*, em que os textos são de um tipo específico, representando geralmente gêneros ou registros definidos; *regional ou dialetal*, no qual os textos pertencem a variedades sociolinguísticas específicas e; *multilíngue*, quando comporta vários idiomas.

Quanto à **autoria**, o *corpus* pode ser classificado em: *de aprendiz*, quando os autores dos textos não são falantes nativos; *de língua nativa*, quando ocorre justamente o inverso, ou seja, os autores dos textos são falantes nativos.

No tocante à **finalidade**, há o *corpus de estudo*, que se constitui ele mesmo objeto de descrição; *de referência*, o qual é constituído para servir de comparação ao *corpus* de estudo a fim de se verificar, por exemplo, as palavras-chave do *corpus* que se está descrevendo e; de treinamento ou teste, cujo objetivo é permitir o desenvolvimento de ferramentas e aplicações de análise.

Por fim, no que se refere à **temporalidade**, Sardinha (2000) divide os *corpora* entre os que se propõem a tratar de períodos históricos de tempo e aqueles que não o fazem.

Desse modo, a Linguística de *Corpus* trabalha com diferentes tipos de *corpus*, de acordo com critérios também distintos. Mas, em todo o caso, devem ser textos coletados de forma sistemática e armazenados em arquivos que permitam o uso do computador, representando uma língua ou uma variedade dessa língua.

## CAPÍTULO 4 – METODOLOGIA

Neste capítulo, apresentam-se o contexto geral da pesquisa, a metodologia e descrição de suas etapas, bem como os recursos utilizados para elaboração do glossário.

### 4.1 NATUREZA DA PESQUISA

A presente pesquisa caracteriza-se por uma abordagem quali-quantitativa. Assim, os dados receberam tratamento tanto quantitativo, em virtude dos procedimentos orientados pela Linguística de *Corpus* (SARDINHA, 2000, 2004; TAGNIN, 2011), no que se refere à quantificação dos textos, contagem de palavras, frequência de ocorrência e de concordância, quanto qualitativo, na medida em que os dados também foram submetidos a critérios analíticos que põem em relação fatores de ordem mais subjetiva, como os diferentes graus de opacidade/idiomaticidade, cristalização e polilexicalidade dos fraseologismos, além da descrição dessas unidades.

### 4.2 DESCRIÇÃO DAS ETAPAS DE PESQUISA

A presente investigação foi dividida em cinco etapas principais, a saber: (i) revisão da bibliografia da área; (ii) constituição e tratamento do *corpus* de estudo; (iii) seleção do *corpus* de referência; (iv) procedimentos de análise dos resultados e; v) elaboração do glossário fraseológico. Por sua vez, cada uma dessas etapas desdobrou-se em outras conforme a necessidade observada.

#### 4.2.1 Recensão da literatura especializada

Nesta primeira etapa, procedeu-se a um levantamento dos trabalhos realizados na área da Fraseologia, de modo geral, e especificamente no domínio da fraseologia da política brasileira. Além de livros impressos, foram consultados trabalhos disponíveis na *internet*, como por exemplo, o banco de dissertações e teses da CAPES e demais acervos virtuais dos programas de Pós-graduação

em Letras do Brasil, como também os projetos de pesquisa relacionados à área em foco.

De início, foi feita uma consulta ao *site* do Diretório dos Grupos de Pesquisa do CNPq<sup>45</sup>, inserindo-se o descritor *fraseologia* na caixa de pesquisa dessa plataforma, a fim de verificar quais e quantos projetos estão cadastrados nessa linha de investigação, na referida base de dados. Obteve-se um registro de 15 projetos cadastrados, os quais estão organizados no quadro abaixo:

Quadro 3 – Projetos ligados à fraseologia cadastrados no diretório do CNPq.

INSTITUIÇÃO A QUE SE VINCULA	NOME DO PROJETO	COORDENADOR(A)
Universidade de São Paulo (USP)	1- COMET - Corpus Multilíngue para Ensino e Tradução	Stella Esther Ortweiler Tagnin
	2- Ensino e Aprendizagem de Espanhol	Isabel Gretel María Eres Fernández
Universidade de Brasília (UNB)	A fraseologia e sua equação nas subáreas da Linguística Aplicada	Maria Luísa Ortíz Alvarez
Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP)	1- Grupo de Pesquisa Estudos do Léxico: descrição e ensino	Odair Luiz Nadin da Silva
	2- Grupo de Pesquisa Teorias da Música	Marcos Fernandes Pupo Nogueira
	3- Lexicologia e Lexicografia contrastiva	Claudia Maria Xatara
	4- Pedagogia do Léxico e da Tradução a partir de Corpora	Marilei Amadeu Sabino e Adriane Orenha Ottaiano
Universidade estadual do Amapá (UEAP)	CELL - Complexidade, estudos em Linguística Aplicada e Linguagem	Fábio Xavier da Silva Araújo (UFPA)
Universidade Federal de Uberlândia (UFU)	GECon - Grupo de Estudos Contrastivos	Ariel Novodvorski
Universidade Estadual de Feira de Santana (UEFS)	Grupo de Estudos do Léxico do Português	Rita de Cássia Ribeiro de Queiroz
Universidade Federal do Mato Grosso do Sul (UFMS)	Grupo de estudos em fraseologia	Elizabete Aparecida Marques e Maria Eugênia Olímpio de Oliveira Silva

<sup>45</sup> [http://dgp.cnpq.br/dgp/faces/consulta/consulta\\_parametrizada.jsf](http://dgp.cnpq.br/dgp/faces/consulta/consulta_parametrizada.jsf). A consulta foi realizada em novembro de 2016.

Universidade Estadual de Goiás (UEG)	LexFras: Lexicologia, Lexicografia, Terminologia e Fraseologia	Huélinton Cassiano Riva e Kely Araújo Melo
Universidade Federal do Rio Grande do Norte (UFRN)	Linguística do texto e do discurso - operações de textualização	Luís Álvaro Sgadari Passeggi e Maria das Graças Soares Rodrigues
Universidade Federal do Rio Grande do Sul (UFRGS)	TERMISUL - Projeto Terminológico Cone Sul	Cleci Regina Bevilacqua e Patrícia Chittoni Ramos Reuillard
Universidade Federal da Grande Dourados (UFGD)	TRALEX- Ensino - Tradução, Léxico e Ensino de Línguas Estrangeiras	Rosana Budny e Lucilia Teodora Villela de Leitgeb Lourenço

Fonte: organizado pelo autor.

Como demonstrado no quadro acima, universidades de distintas regiões do país sediam programas ou projetos de pesquisa a respeito da fraseologia. Dentre essas instituições, destacam-se, pela produção dos pesquisadores envolvidos e pela quantidade de projetos na área, a UNESP, USP, UNB, UFMS e a UFRGS. Por outro lado, no Norte e Nordeste, apesar de ainda serem poucas as universidades e não muitos os pesquisadores que desenvolvem trabalhos no domínio em foco, tem-se conhecimento de que as seguintes instituições sobressaem-se no desenvolvimento de pesquisas na área, a saber: a Universidade Federal do Ceará, a Universidade Estadual do Ceará, a Universidade Federal do Maranhão, a Universidade Federal da Bahia, a Universidade Estadual do Amapá e a Universidade Federal do Pará, no âmbito da qual o projeto GeoLinTerm passou recentemente a realizar pesquisas no campo em tela.

Embora só tenham aparecido 15 projetos registrados no Diretório do CNPq, acredita-se que existam muitos outros em diferentes fases de desenvolvimento nas muitas universidades brasileiras. Há, por exemplo, o projeto Variação Lexical, Teorias, Recursos e Aplicações – VALEXTRA (CAPES/COFECUB n° 838/150), projeto celebrado entre a universidade francesa Paris 13 e a Universidade Federal da Bahia (UFBA).

Na realidade, Xatara & Ortiz Alvarez (2012) mencionam, além de alguns projetos já citados, muitos outros grupos de estudo e projetos de pesquisa relacionados ao estudo da fraseologia. Entre os grupos de pesquisa, as autoras destacam: *O arco fraseológico e a sua equação com as subáreas da Linguística*

*Aplicada* (UNB), coordenador por Maria Luisa Ortíz Alvarez e; os já mencionados *Grupo de Estudos em Fraseologia* (UFMS), dirigido por Elizabete Aparecida Marques e Maria Eugênia Olímpio Oliveira Silva, e *Lexicologia e Lexicografia contrastivas* (UNESP), coordenador por Claudia Xatará.

Entre os projetos de pesquisa, as autoras aludem a(o): *Dicionário eletrônico alemãoportuguês/português-alemão de construções com verbo de suporte e colocações*, de Eva Maria Ferreira Glenk – USP; *Uma visão panorâmica da fraseologia no Brasil nos últimos 10 anos*, de Elizabete Aparecida Marques – UFMS; *Vocabulário erótico-obsceno português-italiano e italiano-português*, de Claudia Zavaglia – UNESP; *Dictionnaire en ligne d'expressions idiomatiques du portugais du Brésil et du Portugal et de français de la France et du Québec*, de Claudia Maria Xatará – UNESP; *As expressões idiomáticas no ensino de língua estrangeira e na prática lexicográfica*, dirigido por Maria Luisa Ortiz Alvarez – UnB; *Processamento das unidades fraseológicas e o Estudo teórico-metodológico de conceitos-chave na fraseologia brasileira*, de Rosemeire Selma Monteiro-Plantin – UFC; *Estudo contrastivo de expressões idiomáticas português-espanhol*, de Livia Marcia Tiba Radis Baptista – UFC; *Fraseografia teórica e prática: bases para a elaboração de um dicionário de locuções*, de Maria Eugênia Olímpio Oliveira Silva – UFBA; *Combinatórias léxicas especializadas da linguagem legal, normativa e científica*, Cleci Regina Bevilacqua – UFRGS e; *Dicionário brasileiro de fraseologia*, de José Pereira da Silva – UERJ.

Paralelamente à consulta ao diretório do CNPq, realizou-se uma pesquisa bibliográfica sobre os trabalhos desenvolvidos no campo fraseológico no Brasil e no mundo. Foram, então, selecionados 24 textos, produzidos<sup>46</sup> por 11 autores que, a seu modo e em consonância com a abordagem teórica a que se filiam, tratam do fenômeno fraseológico, focalizando diferentes aspectos das combinatórias sintagmáticas. São textos que abordam majoritariamente a fraseologia na perspectiva da língua comum, embora haja trabalhos, como os de Bevilacqua (1999; 2004), que também discutem a ocorrência e o funcionamento das chamadas fraseologias especializadas, no contexto de domínios temáticos específicos. Ademais, convém esclarecer que esses textos foram publicados

---

<sup>46</sup> Além desses 24 textos, outros trabalhos foram utilizados como leitura complementar.

entre 1988 a 2018, sendo alguns em português, outros em espanhol e os demais em francês.

Dentre os muitos textos obtidos, destacam-se os seguintes:<sup>47</sup>: Alberto Zuluaga (1975, 1980); Antonia María Tristán (1988); Antônio Luciano Pontes (2010-2011); Antonio Pamies (2014); Beatrice Lamiroy (2008); Gaston Gross (1988, 1996); Charles Bally (1951); Cláudia Maria Xatara (1994, 1997, 1998a, 1998b, 2011, 2017); Corpas Pastor (1996, 2001, 2017); Esteban Tomás Montoro Del Arco (2004); Cleci Regina Bevilacqua (1996, 1999, 2004, 2005); Leonor Ruiz Gurillo (1997); Maria Auxiliadora Castillo Carballo (1997-1998); Inès Sfar (2010, 2013, 2015); Maurice Gross (1982, 1993); Marilei Amadeu Sabino (2011); Salah Mejri (1997a, 1997b, 1998, 1999, 2002, 2004, 2005, 2008, 2012, 2018); Maria Luísa Ortiz Alvarez (2000, 2011, 2012a, 2012b); Maria Tereza Camargo Biderman (1999, 2001, 2005); Maria Aparecida Barbosa (2012); Oto Araújo Vale (1999); René G. Strehler (2009); Rosemeire Selma Monteiro-Plantin (2011, 2014, 2017); Stella E. O. Tagnin (1988, 2005, 2011).

Dos autores selecionados, a maioria (cerca de 11) desenvolve pesquisas na área da fraseologia, tomando como base o português brasileiro, seja para efeito de contraste com outra língua, ou não. São estes: Antônio Luciano Pontes, Cláudia Maria Xatara, Cleci Regina Bevilacqua, Marilei Amadeu Sabino; Maria Luisa Ortiz Alvarez, Maria Aparecida Barbosa, Maria Tereza Camargo Biderman, Rosemeire Selma Monteiro-Plantin, Oto Araújo Vale; René G. Strehler; e Stella E. O. Tagnin. Os demais direcionam seus estudos ao espanhol (Zuluaga, Corpas Pastor e Carballo) e à língua francesa (Gaston Gross, Salah Mejri e Beatrice Lamiroy).

Já com relação especificamente a trabalhos que tratem sobre a fraseologia no domínio da política, objeto da pesquisa de mestrado, identificaram-se, até o momento, apenas quatro estudos. Trata-se do artigo *La fraseología en el discurso político y económico de los medios de comunicación*, de Carmem Navarro (2002); dos artigos *A denominação fraseológica no humor e na política* (2005a), *A dinâmica e potencialidade da denominação fraseológica no discurso político e humorístico* (2005b), ambos de autoria de Maria Luísa Ortiz Alvarez; da obra *A CPI das palavras*, de Ari Riboldi (2008) e; do livro *El lenguaje*

---

<sup>47</sup> Por questão de espaço, optou-se por informar os títulos completos dos textos apenas nas referências bibliográficas.

*político: características y análisis del discurso político con ejercicios y clave* (2016), de Luísa Messina Fajardo. O que se sabe, porém, até o presente momento, é que ainda não foi feito nenhum glossário fraseológico com unidades fraseológicas utilizadas no discurso político brasileiro, como o que se produziu nesta pesquisa de mestrado, o qual resulta de abordagem metodológica oriunda da Linguística de *Corpus*.

#### **4.2.2 Constituição e delimitação do *corpus* de estudo**

Tagnin (2005) define *corpus* como coletânea de textos, em formato eletrônico, organizados conforme os objetivos da pesquisa para a qual ele é proposto. Neste trabalho, o *corpus* da pesquisa foi constituído de acordo com a orientação geral da Linguística de *Corpus* (BERBER SARDINHA, 2004), considerando a relação entre esta última e a Fraseologia (TAGNIN, 2011). Desse modo, decidiu-se por compilar um *corpus* de estudo com textos escritos, publicados em língua portuguesa, e disponibilizados na *WEB*, de onde foram copiados e organizados atentando-se para as recomendações metodológicas de Aluisio & Almeida (2006).

Esta fase, por sua vez, foi dividida em duas etapas: levantamento dos textos escritos sobre política e tratamento do *corpus*.

##### **4.2.2.1 Levantamento dos textos escritos**

Para selecionar os textos e as fontes a serem consultadas, estabeleceram-se os seguintes critérios: (i) os textos a serem extraídos deveriam ter sido publicados entre janeiro de 2014 e dezembro de 2016<sup>48</sup>; (ii) o colunista deveria estar vinculado a revistas/jornais de grande circulação no país; (iii) deveria ser, preferencialmente, engajado em discutir assuntos relacionados à política e; (iv) na medida do possível, representar ideologicamente uma ou mais tendência política, de direita, esquerda, centro etc. Ressalte-se que este último critério, pela dificuldade de aferir com precisão a orientação política do colunista,

---

<sup>48</sup> Escolheu-se essa faixa temporal porque ela abrange, pelo menos, dois eventos de forte impacto no âmbito da política brasileira, quais sejam: o início da Operação Lava Jato (março de 2014) e o Impeachment da ex-presidente Dilma Rousseff, no final de 2016.

não foi utilizado como determinante para a seleção de um ou outro jornalista, embora os colunistas escolhidos expressem implícita e/ou explicitamente, nos textos, sua preferência ideológica.

Com base nesses critérios, foram selecionados 04 colunistas vinculados, cada um, aos periódicos revistas *Época*, *Istoé*, *Carta Capital* e jornal *Folha de São Paulo*.

Quanto ao perfil dos colunistas<sup>49</sup>, Mino Carta (**M.C**) ocupa o cargo de diretor da Redação da revista *Carta Capital*. Ricardo Boechat (**R.B**) é apresentador do Jornal da *Band* e da rádio *BandNews FM*, colunista da revista *ISTOÉ* e ganhador de três prêmios Esso. Já Ruth de Aquino (**R.A**) é jornalista com mestrado em Mídia na *London School of Economics* e doutorado sobre Ética. Trabalhou na BBC, foi correspondente em Londres e Paris, editora internacional, diretora de redação e redatora-chefe. Por sua vez, Vladimir Safatle (**V.S**) é professor e doutor em filosofia pela Universidade de São Paulo, atuando também como colunista do Jornal *Folha de São Paulo*.

No quadro abaixo, apresentam-se a relação dos colunistas, o tipo de periódico e a quantidade de textos extraídos da *Web*, para a composição do *corpus*.

Quadro 4 – Relação dos colunistas, periódico e quantidade de textos selecionados

<b>COLUNISTA</b>	<b>PERIÓDICO</b>	<b>QUANTID. DE TEXTOS SELECIONADOS</b>	<b>TOTAL DE TEXTOS</b>
M.C	Carta Capital	151	570
R.B	Istoé	147	
R.A	Época	146	
V.S	Folha de São Paulo	126	

Fonte: elaborado pelo autor.

#### 4.2.2.2 Limpeza e armazenamento dos textos

<sup>49</sup> Essas informações podem ser obtidas diretamente do *blog* ou coluna em que cada um dos autores publica seus textos.

Como dito anteriormente, os textos selecionados foram extraídos direta e integralmente dos *blogs* dos colunistas, sendo copiados um a um, de acordo com o dia, mês e ano da publicação, e depois colados em arquivos individuais no programa *Word*. Em seguida, esses arquivos foram organizados em pastas conforme a quantidade dos comentaristas.

Após a extração dos textos, procedeu-se ao processo de limpeza dos arquivos, visto que traziam imagens, propagandas, anúncios e outras informações não relevantes para o foco deste trabalho. Desta forma, as colunas em *Word* foram, uma a uma, lidas e pré-processadas para serem depois utilizadas pelo *software WordsSmith Tools* (SCOTT, 2008), versão 6.0, que, dentre outras funções, realiza busca semiautomática em grandes quantidades de dados.

## Ilustração 1 – Exemplo de um texto do *corpus* em *Word* antes da limpeza.

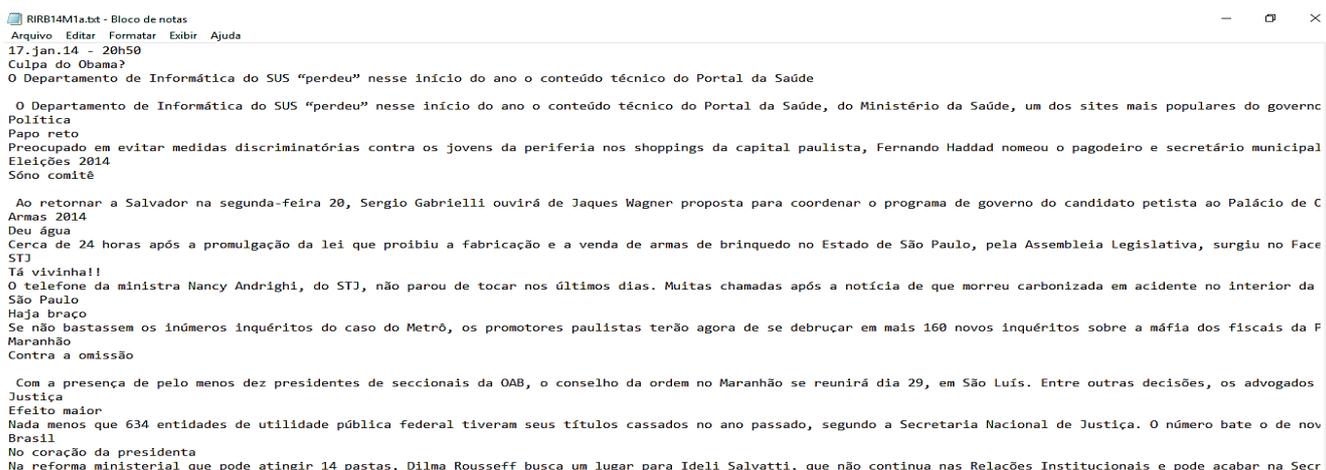
<p>15/01/14 - 2014</p> <h3> Culpa do Obama? </h3> <p>O Departamento de Informática do SUS "perdeu" nesse início do ano o conteúdo técnico do Portal da Saúde</p> <p><b>Governo</b> <b>Culpa do Obama?</b></p>  <p>O Departamento de Informática do SUS "perdeu" nesse início do ano o conteúdo técnico do Portal da Saúde, do Ministério da Saúde, um dos sites mais populares do governo federal. Informações sobre para protocolos sobre como solicitar visitas da Uniao, plantas para Centros de Saúde e dados técnicos para profissionais de saúde sumiram. Quem acessa a página 18 "site em manutenção". Se falta em Brasil a acusaram Obama de <b>hacking</b>, os computadores do <b>Deleuou</b>.</p> <p><b>Política</b> <b>Fape rno</b> Preocupado em evitar medidas discriminatórias contra os jovens de periferia nos shoppings da capital paulista, Fernando Haddad nomeou o pedagogo e secretário municipal Neto do Paula (PCdoB) para dialogar com os marcos dos <b>governos</b>. De o filo na turma, no voto de moçada e com uma missão a cumprir, ele</p>	<p>massacrará a cadeira do venador, hoje com o ex-ministro do Exporto Orlando Silva. Além, os dois com unhas dispostas em uma vaga na Câmara dos Deputados em outubro: um com o legião da classe C, outro com a Copa.</p> <p><b>Crime</b> <b>Da ficção à realidade</b></p>  <p>Quem assistiu há dias <b>hacking</b>, na sala 4DX localizada no shopping JK, em São Paulo, sentiu na pele as aventuras do filme. E que em dado momento, graças a um defeito, todos os efeitos proporcionados pela moderna tecnologia entraram em funcionamento, como o giro brusco das cadeiras, fumaça, respingo de água, vento, sons, etc. Medidos e surpresos, o público recebeu o dinheiro de volta e um ingresso de cortesia. Detalhe: o investimento total nesse tipo de sala ronda US\$ 2 milhões.</p> <p><b>Relações 2014</b> <b>Só no centô</b></p>
 <p>Ano retornar a Salvador na segunda-feira 20, Sérgio Gabriel ouvirá de Jaques Wagner proposta para coordenar o programa de governo de parcerias entre o Palácio de Freixo, Rio Costa, O etc.</p>	<p>testemunhas – igualmente mais. No processo do Judiciário e do Legislativo, fatos irrelevantes ou falsos ganham importâncias importantes.</p> <p><b>São Paulo</b> <b>Maja brago</b> Se não bastassem os inúmeros inquiridos do caso do Mielô, os promotores paulistas terão agora de se debruçar em mais 160 novos inquiridos sobre a máfia dos fiscais da Prefeitura de São Paulo. A artilharia do MP neste caso mira as construtoras e colegas de órgão dos servidores já flagrados.</p> <p><b>Atuação</b> <b>Contra a corrupção</b></p> 

Posteriormente, os textos em *Word*<sup>50</sup> foram convertidos para o formato *txt.*, isto é, texto sem formatação, para poderem ser submetidos à análise semiautomática. Com isso, pretendeu-se identificar padrões de combinações lexicais e sua frequência, além dos contextos de uso.

<sup>50</sup> Os textos em *Word* estão organizados em uma única pasta, contendo toda a configuração antes da conversão para TXT.

A tarefa de conversão para o formato *txt*. é fundamental no processo de tratamento do *corpus*, visto que permite a organização dos arquivos de acordo com a codificação previamente estabelecida de maneira lógica. Isso é necessário para que o programa leia os textos e possibilite a recuperação de informações importantes, vinculadas ao contexto de produção, como a revista em que o texto foi publicado, o autor, o ano, o mês e o dia da publicação. A título de exemplificação, cada texto foi sendo copiado e colado em arquivo separado, obtendo-se a seguinte renomeação, conforme o código estabelecido: R para revista (ou J, no caso do Jornal) + Primeira letra do nome da revista ou jornal+ Iniciais do colunista + ano de publicação + mês + vogais para diferenciar as edições semanais. Assim, para o texto de Ricardo Boechat, publicado em sua coluna, na revista *Istoé*, em 17 de janeiro de 2014, adotou-se o seguinte: RIRB14m1a.

#### Ilustração 2 – Exemplo de texto em formato *TXT* após o processo de limpeza



Criou-se, assim, o *corpus* de estudo da pesquisa, como se vê adiante, constituído de 570 textos que tratam acerca de questões relacionadas à política, seja diretamente ou de forma transversal.

Quadro 5 – Dimensões do *corpus* de estudo

Textos	570
Formas ( <i>tokens</i> )	457, 217
Itens ( <i>types</i> )	37, 188

Fonte: elaborado pelo autor.

### 4.2.3 Seleção do *corpus* de referência

Para fins de extrair a lista de palavras-chave (*KeyWords*) do *corpus* de estudo, houve necessidade de comparar os dados com outro *corpus*, chamado de referência. Segundo Berber Sardinha (2009), o *corpus* de referência (ou de controle) serve para oferecer uma norma que poderá ser comparada às frequências do *corpus* de estudo, demonstrando assim as palavras significativamente mais frequentes no conjunto de textos compilados para a pesquisa. Para tanto, segundo a orientação da Linguística de *Corpus* (BEBER SARDINHA, 2004, 2006; TAGNIN, 2011), tal *corpus* de contraste deve ser pelo menos de três a cinco vezes maior que o *corpus* de estudo.

Dentre as opções de *corpora* disponíveis *on-line*, escolheu-se o *Corpus Brasileiro*<sup>51</sup>, composto por um bilhão de palavras do português brasileiro contemporâneo. O referido *corpus* dispõe ao usuário as frequências de ocorrência dos termos de sua busca e as linhas de concordância onde os termos ocorrem. Além disso, disponibiliza a lista de frequência das palavras do *corpus*, bem como a licença para se obter o *corpus* na íntegra, mediante uma solicitação e cadastro virtual.

O *corpus* brasileiro é composto por textos escritos e falados, conforme se vê no quadro abaixo, extraído de Berber Sardinha (2010):

Quadro 6 – Dimensões do *corpus* de referência.

Componente	<i>Tokens</i>	
Escrito	1.005.163.599	92.4%
Falado	83.055.313	7.6%
Total	1.088.218.912	100%

Fonte: Berber Sardinha (2010, p. 2).

Neste trabalho, porém, utilizou-se somente o componente escrito do *Corpus Brasileiro*, que representa a grande maioria dos dados compilados. Esses textos tratam sobre assuntos diversos (Cinema e TV, Jornalismo, Política,

<sup>51</sup> Para mais informações, acessar o site: <http://corpusbrasileiro.pucsp.br/cb/Inicial.html>.

Academia, Esporte, Informática, Legislação etc.) e integram diferentes gêneros discursivos<sup>52</sup>.

#### 4.2.4 Procedimentos de análise

Nesta seção, descrevem-se os procedimentos de elaboração das listas de frequência e lista de palavras-chave, com o auxílio do *WordSmith Tools*, e as etapas de levantamento de candidatos a fraseologismos e sua posterior validação.

##### 4.2.4.1 Elaboração das listas de frequência e de palavras-chave

Um passo inicial, antes de buscarem-se os prováveis fraseologismos no *corpus*, foi verificar quais palavras eram mais frequentes e quais estavam semanticamente mais relacionadas ao universo em estudo. Para tanto, em relação ao primeiro caso, fez-se o uso do programa *WordsList*, que está integrado ao *software WordSmith Tools*, juntamente com mais dois programas principais: *KeyWords* e *Concord*. Veja-se a ilustração abaixo, que exhibe a tela inicial do *software*.

Ilustração 3 – Tela inicial do programa *WordSmith Tools* 6.0.



<sup>52</sup> Para uma visão mais detalhada do *corpus* e suas divisões, sugere-se a leitura do *Manual Corpus Brasileiro* (BERBER SARDINHA, 2010).

Para gerar a lista de palavras, seguiram-se os passos: clicou-se no ícone *WordList*, que por sua vez exibiu sua própria tela; já na tela deste programa, clicou-se em “file”; depois, “new”; em seguida, selecionaram-se os textos já previamente convertidos para *txt.*, organizados em pasta, e deu-se um “OK”, cujo comando leva a uma nova tela, na qual se clica em “Make a word list now” para finalmente gerar a lista, como se vê na ilustração abaixo.

Ilustração 4 – Janela de trabalho do *WordList*, contendo um exemplo do *corpus* de estudo

N	Word	Freq.	%	Texts	%	Lemmas	Set
1	de	21,270	4.65	570	100.00		
2	a	14,613	3.20	570	100.00		
3	o	11,581	2.53	570	100.00		
4	e	10,481	2.29	570	100.00		
5	que	10,097	2.21	570	100.00		
6	#	8,920	1.95	570	100.00		
7	do	8,365	1.83	569	99.82		
8	da	7,648	1.67	569	99.82		
9	em	5,477	1.20	570	100.00		
10	para	4,326	0.95	565	99.12		
11	não	4,217	0.92	567	99.47		
12	um	4,026	0.88	561	98.42		
13	se	3,977	0.87	566	99.30		
14	é	3,791	0.83	565	99.12		
15	no	3,737	0.82	559	98.07		
16	com	3,636	0.80	560	98.25		
17	na	3,520	0.77	556	97.54		
18	uma	3,164	0.69	548	96.14		
19	os	3,120	0.68	560	98.25		
20	por	2,833	0.62	554	97.19		
21	O	2,648	0.58	527	92.46		
22	dos	2,241	0.49	525	92.11		
23	ao	2,153	0.47	525	92.11		
24	mais	2,149	0.47	540	94.74		
25	A	2,127	0.47	508	89.12		
26	como	2,085	0.46	544	95.44		
27	as	1,785	0.39	526	92.28		
28	à	1,705	0.37	504	88.42		
29	Brasil	1,444	0.32	470	82.46		
30	Dilma	1,320	0.29	333	58.42		
31	das	1,262	0.28	490	85.96		

Como já era de se esperar, as palavras gramaticais (preposições, artigos, conjunções) encabeçaram a *wordlist*, sendo “Brasil” a primeira palavra de conteúdo a aparecer na vigésima nona posição, pela ordem de frequência, seguida da palavra “Dilma”. A posição ocupada pela preposição “de”, liderando a lista acima, vai ao encontro do que observa Finatto (2017, p. 275), para quem: “Na escrita, DE é a palavra de maior frequência de uso no português do Brasil”.

A fim de limpar a *wordlist*, para que se destacassem as palavras de conteúdo, elaborou-se uma *stoplist*, que é um recurso usado para “especificar lista de palavras a serem excluídas do processamento dos textos” (BERBER SARDINHA, 2009, p. 251). Assim, depois de inserida a *stoplist* juntamente com os 570 textos do *corpus*, gerou-se uma nova *wordlist*, demonstrada a seguir:

Ilustração 5 – *wordList* elaborada após inserção da *stoplist*.

N	Word	Freq.	%	Texts	%Lemmas	Set
1	BRASIL	1,450	0.32	471	82.63	
2	DILMA	1,320	0.29	333	58.42	
3	GOVERNO	1,064	0.23	393	68.95	
4	PAÍS	980	0.21	413	72.46	
5	PRESIDENTE	944	0.21	360	63.16	
6	LULA	827	0.18	234	41.05	
7	POLÍTICA	730	0.16	347	60.88	
8	PT	559	0.12	234	41.05	
9	EX	522	0.11	252	44.21	
10	PAULO	522	0.11	257	45.09	
11	ESTADO	513	0.11	252	44.21	
12	MINISTRO	511	0.11	250	43.86	
13	MUNDO	487	0.11	272	47.72	
14	TODOS	485	0.11	288	50.53	
15	CASA	484	0.11	252	44.21	
16	GRANDE	474	0.10	253	44.39	
17	HOJE	465	0.10	286	50.18	
18	PODE	463	0.10	255	44.74	
19	AGORA	457	0.10	301	52.81	
20	BEM	453	0.10	272	47.72	
21	FEDERAL	451	0.10	222	38.95	
22	ERA	450	0.10	259	45.44	
23	PODER	428	0.09	258	45.26	
24	JUSTIÇA	407	0.09	206	36.14	
25	RIO	407	0.09	187	32.81	
26	SERÁ	407	0.09	253	44.39	
27	DEPOIS	400	0.09	248	43.51	
28	FEIRA	396	0.09	195	34.21	
29	TUDO	392	0.09	257	45.09	
30	NADA	391	0.09	260	45.61	
31	MII	384	0.08	212	37.19	

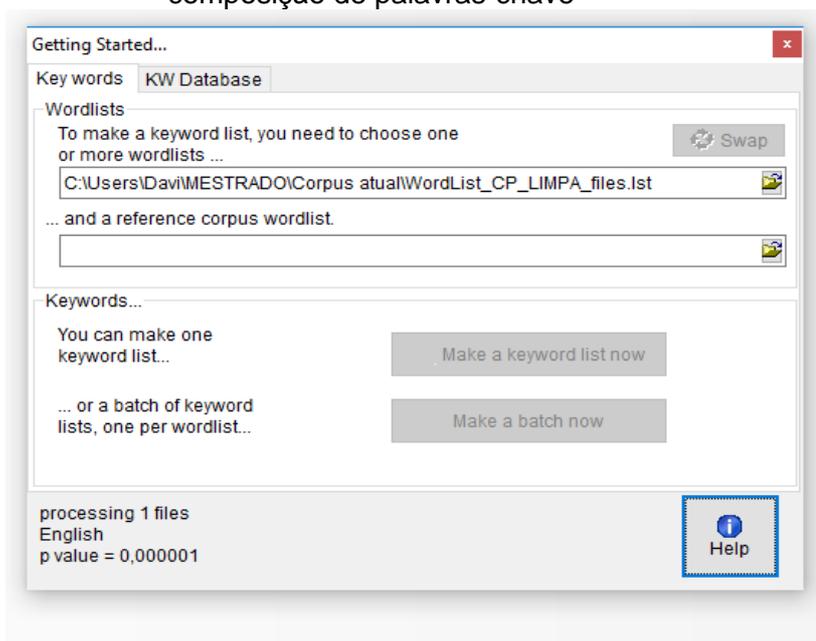
frequency alphabetical statistics filenames notes

32,723 Type-in

Com a retirada de preposições, conjunções, artigos e de símbolos diversos (como o asterisco), o programa gerou uma nova *wordList*, destacando as palavras de conteúdo mais frequentes no *corpus*, as quais são pertinentes ao discurso político.

Após criar a *wordlist* do *corpus* de estudo, o próximo passo foi gerar a lista de palavras-chave, na tela do programa *Keywords*. Para isso, além da lista de frequência do *corpus* de estudo, foi necessária a lista de frequência do *corpus* de referência, uma vez que esse recurso cruza informações de ambas as partes e gera uma relação de palavras que possuem maior frequência no *corpus* em foco do que no de controle.

Ilustração 6 – Tela do *Keywords* indicando a seleção de duas *wordlists* para a composição de palavras-chave



Os passos percorridos podem ser sintetizados nesta ordem: clicar em “file” na tela do programa *KeyWords*; depois clicar em “new”; em seguida, selecionar as *wordLists* do *corpus* de estudo e do de referência (ver. ilustração 6); e, finalmente, acionar o comando “make a keyword list now”. Tendo feito isso, obteve-se uma lista com 500 palavras-chave. Na ilustração a seguir, apresentam-se as 30 primeiras palavras-chave indicadas pela ordem de frequência.

Ilustração 7 – *KeyWords* com as 30 primeiras palavras-chave

KeyWords_LIMPA_files_1st.kws										
File Edit View Compute Settings Windows Help										
N	Key word	Freq.	%	RC. Freq.	RC. %	Keyness	P	Lemmas	Set	
1	DILMA	1,320	0.29	1		17,606.02	0.0000000000			
2	GOVERNO	1,064	0.23	200		11,619.75	0.0000000000			
3	PRESIDENTE	944	0.21	129		10,857.14	0.0000000000			
4	LULA	827	0.18	22		10,705.18	0.0000000000			
5	PAÍS	980	0.21	273		9,806.71	0.0000000000			
6	BRASIL	1,450	0.32	1,324	0.02	8,784.10	0.0000000000			
7	MINISTRO	511	0.11	49		6,126.95	0.0000000000			
8	POLÍTICA	730	0.16	355		6,083.36	0.0000000000			
9	SERÁ	407	0.09	33		4,955.12	0.0000000000			
10	AGORA	457	0.10	91		4,928.75	0.0000000000			
11	ROUSSEFF	371	0.08	1		4,926.36	0.0000000000			
12	PODE	463	0.10	114		4,770.55	0.0000000000			
13	HOJE	465	0.10	117		4,766.64	0.0000000000			
14	ERA	450	0.10	115		4,595.25	0.0000000000			
15	TODOS	485	0.11	181		4,448.17	0.0000000000			
16	MILHÕES	378	0.08	48		4,383.18	0.0000000000			
17	FEDERAL	451	0.10	141		4,366.87	0.0000000000			
18	TEMER	329	0.07	6		4,290.34	0.0000000000			
19	JATO	334	0.07	12		4,270.02	0.0000000000			
20	FEIRA	396	0.09	82		4,236.01	0.0000000000			
21	GOLPE	329	0.07	12		4,203.32	0.0000000000			
22	JUSTIÇA	407	0.09	113		4,069.45	0.0000000000			
23	LAVA	326	0.07	20		4,052.06	0.0000000000			
24	VEZ	344	0.08	43		3,996.00	0.0000000000			
25	FORAM	370	0.08	74		3,984.71	0.0000000000			
26	GRANDE	474	0.10	231		3,943.13	0.0000000000			
27	VAI	376	0.08	87		3,927.49	0.0000000000			
28	DEPOIS	400	0.09	123		3,890.09	0.0000000000			
29	NADA	391	0.09	112		3,877.04	0.0000000000			
30	CRISE	342	0.07	54		3,840.65	0.0000000000			
31	DOIS	366	0.08	90		3,769.82	0.0000000000			

Embora não possa ser considerado um *corpus* específico do domínio político, pelo fato de os colunistas não terem formação específica em Ciências Políticas, a *keywords* acima demonstra a pertinência de palavras que estão, sobretudo no contexto atual, intimamente relacionadas ao universo político, especialmente o brasileiro, como Dilma, Lula, Temer, Jato, dentre outras.

#### 4.2.4.2 Levantamento de candidatos a fraseologismos

Diante das palavras mais frequentes e significativas para o enfoque aqui proposto, passou-se a uma outra etapa fundamental do trabalho, qual seja: o levantamento de candidatos a fraseologismos na plataforma do programa *Concord*. De acordo com Berber Sardinha (2009, p. 11), concordâncias são “recortes do conteúdo de vários textos ao mesmo tempo”. Desse modo, em vez de ler cada texto do *corpus* um por um, o programa *Concord* exhibe todas as ocorrências em contexto para uma determinada palavra ou expressão de busca, tal como observado para a palavra *Dilma*, que obteve 1.320 concordâncias.

Ilustração 8 – Concordâncias para a palavra *Dilma*.

Concord														
File Edit View Compute Settings Windows Help														
N	Concordance	Set	Tag	Word #	Sen	Sen	Para	Para	lead	lead	Sec	Sec	File	%
1	sobre os bastidores da deposição de <b>Dilma</b> Rousseff envolvem cifras que vão			719	3551%	059%					059%		RIRB16M9d.txt	69%
2	da Justiça no governo petista de <b>Dilma</b> Rousseff, Eugênio Aragão.			702	3234%	070%					070%		RIRB16M9b.txt	72%
3	pública. Eleições 2016 Ele & Ela Com <b>Dilma</b> Rousseff fora de combate, Lula			489	2240%	052%					052%		RIRB16M9a.txt	53%
4	que preservaria os direitos políticos de <b>Dilma</b> , sob os auspícios do PMDB, foi o			243	1056%	026%					026%		RIRB16M9a.txt	27%
5	no capítulo final do julgamento de <b>Dilma</b> Rousseff no Senado. Segundo			189	836%	020%					020%		RIRB16M9a.txt	21%
6	ela se acovardou, nada perguntando a <b>Dilma</b> e nem falando na sessão de			558	2748%	050%					050%		RIRB16M9a.txt	60%
7	defesa no Senado na segunda-feira 29, <b>Dilma</b> recebeu apoiadores no Alvorada.			848	4278%	031%					031%		RIRB16M9a.txt	91%
8	OEA cerca de US\$ 27 milhões. Quando <b>Dilma</b> perdeu o cargo, o Conselho			645	3316%	059%					059%		RIRB16M9a.txt	69%
9	na Hora H disse "sim" para a saída de <b>Dilma</b> . OEA Tempo para pensar Dos			580	2830%	052%					052%		RIRB16M9a.txt	62%
10	acontecimentos celestiais que cercam <b>Dilma</b> , "tudo indica o final de um ciclo".			166	675%	019%					019%		RIRB16M8d.txt	19%
11	, na quinta-feira 25 vários retratos de <b>Dilma</b> foram retirados das paredes do			81	343%	0 9%					0 9%		RIRB16M8d.txt	9%
12	A estrela que não brilha Quando <b>Dilma</b> Rousseff colocar os pés no			15	038%	0 2%					0 2%		RIRB16M8d.txt	2%
13	Nova perto de Marte, em sua casa 12. " <b>Dilma</b> Rousseff deverá sofrer um golpe			278	1320%	032%					032%		RIRB16M8d.txt	30%
14	conduzindo o julgamento de <b>Dilma</b> Rousseff no Senado. Mas a			764	4130%	037%					037%		RIRB16M8d.txt	87%
15	individual e sem mensagem". Com <b>Dilma</b> Rousseff condenada, Guedes			424	2223%	048%					048%		RIRB16M8d.txt	47%
16	para a sua saída definitiva do cargo. Se <b>Dilma</b> vai virar uma estrela um dia na			325	1610%	037%					037%		RIRB16M8d.txt	36%
17	comprados em Israel, no Governo <b>Dilma</b> , praticamente ficaram no chão. A			351	1975%	039%					039%		RIRB16M8c.txt	39%
18	como uma pessoa diferente de <b>Dilma</b> . Pode provar o que pensa não			458	2130%	049%					049%		RIRB16M8b.txt	49%
19	contra a presidente afastada <b>Dilma</b> Rousseff, na madrugada da			649	2957%	070%					070%		RIRB16M8b.txt	68%
20	4,5 milhões de caixa 2 na campanha de <b>Dilma</b> , em 2010, que mentiram nos			79	259%	0 8%					0 8%		RIRB16M8a.txt	8%
21	no Congresso terminem quando <b>Dilma</b> Rousseff for de vez afastada da			605	3457%	054%					054%		RIRB16M8a.txt	63%
22	os principais projetos do governo, <b>Dilma</b> Rousseff enfraqueceu os			21	072%	0 2%					0 2%		RIRB16M7c.txt	2%
23	virou uma espécie de ventríloco de <b>Dilma</b> Rousseff. Por onde passa repete			193	737%	017%					017%		RIRB16M7c.txt	16%
24	depois que seu mandato terminar. <b>Dilma</b> Rousseff tem três apartamentos			636	3511%	057%					057%		RIRB16M7b.txt	68%
25	tudo dar certo. Brasil Bateu o martelo <b>Dilma</b> Rousseff não levará nenhum			473	2721%	050%					050%		RIRB16M7b.txt	51%
26	. No terceiro e mais antigo imóvel, <b>Dilma</b> montará um escritório. Nele			708	4073%	075%					075%		RIRB16M7b.txt	75%
27	segundo mandato, um ministro de <b>Dilma</b> Rousseff dava a nítida impressão			836	5033%	039%					039%		RIRB16M7b.txt	89%
28	obras desejam. Nos últimos tempos <b>Dilma</b> ficou craque no Spotify, dando			749	4436%	079%					079%		RIRB16M7b.txt	80%
29	sobre o processo de impeachment de <b>Dilma</b> Rousseff. Uma derrota para o			35	037%	0 4%					0 4%		RIRB16M7a.txt	4%
30	23 anos. Sob ameaça de impeachment, <b>Dilma</b> Rousseff nomeou o pastor José			237	1240%	029%					029%		RIRB16M7a.txt	29%
31	do processo de impeachment de <b>Dilma</b> Rousseff. Assim, decidirá se			101	1937%	038%					038%		RIRB16M6d.txt	37%

concordance collocates plot patterns clusters filenames follow up source text notes

1,320 Set

Além de mostrar as linhas de concordância, o programa permite visualizar termos que coocorrem mais frequentemente com a palavra *Dilma*, são os chamados “colocados”.

No ícone “clusters”, situado na parte inferior, ao lado direito de “patterns”, é possível obter facilmente os agrupamentos entre duas ou mais palavras, tendo como componente o termo de busca. A ilustração abaixo demonstra os *clusters* observados para as seguintes especificações: agrupamento de três palavras e uma frequência mínima de duas ocorrências.

Quadro 7 – Listagem dos 30 primeiros agrupamentos com a palavra *Dilma*

<b>POSIÇÃO DO CLUSTER EM FUNÇÃO DA FREQUÊNCIA</b>	<b>CLUSTERS DE TRÊS PALAVRAS</b>
1	DE DILMA ROUSSEFF
2	IMPEACHMENT DE DILMA
3	LULA E DILMA
4	PRESIDENTE DILMA ROUSSEFF
5	A PRESIDENTE DILMA
6	DILMA ROUSSEFF E
7	A PRESIDENTA DILMA
8	O GOVERNO DILMA
9	DILMA E LULA
10	DE DILMA E
11	QUE DILMA ROUSSEFF
12	DE LULA E
13	O IMPEACHMENT DE
14	DILMA ROUSSEFF A
15	DILMA NÃO É
16	POR DILMA ROUSSEFF
17	PROCESSO DE IMPEACHMENT
18	A DILMA ROUSSEFF
19	COM DILMA ROUSSEFF
20	DE DILMA A
21	DILMA ROUSSEFF NA
22	DILMA E MARINA
23	DA PRESIDENTE DILMA
24	DE IMPEACHMENT DE
25	PRESIDENTA DILMA ROUSSEFF
26	DILMA ROUSSEFF EM
27	DO GOVERNO DILMA
28	DILMA ROUSSEFF NÃO
29	DILMA ROUSSEFF NO

30	CONTRA DILMA ROUSSEFF
----	-----------------------

Fonte: elaborado pelo autor.

Nota-se que o programa não determina, a priori, que todos os agrupamentos sejam fraseologismos, pois, como se pode observar, o par *De Dilma Rousseff* (1), *Dilma e Marina* (22), entre tantos outros agrupamentos, não passam de combinações aleatórias geradas pelo programa em virtude da frequência com o que coocorrem. Por esse motivo, convém observar que a frequência, embora seja um critério importante para delimitar o uso de fraseologismo, não pode ser tomada isoladamente de outras propriedades.

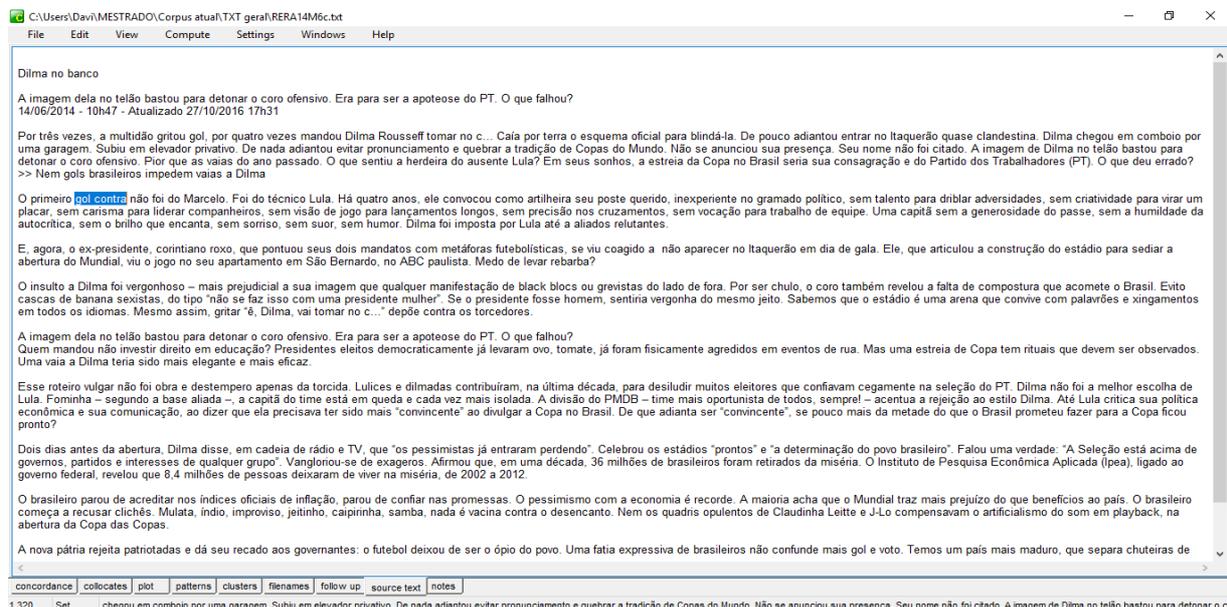
Ressalte-se que, até este momento da análise, muitos dos agrupamentos observados não constituem fraseologismos, conforme os critérios que estão sendo adotados nesta pesquisa. Isso tem feito com que a leitura das linhas de concordância seja muito mais atenta, pois, muitas vezes, há algumas unidades fraseológicas sendo usadas no entorno do contexto em que se encontra o *cluster*, como se exemplifica a seguir.

Ilustração 9 – Concordância para *Dilma*, ordenada pela primeira palavra à esquerda

Concord														
File Edit View Compute Settings Windows Help														
N	Concordance	Set	Tag	Word #	Sen	Sen	Para	Para	lead	lead	Sec	Sec	File	%
60	. Medo de levar rebarba? O insulto a Dilma foi vergonhoso – mais prejudicial			304	2420%		036%				036%		RERA14M6c.txt	36%
61	responsabilidade, o mesmo atribuído a Dilma Rousseff sem prova. Itamar			388	1532%		037%				037%		RCMC16M4c.txt	37%
62	Nem gols brasileiros impedem vaias a Dilma O primeiro gol contra não foi do			154	1653%		018%				018%		RERA14M6c.txt	17%
63	de então, uma agenda diferente. Caso a Dilma consiga recuperar gradativamente			289	2218%		022%				022%		RCMC15M9c.txt	21%
64	vai ter faixa contra o Temer, contra a Dilma, contra o Lula, a favor do Lula,			688	4720%		035%				035%		RERA16M8b.txt	87%
65	parte abandonada para deixar espaço a Dilma Rousseff. Quem supõe que, ao			493	2836%		033%				033%		RCMC14M5c.txt	85%
66	nossa grotesca tradição. Caberia a Dilma partir de imediato para o mesmo			218	913%		026%				026%		RCMC16M1c.txt	25%
67	das redes. Uma irreverência útil a Dilma, pois tirava o foco de um bando			110	715%		014%				014%		RERA15M1a.txt	14%
68	a resignação. A quem mais, senão a Dilma, compete salvar o País? Creio			371	1634%		044%				044%		RCMC16M1c.txt	42%
69	de desabafo que Michel Temer enviou a Dilma Rousseff, narrando o seu			16	055%	0 3%				0 3%		RIRB15M12c.txt	2%	
70	se arrepender de não ter permitido a Dilma se afundar em seu próprio			646	3552%		078%				078%		RERA16M4c.txt	78%
71	como crime 02/09/2016 02h07 Coube a Dilma Rousseff a oração fúnebre da			18	072%	0 2%				0 2%		JFVS16M9a.txt	2%	
72	, não deve ser atribuído apenas a Dilma. Mas não chamem de "acidente"			417	2230%		053%				053%		RERA15M11c.txt	54%
73	Será que alguém esqueceu de avisar a Dilma Rousseff? Eleições 2016			280	1530%		045%				045%		RIRB15M11d.txt	43%
74	que votarão pelo adeus definitivo a Dilma Rousseff. Cultura Sem estresse			860	3636%		032%				032%		RIRB16M6c.txt	92%
75	mais afinado com a elite. Mas a Dilma é a mais preparada. Fico triste			4,048	25850%		057%				057%		RCMC14M6a.txt	67%
76	de condições. O que eu dizia e a Dilma diz: nós governamos este País			2,242	14427%		033%				033%		CMC14M10b.txt	85%
77	uma vantagem abnorme em relação a Dilma Rousseff. Segundo aspecto do			171	637%		018%				018%		CMC14M10a.txt	18%
78	mesmo protagonizar a oposição a Dilma, Cunha tenha "rompido" com o			208	1271%		026%				026%		RERA15M7d.txt	27%
79	está em disputa nesta eleição. Não é a Dilma contra o Aécio, não é PSDB			1,786	11128%		056%				056%		CMC14M10b.txt	69%
80	os 3,4 milhões de votos que deram a Dilma Rousseff vitória sobre Aécio			17	058%	0 3%				0 3%		RIRB15M4e.txt	3%	
81	Jacques Wagner tem insistido junto a Dilma Rousseff para ela suspender o			409	2025%		056%				056%		RIRB15M12c.txt	67%
82	focado em Lula. Nenhuma menção a Dilma Rousseff. De olho no governo do			312	1433%		052%				052%		RIRB14M7a.txt	61%
83	mobilizam atos diários de apoio a Dilma. Mesmo que os eventos reúnam			684	3930%		076%				076%		RIRB16M4b.txt	77%
84	1 Pressão em Washington Graças a Dilma Rousseff, a ex-ministra Ideli			458	2538%		045%				045%		RIRB16M5b.txt	45%
85	em 2018, com o intuito de abater Dilma pelo caminho, e o obstáculo			974	4753%		037%				037%		RCMC16M4d.txt	97%
86	pronto? Dois dias antes da abertura, Dilma disse, em cadeia de rádio e TV,			584	4535%		059%				059%		RERA14M6c.txt	68%
87	que, ainda hoje, com o país no abismo, Dilma negocia, em troca de votos de			214	1127%		026%				026%		RERA16M4b.txt	27%
88	, que nunca votou no PT, mas achou Dilma "muito humana". A produtora			982	3933%		038%				038%		RIRB16M10a.txt	98%
89	crise com o Legislativo, se aconselha Dilma Rousseff a ter paciência com os			315	1134%		031%				031%		RIRB15M8a.txt	31%
90	Não vão condaná-lo? DD- Não acredita. Dilma saiu de governo, diminuiu a carga			1 630	12518%		077%				077%		RCMC16M11a.txt	77%

Nas concordâncias arroladas acima, salta aos olhos o *cluster* “a Dilma”, o qual, numa primeira observação, é muito improvável que chegue a formar um fraseologismo. Todavia, se lida com atenção, a linha 62 apresenta próximo ao *cluster* “a Dilma” o sintagma *gol contra*, o qual, apesar de ser oriundo do domínio do futebol, alcança sentido diferente e cristalizado no uso geral da língua. Para confirmar a hipótese, basta clicar duas vezes na linha 62 e assim poder visualizar o texto, em *txt*. em que *gol contra* é utilizado.

### Ilustração 10 – Visualização do fraseologismo *gol contra* no seu contexto real de uso, exibida na plataforma do *Concord*.



#### 4.2.4.3 Validação dos fraseologismos identificados no corpus

Esta etapa fundamental da pesquisa demandou a adoção de procedimentos que, embora tenham tornado o trabalho um pouco mais árduo, visaram assegurar o seu caráter científico. Em um primeiro momento, todas as sequências selecionadas diretamente do *corpus* de estudo foram submetidas aos critérios previstos na teoria fraseológica, particularmente os critérios de fixidez e congruência propostos por Salah Mejri (1997, 2012).

Posteriormente, devido ao fato de o *corpus* não ser estritamente especializado, por razões já explicadas alhures, e, sobretudo, pelo enfoque deste estudo não ser terminológico, mas fraseológico, foram adotados dois procedimentos para validar os fraseologismos identificados, como se descreve a seguir.

Realizou-se primeiramente uma consulta a dicionários gerais de língua portuguesa e a dicionários especiais fraseológicos, a fim de se atestar o registro das sequências selecionadas. Constituíram o *corpus* de consulta: o Novo Dicionário Aurélio (versão eletrônica, publicada em 2004), o Dicionário Houaiss (versão eletrônica, publicada em 2009), a versão on-line do Dicionário Michaelis, ancorada no sítio (<http://michaelis.uol.com.br/>), o Dicionário da Língua

Portuguesa (versão on-line), o Dicionário on-line de expressões, o Dicionário brasileiro de fraseologia (SILVA, 2013) e o Dicionário do futebol (SALVADOR, 2017). Para os objetivos desta dissertação, considerou-se como válida, isto é, admitiu-se como fraseologismo a unidade que se adequa aos critérios linguísticos da análise e que possui registro, com definição ou não, nessas obras de referência.

Dessa primeira etapa de consulta às obras de referência, obtiveram-se, entre as sequências para as quais não se encontrou registro, dois grupos: a) unidades fraseológicas de uso geral ou de outros domínios que não o político e b) fraseologismos considerados serem pertencentes ao discurso político, apesar do reconhecimento de que o fenômeno político possui caráter interdisciplinar (CHARAUDEAU, 2006).

Assim, para validar as sequências do primeiro grupo, elas foram submetidas ao teste de reconhecimento por um grupo de 4 (quatro) colaboradores, falantes nativos de português, sendo dois do sexo masculino, dois do sexo feminino, distribuídos equitativamente com relação ao nível de escolaridade (ensino médio completo e ensino superior completo). Os sujeitos, após concordarem e darem ciência de sua participação na pesquisa, mediante assinatura do Termo de Compromisso Livre e Esclarecido, receberam em suas mãos ou por *e-mail* uma lista com esses fraseologismos, sem o contexto de ocorrência, apenas com as combinatórias lematizadas e organizadas alfabeticamente. Os colaboradores deveriam reconhecer (ou não) as sequências, atribuindo o(s) significado(s) conforme sua competência fraseológica (MEJRI, 2012). A coincidência nas respostas deveria ser de pelo menos 50% para que a unidade pudesse ser considerada validada<sup>53</sup>.

Quadro 8 – Perfil dos colaboradores do **grupo A** que validaram os fraseologismos de uso geral e de outros domínios diferentes do político

GRUPO A				
código do colab.	sexo	idade	profissão	nível de escolaridade

<sup>53</sup> Em nenhum momento durante a pesquisa, os colaboradores trocaram informações a respeito das unidades que estavam avaliando. Na verdade, todos moram distantes uns dos outros e não mantêm contato entre si.

RR	fem.	36	Professora de séries iniciais	Ensino superior completo – graduada em Pedagogia
EP	fem.	30	Autônoma	Ensino médio completo
DM	masc.	35	Professor de Geografia	Ensino superior completo – graduado em Geografia pela UEPA
DC	masc.	35	Mecânico	Ensino médio completo

Fonte: elaborado pelo autor.

Em se tratando dos fraseologismos que integraram o segundo grupo, isto é, o de unidades consideradas pertencentes especificamente ao domínio da política, a validação foi realizada por outros colaboradores, com perfil diferente dos demais. Para tanto, estabeleceu-se contato com dois cientistas políticos e um cientista social, todos falantes nativos de português, formados pela Universidade Federal do Pará. Esses colaboradores especialistas receberam uma lista de fraseologismos do discurso político, dissociados do contexto de uso real, para que procedessem ao teste de reconhecimento e pudessem, também, formular as definições técnicas ou científicas. Ademais, tais unidades foram submetidas a um pequeno *corpus* de caráter mais especializado formado apenas por textos de cientistas políticos e dicionários políticos, como o dicionário de Política, organizado por Norberto Bobbio, Nicolau Matteucci e Gianfranco Pasquino, publicado pela editora da UNB em 1998.

Quadro 9 – Perfil dos colaboradores do **grupo B** que validaram os fraseologismos do discurso político

<b>GRUPO B</b>				
<b>código do colab.</b>	<b>sexo</b>	<b>idade</b>	<b>Profissão</b>	<b>formação acadêmica</b>
MN	masc.	26	Analista ambiental	Bacharel em Ciências Sociais e Mestre em Ciência Política, todos pela UFPA
AM	masc.	28	Prof. de sociologia	Bacharel em Ciências Sociais

				e atualmente graduando em Filosofia pela UFPA
JS	masc.	25	Prof. de sociologia	Graduado em Ciências Sociais e Mestre em Ciência Política, todos pela UFPA

Fonte: elaborado pelo autor.

#### 4.2.5 Elaboração do glossário fraseológico

O glossário produzido destina-se ao público em geral de falantes de português brasileiro que lidam ou se interessam pela política. Não é, portanto, um repertório direcionado para especialistas, embora estes também possam estar incluídos naquele grande público.

Sua confecção ocorreu em três fases, como se descreve a seguir.

##### 4.2.5.1 Organização macroestrutural do glossário

O glossário foi produzido em duas versões (uma impressa e outra eletrônica). Trata-se de um repertório monolíngue com unidades fraseológicas utilizadas **no** discurso político. Em princípio, constam na sua nomenclatura todos os fraseologismos identificados no *corpus*, incluindo, portanto, sequências da língua comum e unidades que circulam mais frequentemente no discurso político, especializados ou não. Assim, mesmo que a frequência seja de apenas uma ocorrência, como nos casos a serem discutidos a seguir, tal fraseologismo figura como entrada no glossário se seu uso produzir um efeito expressivo no texto no bojo das discussões políticas.

A nomenclatura está organizada alfabeticamente, registrando-se as unidades pela forma lematizada, salvo nos casos em que o próprio fraseologismo não admite a lematização. Neste aspecto, concorda-se com a observação de Welker (2011), a qual defende que não se devem registrar formas inexistentes no uso.

##### 4.2.5.2 Organização microestrutural do glossário

A microestrutura dos verbetes, baseando-se nas orientações de Montoro Del Arco (2004) e Welker (2011), e adaptando-as para os objetivos desta pesquisa, apresenta o seguinte esquema:

#### Esquema 1 – Microestrutura do glossário

ENTRADA + CATEGORIA GRAMATICAL (estrutura sintagmática) + DEFINIÇÃO (±Referência) + CONTEXTO (Referência) ± VARIANTE FRASEOLÓGICA ± REMISSIVA ± NOTA.

Fonte: elaborado pelo autor.

Os elementos com o símbolo + são obrigatórios, ao passo que os demais, identificados pelo sinal ±, dependerão do fraseologismo que estiver sendo registrado, já que nem todos apresentam variantes, por exemplo.

A *entrada*, também conhecida como lema, encabeça o verbete e é constituída pela própria unidade fraseológica repertoriada. Diferentemente dos dicionários gerais, que tradicionalmente registram os fraseologismos em subentradas, no interior do verbete do vocábulo-base da combinatória, os repertórios fraseológicos costumam inserir os fraseologismos na nomenclatura, na forma lematizada do primeiro vocábulo da unidade. Adotou-se tal orientação no glossário confeccionado, mas convém esclarecer e justificar algumas escolhas metodológicas.

Em primeiro lugar, procurou-se ser o mais fiel possível à maneira e à forma como o fraseologismo ocorre nos textos coletados. Assim, embora se saiba, pela consulta às obras de referências ou ao registro em outros textos correntes da língua portuguesa, que a forma mais cristalizada de uma determinada unidade se apresenta sob uma configuração consagrada, preferiu-se registrar o fraseologismo exatamente como ele ocorre no *corpus*. Desse modo, registrou-se a forma <adotar dois pesos e duas medidas>, e não a outra variante mais institucionalizada, <ter dois pesos e duas medidas>, pois esta última não ocorreu no *corpus*, mas sua existência foi mencionada, no verbete daquela, em nota. O mesmo procedimento foi adotado para o registro da

sequência <assalto de colarinho branco>, em vez de <crime de colarinho branco>.

Por outro lado, há casos em que “completou-se” a sequência com um dos componentes ausentes, procurando-se, sempre, nestes casos, a forma mais cristalizada do fraseologismo. Por esse motivo, embora só tenha ocorrido a forma *faces da moeda* no *corpus*, a competência fraseológica (MEJRI, 2012; ORTIZ ALVAREZ, 2014) justifica a inclusão do elemento elíptico. Desse modo, a forma registrada foi < faces da **mesma** moeda>.

Quanto à *categoria gramatical*, ela informa, no caso do glossário, a estrutura sintagmática das unidades fraseológicas, estando logo após a entrada. Utilizaram-se, para esse campo do verbete, as abreviações *sv.*, *sn.*, *sadj.*, *sp.*, para indicar respectivamente sintagma verbal, sintagma nominal, sintagma adjetival e sintagma preposicional.

A *definição*, por sua vez, exprime o significado ou conceito dos fraseologismos. No caso das unidades registradas nas obras de referência, transcreveu-se a definição apresentada, mencionando-se a fonte logo após o enunciado definitório, entre parênteses. Já para os fraseologismos sem registro nas obras consultadas, os quais foram submetidos à validação pelos grupos de colaboradores, elaborou-se, a partir das definições fornecidas pelos sujeitos, um enunciado definitório que mantivesse o significado cristalizado da sequência, observando-se o funcionamento e o efeito de sentido provocado nos textos quando do uso do fraseologismo. Além disso, buscou-se, sempre que possível, estabelecer um paralelo entre a categoria gramatical da primeira palavra da sequência e o início da definição redigida. Assim, para o fraseologismo <**andar** na contramão>, elaborou-se a seguinte definição, também iniciada por verbo: **agir** de maneira contrária ao que se espera.

Por seu turno, o *contexto*, constituído por um excerto extraído do próprio *corpus*, desempenha função fundamental ao abonar o uso do fraseologismo, demonstrando assim como tal unidade é utilizada no universo repertoriado. No glossário produzido nesta dissertação, o contexto exerce um papel crucial, pois é ele quem melhor justifica a realização de um repertório lexical sobre fraseologismos utilizados no discurso político. De fato, muitas das unidades inventariadas, sobretudo as que se situam na língua comum, já foram objeto de outros trabalhos e descrições, mas aqui assumem um colorido especial ao se

integrarem ao discurso político, e só o contexto autêntico pode oferecer o funcionamento real dessas unidades nos textos coletados.

Em decorrência disso, deve-se justificar o porquê de alguns trechos serem um pouco mais extensos do que se costuma notar em outros trabalhos lexicográficos/terminográficos/fraseológicos. Nos textos analisados, observou-se, por exemplo, o emprego de uma sequência para se referir a toda uma situação descrita em um parágrafo inteiro. Desse modo, preferiu-se, em certos casos, recortar um trecho relativamente extenso, a fim de evitar frases truncadas que não permitissem ao leitor uma condição melhor para compreender o fraseologismo. Por esse motivo, no verbete da sequência <a cobra vai fumar>, consta o trecho a seguir:

<A cobra vai fumar>? O Tribunal de Justiça de SP julgará na quarta-feira 28 uma ação histórica – 20 anos depois. A Philip Morris e a Souza Cruz estão no banco dos réus. Trata-se da primeira ação coletiva em que se inverteu o ônus da prova. Elas precisam provar que o cigarro não causa dependência e a sua publicidade não é abusiva e enganosa. A Associação de Defesa da Saúde dos Fumantes quer reconhecida a responsabilidade das empresas pelos danos a seus consumidores. Vencedora, as portas da Justiça abrem-se para fumantes e ex-fumantes buscarem indenizações por danos morais e patrimoniais contra a indústria do tabaco (RIRB15M1a).

Ao final de cada excerto utilizado como contexto, menciona-se a fonte de onde foi retirado o texto. Tal referência encontra-se codificada do seguinte modo: R para revista ou J para jornal + Primeira letra do nome do periódico + Iniciais do colunista + ano de publicação + mês + vogais para diferenciar as edições semanais. Assim, para o texto de Ricardo Boechat, publicado em sua coluna, na revista *Isto é*, em 17 de janeiro de 2014, adotou-se: RIRB14m1a.

Já as *variantes fraseológicas*, pertinentes ao caráter dinâmico da língua e, conseqüentemente, dos fraseologismos, dizem respeito às diferentes formas em que a unidade pode se manifestar, sem perder de vista seu significado fraseológico. Constituem, conforme assinala Montoro Del Arco (2004)<sup>54</sup>, alternativas léxicas. Levam-se em conta, também, as observações de Mejri (2012) quanto à existência de variantes que não revogam o caráter cristalizado

---

<sup>54</sup> O autor estabelece as seguintes premissas para as variantes fraseológicas: a) não podem apresentar diferenças de significado; b) são livres, independentes do contexto; c) são parcialmente idênticas em sua estrutura e em seus componentes e; d) sua substituição é fixa (MONTORO DEL ARCO, 2004).

do sentido e de outras que apresentam nuances semânticas e de níveis da língua. Algumas, por exemplo, embora sejam variantes, são frequentemente empregadas em contextos formais ou informais, decorrendo da situação de comunicação possíveis nuances.

No tocante ao sistema de remissivas, ele está associado às relações semânticas e conceituais entre os fraseologismos, podendo indicar hipônimos, antônimos etc. Apesar de o glossário em pauta não ter sido organizado onomasiologicamente, utilizaram-se as remissivas para agrupar semanticamente as sequências inventariadas. Dessa maneira, como se referem a situações de apuros, os fraseologismos <estar com a corda no pescoço> e <estar no olho do furacão> remetem um ao outro.

Por seu turno, as *notas* servem para indicar ou destacar informações relevantes sobre o fraseologismo. Podem, então, estar relacionadas a aspectos de ordem sintática, semântica, pragmática ou cultural.

#### 4.2.5.3 O uso do *Lexique Pro* na organização do glossário

O programa *Lexique Pro* tem sido muito útil na elaboração de repertórios lexicais, diminuindo sobremaneira o trabalho extenuante de quem produz dicionário ou glossário, auxiliando-o na organização macro e microestrutural do repertório. Trata-se de um *software* de distribuição livre, desenvolvido pelo SIL (*Summer Institute of Linguistics*) com o intuito de “criar bases de dados, gerenciar arquivos e gerar documentos em formato de dicionário para *Word* ou para *Web*, conforme a escolha do pesquisador” (LIMA & MARTINS, 2014, p. 259).

O funcionamento do programa obedece a comandos inteligentes, possuindo uma plataforma interativa. Utiliza um sistema de etiquetas pré-definidas com as quais o pesquisador deve se familiarizar a fim de preencher, no campo adequado, as informações relativas à estruturação do verbete. Neste aspecto, o *software* constitui um recurso indispensável na elaboração do glossário de fraseologismos utilizados no discurso político, visto que auxiliou fundamentalmente no preenchimento da ficha fraseológica, que foi alimentada na própria plataforma do programa.

Após instalada a versão 3.6 do *software* (SIL, 2004-2012) e criado o glossário, o primeiro passo foi alimentá-lo com as sequências cristalizadas

coletadas no *corpus*. Posteriormente, as demais informações que constam no verbete foram adicionadas, de acordo com a etiquetagem pré-definida. Uma das vantagens do programa é que, enquanto não se fecha o léxico, é possível continuar editando-o. O sistema de etiquetas adotado reflete a microestrutura estabelecida para os verbetes do glossário, como se vê:

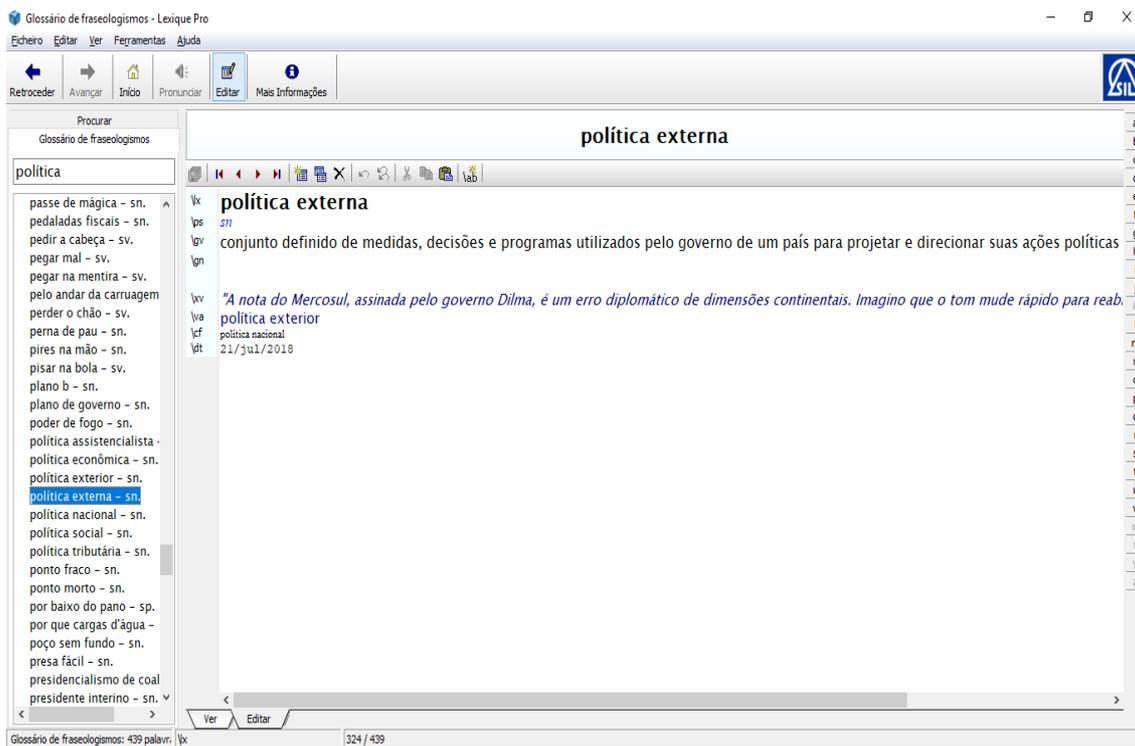
Quadro 10 – Etiquetas para o verbete principal.

etiquetas do <i>Lexique Pro</i>	campos correspondentes do verbete
\lx – lexeme	Entrada
\ps – part of speech	categoria gramatical
\gv – gloss	enunciado definatório
\xv – example	contexto retirado do corpus (FONTE)
\va – variant form	variante fraseológica
\cf – cross reference	remissiva
\nt – notes (general)	notas em geral
\pc – picture	imagem
\dt – date	data da última edição

Fonte: elaborado pelo autor.

Na ilustração abaixo, ilustra-se a composição do verbete para a sequência *política externa*.

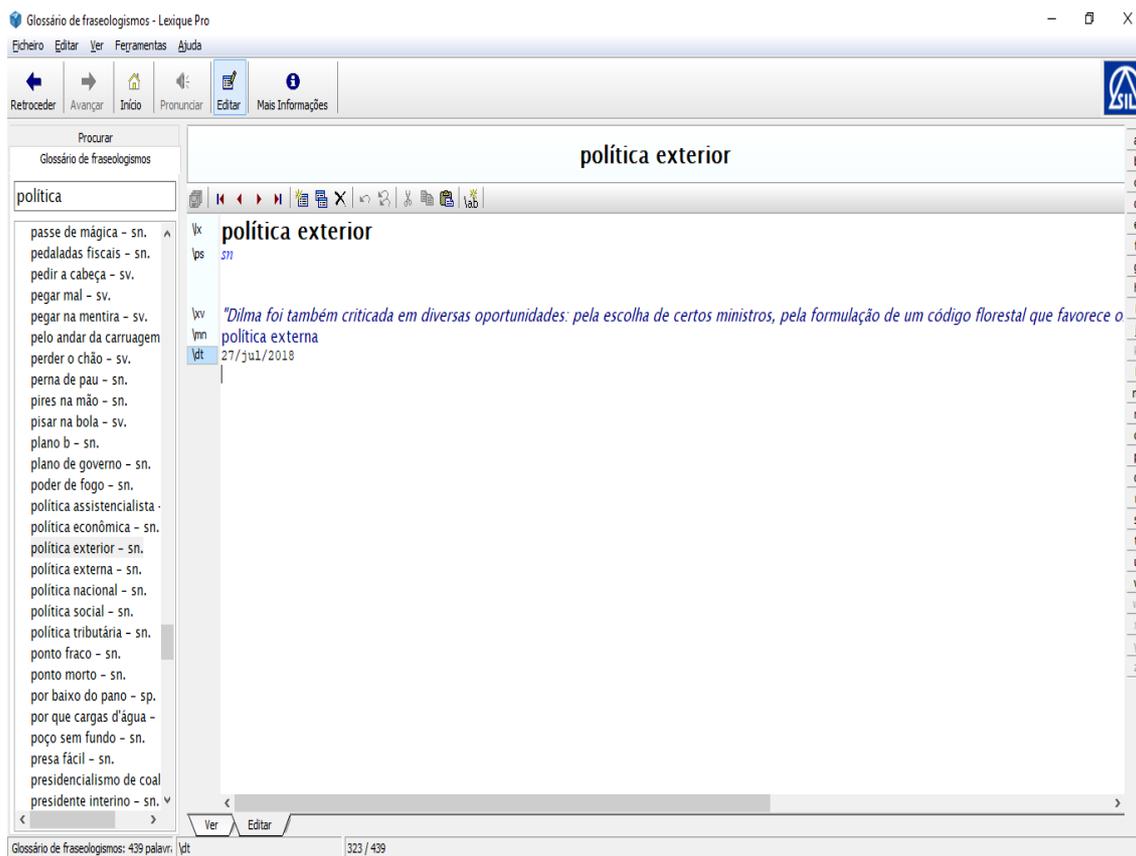
Ilustração 11 – Verbetes de *política externa* no modo editar.



Fonte: extraído da plataforma do Lexique Pro.

No caso de verbetes para variantes, dispensa-se a etiqueta `\lx` (definição), adotando-se, por outro lado, a etiqueta `\mn` (*main entry cross reference*), a qual é utilizada para fazer referência cruzada de uma variante secundária a uma entrada principal. Na versão eletrônica do glossário, essas remissivas funcionam como *hiperlinks*, com os quais o usuário poderá ir de uma entrada a outra com apenas um clique. A seguir, tem-se um exemplo de uma entrada composta por variante.

Ilustração 12 – Verboete de *política exterior*, no modo editar.



Fonte: extraída do Lexique Pro.

À medida que se incluem as informações necessárias, o programa vai automaticamente organizando a macroestrutura do glossário, dispondo as entradas em ordem alfabética, no canto esquerdo do léxico. Ao fim ou durante o processo de elaboração do repertório, o programa disponibiliza a opção de exportar o léxico como documento em formato *doc.* ou diretamente para a WEB, bastando, para isso, adotar os seguintes passos: abrir, no menu de ferramentas do *software*, o ficheiro e clicar na opção “exportar como Documento...”, seguindo-se posteriormente algumas telas que permitirão configurar a maneira como o repertório irá aparecer no texto em *Word* (se em uma ou duas colunas, com fontes e cores variadas, tamanho e espaçamento de letras etc.).

Por tudo isso, o *Lexique Pro* demonstrou-se um recurso extremamente importante na produção do glossário, otimizando tempo e reduzindo dispêndio de energia por parte do pesquisador. Constitui, de fato, um programa indispensável para quem produz repertórios lexicais.

## CAPÍTULO 5 – DISCUSSÃO E ANÁLISE DOS RESULTADOS

Neste capítulo, apresentam-se os resultados da análise, com a descrição geral das sequências cristalizadas coletadas no *corpus* e, no volume II, o glossário de fraseologismos utilizados no discurso político.

### 5.1 DESCRIÇÃO GERAL DOS RESULTADOS

A partir dos critérios apresentados por Gaston Gross (1996) e fundamentalmente com base nos critérios de fixidez e de congruência, propostos por Mejri (2012), foram identificadas **438** sequências cristalizadas, as quais exibem graus distintos de cristalização sintática, semântica e pragmática, idiomaticidade, frequência de uso e congruência.

Embora as fronteiras discursivas entre o uso da língua geral e os domínios especializados não sejam rigidamente delineadas, sendo muito frequente o fluxo entre uma instância e outra, tentou-se identificar unidades próprias ou características do discurso político e unidades da língua comum e de domínios diferentes do político. Assim, verificou-se que, do total de 438 sequências, apenas **80** foram consideradas de uso próprio ou mais natural ao discurso político, como *acordo de delação premiada, ajuste fiscal, analfabeto político, assalto de colarinho branco, bancada da bala, base aliada, base do governo, governo de esquerda, cabo eleitoral*, dentre outras. A ilustração a seguir exhibe, em forma de nuvem, os componentes lexicais dessas 80 sequências cristalizadas, com destaque para a palavra política, que aparece em 13 unidades.

Ilustração 13 – Visualização em nuvem dos componentes lexicais dos fraseologismos que remetem à política.



inventariar fraseologismos utilizados **no** discurso político. Além disso, como assinalou Charaudeau (2006), o fenômeno político é de natureza interdisciplinar, situando-se no cruzamento de fatos pertencentes a outros campos, como as ciências políticas, as ciências sociais, o direito, a filosofia etc. A política, portanto, alcança muitos aspectos da vida social, estabelecendo, nesse percurso, diferentes relações com outras áreas.

Esse movimento e dinamicidade das fraseologias transitando de um uso mais geral a outro(s) mais especializado(s), e vice-versa, é absolutamente comum, natural e compreensivo dentro de uma visão de língua também dinâmica e heterogênea, adotada no presente estudo.

Com isso, corrobora-se a afirmação de Ortiz Alvarez & Unternbäumen (2011, p. 7) de que a “fraseologia descreve o mundo real, as experiências quotidianas, o colorido e a sabedoria de um povo e por isso torna-se um importantíssimo veículo de identidade e de cultura”. Com efeito, o caudal fraseológico está a serviço das múltiplas necessidades de comunicação e expressão, diversificando-se e transitando entre diferentes domínios.

### **5.1.1 Quanto ao registro nos dicionários de língua portuguesa**

Das 438 unidades identificadas, um pouco mais da metade (231) não se encontra registrada nos cinco dicionários de referência da língua portuguesa utilizados para consulta. Desse universo de unidades sem registro, 67 remetem ao universo político. No entanto, as 207 sequências que possuem registro em dicionários gerais não constituem entradas nessas obras, mas estão localizadas no interior dos verbetes da palavra considerada principal, como se vê na ilustração abaixo.

Ilustração 14 – Consulta para *bode expiatório* na versão eletrônica do dicionário Aurélio

The screenshot shows the 'Novo Dicionário Aurélio' website. The search bar contains 'bode' and the results are displayed in a list. The entry for 'bode' is highlighted in yellow. The entry includes 15 numbered definitions and three idiomatic expressions. The definitions are:

1. O macho da cabra; cabrão. [Sin., poét.: capro.]
2. Caprino em geral.
3. Pej. V. *mestiço* (3).
4. Caixinha envernizada, em geral para guardar dinheiro.
5. Matula, farnel.
6. Confusão, complicação, encrenca.
7. V. *rolo*<sup>1</sup> (16).
8. Fig. Homem muito feio.
9. Bras. Indivíduo libidinoso; sático.
10. Bras. N.E. Pop. V. *protestante* (6).
11. Pop. Cada uma das figuras do baralho, esp. o valete.
12. Gír. Ant. Mil-réis.
13. Bras. Gír. Estado depressivo, ou de sonolência, provocado por droga (4), ou não.
14. Bras. Gír. Situação embaraçosa, difícil, complicada ou deprimente.
15. Bras. Papagaio (5) grande.

The idiomatic expressions are:

- ◆ **Bode expiatório.** 1. Pessoa sobre quem se faz recair as culpas alheias ou a quem são imputados todos os reveses.
- ◆ **Amarrar o bode.** Bras. Fam. 1. Ficar de cara amarrada; ficar sério. 2. Ficar mal-humorado.
- ◆ **Amarrar um bode.** 1. Bras. Gír. Deprimir-se, ger. sob efeito de droga (4); entrar em fossa (6).
- ◆ **De bode amarrado.** 1. Bras. Fam. Amuado, mal-humorado: *estar, andar, ficar de bode amarrado.*

The search results also include a list of related words (bochornoso, bócio, boco, bocó<sup>1</sup>, bocó<sup>2</sup>, bocório<sup>1</sup>, bocório<sup>2</sup>, boçoroca, boçu, bocuba, bocuda, bocué, bocuiabá, bocuuvaçu, boda, bodar, bodas, bode<sup>1</sup>, bode<sup>2</sup>, bodeado, bodeco, bodega, bodegão, bodegueiro, bodeguice, bodeguim, bodejar, bodejo, bodelha, bode-preto, bodiano, bodião<sup>1</sup>, bodião<sup>2</sup>, bodião-papagaio, bodião-sabonete) and a historical section (bode<sup>2</sup>, bode<sup>1</sup>, b, boca<sup>1</sup>, costas).

Fonte: Novo Dicionário Aurélio (2004) – versão eletrônica.

Observa-se que o fraseologismo em pauta encontra-se como subentrada após as 15 acepções registradas para a palavra “bode”.

Esse tratamento lexicográfico dispensado às unidades fraseológicas tem sido frequentemente discutido e criticado principalmente por fraseólogos. Welker (2011) apresenta uma reflexão interessante a esse respeito. Veja-se, também, Mel’cuk (2001), que defende o registro dos frasemas nos dicionários gerais da mesma maneira que é feito com relação às unidades simples.

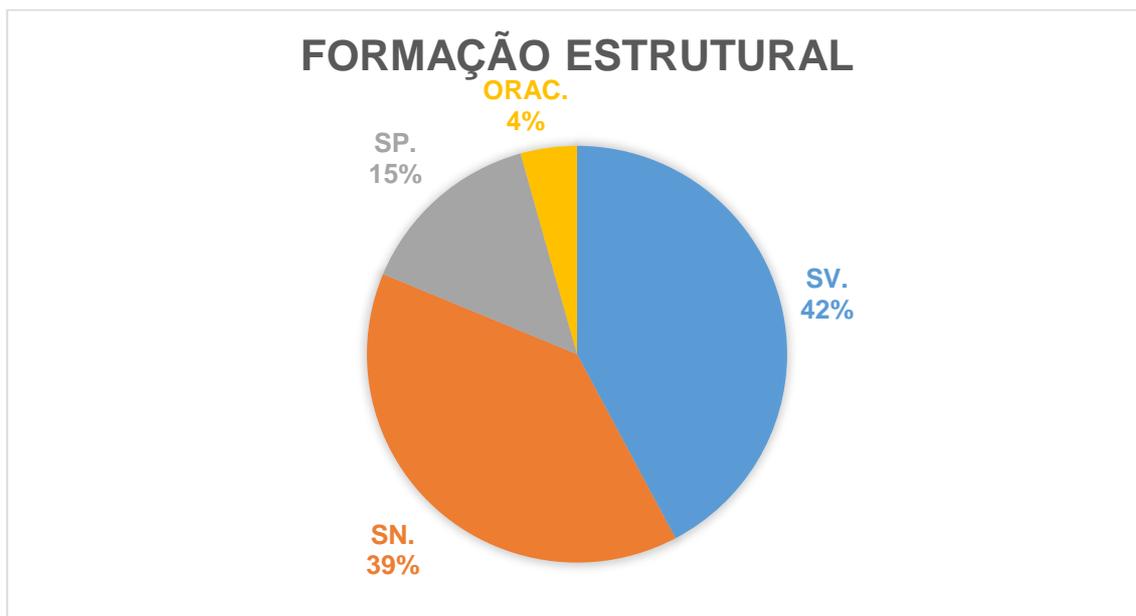
Quanto à lacuna observada nos dicionários, no que se refere à falta de registro das 232 unidades, pode estar relacionada a dois motivos. O primeiro diz respeito ao problema nevrágico da lexicografia brasileira que, tradicionalmente, centraliza sua atenção no vocábulo simples, deixando de repertoriar unidades polilexicais; e quando o faz, geralmente trata as combinatórias sintagmáticas como subentradas, inserindo-as no verbete de uma das palavras considerada a mais importante, por razões nem sempre claramente justificadas.

O segundo motivo a ser apontado refere-se ao fato de muitas das unidades ausentes serem associações sintagmáticas ligadas temática e funcionalmente a um domínio mais especializado, cujas formas de expressão muitas vezes representam e veiculam conhecimento específico. Por essa razão, comumente ficam de fora dos dicionários de língua geral.

### 5.1.2 Padrões formais de combinação

Quanto à estrutura combinatória das sequências, notou-se a predominância de unidades de base verbal, seguida por fraseologismos de base nominal e preposicional, sendo menos frequente as de caráter oracional e textual. O gráfico abaixo resume percentualmente esse resultado.

Gráfico 1 – Estrutura sintagmática dos fraseologismos



Fonte: elaborado pelo autor.

Entre as sequências de base verbal, predomina a estrutura sintática de verbo e objeto. Há sintagmas formados por verbo + substantivo (*abrir mão, bater boca, bater cabeça*); por verbo + sintagma nominal (*apertar o cinto, baixar a bola, cair a ficha, fechar os olhos, tomar as rédeas, pagar o pato, dar as cartas, virar as costas*); por verbo + sintagma preposicional (*pisar na bola, mandar às favas, pular do barco, estar no vermelho, bate às portas, mergulhar de cabeça*); verbo + sintagma nominal + adjetivo (*lavar a roupa suja*); verbo + sintagma nominal + sintagma preposicional (*colocar o bode na sala*).

Quanto as de base nominal, observou-se que elas são formadas por substantivo, adjetivo e locução adjetival. A maioria é constituída por sintagmas formados por substantivo + adjetivo, como: *base aliada, bode expiatório, cabo eleitoral, rabo preso, dinheiro sujo, condução coercitiva, classe política, cartão vermelho, Governo interino, Governo tucano, saia justa, política econômica, política exterior, política externa, política nacional, política social, ponto morto, presa fácil, presidente interino, programas sociais, reforma política, terreno minado*.

Outros sintagmas nominais são formados por dois substantivos, em que o segundo desempenha função adnominal em relação ao primeiro substantivo, como: *República Bananeira, plano B, classe C*.

Notam-se, também, combinações maiores formadas por: substantivo + sintagma preposicional + adjetivo (*acordo de delação premiada, Chefe da Casa Civil, Lei da Ficha Limpa, assalto de colarinho branco*); por substantivo + substantivo + adjetivo (*Operação Sangue Negro*); substantivo + verbo + substantivo (*Operação Lava Jato*); substantivo + advérbio + sintagma preposicional (*carta fora do baralho*); substantivo+sintagma adverbial (*conversa para boi dormir*).

No caso das sequências que são sintagmas preposicionais, observou-se estruturas constituídas por: preposição + substantivo (*a rodo*); preposição + adjetivo (*à larga*); preposição + sintagma nominal (*à deriva, a céu aberto, a duras penas, ao deus-dará, na marra, na corda bamba*); preposição + sintagma nominal + sintagma preposicional (*à beira de um abismo*)

Igualmente, embora tenham sido em menor quantidade, as unidades de valor oracional também apresentam diferentes estruturas. Algumas estabelecem relações sintáticas de coordenação, com orações sindéticas (*a justiça tarda, mas*

*não falha; morde e assopra*), assindéticas (*escreveu, não leu, o pau comeu*), verbos em construções perifrásticas (*a cobra vai fumar*), verbos copulativos (*a voz do povo é a voz de Deus*), entre outras estruturas, como *quem com ferro fere com ferro será ferido*.

### 5.1.3 Distribuição dos fraseologismos em agrupamentos temáticos

Embora o glossário não tenha sido organizado onomasiologicamente, a maioria das unidades fraseológicas podem ser organizadas em agrupamentos temáticos, pois alguns domínios apresentaram uma produtividade significativa de sequências cristalizadas. No entanto, para efeito de ilustração, deve-se enfatizar que essa distribuição geral leva em conta uma orientação referencial, mencionada por Ortiz Alvarez (2000)<sup>55</sup> como uma das possibilidades de cotejamento das expressões idiomáticas. Desse modo, o critério da organização em agrupamentos lexicais depende da inclusão, nas sequências, de um ou mais elementos que fazem referência a determinado domínio.

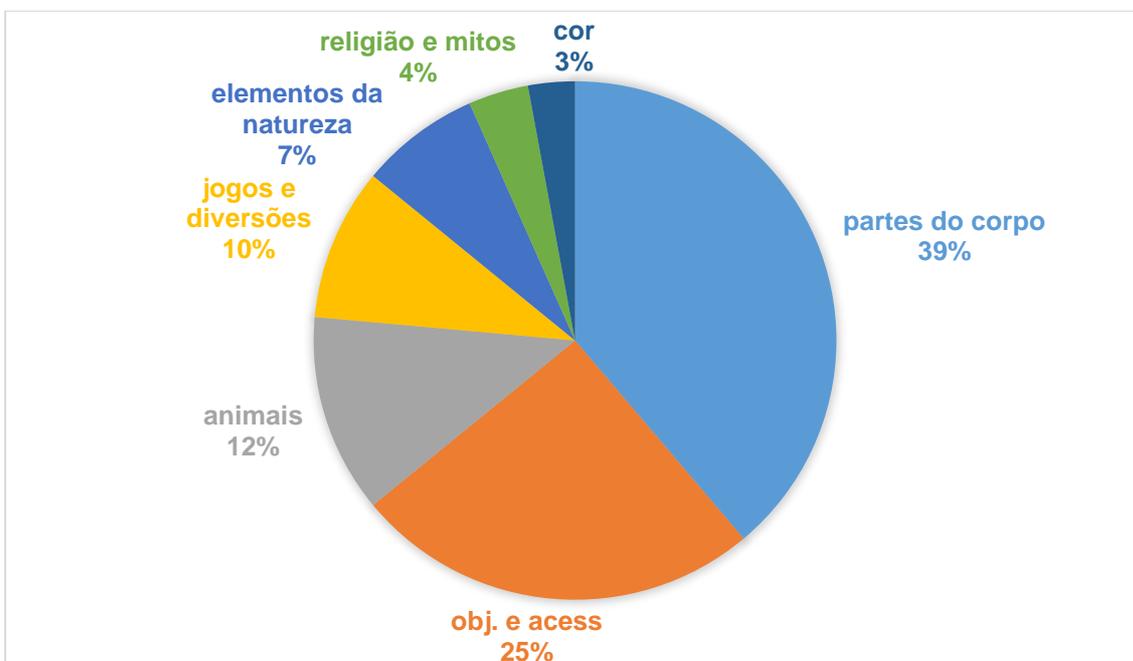
A adoção desta orientação referencial justifica-se, neste trabalho, apenas para efeito de uma primeira forma de sistematização das unidades fraseológicas analisadas. Em outras palavras, a presente dissertação não nega o fato de o fraseologismo comumente apresentar um significado global, em geral opaco, o que exigiria um tratamento semântico mais adequado à sua natureza idiomática, conforme defendem Ortiz Alvarez (2000) e Fulgêncio (2014). Contudo, para os objetivos desta seção de resultados, a sistematização por agrupamentos temáticos mostrou-se útil, pois permitiu demonstrar, de uma maneira geral, a produtividade dos agrupamentos como base de referência para a criação e uso de fraseologismos. Além disso, conforme ressalta Mejri (1997, 1998), as sequências cristalizadas possuem uma dupla estruturação semântica, que possibilita dois tipos de decodificação, sendo um o mais frequentemente utilizado, baseado no significado global da sequência, e o outro que recupera os significados individuais dos componentes lexicais da combinatória. Segundo o autor, este último caso é sempre possível, uma vez que o fraseologismo possui

---

<sup>55</sup> A autora distingue ainda uma orientação semântica, a qual “engloba uma organização em torno de um conceito mais geral e abrangente e que sugere uma interpretação semântica das Eis (habilidade/ inteligência; ignorância/ conhecimento, etc.)” (ORTIZ ALVAREZ, 2000, p. 175).

uma espécie de memória lexical (MEJRI, 1997, 1998). O gráfico abaixo permite uma visualização dos agrupamentos temáticos envolvidos e o uso de fraseologismos. A título de ilustração, serão demonstrados e comentados os fraseologismos dos três agrupamentos que se mostraram mais produtivos.

Gráfico 2 – Distribuição dos fraseologismos por agrupamentos temáticos.



Fonte: elaborado pelo autor.

Curiosamente, o domínio *partes do corpo* contém a maior quantidade de fraseologismos dentre os demais agrupamentos temáticos. São 93 unidades que, na sua composição lexical, fazem alusão a 19 partes diferentes do corpo, como se vê na ilustração 15.

Ilustração 15 – Nomes relativos a partes do corpo que compõem fraseologismos.



Fonte: elaborado pelo autor.

Da parte do corpo que serviu de base metonímica e metafórica para as sequências cristalizadas, da mais produtiva a menos produtiva, tem-se: para **mão** (*abrir mão, botar a mão no fogo, cair nas mãos de, com a mão na massa, conhecer na palma da mão, de mão beijada, de mão em mão, fazer justiça pelas próprias mãos, jogar nas mãos de, lançar mão de, meter a mão, mão de ferro, mãos à obra, nas mãos de, passar a mão na cabeça, pires na mão*); para **olho** (*de olho, de olho aberto, de rabo de olho, fechar os olhos, menina dos olhos, olhar com o rabo do olho, olho clínico, olho gordo, olho no olho, saltar aos olhos, ver com bons olhos*); para **cabeça** (*bater cabeça, de ponta-cabeça, dor de cabeça, fazer a cabeça de, mergulhar de cabeça, passar a mão na cabeça, passar pela cabeça, pedir a cabeça*); para **cara** (*cara de pau, cara lavada, dar a cara a tapa, dar com a cara na porta, dar na cara, livrar a cara, mostrar a cara*), para **pé** (*ao pé da letra, ao pé do ouvido, bater o pé, dar um tiro no pé, de pé, tiro no pé, um pé cá outro lá*); para **pescoço** (*atolar-se até o pescoço, carne de pescoço, envolver-se até o pescoço, estar com a corda no pescoço, livrar o pescoço*); para **braço** (*braço direito, braço forte, cair nos braços de, cruzar os braços, queda de braço*); para **sangue** (*com gosto de sangue na boca, dar o sangue, ferver o sangue, irmão de sangue, Operação Sangue Negro*); para **costas** (*jogar nas costas, nas costas, virar as costas, virar as costas para*); para

**boca** (*bater boca, com gosto de sangue na boca, morrer pela boca*); para **ouvido** (*ao pé do ouvido, as paredes têm ouvidos, dar ouvido a*); para **barba** (*nas barbas de, pôr as barbas de molho*); para **dedo** (*dar os anéis para conservar os dedos, ficar chupando o dedo*); para **ombro** (*carregar nos ombros, dar de ombros*); para **perna** (*bater pernas, perna de pau*); para **testa** (*testa de ferro*); para **pele** (*sentir na pele*); para **língua** (*sem papas na língua*); para **goela** (*meter goela abaixo*); para **barriga** (*empurrar com a barriga*); para **peito** (*do peito*); para **coração** (*cortar o coração*); para **calcanhar** (*calcanhar de aquiles*); para **cabelo** (*até a raiz dos cabelos*); para **medula** (*até a medula*) e; para **dente** (*armado até os dentes*).

Algumas dessas sequências utilizam mais de um componente que faz referência ao corpo, razão por que constam tanto em um grupo quanto em outro. É o caso, por exemplo, *de passar a mão na cabeça, olhar com o rabo do olho, ao pé do ouvido, com gosto de sangue na boca*, entre outros.

Nota-se, também, que, em determinados casos, a variante mais informal possui preferência na produção de fraseologismos. No *corpus* de estudo, não se encontrou, por exemplo, nenhuma unidade fraseológica formada com a palavra *rosto*<sup>56</sup>. No entanto, a sua variante “cara” mostrou-se muito produtiva, constituindo 7 (sete) sequências cristalizadas (*cara de pau, cara lavada, dar a cara a tapa, dar com a cara na porta, dar na cara, livrar a cara, mostrar a cara*).

Quanto ao segundo agrupamento temático mais produtivo, ou seja, o de *objetos e acessórios*, trata-se de um conjunto muito variado de coisas que inclui instrumentos e ferramentas de trabalho, peças de vestuário, utensílios de decoração doméstica, móveis, munições e tipos de embarcação. Na ilustração 16, ilustram-se os 61 fraseologismos agrupados nesse conjunto.

#### Ilustração 16 – Nomes de objetos e acessórios que formam fraseologismos.

---

<sup>56</sup> Deve-se ressaltar que essa observação diz respeito especificamente ao *corpus* utilizado, pois, na língua comum, há também uma infinidade de sequências cristalizadas com a palavra *rosto*. Silva (2013) registra, por exemplo: *maçã do rosto, folha de rosto, amassar/ganhar o pão com o suor do rosto, lançar em rosto, ter o rosto liso como bundinha de nenê* etc.



Fonte: elaborado pelo autor.

Do mais produtivo ao menos frequente, registrou-se: para **bola** (*pisar na bola, baixar a bola, bola da vez, bola fora, dar bola a, levantar a bola*); para **porta** (*dar com a cara na porta, a portas fechadas, bater à porta de, bater às portas*); para **tapete** (*puxar o tapete, varrer para baixo do tapete, estender o tapete vermelho*); para **moeda** (*casa da moeda, faces da mesma moeda, moeda de troca*); para **colarinho** (*assalto de colarinho branco, colarinho branco*); para **barco** (*tocar o barco para a frente, pular do barco*); para **parede** (*as paredes têm ouvido, encostar na parede*); para **cortina** (*cortina de ferro, cortina de fumaça*); para **telhado** (*telhado de vidro*); para **martelo** (*bater o martelo*); para **cinto** (*apertar o cinto*); para **bala** (*bancada da bala*); para **barril** (*barril de pólvora*); para **tecla** (*bater na mesma tecla*); para **foice** (*briga de foice*); para **berço** (*berço político*); para **pá** (*pá de cal*); para **luva** (*dar com luva de pelica*); para **saia** (*saia justa*); para **bastão** (*passar o bastão*); para **roupa** (*lavar roupa suja*); para **toalha** (*jogar a toalha*); para **lençol** (*em maus lençóis*); para **carteirinha** (*de carteirinha*); para **foice** (*briga de foice*) e; para **canoas** (*canoas furadas*).

Algumas poucas unidades organizadas em torno do agrupamento *objetos e acessórios* apresentam em seus componentes lexicais palavras que remetem

a outros agrupamentos, como o de *partes do corpo*. Por exemplo: *dar com a cara na porta e as paredes têm ouvido*.

Por sua vez, o agrupamento em torno das figuras de animais mostrou-se igualmente relevante para a produção de fraseologismos, como se vê na ilustração a seguir.

Ilustração 17 – Visualização em nuvem de palavras do reino animal que se mostraram produtivas em relação aos fraseologismos coletados no *corpus*.



Fonte: elaborado pelo autor.

Adotando-se a mesma orientação anterior, seguem os fraseologismos do agrupamento sobre animais, destacando-se as palavras mais produtivas, em ordem decrescente. Assim, registraram-se: para **rabo** (*olhar com o rabo do olho, república dos rabos presos, de rabo de olho, rabo preso*); para **bode** (*bode expiatório, colocar o bode na sala, tirar o bode da sala*); para **gato** (*fazer gato e sapato de, gato escaldado, gato por lebre*); para **vaca** (*ir a vaca para o brejo, nem que a vaca tussa, voltar à vaca-fria*); para **mosca** (*acertar na mosca, ser picado pela mosca azul*); para **papo** (*fim de papo, papo reto*); para **pato** (*pagar o pato*); para **cobra** (*a cobra vai fumar*); para **vira-lata** (*complexo de vira-lata*);

para **boi** (*conversa para boi dormir*); para **fera** (*cutucar a fera com vara curta*); para **sapo** (*engolir sapo*); para **lebre** (*gato por lebre*); para **tucano** (*governo tucano*); para **leão** (*ser dose para leão*); para **cão** (*não despertar o cão que dorme*); para **tetas** (*mamar nas tetas públicas*) e; para **caranguejo** (*república dos caranguejos*).

De fato, assim como os dois primeiros grupos temáticos mencionados, o campo zoológico desempenha um papel relevante no processo de formação de unidades fraseológicas. Revela-se profícuo como fonte de relações metafóricas, para expressar diversos efeitos de sentido que realçam uma das funções mais importantes da linguagem, isto é, sua expressividade. Observou-se, por exemplo, sequências compostas por mais de uma palavra referente a animal, como *gato por lebre*, cuja forma plena e mais cristalizada é certamente *comprar gato por lebre* (ser enganado). Nota-se, ainda, o uso da figura de uma ave específica, considerada símbolo do Partido da Social Democracia (PSDB), para designar os governos presididos por candidatos filiados a esse partido. Trata-se da sequência *governo tucano*, em que o nome do animal funciona como modificador do substantivo, atuando como traço de identidade ideológica.

## 5.2 CARACTERIZAÇÃO DAS UNIDADES FRASEOLÓGICAS

### 5.2.1 Quanto à polilexicalidade

Sendo a polilexicalidade uma das propriedades fundamentais dos fraseologismos, observou-se, entre as unidades coletadas, uma relação inversamente proporcional: as sequências com menos componentes são exatamente as mais produtivas, enquanto as unidades maiores são as menos frequentes. A menor sequência é formada por dois componentes, e a maior unidade, por nove componentes.

Gráfico 3 – Relação entre a estrutura polilexical e as ocorrências no *corpus*.



Fonte: elaborado pelo autor.

Como se vê no gráfico 3, do ponto de vista da polilexicalidade, à medida que a quantidade de componentes da unidade aumenta, diminui a frequência desse fraseologismo no *corpus*. No caso específico deste trabalho, os fraseologismos mais produtivos foram os que possuem estrutura sintagmática composta por três elementos (*a céu aberto*, *a duras penas*, *aliviar a barra*, *baixar a bola*), seguidos pelos que têm dois elementos (*base aliada*, *bater boca*, *cabo eleitoral*, *caixa dois*). Por outro lado, as sequências mais extensas só tiveram uma ocorrência, como por exemplo *quem com ferro fere com ferro será ferido* e *a voz do povo é a voz de Deus*.

### 5.2.2 Quanto à fixidez

Conforme já se viu anteriormente, a fixidez, tal como proposta por Mejri (2012), pode ocorrer tanto na sintaxe, semântica e pragmática, como nos eixos sintagmático e paradigmático, descrevendo um *continuum*, indo da sequência mais fixa à menos fixa, e vice-versa.

Com respeito à fixidez sintática, que age no plano da combinatória sintagmática, pôde-se observar a existência de fraseologismos que apresentam diferentes graus de restrição, como demonstrado nos quadros abaixo.

Escolheram-se apenas alguns exemplos de unidades verbais, para demonstrar a aplicação dos critérios adotados que basearam a seleção dos fraseologismos que compõem o glossário. Os símbolos \* (asterisco) e  $\pm$ , utilizados no início das sequências, indicam, respectivamente, que *a reestruturação sintática (ou paradigmática) não é permitida pelas regras de congruência, e que é possível que a reestruturação encontre respaldo no uso, em determinados contextos específicos*<sup>57</sup>.

Quadro 11 – Análise de fraseologismos verbais quanto à fixidez sintática.

Fraseologismos	acréscimo de determinantes e de intensificadores	acréscimo de modificadores
<b>abrir mão</b>	*abrir <i>uma/a</i> mão *abrir <i>muito a</i> mão	*abrir mão <i>pesada</i> *abrir <i>pesada</i> mão
<b>bater boca</b>	*bater <i>uma/a</i> boca *bater <i>muito a</i> boca	*bater boca <i>grande</i> *bater <i>grande</i> boca
<b>bater cabeça</b>	*bater <i>uma/a</i> cabeça *bater <i>muito a</i> cabeça	*bater cabeça <i>grande/pequena</i> *bater <i>grande/pequena</i> cabeça
<b>bater pernas</b>	*bater <i>umas/as</i> pernas *bater <i>muito as</i> pernas	*bater pernas <i>bonitas/grandes</i> *bater <i>bonitas/grandes</i> pernas
<b>dar bola a</b>	*dar <i>uma/a</i> bola a dar <i>muita</i> bola	*dar bola <i>vazia a</i> *dar [ <i>uma</i> ] <i>vazia</i> bola a
<b>dar palpite</b>	dar <i>um</i> palpite $\pm$ dar <i>muito</i> palpite	$\pm$ dar [ <i>um</i> ] palpite <i>bom/ruim</i> $\pm$ dar [ <i>um</i> ] <i>bom/ruim</i> palpite
<b>dar sinal</b>	dar <i>um/o</i> sinal $\pm$ dar <i>muito</i> sinal	*dar sinal <i>bom</i> $\pm$ dar <i>bom</i> sinal
<b>engolir sapo</b>	*engolir <i>um/o</i> sapo engolir <i>muito</i> sapo	*engolir sapo <i>vivo/morto</i>
<b>tirar vantagem</b>	tirar <i>uma</i> vantagem *tirar <i>a</i> vantagem tirar <i>muita</i> vantagem	*tirar vantagem <i>boa</i> $\pm$ tirar <i>boa</i> vantagem

<sup>57</sup> Como os objetivos desta pesquisa não previam a realização de teste específico com falantes para averiguar as possibilidades de reestruturações sintáticas, sintagmáticas e paradigmáticas nos fraseologismos, as mudanças indicadas como possíveis se baseiam no conhecimento experiencial do autor, na condição de falante de português e estudioso do assunto, assim como nas leituras de outros trabalhos e obras de referência. Longe de serem um empecilho para o alcance do objetivo principal da dissertação, essas observações empíricas constituem dados e problemáticas a serem investigadas com mais profundidade em pesquisas posteriores.

<b>lançar mão de</b>	*lançar <i>uma/a</i> mão de ±lançar <i>muita</i> mão de	*lançar mão <i>pequena/delicada/suja</i> *lançar <i>pequena/delicada/suja</i> mão
<b>lavar dinheiro</b>	*lavar <i>um/o</i> dinheiro ±lavar <i>muito</i> dinheiro	lavar dinheiro <i>sujo</i> *lavar <i>sujo</i> dinheiro *lavar dinheiro <i>novo</i> *lavar <i>novo</i> dinheiro
<b>levar vantagem</b>	±levar <i>uma/a</i> vantagem levar <i>muita</i> vantagem	*levar vantagem <i>boa</i>  levar <i>boa</i> vantagem
<b>tomar medidas</b>	tomar <i>umas/algumas/as</i> medidas  tomar <i>muitas</i> medidas	tomar medidas <i>necessárias</i> ±tomar <i>necessárias</i> medidas
<b>trocar farpas</b>	trocar <i>umas</i> farpas ±trocar <i>as</i> farpas trocar <i>muitas</i> farpas	±trocar farpas <i>afiadas</i> ±trocar <i>afiadas</i> farpas
<b>ter juízo</b>	*ter <i>um/o</i> juízo ±ter <i>muito</i> juízo	*ter juízo <i>bom/ruim/mau</i> ter <i>bom</i> juízo *ter <i>ruim/mau</i> juízo
<b>virar moda</b>	*virar <i>uma/a</i> moda *virar <i>muita</i> moda	*virar moda <i>boa</i> ±virar <i>boa</i> moda
<b>virar poeira</b>	*virar <i>uma/a</i> poeira *virar <i>muita</i> poeira	*virar poeira <i>cósmica</i>
<b>virar pó</b>	*virar <i>um/o</i> pó *virar <i>muito</i> pó	*virar pó <i>branco</i> *virar <i>branco</i> pó
<b>correr solto</b>	±correr <i>muito</i> solto	±correr solto e <i>livre</i> correr <i>livre</i> e solto
<b>falar grosso</b>	±falar <i>muito/mais</i> grosso	*falar grosso e <i>alto</i> *falar <i>alto</i> e grosso
<b>fazer jus</b>	±fazer <i>muito</i> jus	*fazer jus e <i>correto</i> ±fazer <i>correto</i> e jus
<b>pagar caro</b>	pagar <i>muito</i> caro	*pagar caro e <i>alto</i> ±pagar <i>alto</i> e caro
<b>pegar mal</b>	pegar <i>muito</i> mal	±pegar mal e <i>feio</i> ±pegar <i>feio</i> e mal
<b>ser dose</b>	ser <i>uma</i> dose ±ser <i>a</i> dose *ser <i>muita</i> dose	*ser dose <i>boa/ruim</i> *ser <i>boa</i> dose

Fonte: elaborado pelo autor.

Das sequências verbais analisadas no quadro acima, convém destacar um fato que se estende a todas as unidades: os verbos não estão fixos em apenas um modo, tempo e pessoa verbais, mas admitem diferentes flexões. No

tocante ao acréscimo de determinantes e intensificadores, percebeu-se que as unidades apresentam maior restrição aos primeiros, enquanto as modificações decorrentes da introdução de intensificadores são menos restritas, independentemente da natureza do verbo.

Quanto aos modificadores, nota-se que a maioria das sequências interdita a introdução de um adjetivo, seja como parte do mesmo sintagma (\*abrir mão *pesada*), seja como elemento coordenado (\*falar grosso e *alto*). A posição do adjetivo deve ser levada em conta em alguns casos, como em *lavar dinheiro*, em que o acréscimo do adjetivo *sujo* só é aceitável em posição posposta ao substantivo (*lavar dinheiro sujo*). Contrariamente, em *tirar vantagem*, é a posposição do adjetivo *boa* que não parecer ser aceitável (\*tirar vantagem *boa*), sendo sua anteposição muito mais natural e cristalizada (*tirar boa vantagem*). O mesmo se aplica à variante *levar vantagem*.

Com relação à fixidez paradigmática, que apresenta restrições no eixo das comutações lexicais, utilizaram-se como exemplos as mesmas unidades descritas acima quando se tratou da fixidez sintática. Segue o quadro com algumas (im)possibilidades de substituição lexical.

Quadro 12 – Análise de fraseologismos verbais quanto à fixidez paradigmática.

Fraseologismos	comutações paradigmáticas	fraseologismos	comutações paradigmáticas
<b>abrir mão</b>	*abrir <i>pé/palma</i>  * <i>separar/distender</i> mão	<b>tomar medidas</b>	tomar <i>providências</i> * <i>pegar/arrancar/assumir</i> medidas ± <i>adotar</i> medidas
<b>bater boca</b>	*bater <i>lábios/pé/cabeça/olho</i>  * <i>colidir/atingir/golpear</i> boca	<b>trocar farpas</b>	* <i>trocar pontas metálicas penetrantes/estilhas de madeira que acidentalmente se introduz na pele ou na carne</i>  * <i>permutar/inverter/substituir</i> farpas
<b>bater cabeça</b>	*bater <i>pé/olho/coração</i> * <i>colidir/atingir/golpear</i> cabeça	<b>ter juízo</b>	* <i>ter julgamento/opinião</i>  ±ter tino ±possuir juízo
<b>bater pernas</b>	*bater <i>braço/perna</i>  * <i>colidir/atingir/golpear</i> pernas	<b>virar moda</b>	* <i>virar uso</i> ± <i>virar costume/padrão</i> ± <i>tornar-se/transformar-se [em]</i> moda

<b>dar bola a</b>	*dar ? *doar bola	<b>virar poeira</b>	virar <i>pó</i> ± <i>tornar-se/transformar-se [em] poeira</i>
<b>dar palpite</b>	dar <i>opinião</i> *doar/oferecer palpite	<b>virar pó</b>	virar <i>poeira</i> ± <i>tornar-se pó</i> <i>transformar-se [em] pó</i>
<b>dar sinal</b>	±dar <i>indício/pista</i> *doar/oferecer sinal	<b>correr solto</b>	±correr <i>livre</i> *apressar-se/andar solto
<b>engolir sapo</b>	*engolir <i>rã/pereca/anfíbio</i> *deglutir/tragar/comer sapo	<b>falar grosso</b>	*falar <i>grave</i> *dizer/pronunciar/proferir grosso
<b>tirar vantagem</b>	*tirar <i>benefício/ganho</i> tirar <i>proveito</i> *extrair/arrancar/puxar vantagem	<b>fazer jus</b>	*fazer <i>direito</i> *criar/construir jus
<b>lançar mão de</b>	*lançar <i>palma/pé de</i> *atirar/jogar mão de	<b>pagar caro</b>	*pagar <i>alto</i> *remunerar/retribuir/quitacar caro
<b>lavar dinheiro</b>	*lavar <i>moeda/cédula/recurso financeiro</i> *limpar/banhar dinheiro	<b>pegar mal</b>	*pegar <i>ruim/péssimo</i> *agarrar/prender/segurar mal
<b>levar vantagem</b>	*levar <i>benefício/ganho</i> levar <i>proveito</i> *portar/conduzir vantagem	<b>ser dose (é dose!)</b>	*ser <i>quantidade/porção</i> *tornar-se dose

Fonte: elaborado pelo autor.

Como se vê, a fixidez paradigmática age de forma intensa na grande maioria das sequências do quadro 12, impedindo comutações lexicais até com as palavras sinônimas ou que estão em relação de hiperonímia. É particularmente interessante a maneira como se dá essa fixidez em fraseologismos que apresentam estrutura sintagmática semelhante, compartilhando uma mesma base, como o verbo *bater*, presente nas sequências *bater boca*, *bater cabeça* e *bater pernas*. Mesmo sendo de um mesmo domínio (partes do corpo), a mera substituição de um substantivo por outro “quebra” a unidade cujo sentido se encontra cristalizado, podendo, no entanto, produzir

outro fraseologismo, e não uma variante daquela. Assim, *bater boca* é discutir, mas *bater pernas* é andar ociosamente.

Diferentemente, há unidades nas quais a comutação lexical por unidade sinonímica, além de não desfazer o fraseologismo, produz, na verdade, variantes. É o caso de *virar pó* e *virar poeira*, de *tirar vantagem* e *levar vantagem*. Por outro lado, *falar grave* e *falar grosso* não constituem variantes, embora, no uso comum, o termo *grosso* seja entendido geralmente como sinônimo de *grave* (voz grave, voz grossa). No entanto, *falar grosso*, como unidade fraseológica, é "Falar com autoridade; agir com coragem, desassombro, prepotência" (SILVA, 2013, p. 604).

Há ainda que se destacar o fato de algumas comutações serem impossíveis pela própria inexistência de um termo sinonímico correspondente no eixo paradigmático. Assim, em *dar bola a* (dar atenção a; dar importância a), pode-se substituir o verbo *dar* por *doar*, por exemplo, mesmo que isso resulte na "quebra" do fraseologismo; contudo, qual seria o sinônimo a ser comutado com a palavra *bola*? Esfera? Ainda que haja um sinônimo a ser substituído, a fixidez paradigmática conservará a forma cristalizada.

Para Mejri (1997), essa forte restrição no eixo das comutações se deve à saturação lexical dos componentes da sequência. Desse modo, *engolir sapo*, isto é, "tolerar coisas ou situações desagradáveis sem responder por incapacidade ou conveniência" (HOUISS, 2009), não aceita comutação de nenhum de seus constituintes, e a razão disso é que, no âmbito dessa sequência cristalizada, tanto o verbo *engolir* quanto o substantivo *sapo* estão lexicalmente saturados, o que impede, por exemplo, a substituição por elementos de valor sinonímico.

Portanto, as restrições paradigmáticas e sintáticas têm relação com questões de natureza semântica. Com efeito, Mejri (1997, 1998) defende que o comportamento sintático dos fraseologismos deriva de sua organização semântica, de seus traços categoriais e de sua saturação lexical. Em se tratando de opacidade e de idiomaticidade, notam-se, por exemplo, que as sequências coletadas apresentam-nas em diferentes graus, havendo unidades mais opacas que outras. Por exemplo, as sequências *a cobra vai fumar* e *cutucar a fera com vara curta* apresentam maior grau de idiomaticidade do que *ajuste fiscal* e *condução coercitiva*. Quanto mais idiomática a estrutura, mais opaco seu

sentido, daí a importância do contexto para esclarecer o uso da unidade, como será demonstrado no glossário.

Saliente-se, porém, que a fixidez semântica não se resume à idiomaticidade ou opacidade do sentido. Como se viu no capítulo teórico, fixidez semântica também leva em consideração a relação entre o fraseologismo e a situação de uso, situando-se, portanto, também no nível pragmático. Assim, unidades gíricas como *papo reto*, *na marra* e *meter o pau em* geralmente são utilizadas em situações informais de uso, em conversas entre jovens, no cotidiano. Por sua vez, *acordo de delação premiada*, *ajuste fiscal*, *base do governo*, *cabo eleitoral*, *caixa dois*, *estado de direito*, entre outros, costumam aparecer mais frequentemente no meio político, judiciário ou jornalístico.

### 5.2.3 Quanto à congruência

Como se viu no capítulo teórico, Mejri (2012) defende o cruzamento das noções de fixidez e de congruência para que se possa, de um lado, delimitar as sequências cristalizadas, diferenciando-as das combinatórias livres e, de outro lado, descrever o funcionamento dos fraseologismos, em seu aspecto sintático, semântico e pragmático. Segundo o autor, a congruência descreve o processo de adaptação das sequências às regras da combinatória. Desse modo, as unidades que não se adequam a essas normas tornam-se incongruentes. Convém enfatizar, porém, que o fato de uma sequência violar uma ou mais das normas da combinatória não significa, de imediato, que a unidade não seja um fraseologismo, mas permite indicar o grau de cristalização da sequência. No quadro abaixo, analisam-se alguns exemplos de sintagmas nominais quanto a algumas tentativas de reestruturação.

Quadro 13 – Análise de sequências nominais quanto ao cruzamento das noções de fixidez e congruência.

FRASEOLOGIS MO	REESTRUTURAÇÕES		NÍVEL DE CRISTALIZAÇÃO
	no plano sintagmático	no plano paradigmático	
	- nominalização	comutação sinonímica	-cristalizado -semicristalizado

	- adição de advérbios intensificadores - predicação		<b>-totalmente incongruente</b>
<b>ajuste fiscal</b>	*fiscalização do ajuste *ajuste muito fiscal ±o ajuste é fiscal	*ajuste tributário/financeiro  *reparo fiscal  ±arrocho fiscal	semicristalizado
<b>bode expiatório</b>	*expição do bode *bode muito expiatório  ±o bode é expiatório	*bode <i>perdoador</i>  *cabrito expiatório	cristalizado
<b>braço direito</b>	*direita do braço  *braço muito direito  *o braço é direito	*braço <i>destro</i>  *membro superior direito	cristalizado
<b>cabo eleitoral</b>	*eleição do cabo *cabo muito eleitoral *o cabo é eleitoral	*cabo <i>competitivo?</i>  *aliciador eleitoral	cristalizado
<b>canoa furada</b>	*furação da canoa  ±canoa muito furada  a canoa é furada	*canoa <i>esburacada</i>  *embarcação/casc o furado(a)  <i>barca furada</i>	semicristalizado
<b>carta branca</b>	*brancura da carta  *carta muito branca  ±a/esta carta é branca	*carta <i>alva/clara</i>  *baralho branco	cristalizado
<b>dinheiro sujo</b>	*sujeira do dinheiro ±dinheiro muito sujo ±o dinheiro é sujo	*dinheiro <i>imundo/manchado</i>  ±grana suja	semicristalizado
<b>ficha limpa</b>	*limpeza da ficha  *ficha muito limpa	*ficha <i>asseada/lavada/livre</i>	cristalizado

	*a ficha é limpa	*currículo limpo	
<b>lista negra</b>	*negritude da lista *lista muito negra  ±a lista é negra	*lista <i>preta/escura</i>  *relação/rol negro(a)	cristalizado
<b>olho gordo</b>	*gordura do olho  olho muito gordo  ±o olho é gordo	*olho <i>gorducho</i>  ? gordo	semicristalizado

Fonte: elaborado pelo autor.

Dentre as tentativas transformacionais no plano sintagmático, percebe-se claramente maior fixidez na aplicação da nominalização, a qual foi interdita por todos os fraseologismos acima descritos. Já quanto à adição de advérbios intensificadores, algumas sequências obviamente a admitem, como *olho muito gordo*, podendo-se também aplicar-se, em determinados contextos, a *±dinheiro muito sujo* e *±canoa muito furada*. Por fim, no caso da predicação, a maioria das sequências poderiam admiti-la, mas, em outros sintagmas, tal reestruturação significaria a perda do *status* fraseológico (*\*o braço é direito*, *\*a ficha é limpa*, *\*o cabo é eleitoral*). Assim, pode-se afirmar, quanto às regras da fixidez sintática, que seriam incongruentes as transformações contrárias às normas da combinatória, o que está relacionado ao nível de cristalização da sequência.

No plano paradigmático, no qual a fixidez pode manifestar-se parcial ou totalmente, observa-se uma restrição mais intensa, impedindo, na maioria das sequências, a constituição de um paradigma sinonímico, como se vê na análise dispensada aos sintagmas cristalizados *bode expiatório*, *braço direito*, *cabo eleitoral*, *carta branca*, *ficha limpa* e *lista negra*. No entanto, nos demais casos, há que se destacar dois fatos importantes: a forte restrição paradigmática dos discursos especializados e a variação fraseológica.

No que se refere às especialidades, Mejri (2012) enfatiza que as colocações especializadas, embora geralmente transparentes semanticamente, possuem restrição paradigmática acentuada. É o caso de *ajuste fiscal*. Proveniente da área da economia, situando-se, portanto, no campo político (cf. CHARAUDEAU, 2006), *ajuste fiscal* apresenta forte restrição à substituição do adjetivo por outro sinônimo. Embora relacionados, os adjetivos *fiscal*, *tributário* e *financeiro* não podem ser simplesmente substituídos um pelo outro no domínio

estritamente técnico da economia. Ajuste fiscal não deixa de implicar tributos e, conseqüentemente, finanças, entretanto, designa um conjunto específico de políticas econômicas adotadas pelo Estado para conter gastos públicos, equilibrando o fisco. Segundo o colaborador AM<sup>58</sup>, trata-se de um “conjunto de medidas [...] que busquem promover o equilíbrio orçamentário do Estado, é a ação de adequar os gastos públicos aos valores arrecadados”.

Nessa direção, também podem ser mencionadas, a título de ilustração, algumas das unidades fraseológicas pertencentes ao universo político que não se prestam à comutação paradigmática. Por exemplo: as unidades *Estado mínimo*, *Estado de exceção*, *Estado de direito*, *analfabeto político*, *colarinho branco*, *caixa dois*, *golpe de Estado* não aceitariam substituições como \*Estado pequeno/minúsculo, \*Estado de restrição, \*Estado de garantia, \*analfabeto funcional, \*colarinho alvo/claro, \*caixa três, \*ferimento/corte de Estado.

Em se tratando de variação fraseológica, Mejri (2012) explica que, muitas vezes, a fixidez paradigmática parcial permite determinadas substituições que produzem variantes fraseológicas. Desse modo, as formas *barca furada* e *grana suja* poderiam ser consideradas variantes das sequências *canoa furada* e *dinheiro sujo*, visto que não revogam o caráter cristalizado destas unidades. O autor distingue, no entanto, um outro tipo de variação fraseológica em que a substituição lexical implica diferenças de nuances semânticas ou relativas a níveis de língua. Poderia ilustrar este caso a variante *arrocho fiscal*. Segundo os colaboradores especialistas da área da política, *ajuste fiscal* e *arrocho fiscal* podem ser considerados sinônimos, mas explicam que o termo “arrocho” é normalmente utilizado com o sentido pejorativo, para enfatizar as medidas de contenção do Estado em relação a investimentos públicos.

### 5.3 A COBRA VAI FUMAR NA POLÍTICA OU AINDA É POSSÍVEL DAR A VOLTA POR CIMA? ASPECTOS DO USO DE FRASEOLOGISMOS NO DISCURSO POLÍTICO

---

<sup>58</sup> AM é um dos três especialistas que validou as unidades referentes ao discurso político. (ver descrição completa do perfil dos colaboradores no capítulo de metodologia, seção sobre validação dos fraseologismos).

Serão analisados neste tópico alguns aspectos relativos ao uso de fraseologismos no discurso político, evidenciando os efeitos de sentido decorrentes desse uso nos contextos em que foram empregados.

### 5.3.1 A capacidade expressiva e de concisão do fraseologismo

Das unidades coletadas no *corpus*, 19 (dezenove) são utilizadas como título e manchetes de textos inteiros ou tópicos de parágrafos. Favorece esse uso o caráter notadamente expressivo das sequências cristalizadas, bem como a capacidade de síntese observada em algumas unidades, que se encaixam adequadamente aos propósitos textuais, permitindo uma concisão eficiente para situações que normalmente exigiriam um esforço linguístico muito maior, caso não existissem fraseologismos à disposição. Nos excertos abaixo, extraídos do *corpus*, pode-se notar essa função atribuída às combinatórias.

<**A cobra vai fumar**>? O Tribunal de Justiça de SP julgará na quarta-feira 28 uma ação histórica – 20 anos depois. A Philip Morris e a Souza Cruz estão no banco dos réus. Trata-se da primeira ação coletiva em que se inverteu o ônus da prova. Elas precisam provar que o cigarro não causa dependência e a sua publicidade não é abusiva e enganosa. A Associação de Defesa da Saúde dos Fumantes quer reconhecida a responsabilidade das empresas pelos danos a seus consumidores. Vencedora, as portas da Justiça abrem-se para fumantes e ex-fumantes buscarem indenizações por danos morais e patrimoniais contra a indústria do tabaco" (RIRB15M1a). (grifo nosso).

A sequência *a cobra vai fumar* anuncia que haverá perigo ou problemas pela frente em uma determinada situação. Ao ser utilizada pelo colunista Ricardo Boechat, da Revista *Época*, no início do texto que trata a respeito do conflito judicial envolvendo os interesses da Associação de Defesa da Saúde dos Fumantes e das distribuidoras de cigarro, o fraseologismo permite, a um só tempo, chamar a atenção do leitor pela sua expressividade e sintetizar o assunto do texto. Convém ainda destacar o fato de o próprio fraseologismo conter a palavra “fumar”, com a qual o autor estabelece uma relação entre o sentido literal desse verbo e o tema principal do texto, isto é, o problema judicial em torno da comercialização de cigarro.

Outro caso similar envolve o uso da sequência *barril de pólvora*, no trecho:

<**Barril de pólvora**> A 900 km de Belém, Xinguara está prestes a ver intensificadas as brigas de fazendeiros com sem terra. Chegou à cidade o líder da Frente Nacional de Luta Campo e Cidade, José Rainha. Na região, em que cerca de 800 pessoas morreram desde 70 em conflitos agrários, ele só anda em carro blindado. Dirigentes de sindicatos no sul do Estado estão apreensivos com o governador por não cumprir 200 mandados de reintegração" (RIRB16M8b). (grifo nosso).

Segundo o dicionário Aurélio (2004), *barril de pólvora* designa uma circunstância “delicada, tensa, de perigo latente, que está a ponto de explodir, com consequências imprevisíveis”. Em função desse significado, o fraseologismo mostra-se muito pertinente para referir-se ao clima hostil vivenciado em Xinguara, como consequência dos conflitos relacionados à posse da terra. Desse modo, consegue apreender sinteticamente o conteúdo do texto, além de chamar a atenção do leitor desde o início.

Esse recurso textual possibilitado pela sequência cristalizada também pode ser percebido no excerto iniciado pela sequência *signal vermelho*, como segue:

<**Signal vermelho**> Acusado de envolvimento no escândalo do mensalão, Paulo Rocha renunciou ao mandato de deputado federal. Para quem acredita que a Justiça tarda mas não falha, o petista acaba de ver o Tribunal Regional Eleitoral do Pará impugnar o registro de sua candidatura ao Senado, justamente pelo ato que praticou em 2005 (RIRB14M7b). (grifo nosso).

Proveniente do domínio dos sistemas de trânsito, *signal vermelho* indica que o veículo ou pedestre deve parar. Por extensão metafórica, a sequência significa, na língua comum, impedimento ou interrupção de algo ou processo. Baseado neste último sentido, Ricardo Boechat faz uso da unidade para se referir ao fato de o ex-deputado federal ter tido sua candidatura ao Senado impugnada por crimes praticados em 2005. Assim, o fraseologismo, ao iniciar o texto, já anuncia por meio de uma síntese o assunto referido. Nota-se, também, neste caso, que o autor explora a relação entre o fraseologismo e a cor-símbolo do partido ao qual Paulo Rocha é filiado, o Partido dos Trabalhadores (PT), cuja bandeira é vermelha.

Os exemplos dados acima trazem o fraseologismo no início do texto, servindo como uma espécie de título ou tópico-síntese do assunto tratado. Mas essa posição não é única, há casos em que a sequência cristalizada é posta ao

final do texto ou parágrafo, como forma de conclusão da discussão feita, tal como se verifica no excerto:

Como conciliar os interesses dos golpistas com a sanha dos magistrados curitibanos, o nihil obstat do Ministério Público Federal e o clangoroso silêncio do Supremo Tribunal? Apresenta-se, isto sim, um conflito de árdua composição, se não impossível. <**Beco sem saída**> (RCMC16M7e). (grifo nosso).

Neste caso, após demonstrar os atores e interesses envolvidos no contexto do que ele chama de golpe, Mino Carta desloca para o final de sua reflexão o fraseologismo *beco sem saída* (situação muito embaraçosa; de grande aperto). Com isso, considerando o significado global arribuído à sequência, pode-se perceber que a unidade serve para indicar a conclusão pessimista do autor diante do desfecho da questão, servindo como uma síntese da argumentação.

### 5.3.2 Os efeitos de ironia implicados no uso de fraseologismos

Uma das marcas do discurso político é certamente a ironia. Para este aspecto, há numerosos exemplos no *corpus*, mas acredita-se que a análise de apenas alguns será suficiente para observar de que maneira o fraseologismo se encaixa num projeto de ironia estabelecido pelos autores.

Assim, ao criticar ironicamente a observação (atrasada) de Michel Temer sobre os efeitos negativos do *impeachment* da ex-presidente Dilma para a imagem do Brasil, o jornalista Mino Carta, fundador e diretor da revista Carta Capital, lança mão do fraseologismo *além da conta* (*em excesso, demasiadamente*), para acentuar a ação desses efeitos no país, como se vê no trecho:

A rapidez de raciocínio de Michel Temer pareceu-me demonstrada na quarta-feira 20. Disse ele que a história do golpe prejudica a imagem do Brasil no exterior. Observação impecável. Prejudica muito. <**Além da conta**>" (RCMC16M5a). (grifo nosso).

A escolha e a disposição da sequência para o final do parágrafo também indicam que o fraseologismo foi usado estrategicamente para intensificar o efeito irônico produzido.

Outra unidade fraseológica utilizada para produzir ironia no texto sobre política refere-se à combinatória *assim que a banda toca* (as leis e regras seguem o protocolo estabelecido), empregada no excerto: “É uma ironia que possa [Eduardo Cunha] ser cassado apenas por 'quebra de decoro', mas é **<assim que a banda toca>**” (RERA16M9b). (grifo nosso). Neste caso, a jornalista Ruth de Aquino, colunista da Revista Época, insere o fraseologismo num jogo de ironia, para demonstrar seu descontentamento com a (péssima!) cultura brasileira de passar a mão na cabeça dos políticos, tratando-os sem o rigor esperado para os atos e crimes que eles cometem.

Outro exemplo consiste no uso do fraseologismo *armado até os dentes* (com muitas armas), feito pelo jornalista Mino Carta, em crítica irônica à maneira como foi realizada a condução coercitiva do ex-presidente Lula. Segundo o jornalista:

Não me surpreende, de todo modo, a serenidade que Lula conserva em meio à tempestade, como se deu quando na dita 'condução coercitiva', operação anticonstitucional e de inaudita violência executada por 200 policiais **<armados até os dentes>**” (RCMC16M3d).

Nota-se, portanto, que a sequência em destaque foi inserida no jogo de ironia estabelecido, para acentuar a crítica à legitimidade do mecanismo de condução coercitiva, posta em xeque pelo jornalista, uma vez que a considerou, no caso do Lula, “operação anticonstitucional e de inaudita violência”. E, para justificar seu ponto de vista, destaca o fato da ação ter sido operada por 200 policiais demasiadamente armados, ou seja, *armados até os dentes*.

Um último exemplo a ser destacado diz respeito ao fraseologismo *atolados até o pescoço*, variante da sequência *envolvido até o pescoço* (até o máximo). Ao criticar o repentino interesse dos senadores Renan Calheiros e Romero Jucá em aprovar a alteração na Lei de Abuso de Autoridade, Ricardo Boechat ressalta o fato de os dois políticos estarem em condições que os torna ligeiramente parciais na aprovação da proposta legislativa. Por esse motivo, ironiza tal situação, enfatizando-a a partir do fraseologismo em tela: “Mais uma pergunta se impõe: é aceitável que dois senadores **<atolados até o pescoço>** em suspeitas diversas protagonizem tal manobra legislativa?” (RIRB16M10b) (grifo nosso).

Acredita-se que os poucos exemplos mencionados permitem ilustrar como os fraseologismos se prestam ao jogo da ironia e da crítica realizado no discurso político. De fato, revelam-se, nesse domínio, recursos indispensáveis para promover variados efeitos de sentido, levando o leitor/ouvinte a aderir a posicionamentos, a refutar teses etc.

### 5.3.3 Efeitos de ambiguidade implicados no uso de fraseologismos

Embora os fraseologismos possuam significado global, é possível acessar o sentido literal dos itens lexicais que os formam devido à dupla estruturação semântica sobre a qual se funda a sequência cristalizada (MEJRI, 1997, 2004, 2018). De fato, em muitos dos casos observados, pode-se notar o efeito de ambiguidade que se produziu ao retomar, simultaneamente, os sentidos literal e global da unidade, como se vê nos exemplos a seguir:

Diante do cenário de crise política e econômica que acabou desencadeando fragilidade e falta de governabilidade em torno da última gestão da ex-presidente Dilma Rousseff, Ruth de Aquino, colunista de *Época*, alerta aos cidadãos brasileiros que “<**Apertem os cintos**>: o piloto do Brasil sumiu” (RERA15M3e). No uso desse fraseologismo, nota-se, ao mesmo tempo, a referência sutil ao sentido literal (prender-se à poltrona atracando o cinto à fivela) que se esperaria para uma situação de avião em queda livre e o sentido global relativo à redução de gastos, já que, como fraseologismo, *apertar os cintos* designa "reduzir os gastos ou fazer economia" (MICHAELIS, versão on-line).

Por sua vez, Ricardo Boechat, da Revista *Istoé*, utiliza em sua análise a sequência *bater o martelo*, prevendo que o Supremo Tribunal Federal demoraria para decidir sobre o pedido de cassação da ex-presidente Dilma Rousseff. De acordo com o jornalista, “É mínima a chance de o TSE <**bater o martelo**> até meados de outubro, na ação que pede a cassação do mandato de Dilma Rousseff” (RIRB15M9d). Neste caso, embora prevaleça o sentido fraseológico de *apresentar decisão final sobre determinada questão posta em debate*, o sintagma verbal *bater o martelo* remete literalmente à prática observada no âmbito jurídico de um juiz bater o martelo na mesa para indicar o encerramento de uma sessão ou dar o veredicto sobre a questão.

Também explora-se a ambiguidade da sequência *berço político* no modo como foi empregada pelo jornalista Ricardo Boechat, no trecho:

Já não existem os 3,4 milhões de votos que deram à Dilma Rousseff vitória sobre Aécio Neves no ano passado. Pelo menos na ótica do eleitor mineiro, estado onde o tucano sofreu a mais emblemática derrota na corrida ao Planalto. Segundo pesquisa do Instituto Sensus concluída na terça-feira 28, se a disputa fosse hoje ele venceria com frente de 6,6 milhões de votos. Foram 1 mil entrevistados em MG: Aécio faturou 79% dos votos (47,6% em 2014) e Dilma 21% (52,4%). Em seu <**berço político**>, o senador já pode dormir em paz. (RIRB15M4e).

No âmbito do discurso político, o sintagma *berço político* remete à origem e à herança política que um determinado candidato a cargo eletivo recebe de sua família, que já possui prestígio e reconhecimento político em determinado local. De fato, esse sentido está presente na citação acima, uma vez que realmente o Senador Aécio Neves, neto do ex-presidente Tancredino Neves, é mineiro, tendo recebido o legado e o capital político de sua família nas terras mineiras. Entretanto, além desse sentido global, pode-se perceber que o autor também associa, ironicamente, o sintagma em destaque ao sentido literal da palavra “berço” (tipo de cama apropriado para crianças de colo), já que, segundo o jornalista, o Senador, então, poderia dormir tranquilo.

Outro caso em que recorreu-se à ambiguidade do fraseologismo refere-se a uma sequência originalmente do futebol, *gol contra*, como se vê no excerto:

O primeiro <**gol contra**> não foi do Marcelo. Foi do técnico Lula. Há quatro anos, ele convocou como artilheira seu poste querido, inexperiente no gramado político, sem talento para driblar adversidades, sem criatividade para virar um placar, sem carisma para liderar companheiros, sem visão de jogo para lançamentos longos, sem precisão nos cruzamentos, sem vocação para trabalho de equipe. Uma capitã sem a generosidade do passe, sem a humildade da autocrítica, sem o brilho que encanta, sem sorriso, sem suor, sem humor. Dilma foi imposta por Lula até a aliados relutantes (RERA14M6c). (grifo nosso)

A partir do domínio do futebol, Carlene Nunes Salvador (2017, p. 332) define o fraseologismo *gol contra* como um “Lance em que o jogador, num momento de extrema falta de sorte, marca um gol contra o seu próprio time”. Tal acepção é claramente acionada no excerto destacado acima, ao fazer alusão ao fato de Marcelo, o lateral-esquerdo dono da camisa 6 da seleção brasileira, ter

feito o primeiro gol contra, na Copa de 2014, durante o jogo entre Brasil e Croácia. Entretanto, a esse sentido especializado, superpõe-se outro, de caráter mais geral, sobre o qual a jornalista Ruth de Aquino tece suas críticas ao ex-presidente Lula, responsável, segundo ela, pela má escolha de Dilma Rousseff para governar o país. Neste sentido, *gol contra* também pode ser definido como atitude ou medida cujos efeitos prejudicam a própria pessoa, movimento ou instituição que a tomou. O interesse deste exemplo reside no fato de ele demonstrar que a ambiguidade do fraseologismo pode ser explorada tanto em relação à coexistência simultânea do sentido literal e global do fraseologismo quanto no que tange à relação entre acepções distintas que correspondem a domínios também distintos, neste caso, o do futebol e o uso da língua comum.

Como se disse anteriormente, essa possibilidade de explorar a ambiguidade da sequência deve-se à sua dupla estruturação semântica que subjaz ao significante plural fixado pela saturação lexical. Mejri (1997, 1998, 2018) explica que essa dupla estruturação semântica agencia os significados literal dos componentes da sequência e o sentido global da estrutura sintagmática. Devido a esse traço de origem, o autor defende que os fraseologismos guardam uma memória lexical, cujos sentidos podem ser acessados quando se utilizam as sequências em determinados contextos.

Desse modo, pode-se concluir que os fraseologismos, além de sua significativa presença nos textos, revelam-se recursos lexicais indispensáveis na comunicação. Isso se deve ao seu caráter expressivo, à sua capacidade de sintetizar conceitos ou situações muito complexas e, ainda, por alimentarem e estruturarem o discurso político, moldando-se facilmente aos jogos argumentativos, às ironias, às tensões ideológicas etc.

Os resultados da pesquisa em foco demonstram, portanto, que as unidades fraseológicas estão presentes no discurso político, conferindo expressividade e um colorido local para as expressões utilizadas. Longe de se reduzirem a clichês e a lugares-comuns, esses fraseologismos representam concepções e modos de expressão de um povo e de uma cultura, servindo para caracterizar linguisticamente o domínio da política brasileira na mídia impressa. Apesar de apresentarem graus diferentes de fixidez, frequência de uso e polilexicalidade, todas essas unidades inventariadas partilham de certos

aspectos que permitem classificá-las como fraseologismos, como já se demonstrou na análise.

## CONSIDERAÇÕES FINAIS

O objetivo principal desta dissertação consistiu em produzir um glossário, em versão impressa e eletrônica, dos fraseologismos utilizados no discurso político veiculado em textos escritos de ampla circulação no Brasil, no período compreendido entre janeiro de 2014 e dezembro de 2016. Para isso, foi constituído um *corpus* com 570 textos, 37.188 palavras, extraídos diretamente das colunas e *blogs* pessoais de quatro colunistas vinculados às revistas *Época*, *Istoé*, *Carta Capital* e jornal *Folha de São Paulo*.

Após extração, tratamento e análise semiautomática dos textos, verificaram-se 438 unidades fraseológicas que constituem a nomenclatura do glossário. Essas unidades, na sua grande maioria, não são específicas do discurso político, mas provêm do uso mais geral da língua. Contudo, estão relacionadas a esse âmbito por serem utilizadas frequentemente como meios de expressão de conceitos, práticas e relações imbrincadas no dizer da política brasileira, revelando os jogos de ironia e as tensões ideológicas que comumente se estabelecem em torno do poder.

Com a produção do glossário, baseado na análise dos dados, pôde-se alcançar os objetivos específicos delineados, quais sejam: a) *descrever os fraseologismos que caracterizam o discurso político brasileiro*; b) *identificar padrões de combinatórias sintagmáticas recorrentes no corpus*; c) *verificar e descrever possíveis variantes fraseológicas*. Em relação ao primeiro, os resultados demonstraram que, pelo menos no *corpus* de estudo, o discurso político apresenta majoritariamente muitas sequências provenientes do uso da língua comum, caracterizando-se, no entanto, por agenciar essas unidades com o fim de fazê-las proveitosas à finalidade discursiva de convencimento e de ação sobre o outro. Isso não significa que o sentido básico de cada fraseologismo tenha mudado em função de sua ocorrência nesses textos, mas o contexto de uso de algum modo contribui para pôr em evidência a expressividade, a praticidade e a eficácia dessas combinatórias para os propósitos da comunicação política.

Quanto ao segundo objetivo específico, identificou-se a predominância de sintagmas verbais, seguida de sintagmas nominais e preposicionais. As sequências frásticas apareceram em número mais reduzido. Verificou-se, ainda,

que os fraseologismos mais extensos são justamente os que possuem menor ocorrência, diferenciando-se dos sintagmas mais reduzidos, sobretudo os sintagmas formados por apenas três componentes, que apresentaram frequência significativamente muito mais elevada.

Por fim, com relação ao registro de variantes fraseológicas, foram identificados 25 casos, os quais figuram como entradas no glossário. Notou-se, também, que, para algumas das sequências inventariadas, há variantes que, embora não tenham ocorrido no *corpus*, encontram-se registradas nas fontes consultadas ou foram sugeridas pelos colaboradores da pesquisa que, na condição de falantes nativos de português, (re)conhecem muito mais fraseologismos que as obras de referência possam ter inventariado. Essas variantes sem ocorrência nos textos coletados não constituem entradas no glossário, mas são indicadas, em nota, no verbete da variante correspondente no *corpus*.

Importa ressaltar, nesse contexto, a contribuição da Linguística de *Corpus* como metodologia para a realização da pesquisa fraseológica. Sem o instrumental fornecido pela LC, o trabalho teria sido extremamente penoso. Desse modo, foram fundamentais, desde o processo de constituição, pré-processamento e análise do *corpus*, e da organização e produção do glossário, os *softwares WordSmith Tools* (versão 6.0) e *Lexique Pro* (3.6). Por esse motivo, concorda-se, portanto, com Tagnin (2011), para quem a Linguística de *Corpus* e a Fraseologia foram feitas uma para a outra.

No tocante às dificuldades que se apresentaram durante a realização do trabalho, destacam-se, de um lado, dois fatos de caráter mais pessoal e, de outro, problemáticas levantadas pela própria área de pesquisa. Com relação ao primeiro grupo, podem-se mencionar: o fato de o campo fraseológico ter sido considerado novo pelo autor, embora já constitua um domínio com tradição de pesquisa em outros países e em crescente desenvolvimento no Brasil (a impressão que se tem, para o iniciante na área, é de que o campo fraseológico é o lugar da falta de consenso em que os conceitos e termos situam-se num emaranhado de coisas); a falta de proficiência adequada do pesquisador nas línguas inglesa e francesa, já que o trabalho segue a vertente francesa da Fraseologia, fato que motivou (e tem motivado) a busca pelo estudo de outros idiomas.

Do segundo grupo de dificuldades, convém salientar três problemáticas. A primeira se refere à delimitação dos fraseologismos (quais critérios considerar? Frequência? Idiomaticidade? Fixidez sintática e semântica? Devem ser tomados em conjunto ou privilegiar alguns? Onde segmentar a unidade fraseológica?). O terceiro fato diz respeito ao registro de fraseologismos nas obras de referência, particularmente os dicionários gerais de língua. A lexicografia brasileira ainda não trata de forma adequada, ampla e sistematicamente as sequências cristalizadas. Por outro lado, apesar de a fraseografia ter apresentado crescimento considerável ao longo dos anos, o acesso aos repertórios fraseográficos ainda pode ser considerado limitado. Uma terceira problemática está relacionada à validação de fraseologismos por falantes nativos. Considera-se veementemente que os falantes, devido à competência fraseológica (MEJRI, 2012; ORTIZ ALVAREZ, 2014) que possuem, são as principais “autoridades” a serem consultadas no processo de validação e reconhecimento das sequências cristalizadas, no entanto, nem todos compartilham o mesmo nível de competência, além do fato de ser da própria unidade fraseológica apresentar-se em graus distintos de cristalização.

Apesar dos percalços enfrentados, acredita-se que foram alcançados os objetivos delineados para a pesquisa, sendo a produção do glossário sua principal conquista. O repertório poderá contribuir, do ponto de vista do público ao qual se destina (falantes de português com interesse em assuntos políticos), para constituir um acervo de expressões e sequências cristalizadas, específicas ou não ao discurso político, mas em cujo contexto se mostram produtivas, servindo aos propósitos comunicativos desse domínio. Do ponto de vista acadêmico, o glossário se soma aos esforços de numerosos pesquisadores brasileiros que buscam descrever os fraseologismos do português, preenchendo, portanto, parte da lacuna fraseográfica. Deve-se ressaltar, ainda, que os resultados da pesquisa podem ensejar novos trabalhos na área, explorando de forma mais exaustiva o *corpus* coletado ou ampliando-o para abranger novos fatos e períodos da língua portuguesa e da política brasileira.

## REFERÊNCIAS

- ALVES, Ieda M. A terminologia política no período pré-eleitoral. **Revista Alfa**, nº 27, p. 39-46, 1983.
- ARAGÃO, M. do S. S. de. A fraseologia como marca do léxico regional popular. In: COSTA, Daniela de S.S; BENÇAL, Dayme R. (orgs.) **Nos caminhos do léxico**. Campo Grande, MS: Ed. UFMS, 2016, p. 33-49.
- ARENDET, Hannah. **Entre o passado e o futuro**. Trad. Mauro W. Barbosa de Almeida. São Paulo: Perspectiva, 1972.
- ARISTÓTELES. **Política**. Tradução de Pedro Constantin Tolens. São Paulo: Martin Claret, 2001.
- BARBOSA, M. A. Aspectos conceituais, terminológicos e políticos das denominações: discurso científico e discurso político.. In: 57º Reunião Anual da SBPC, 2005, Fortaleza - CE. **Anais/Resumos da 57º Reunião Anual da SBPC**. São Paulo - SP: SBPC/UECE, 2005. v. 1. p. 1-2. Disponível em: <[http://www.sbpcnet.org.br/livro/57ra/programas/CONF\\_SIMP/textos/mariaapbarbosa.htm](http://www.sbpcnet.org.br/livro/57ra/programas/CONF_SIMP/textos/mariaapbarbosa.htm)>. Acesso em: nov. 2016.
- BARBOSA, M. A. A fraseologia no percurso gerativo de enunciação de codificação: no sistema, nas normas, no falar concreto. In: ORTIZ ALVAREZ, M. L. (org.). **Tendências atuais na pesquisa descritiva e aplicada em fraseologia e paremiologia**. V.1. Campinas, SP: Pontes Editores, 2012a.
- BALLY, Charles. **Traité de stylistique française**. 2 ed. Paris: Klincksieck, 1951.
- BART, Christian Le. Parler en politique. **Mots**. Les langages du politique, 94 | 2010.
- BEVILACQUA, Cleci Regina. Fraseologia: perspectiva da língua comum e da língua especializada. **Revista Língua & Literatura**. v.6/7, nº 10/11, p. 73-80. 2004.
- BEVILACQUA, C. R.; FINATTO, M. J. B. Lexicografia e Terminografia: alguns contrapontos fundamentais. **Alfa**, São Paulo, v. 50 (2), p. 43-54. 2006.
- BERBER SARDINHA, T.  
Linguística de Corpus: histórico e problemática. **D.E.L.T.A.** São Paulo, v. 16, n. 2, p.323-367, 2000. Disponível em:<[http://www.scielo.br/pdf/delta/v16n2\\_/a05v16n2.pdf](http://www.scielo.br/pdf/delta/v16n2_/a05v16n2.pdf)>. Acesso em: 29 jun. 2017.
- BERBER SARDINHA, T. **Linguística de Corpus**. Barueri, SP: Editora Manole, 2004.

- BERBER SARDINHA, T. **Pesquisa em Linguística de Corpus com WordSmith Tools**. 1. ed. Campinas: Mercado de Letras, 2009. v. 1.
- BERBER SARDINHA, T.; BARCELLOS ALMEIDA, G. M. A Linguística de *Corpus* no Brasil. In: TAGNIN, S.; VALE, O. A. (Orgs.). **Avanços da Linguística de Corpus no Brasil**. São Paulo: Humanitas, 2008. p. 17-40.
- BIDERMAN, M. T. C. Unidades complexas do léxico. In: Rio-Torto, G.; Figueiredo, O. M.; SILVA, F. (Org.) **Estudos em homenagem ao professor doutor Mário Vilela**. 1ª ed. Porto: Portugal: Faculdade de Letras da Universidade do Porto, v. II, p.747-757, 2005.
- BOBBIO, Norberto; MATTEUCCI, Nicola; PASQUINO, Gianfranco. **Dicionário de política**. Tradução de Carmen C. Varriali et al. Brasília: Editora Universidade de Brasília, 1998.
- BONAVIDES, Paulo. **Ciência política**. 10 ed. São Paulo: Malheiros Editores, 2000.
- CAMPOS, Paula J. de Souza. **Uma terminologia do discurso neoliberal?** 2007. 280 f. Dissertação (Mestrado) – Universidade Federal do Mato Grosso, Cuiabá, 2007.
- CARBALLO, M. A. C. El concepto de unidad fraseológica. **Revista de Lexicografía**, Volume IV, 1997-1998, p. 67-79.
- CHARAUDEAU, Patrick. O discurso político. In: EMEDIATO, Wander; MACHADO, Ida Lúcia; MENEZES, William (orgs.). **Análise do discurso: gêneros, comunicação e sociedade**. Belo Horizonte: Núcleo de Análise do Discurso, Programa de Pós-graduação em Estudos Linguísticos, Faculdade de Letras da UFMG, 2006, p. 251-268.
- CHARAUDEAU, Patrick. Pathos et discours politique. In: M. Rinn (éd.), **Émotions et discours**. L'usage des passions dans la langue. Rennes: Presses universitaires de Rennes, 2008, p. 49-58 [en ligne]. URL: <<http://www.patrick-charaudeau.com/Pathos-et-x.html>>.
- CLAS, A. & GROSS, G. Classes de figement des locutions verbales. In: S. Mejri, G. Gross, A. Clas & T. Baccouche (éds). **Le figement lexical**, Tunis, CERES, 1998. p. 11-18.
- CORPAS PASTOR, G. **Manual de fraseologia española**. Madrid: Gredos, 1996.
- CORPAS PASTOR, G. ; ORTIZ ALVAREZ, M. L. Fraseologia e Paremiologia: uma entrevista com Gloria Corpas Pastor. **ReVEL**, vol. 15, n. 29, 2017. Tradução de Ana Carolina Spinelli. Revisão técnica de Gabriel de Ávila Othero. Disponível em: <[www.revel.inf.br](http://www.revel.inf.br)>. Acesso em: 25 out. 2017.

CORREIA, Emanuel de Moura; TEIXEIRA, Persília de Merlim. **Dicionário Prático de Locuções e Expressões Correntes**. Santa Catarina: Papiro Editora, 2007.

DICIONÁRIO DA LÍNGUA PORTUGUESA. Porto: Porto Editora, 2017.

DICIONÁRIO DE EXPRESSÕES. Disponível em:  
<<http://www.dicionariodeexpressoes.com.br/index.do>>. Acesso em: 28 jul. 2018.

DOBROVOL'SKIJ, D. Phraseology: Historical Development and Theoretical Aspects. In. ALVAREZ, M. L. O. (Org.) **Tendências atuais na pesquisa descritiva e aplicada em fraseologia e paremiologia** - Anais – Vol. 1. Campinas, SP: Pontes Editores, 2012.

DORNA, Alexandre. Les effets langagiers du discours politique. **Hermès, La Revue**, 1995/2 (n° 16), p. 131-146.

DUSSEL, Henrique. **20 tesis de política**. México: Siglo XXI: Centro de Cooperación Regional para la Educación de Adultos en América Latina y el Caribe, 2006.

FAJARDO, M. L. A. **El lenguaje político**: características e análisis del discurso político com ejercicios y clave. Santarcangelo di Romagna, RN: Maggioli Editore, 2016.

FAULSTICH, Enilde. Socioterminologia: mais que um método de pesquisa, uma disciplina. **Ciência da Informação**, Brasília, v. 24, n. 3, p. 281-288, set./dez. 1995.

FERREIRA, Aurélio Buarque de Holanda. **Novo dicionário eletrônico Aurélio**. Editora Positivo, 2004.

FIORIN, J. L. Língua, discurso e política. **Revista ALEA**. v 11. nº 01, 2009, p. 148-165.

FULGÊNCIO, Lúcia. **Expressões fixas e idiomatismos do português brasileiro**. 502f. Tese de Doutorado – Pontifícia Universidade Católica de Minas Gerais. Belo Horizonte, 2008, 502p.

FULGÊNCIO, Lúcia. Zoomorfismos, botanismos, gastronomismos: é assim que devem ser classificados os fraseologismos? **Calígrama**, Belo Horizonte, v.19, n.2, p. 179-196, 2014.

GROSS, G. Degré de figement des noms composés. **Langages**, v. 90, 1988, p. 57-72.

GROSS, G. **Les expressions figées en français**. Noms composés et autres locutions. Paris: Ophrys, 1996.

GROSS, M. Une classification des phrases figées du français. **Revue québécoise de linguistique**, v. 11, n. 2, 1982, pp. 151-185.

GROSS, Maurice. Les phrases figées en français. **L'Information Grammaticale**, N. 59, 1993. pp. 36-41.

HOUAISS, A. **Dicionário eletrônico Houaiss da língua portuguesa**. Rio de Janeiro: Objetiva. Versão 3.0, 2009.

KLEIN, Jean René; LAMIROY, Béatrice. Collocations, expressions figées, phrases situationnelles, proverbes. **L'Information grammaticale** n° 148, janvier 2016, p. 15-20.

KRIEGER, Maria da Graça; FINATTO, Maria José Bocorny. **Introdução à Terminologia: teoria e prática**. São Paulo: Contexto, 2004.

DEL ARCO MONTORO, Steban Tomás. La variación fraseológica y el diccionario. In: De lexicografía. **Actas del Symposium Internacional de Lexicografía**. Barcelona, 2004. Institut Universitari de Lingüística Aplicada.

LAMIROY, Béatrice et al. Expressions verbales figées et variation en français: le projet "BFQS". **Cahiers de lexicologie**. 2003, 83-2: 153-172.

LAMIROY, Béatrice; KLEIN, Jean René. Le problème central du figement est le semifigement. **Linx**, 53, 2005.

LAMIROY, Béatrice. Le figement: à la recherche d'une définition. **Zeitschrift für Französische Sprache und Literatur**; 2008; Vol. 36; pp. 85 – 99.

LAMIROY, Béatrice. KLEIN, Jean René. Le Figement. Unité et diversité. Collocations, expressions figées, phrases situationnelles, proverbes. **L'Information Grammaticale**; 2016; Vol. 148; pp. 15 – 20.

LIMA, Alcides F. de; MARTINS, Arlon F. C. Utilização do programa Lexique Pro na elaboração de glossários e dicionários terminológicos. In: RAZKY, A. et al. (Orgs.). **Estudos Sociodialetais do Português Brasileiro**. Campinas: Pontes, 2014. p. 257-277.

MAAR, Wolfgang Leo. **O que é política**. São Paulo: Brasiliense, 2006.

MONTEIRO-PLANTIN, R. S. **Fraseologia: era uma vez um patinho feio no ensino de língua materna (volume I)**. Fortaleza: Imprensa Universitária, 2014.

MEJRI, Salah. **Le figement lexical: descriptions linguistiques et structuration sémantique**. Tunis: Publications de la faculté des lettres de la Manouba, 1997.

MEJRI, Salah. Le figement lexical: descriptions linguistiques et structuration sémantique. In: **L'Information Grammaticale**, n. 76, 1998, p. 50-51.

MEJRI, Salah. Unité polylexicale et polylexicalité. **Linx** n. 40, 1999, p. 79-93.

MEJRI, Salah. Le figement lexical: nouvelles tendances, **Cahiers de lexicologie** n. 80, 2002, p. 213-223.

MEJRI, Salah. Figement absolu ou relatif: la notion de degré de figement. **Linx** [En ligne], n. 53, 2005.

MEJRI, Salah. La traduction des textes spécialisés: le cas des sciences du langage. Colloque du 50e anniversaire de l'ISTI, Oct 2008, Belgique. Editions du Hazard, pp.117-144, 2008.

MEJRI, Salah. Délimitation des unités phraséologiques. In: ORTIZ ALVAREZ, M. L. (org.). **Tendências atuais na pesquisa descritiva e aplicada em fraseologia e paremiologia**. V.1. Campinas, SP: Pontes Editores, 2012a.

MEJRI, Salah. La phraséologie française: synthèse, acquis théoriques et descriptifs. **Revue de linguistique française**, n. 1, 2018.

MEL'CUK, Igor. Fraseología y diccionario en la lingüística moderna. In: Isabel Uzcanga et al. (org.). **Presencia y renovación de la lingüística francesa**, Salamanca, 2001.

NAVARRO, Carmen. La fraseología en el discurso político y económico de los medios de comunicación. In: **AISPI Actas del XX Congreso** [Associazione Ispanisti Italiani] pp. 199-211, 2002.

OLIVEIRA SILVA, M. E. O. de. Dicionários: armas de dois gumes no estudo da fraseologia. O caso das locuções. In: ORTIZ ALVAREZ, M. L.; UNTERNBÄUMEN, Enrique H. (orgs.). **Uma (re)visão da teoria e da pesquisa fraseológicas**. Campinas, SP: Pontes Editores, 2011, p. 161-182.

ORTIZ ALVAREZ, M. L. **Expressões idiomáticas do português do Brasil e do espanhol de Cuba**: Estudo contrastivo e implicações para o ensino de português como língua estrangeira. 2000. 334f. Tese (Doutorado) – Universidade Estadual de Campinas, Instituto de Estudos da Linguagem. UNICAMP, Campinas, SP, 2000.

ORTIZ ALVAREZ, M. L.; UNTERNBÄUMEN, Enrique H. (orgs.). **Uma (re)visão da teoria e da pesquisa fraseológicas**. Campinas, SP: Pontes Editores, 2011.

ORTIZ ALVAREZ, M. L. (org.) **Tendências atuais na pesquisa descritiva e aplicada em fraseologia e paremiologia**. V.1. Campinas, SP: Pontes Editores, 2012a.

ORTIZ ALVAREZ, M. L. A denominação fraseológica no humor e na política. **Revista Brasileira de Linguística**, São Paulo, v. 13, p. 131-141, 2005a.

ORTIZ ALVAREZ, M. L. A dinâmica e potencialidade da denominação fraseológica no discurso político e humorístico. In: 57 Reunião Anual da SBPC, 2005, Fortaleza. **Anais da 57 Reunião Anual da SBPC**, 2005b.

PAMIES, Antonio. A metáfora gramatical e as fronteiras (externas e internas) da fraseologia. **Revista de Letras**. nº 33, v. 1 – jan./jun, p. 51-77, 2014.

RAZKY, A.; LIMA, A. F. de; OLIVEIRA, M. B. de. **Geossociolinguística e Socioterminologia no Brasil** – GeoLinTerm (projeto de pesquisa). UFPA, 2010. 19f.

RIBEIRO, João Ubaldo. **Política**: quem manda, por que manda, como manda. 3 ed. Rio de Janeiro: Nova Fronteira, 1998.

RIBOLDI, Ari. **A CPI das palavras**: origem de palavras e expressões da linguagem política. 2 ed. Caxias do Sul, RS: Lorigraf, 2008.

SALVADOR, Carlene N. **Fraseologismos do futebol brasileiro**: proposta de um dicionário eletrônico. 2017. 500f. Tese (Doutorado em Letras) – Universidade Federal do Pará. Belém, 2017.

SAUSSURE, Ferdinand de. **Curso de linguística geral**. Tradução de Antônio Cheline, José Paulo Paes, Izidoro Blinkstein. 27 ed. São Paulo: Cultrix, 2006.

SCOTT, M. **WordSmith Tools**, versão 5.0. Liverpool: Lexical Analysis Software, 2008.

SIL International. **Lexique Pro 3.6**. Software SIL IVB/Mali, Copy-right 2004-2012. Disponível em: <http://www.lexiquepro.com/download.htm>. Acesso em: 13 fev. 2018.

SILVA, José Pereira da. **Dicionário brasileiro de fraseologia** (versão preliminar). Rio de Janeiro: [s.n], 2013.

TAGNIN, S. E. O. **O jeito que a gente diz**: expressões convencionais e idiomáticas. São Paulo: Disal, 2005.

TAGNIN, S. E. O. Linguística de corpus e fraseologia: uma feita para a outra. In: ORTIZ ALVAREZ, M. L.; UNTERNBÄUMEN, Enrique H. (orgs.). **Uma (re)visão da teoria e da pesquisa fraseológicas**. Campinas, SP: Pontes Editores, 2011, p. 277-302.

TAGNIN, S. E. O.; TEIXEIRA, E. D. Translator-oriented, corpus-driven technical glossaries: the case of cooking terms. **Corpora**, v. 7, p. 51-67, 2012.

TRISTÁ, A. M. **Fraseología y contexto**. Ciudad de la Habana, Cuba: Editorial de Ciencias sociales, 1988.

WEBER, Max. **Ciência e política** : duas vocações. Tradução Leonidas Hegenberg e Octany Silveira da Mota. 18 ed. São Paulo : Cultrix, 2011.

WELKER, H. A. Colocações e expressões idiomáticas em dicionários gerais. In: ALVAREZ, M. L. O.; UNTERNBÄUMEN, Enrique H. (orgs.). **Uma (re)visão da**

**teoria e da pesquisa fraseológicas.** Campinas, SP: Pontes Editores, 2011, p. 139-159.

XATARA, C. ; PARREIRA, M. C. A elaboração de um dicionário fraseológico. In: ORTIZ ALVAREZ, M. L.; UNTERNBÄUMEN, Enrique H. (orgs.). **Uma (re)visão da teoria e da pesquisa fraseológicas.** Campinas, SP: Pontes Editores, 2011, p. 69-75.

ZULUAGA, A. **Introducción al estudio de las expresiones fijas.** Frankfurt: Verlag Peter D. Lang., 1980.