

De grão em grão a galinha enche o papo: primeiros estudos fraseológicos no Norte do Brasil

De grão em grão a galinha enche o papo: first phraseological studies in North of Brazil

Abdelhak RAZKY*

Universidade Federal do Pará (UFPA)

Carlene Ferreira Nunes SALVADOR**

Universidade Federal do Pará (UFPA)

Davi Pereira DE SOUZA***

Universidade Federal do Pará (UFPA)

Rejane Umbelina Garcez Santos DE OLIVEIRA****

Universidade Federal do Pará (UFPA)

RESUMO: O objetivo deste artigo consiste em apresentar os primeiros trabalhos desenvolvidos no Norte do Brasil na área da Fraseologia, ancorados teoricamente na vertente fraseológica francesa (MEJRI, 1997) e orientados pela metodologia da Linguística de *Corpus* (BERBER SARDINHA, 2004; TAGNIN, 2009). Trata-se de duas teses de doutorado – *Estudo das fraseologias do futebol brasileiro das Séries B, C e D em jornais digitais populares: construção de um dicionário eletrônico* (SALVADOR, 2017), *O léxico especializado do corte bovino: uma abordagem terminológica e terminográfica* (OLIVEIRA, 2018) e de uma dissertação de mestrado – *Fraseologismos no discurso político brasileiro: uma proposta de glossário* (SOUZA, 2018), vinculadas ao Projeto GeoLinTerm – PPGL/UFPA. Juntos, os três trabalhos somam 2.954 fraseologismos descritos referentes aos domínios pesquisados, como ilustram respectivamente os exemplos *gol de placa*, *chute de bucho condenado* e *crime de colarinho branco*.

PALAVRAS-CHAVE: Fraseologismos. Futebol. Corte Bovino. Discurso Político. GeoLinTerm.

ABSTRACT: This article aims to present the first works developed in the North of Brazil in the area of Phraseology, anchored theoretically in the French part of Phraseology (MEJRI, 1997)

*Doutor em Linguística. Professor Titular da Universidade Federal do Pará - UFPA. E-mail: arazky@gmail.com.

** Doutora em Linguística pela Universidade Federal do Pará – UFPA. Professora Classe III, da Secretaria de Educação – SEDUC/PA. Professora substituta da Universidade do Estado do Pará – UEPA. E-mail: carlene.salvador77@gmail.com.

***Doutorando em Linguística pela Universidade Federal do Pará – UFPA. Professor substituto da Universidade do Estado do Pará – UEPA. Professor efetivo da Secretaria Municipal de Mãe do Rio – SEMED/PA. E-mail: davips312@gmail.com

**** Doutora em Linguística pela Universidade Federal do Pará – UFPA. Pesquisadora do GeoLinTerm. E-mail: rejane.garcez@gmail.com

and guided by the methodology Linguistics of *Corpus* (BERBER SARDINHA, 2004; TAGNIN, 2009). These are two doctoral thesis – *Estudo das fraseologias do futebol brasileiro das Séries B, C e D em jornais digitais populares: construção de um dicionário eletrônico* (SALVADOR, 2017), *O léxico especializado do corte bovino: uma abordagem terminológica e terminográfica* (OLIVEIRA, 2018) and a master's thesis – *Fraseologismos no discurso político brasileiro: uma proposta de glossário* (SOUZA, 2018), GeoLinTerm Project – PPGL / UFPA. The three works add up 2.954 phraseologisms, which are described such as (*gol de placa, chute de bucho condenado, crime de colarinho branco*).

KEYWORDS: Phraseologisms; Soccer; Bovine Cutting; Political speech; GeoLinTerm.

Introdução

O léxico de uma língua é uma fonte inesgotável de pesquisa, o qual se renova constantemente, acompanhando, em diferentes estágios, as mudanças sociais, políticas, econômicas, culturais e, naturalmente, linguísticas de uma dada comunidade. Por esse motivo, embora se possa argumentar que a descrição lexical no Brasil já disponha de um acervo significativo de dados, não seria forçoso admitir que as pesquisas realizadas na região Norte do país, acerca das unidades lexicais, ainda são poucas. No âmbito do Projeto Geossociolinguística e Terminologia – GeoLinTerm¹ (RAZKY; OLIVEIRA; LIMA, 2010), as produções científicas documentam parte desse universo, por meio de monografias, atlas linguísticos, glossários e dicionários terminológicos. No entanto, até pouco tempo, esses trabalhos voltavam-se apenas para as unidades lexicais monovocabulares. Todavia, o léxico não se constitui apenas de lexias simples (POTTIER, 1974), ao lado destas atuam, de forma numerosa e expressiva, as chamadas unidades fraseológicas ou fraseologismos², que são sequências polilexicais resultantes de associações sintagmáticas recorrentes, como as define Mejri (1997, 1998, 2012).

Para adentrar o campo fraseológico, faz-se necessário explicitar a distinção existente entre Fraseologia como a ciência que se ocupa do estudo das expressões fixas e que a partir de 1950 passou a receber atenção especial de estudiosos da linguagem, como foi o caso dos franceses e dos russos; e fraseologia como fenômeno, em outras palavras, estruturas que apresentam significação estabelecida a partir de um conjunto de elementos que as formam, o que lhes confere um sentido global, em detrimento do

¹ <http://geolinterm.com.br/>

² Apesar de o uso de nomenclaturas diferentes *expressões idiomáticas, unidade fraseológica, expressões pluriverbais*, dentre muitas outras, optou-se em utilizar a designação *fraseologismo* por ela indicar a nomenclatura preferencial de Mejri (1997).

entendimento isolado de seus constituintes, e o seu elevado grau de lexicalização. Dentre as propriedades que lhe são peculiares, restritas, o que caracteriza um fraseologismo é a convencionalidade de seu agrupamento e a sua memorização como um bloco coeso, sua sequência pode ser recuperada da memória pelo usuário da língua como um todo e reconhecida como uma unidade informacional.

Como *de grão em grão a galinha enche o papo*, as investigações de base fraseológica, *um dos braços* do Projeto GeoLinTerm, começaram a *ganhar fôlego!* Sob a orientação de Abdelhak Razky, as três pesquisas aqui referenciadas, as quais são vinculadas ao Programa de Pós-graduação em Letras – PPGL/UFPA, foram finalizadas no biênio 2017-2018. Nesse contexto, o objetivo deste artigo consiste em apresentar a síntese dessas primeiras pesquisas de cunho fraseológico desenvolvidas na Região Norte do Brasil. Para tanto, convém elucidar que os dados apresentados são fragmentos, extratos das Teses de Doutorado realizadas por Salvador (2017) e Oliveira (2018), assim como da Dissertação de Mestrado de Souza (2018).

Este artigo encontra-se estruturado a partir deste texto introdutório, em três seções: na seção 1, são apresentadas as bases teóricas que fundamentaram este estudo e os critérios para identificação e delimitação de fraseologismos. A seção 2 traz uma explanação sobre as decisões metodológicas gerais adotadas nas pesquisas-base descritas (SALVADOR, 2017; OLIVEIRA, 2018; SOUZA, 2018) e elucida a opção metodológica tomada para a escrita deste artigo. A seção 3 apresenta ao leitor as sínteses das primeiras pesquisas fraseológicas realizadas no âmbito da região Norte do Brasil.

1. Fraseologia ou fraseologismos: faces da mesma moeda?

Nesta seção, são apresentadas as definições dos principais conceitos do campo fraseológico, discutindo-se brevemente a problemática terminológica que caracteriza a área, com destaque para os critérios utilizados na identificação e delimitação dos fraseologismos.

O termo *phraséologie* (fraseologia) foi utilizado pela primeira vez no Ocidente, com o sentido que possui hoje, no início do século XX, precisamente em 1909, por Charles Bally (ORTIZ ALVAREZ; UNTERNBÄUMEN, 2011), discípulo de Saussure que se tornou o *Pai da Fraseologia*. Sob esse termo genérico, Bally incluía as locuções

compostas, isto é, o conjunto formado por várias palavras que perdem seus significados individuais para compor um novo significado global, geralmente figurado, que se estende a todo o bloco sintagmático, caracterizando, assim, o tipo de unidade lexicológica em que “[...] l'unité psychologique excède les limites de l'unité graphique et s'étend stir plusieurs mots” (BALLY, 1951, p. 5).

Por sua vez, Ortiz Alvarez (2000, 2000, p.73) define a fraseologia como a “combinação de elementos lingüísticos 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.” A título de exemplificação, em português, ilustram o fenômeno fraseológico as sequências *quebrar um galho* e *bater as botas* que significam, respectivamente, <ajudar alguém> e <morrer>.

Após os estudos pioneiros de Charles Bally a respeito do fenômeno fraseológico, sucedem-se inúmeros trabalhos de linguistas interessados na temática fraseológica.

Ao longo do desenvolvimento dos estudos fraseológicos, o termo fraseologia passou a ser utilizado de forma polissêmica, sendo, por vezes, alvo de controvérsia (cf. CORPAS PASTOR, 1996; OLIVEIRA SILVA, 2011; TAGNIN, 2011), referindo-se, de modo geral, tanto à unidade/conjunto de unidades fraseológicas de uma língua (fraseologia geral) ou de um domínio temático específico (fraseologia especializada) quanto à disciplina/campo de conhecimento que delimita tais unidades como objeto de investigação. Para diferenciar essas acepções, costuma-se adotar a grafia *fraseologia*, para o primeiro caso, e *Fraseologia*, para o segundo.

Em decorrência da diversidade de pontos de vista sobre a fraseologia, a área possui uma terminologia que apresenta significativa pluralidade denominativa. Porém, os termos mais recorrentes são *unidade fraseológica* e *fraseologismo*, segundo Ortiz Alvarez (2012).

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. Devido a isso, Mejri (2012) propõe um reordenamento terminológico que se baseia na oposição conceitual entre o fenômeno *phraséologie* (fraseologia) e o processo por meio do qual ele se exprime nas línguas naturais, a saber:

le figement (cristalização ou congelamento). Para o autor, *phraséologie* e *figement* designam conceitos distintos, porém 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, 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 tunisiano 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* (frasesmas), tendo esta última designação uma implicação de ordem teórica específica no âmbito da proposição Sentido-Texto, defendida por 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 termos como sequências cristalizadas, expressões cristalizadas etc. (MEJRI, 2012).

2. Critérios para identificação e delimitação de fraseologismos

Embora os fraseologismos sejam muito comuns e altamente produtivos nas línguas naturais, não é tarefa fácil identificá-los e delimitá-los objetivamente. Não há apenas um critério e muitos dos que são propostos nem sempre atendem de forma satisfatória à descrição dessas unidades, principalmente se forem tomados isoladamente. Por esse motivo, optou-se, nos trabalhos aqui descritos, pelos critérios adotados na vertente francesa da fraseologia, sobretudo no que propõe Mejri (1997; 2012), para quem as propriedades identificadoras de um fraseologismo são a fixidez e a

congruência, que, por sua vez, levam em consideração critérios como polilexicalidade, idiomaticidade, frequência de uso, entendidos de forma integrada como fatores complementares, caracterizando um *continuum* no processo de cristalização.

A polilexicalidade, por exemplo, é uma das propriedades mais evidentes de um fraseologismo, sendo-lhe inerente. Essa característica se refere à composição estrutural do fraseologismo, o qual apresenta um bloco sintagmático formado por duas ou mais unidades lexicais que mantêm entre si estreita relação de solidariedade sintática e semântica. Difere-se, portanto, do caráter monolexical, propriedade utilizada de maneira exclusiva para definir a tradicional concepção de palavra, a qual seria a unidade que possui apenas um item lexical. Assim, o verbo *morrer* constituiria uma unidade monolexical, ao passo que o fraseologismo *bater as botas*, mesmo tendo sentido equivalente ao do verbo, possui estrutura polilexical pelo fato de ser uma combinatória formada por mais de um componente. Convém, porém, ressaltar que nem toda sequência polilexical constitui um fraseologismo, mas todo fraseologismo apresenta necessariamente uma estrutura polilexical.

Diferentemente, a idiomaticidade é uma propriedade semântica por meio da qual o fraseologismo pode ter sentido opaco ou transparente, a depender da estrutura semântica e dos tropos convencionalizados em uma língua a partir das relações culturais estabelecidas pela comunidade. Por meio dessa característica, costumam-se identificar as expressões idiomáticas cujo sentido geralmente é opaco e não-composicional, isto é, não resulta meramente da soma dos significados individuais dos elementos que formam a expressão.

Por sua vez, a frequência de uso constitui um critério estatístico que pode indicar o grau de convencionalidade ou de institucionalização dos fraseologismos. Refere-se à ocorrência da unidade em um dado texto ou *corpus*. Embora haja estudos que comprovam a alta frequência das unidades fraseológicas em comparação às unidades simples, esse critério não deve ser tomado de forma isolada para identificar um fraseologismo, pois a frequência depende de outros fatores relacionados ao contexto mais amplo onde se estabelecem as interações verbais, como o domínio discursivo, a esfera de circulação do gênero, o tema do texto, a modalidade oral ou escrita da língua, o nível de formalidade, a dimensão do *corpus* etc. Um domínio especializado, por

exemplo, tende a apresentar nos textos de área maior número de colocações que provérbios, mais presentes no uso cotidiano.

Esses critérios e os demais abordados por Gross (1996) podem ser resumidos e reinterpretados à luz dos conceitos de fixidez e de congruência apresentados por Mejri (2012), que os consideram novos elementos metodológicos para a delimitação de fraseologismos. Para o autor, o cruzamento dessas propriedades permite identificar o fraseologismo e diferenciá-lo de uma sequência livre. Na seção seguinte, são descritas as etapas metodológicas traçadas pelos autores nortistas a respeito de cada pesquisa realizada.

3. Metodologia

Esta seção tem por objetivo explicar tanto sobre as decisões metodológicas gerais adotadas nas pesquisas-base descritas (SALVADOR, 2017; OLIVEIRA, 2018; SOUZA, 2018), mesmo que sucintamente, quanto elucidar a opção metodológica tomada para a escrita deste artigo.

As investigações empreendidas atenderam a etapas específicas de constituição de cada *corpus*. Etapas essas que se referem ao processo de coleta, compilação e extração de dados e à organização dos dicionários e do glossário fraseológico. Para a concretização dos objetivos propostos foi necessário percorrer um caminho metodológico de construção de materiais que subsidiaram a organização das amostras coletadas de modo que os resultados alcançados pudessem ser dispostos nos dicionários e no glossário, a saber: a ficha fraseológica, a organização do dicionário, a estruturação e registro dos fraseologismos.

Um ponto relevante, comum aos três trabalhos apresentados, é o fato de os autores optarem por utilizar a abordagem metodológica da Linguística de *Corpus* (BERBER SARDINHA, 2004), o que significa dizer que houve sistematização na constituição dos *corpora*, pois os textos, orais e escritos, foram preparados para a compilação no *software WordSmith Tools, versão 5.0* (SCOTT, 2008). O uso metodológico dessa abordagem possibilitou, além da agilidade na organização dos dados, o levantamento dos candidatos a fraseologismos, principalmente pelo uso do dispositivo *Cluster*, o qual fornece possibilidades de agrupamentos encabeçados por

palavras-chave, facilitando a procura por fraseologismos em meio a diversidade de dados de cada *corpus*.

Ainda no que se refere ao processo de manipulação e organização das obras sob análise, destaca-se o uso da ferramenta computacional *Lexique Pro* Versão 3.6 (SIL, 2010), a partir da qual foi possível organizar a macro e a microestrutura dos dicionários e do glossário. Os três repertórios oferecem tanto a versão impressa quanto a eletrônica, sendo apenas os dois dicionários de apresentação semi ilustrada.

O processo de transposição dos fraseologismos encontrados para os compêndios elaborados revelou também a necessidade de os autores proporem a validação dos dados, isto é, tanto especialistas de cada área quanto não-especialistas (SOUZA, 2018) foram consultados com o intuito de que pudessem certificar a natureza das definições estabelecidas para cada entrada dos verbetes. A preocupação em contemplar a participação de não-especialistas no processo de validação das definições surgiu a partir da verificação de que em dados oriundos de um *corpus* não especializado, como se mostrou, por exemplo, o domínio político, o informante aciona, mesmo que de maneira inconsciente, a competência fraseológica³, assumindo o papel real de um falante capaz de reconhecer estruturas fraseológicas e seus respectivos sentidos. A contribuição de todos os validadores envolvidos, somada à consulta aos materiais de referência em língua portuguesa, possibilitou assegurar a fidelidade aos conceitos propostos e a reelaboração de eventuais discordâncias encontradas.

No que se refere à escrita deste artigo, os autores optaram metodologicamente por adotar a perspectiva documental (GIL, 1999), uma vez que o objetivo central deste estudo consiste em apresentar, em forma de síntese, as três primeiras investigações fraseológicas produzidas no Norte do país, no âmbito da UFPA, as quais ainda não foram objeto de análise em outras pesquisas. A descrição dos procedimentos metodológicos adotados e das demais especificidades das investigações está registrada

³Ortiz Alvarez (2014) define competência fraseológica: “[...] capacidade de mobilizar saberes e conhecimentos adquiridos e experienciados para conseguir identificar, compreender, reconhecer, interpretar e decifrar uma unidade fraseológica dentro de um determinado contexto; é saber processar a informação e carga cultural registrada nessas expressões, características do povo e comunidade que as criou e institucionalizou e assim poder reutilizá-las em outras situações comunicativas de acordo com os objetivos dos sujeitos agentes da interação e do contexto em que se inserem.” (ORTIZ ALVAREZ, 2014, p. 280 apud SILVA, 2016, p. 45).

na seção seguinte, em que, a partir de cada pesquisa, delineiam-se os passos empreendidos por cada autor.

4. Apresentação e discussão dos resultados

4.1 Primeiros estudos fraseológicos na Região Norte: percorrendo caminhos possíveis

A proposta de registrar, descrever e analisar parte da diversidade fraseológica brasileira, a partir do contexto de pesquisa paraense, motivou as investigações tratadas neste artigo, as quais, por meio de coleta sistemática de dados referentes a três domínios discursivos distintos, *futebol, corte bovino e política*, culminaram na elaboração de dois dicionários e de um glossário, os quais são resumidamente apresentados a seguir.

4.1.1 Estudo das fraseologias do futebol brasileiro das Séries B, C e D em jornais digitais populares: construção de um dicionário eletrônico (SALVADOR, 2017)

O primeiro estudo fraseológico da região Norte do Brasil, temática assumida na tese de Doutorado *Estudo da fraseologia do futebol brasileiro das séries b, c e d em jornais digitais populares: construção de um dicionário eletrônico*, foi desenvolvido por Salvador (2017), bolsista Fapespa/PA, sob a orientação de Abdelhak Razky. Arelada ao Projeto GeoLinTerm – PPGL/UFPA, a pesquisa abordou uma das formas de conhecimento da história do pensamento social no decorrer dos séculos, a qual está presente em um vasto número de construções sintagmáticas, estáveis, os chamados fraseologismos. Tais construções são, segundo o viés teórico assumido por Mejri (1997), “[...] portadoras das vivências de uma ou mais gerações e funcionam como instrumentos de conduta aptos para serem aplicados no cotidiano” (MEJRI, 1997, p. 48).

Construída sob três pilares principais, o futebol, a fraseologia e o dicionário fraseológico, Salvador (2017): i) baseou-se em Guterman (2015) para efetuar o apanhado histórico da trajetória do futebol e a gradual condição adquirida por esse esporte dentro da sociedade brasileira em contextos históricos diferentes; ii) ancorou-se na fundamentação teórica defendida pela vertente francesa de Fraseologia, cuja taxonomia proposta por M. Gross (1982), G. Gross (1996) e Mejri (1997, 1998, 2011, 2012) viabiliza o entendimento do processo de estabilização sintagmática dos

fraseologismos tanto do ponto de vista linguístico quanto semântico; iii) registrou na ferramenta computacional *Lexique Pro* (SIL, 2010) os 1316 fraseologismos encontrados.

Nesse contexto, inicialmente procedeu-se ao apanhado teórico necessário e inerente a qualquer pesquisa, sob a justificativa de que a linguagem do futebol, assim como a dinâmica que cerca esse domínio discursivo favorecem a alta produtividade de fraseologismos *gol de letra, bola da vez*, o que motivou a realização de uma investigação que verificasse o alcance e as ocorrências dessas unidades em jornais classificados como *populares*.

Quanto aos procedimentos metodológicos, a investigação empreendida por Salvador (2017), de caráter descritivo (VERGARA, 2000), apresenta natureza quanti qualitativa (DEMO, 2002). As informações obtidas pela autora são oriundas da técnica de coleta bibliográfica na medida em que se busca o contato com escritos acerca da tríplice temática estudada (LAKATOS; MARCONI, 2001), quais sejam: fraseologia, futebol e dicionário.

Assim, os dados utilizados pela autora paraense foram obtidos a partir de coleta sistemática de textos sobre futebol presentes na coluna *Caderno de Esporte*, na mídia impressa *online*, de cinco jornais populares das seguintes capitais brasileiras: Belém/PA, Goiânia/GO, Porto Alegre/RS, Salvador/BA e Rio de Janeiro/RJ. Os textos publicados regularmente nos *sites* dos referidos jornais são referentes ao recorte temporal compreendido entre os anos de 2008 e 2015. A coleta de dados teve início em janeiro de 2014 e foi finalizada em julho de 2015. A amostra coletada passou pelo processo de limpeza, conversão em documentos de extensão *.txt* e codificação. Em seguida, foi realizada a extração dos candidatos a fraseologismos, conforme Berber Sardinha (2004) e Tagnin (2005).

O processo seletivo das fontes de informação foi determinado, fundamentalmente, pela atualidade e abrangência de cada um dos representantes da mídia. Os chamados *jornais populares* dedicam inúmeras páginas, cadernos ou suplementos ao noticiário esportivo, gerando, assim, uma fonte abundante de dados, fato constatado no decorrer da investigação. O *Caderno de Esportes* é uma das seções em que está dividido um jornal, trazendo informações sobre as mais variadas modalidades esportivas praticadas no Brasil e no mundo, no entanto, para os propósitos

da tese de Salvador (2017), foram coletadas apenas notícias referentes ao futebol de campo masculino.

Os princípios orientadores da Linguística de *Corpus* foram utilizados na pesquisa, como abordagem metodológica para a coleta de dados em termos quantitativos, tomando por base Berber Sardinha (2004). Assim, o *corpus* foi preparado e sistematizado, para emprego das ferramentas presentes no *software WordSmith Tools*, em sua versão 5.0 (SCOTT, 2008). Após as etapas iniciais de tratamento da amostra, realizou-se a limpeza, a conversão dos textos em formato *.txt*, pois esta é uma das extensões lida pelo referido *software*, e a codificação do *corpus* com vistas a agilizar a recuperação das informações coletadas e o processo de referenciação do *corpus*. Essa etapa foi realizada em quatro meses de intenso trabalho.

Berber Sardinha (2004), em seu artigo sobre Linguística de *Corpus*, pontua que a definição mais apropriada para *corpus* é:

Um conjunto de dados lingüísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso lingüístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise (BERBER SARDINHA, 2004, p. 18).

A amostra composta por 2674 textos deu origem a um *corpus* de porte médio, pois, ao se fazer a primeira rodada, o *software* contabilizou um pouco mais de cinco milhões de palavras, critérios que atendem parte das noções descritas por Berber Sardinha (2004). Foram realizadas rodadas no *WordSmith* a partir das quais foi possível elencar os candidatos a fraseologismos para que se pudessem aplicar os testes sugeridos por Mejri (2012) e avaliar a natureza das ocorrências encontradas quanto aos critérios da polilexicalidade, da fixidez, da congruência, da previsibilidade e da idiomaticidade, o que possibilitou realizar a identificação, a descrição e a categorização dos fraseologismos encontrados, além de aferir o grau de estabilidade.

A título de exemplificação da etapa de testes aplicada a cada candidato fraseológico, observe-se o caso de *auxiliar técnico* <ajudante do técnico da equipe>. Em primeiro plano, verificou-se o caráter polilexical da unidade, materializado em dois constituintes; em seguida, verificou-se a natureza de sua fixidez, em que houve a aceitação para a inserção da marca de número *auxiliar (es) técnico (s)*, mas a não

aceitação em caso de deslocamento dos constituintes **técnico auxiliar* sem que o sentido (congruência) do fraseologismo fosse desfeito. A partir dessa informação foi possível efetuar a tipologia da ocorrência, um exemplo de colocação, uma vez que não se observa o caráter conotativo, isto é, os itens lexicais apresentados são classificados como transparentes (idiomaticidade), pois suas acepções individuais estão preservadas. O fato de a ocorrência não apresentar opacidade conduziu à constatação de seu aspecto preditivo, neste caso, aferido pela alta frequência de coocorrência dos dois itens lexicais. Esse processo de verificação que conduzia à tipologia dos eventos encontrados ocorreu em todos os 1316 casos, apontando para uma amostra composta por 57% de fraseologismos semifixos, na sua maioria, sintagmas verbais, com alto índice conotativo que servem tanto ao domínio discursivo futebolístico quanto à língua geral.

Após a etapa de aplicação dos critérios de identificação sugeridos por Mejri (2012), o *software Lexique Pro* (SIL, 2010) foi utilizado como ferramenta na organização do dicionário eletrônico e os verbetes foram semasiologicamente dispostos. As ocorrências dispostas no dicionário fraseológico revelam exemplos emblemáticos do domínio do futebol, tais como: *pisar na bola, gol de bicicleta, lá onde a coruja dorme*. A alta produtividade de fraseologismos de uso geral indica a proximidade do futebol com o cotidiano das pessoas, razão pela qual se faz necessário investigá-los.

4.1.2 O Léxico Especializado do Corte Bovino: uma abordagem Terminológica e Terminográfica (OLIVEIRA, 2018)

A tese intitulada *O Léxico Especializado do Corte Bovino: uma abordagem Terminológica e Terminográfica* (OLIVEIRA, 2018) é um estudo que evidencia, de forma inédita, devido à inexistência de obra terminográfica similar, a variação terminológica e fraseotermológica existentes na pecuária de corte bovino no Pará. Composto os primeiros estudos fraseológicos no Norte do Brasil, foi um estudo que buscou identificar, registrar e analisar nomenclaturas especializadas orais e escritas, oferecendo como produto um dicionário eletrônico cuja finalidade é sistematizar esse léxico especializado, tornando mais viável o acesso à diversidade linguística existente na Cadeia Produtiva do Corte Bovino - CPCB.

Para o desenvolvimento da tese, algumas orientações teóricas foram adotadas a partir da Teoria Comunicativa da Terminologia-TCT, postulada por Cabré (1993; 1996), em relação à Terminologia e sua variação; do arcabouço proposto por

Bevilacqua (1996; 2004) sobre Fraseologia Especializada e das diretrizes de Mejri (1997; 1998; 2000; 2006), que descreve e explica o funcionamento dos fraseologismos nas línguas. Dessa forma, tornou-se possível a identificação e o registro do léxico especializado do corte bovino em contextos orais e escritos frequentemente utilizados nos processos dessa cadeia produtiva no estado do Pará.

Como pesquisa socioterminológica, a metodologia adotada fundamentou-se em registros orais e escritos constituindo um repositório de formas e variantes que justificam o porquê de se privilegiar algumas formas patrimoniais em detrimento de outras originadas da cultura local. A pesquisa de campo possibilitou a recolha dos termos e a constituição do *corpus*. Entrevistas e questionários aplicados aos informantes foram instrumentos de pesquisa que nortearam a pesquisadora durante a seleção e os registros das unidades terminológicas especializadas.

Os fraseologismos identificados no léxico especializado do corte bovino têm a função de nomear, de designar novas realidades científicas ou técnicas, mas, como referentes que são, sempre recebem alguma influência da cultura local, da economia, do contexto, enfim. A formação estrutural dos fraseologismos encontrados compõe-se, basicamente, de sintagmas nominais, geralmente mais frequentes na pecuária de corte, a exemplo de *mesa de inspeção*, *concussão cerebral*, *amaciador de bife*; em menor número de sintagmas verbais, como ocorre em *arriar bucho* e de sintagmas de função adjetiva, com ou sem construções paralelas, como ocorre em *material específico de risco*.

Neste estudo, as categorias analisadas restringem-se às variantes cuja estrutura sintática e tipologia decorrem das diversidades existentes na sociedade e de um conjunto de fatores determinantes desse léxico especializado. Posto isso, as fraseologias encontradas no *corpus* de estudo seguem a classificação de Inez Sfar (2013, p. 4) quanto à i) estrutura sintática, com expansões adjetivas como *abate humanitário*, *abate do rabino*, *carne morta*, *carne sifada*, *carne maturada*, *calha móvel*, *calha de resíduos*, *bucho branqueado*, *bucho condenado*; expansões nominais: *chuveiro da seringa*, *lagartinho da pá*, *músculo do coxão mole*; expansões polilexicais: *toalete da carcaça*, *trânsito das carcaças*, *serra da carcaça*, *câmara de carcaça (expansão + carcaça)*, *desossa de cabeças*, *desossa a quente do sequestro*, *recortes de desossa (desossa + expansão / expansão + desossa)*, *carne desossada de bovino*, *carne de primeira*, *carne*

nobre, carne morta, carne de sangria (carne + expansão). Essas unidades sintagmáticas e outras como *toalete da carcaça, carro de transporte de miúdos sequestrados, lavagem da carcaça bovina, vaca reformada para abate* etc. não são meras expressões ou sintagmas, mas fraseologias especializadas formadas por dois ou mais elementos. Alguns exemplos apresentam termos específicos do corte bovino tal como *abate, miúdos, carcaça bovina, vaca*, fato que os torna mais simples de serem identificados como fraseologias especializadas, enquanto que outros apresentam constituintes que, isoladamente, nada dizem sobre a área do corte bovino, mas em cadeia representam um objeto específico da área em estudo (*chuveiro da seringa, lagartinho da pá, coxão mole, antecâmara de entrada*). ii) Tipologias. No *corpus* analisado, os fraseologismos especializados se apresentaram de forma produtiva. Considerando essa alta frequência, achou-se oportuno tipificá-las de acordo com Mejri (2012). O Dicionário do Corte Bovino do Pará (OLIVEIRA, 2018) é formado por 1055 entradas. Destas, 644 são fraseologismos e 411 são termos. Entre os fraseologismos especializados destacam-se: as colocações, com 23%, os semicristalizados, com 56%; e os cristalizados, com 21% do total. A seguir, algumas unidades sintagmáticas exemplificam a tipologia das Unidades Fraseológicas Especializadas – UFES, identificadas no corte bovino e registradas por Oliveira, em 2018, no Dicionário do Corte Bovino do Pará.

As sequências identificadas como Semicristalizadas, a exemplo de *abate humanitário* ou *vaca reformada para abate*, apresentam fixidez parcial, uma vez que admitem regras sintagmáticas ou parte delas (posposição do adjetivo; coordenação ou agrupamento dos adjetivos; modificação pelos advérbios; acréscimo de adjetivos pospostos ou de modificadores preposicionais) e a comutatividade paradigmática. Foram, portanto, consideradas mais transparentes do ponto de vista da idiomaticidade.

Já sequências como *filé mingnon* ou *bife do vazio* são construções cristalizadas por não aceitarem alterações em sua estrutura sintagmática. Constituem formas opacas, porque sofreram o congelamento de suas regras de combinação sintagmáticas nos planos sintático, semântico e pragmático de tal forma que não foi possível submetê-las aos critérios que permitem medir o caráter de fixidez das unidades (já mencionados no Aporte Teórico). São nomes de cortes que o uso consagrou, uma vez que quando se fala em *filé*, por exemplo, já se sabe que é *filé mingnon*. Portanto, o uso cristalizou essas

formas nesse léxico de especialidade. Se houver permuta em algum dos constituintes, muda o sentido, portanto, altera a congruência da unidade.

Quanto às fraseologias tipificadas como colocações, a princípio considerou-se que nas linguagens especializadas todos os instrumentos são termos técnicos e, por isso, deveriam ser classificados como colocações. Entretanto, ao serem analisadas as ocorrências, percebeu-se que as colocações são estruturas muito próximas, mas tão próximas que se atraem. Funcionam como ímãs e pertencem ao conjunto das unidades fraseológicas. São restrições paradigmáticas fortes. Ao se falar de *fundação esquerda* (referência à lateral esquerda do quarto traseiro do boi), a dissociação dos elementos que compõem o sintagma constitui uma incongruência porque fere o critério pragmático. O mesmo ocorre com *quarto dianteiro* e *quarto traseiro* em que os elementos decompostos também ferem esse critério. As colocações são fraseologias que no corte bovino têm alto grau de cristalização, a ponto de serem consideradas como unidades estáveis e pré-fabricadas no sistema da língua.

De acordo com os modelos fraseológicos mais usuais, estudados e postulados por Mejri (2012), na terminologia do corte bovino, a unidade fraseológica surge como hiperônimo que denomina o objeto de estudo da fraseologia especializada, podendo abranger vários tipos de combinações, desde as compostas por duas palavras até aquelas que constituem frases inteiras, tendo a polilexicalidade e a fixidez como principais características.

Os resultados alcançados com esta investigação convergem, portanto, ao atendimento dos objetivos da pesquisa, mediante a hipótese de que junto aos termos e a partir deles, esse léxico especializado é composto em sua maior parte por fraseologias especializadas para referenciar objetos, seres e conceitos nessa área de atividade; que a nomeação que se faz se dá por meio de variantes as quais nem sempre são transparentes a todos os envolvidos na CPCB e que, em consequência disso, o discurso, seja ele oral ou escrito, pode sofrer comprometimentos relativos aos aspectos sintático, semântico e pragmático.

4.1.3 Fraseologismos no discurso político: uma proposta de glossário (SOUZA, 2018)

Na esteira dos trabalhos pioneiros desenvolvidos na área da Fraseologia no Norte do Brasil, a pesquisa intitulada *Fraseologismos no discurso político: uma*

proposta de glossário (SOUZA, 2018) constitui a primeira dissertação de mestrado defendida nesse eixo de investigação no Programa de Pós-graduação em Letras da UFPA, situando-se como a terceira produção acadêmica no campo fraseológico no âmbito do projeto GeoLinTerm.

O referido trabalho, inscrevendo-se sob a temática fraseologia do discurso político, teve como objetivo principal produzir um glossário, em versão impressa e eletrônica, de fraseologismos utilizados no discurso político brasileiro. Por sua vez, os objetivos específicos consistiram em: descrever os fraseologismos que caracterizam o discurso político brasileiro; identificar padrões de combinatórias sintagmáticas recorrentes no *corpus* e; verificar possíveis variantes fraseológicas.

A justificativa para a realização da 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 os fraseologismos 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. 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; portanto, neste aspecto, a 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 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.

Quanto ao referencial teórico, o trabalho ancorou-se na abordagem francesa da Fraseologia, sobretudo na perspectiva de Salah Mejri (1997; 2012; 2018), no âmbito da qual os fraseologismos, ou unidades fraseológicas, são definidos como combinações sintagmáticas recorrentes (MEJRI, 1997; 2012), caracterizados, 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, 2012). 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. O *corpus* constituído possui 570 textos 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 escolhido um colunista por periódico. Os textos foram publicados entre janeiro de 2014 e dezembro de 2016. Optou-se por esse recorte temporal porque nesse período estão circunscritos o início da Operação Lava-Jato, aproximadamente em março de 2014, e o processo de *Impeachment* da ex-presidente Dilma Rousseff, do Partido dos Trabalhadores (PT), acusada de cometer pedaladas fiscais e, assim, comprometer as finanças públicas do País.

Para tratamento dos textos, foram utilizados os *softwares WordSmith 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*.

O glossário produzido contém 438 entradas, lematizadas e organizadas alfabeticamente pela primeira unidade lexical da sequência. Os resultados demonstram a predominância de fraseologismos semifixos, oriundos 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, dentre 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.

Dessa maneira, o conjunto desses três trabalhos pioneiros no Norte do Brasil no campo da fraseologia representa o *passo inicial*, o *primeiro lance* no desafio mais amplo de documentar, descrever e analisar as línguas naturais, considerando especialmente o contexto amazônico, no qual o escopo fraseológico é praticamente inexplorado. Apesar da enorme tarefa que está posta, os três repertórios fraseológicos produzidos somam, juntos, 2954 fraseologismos devidamente registrados, como sintetiza o quadro a seguir.

Quadro 01 – Fraseologismos extraídos dos *corpora*

| Exemplos extraídos dos <i>corpora</i> | | |
|-------------------------------------------------|------------------------------------------------------|----------------------------------------------|
| Salvador (2017) Futebol (1316) | Oliveira (2018) Corte bovino (1200) | Souza (2018) Política (438) |
| Árbitro auxiliar | Contra filé | Lava jato |
| bomba rasteira | Carne moída | Corrupção passiva |
| Gol de placa | Chute de bucho condenado | Crime de colarinho branco |

Fonte: Elaboração dos autores, 2019.

A variedade listada no Quadro 1, mesmo se tratando de domínios diferentes, exhibe, de um lado, a produtividade fraseológica no âmbito desses domínios e, de outro, a necessidade de novos estudos que auxiliem a descrição do caudal fraseológico existente. Nesse sentido, as pesquisas realizadas pelos autores deste artigo constituem

os primeiros grãos que, somados a outros, poderão *preencher os papos* da pesquisa fraseológica na Amazônia.

CONSIDERAÇÕES FINAIS

O presente artigo teve como objetivo apresentar uma síntese dos primeiros trabalhos acadêmicos em nível de pós-graduação que foram produzidos, no Norte, no âmbito da fraseologia. Sob a orientação do Prof. Abdelhak Razky e co-orientação da Profa. Maria Luísa Ortiz Alvarez (UnB), as três pesquisas sistematizam dados, discutem problemas teóricos, apresentam a diversidade terminológica existente na literatura da área e relatam dificuldades metodológicas encontradas na elaboração de dicionários e glossário no que tange à delimitação, categorização e segmentação das unidades polilexicais e de seu tratamento fraseográfico, experiências características de todo trabalho que está iniciando.

Os três domínios estudados apresentam alta produtividade fraseológica, assertiva comprovada com o número de verbetes listados em cada repertório, o que revela a competência e o caráter estratégico dos falantes no manejo da língua. Os dados coletados nas três pesquisas apontam para o pioneirismo desses trabalhos no registro, descrição e análise de diferentes temáticas, mas que estão interligados dentro do Projeto GeoLinTerm.

Por tratar-se de um campo aberto e em constante atualização, as pesquisas aqui descritas, apesar de terem sido as primeiras nessa nova linha de pesquisa integrada ao GeoLinTerm, não podem ser consideradas exaustivas. Pelo contrário, buscam, em última instância, motivar trabalhos futuros, como os que já estão sendo feitos por graduandos em Letras, sob a orientação e avaliação dos autores deste artigo, os quais, por também estarem ligados ao projeto ALiB, aprenderam a lição com o mestre Serafim da Silva Neto, que se esforçou no início do século passado para criar uma mentalidade dialetológica no país, tendo em vista a realização do Atlas Linguístico Nacional. Guardadas as devidas dimensões, o grupo de fraseologia do GeoLinterm, coordenado pelo professor Abdelhak Razky, tem se comprometido com a difusão da pesquisa da fraseologia no Norte do país, um campo aberto a novas contribuições e a desafios.

REFERÊNCIAS

- BALLY, Charles. *Traité de stylistique française*. 2ed. Paris: Klincksieck, 1951.
- BERBER SARDINHA, T. *Linguística de Corpus*. Barueri: Manole, 2004.
- BEVILACQUA, C.R. *A fraseologia jurídico-ambiental*. Dissertação (Mestrado). Porto Alegre: Federal do Rio Grande do Sul, Curso de Pós-Graduação em Letras, 1996.
- BEVILACQUA, C. R. *Unidades Fraseológicas Especializadas Eventivas: descripción y reglas de formación en el ámbito de la energía solar*. 2004. 242 f. Tese (Doutorado em Linguística Aplicada). Universidade Pompeu Fabra, Instituto Universitário de Linguística Aplicada (IULA), Barcelona, 2004, cap. 1.
- CABRÉ, M. T. *La terminología: teoría, metodología, aplicaciones*. Barcelona: Antártida, 1993, p. 37.
- CABRÉ, M. T. *Importancia de la terminología en la fijación de la lengua*. *Revista internacional de língua portuguesa*. Núm. 15, jul. 96. Lisboa: Editorial Notícias, 1996, p. 9-24.
- CHARAUDEAU, Patrick. O discurso político. In: EMEDIATO, Wander; MACHADO, I da 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.
- DORNA, Alexandre. *Les effets langagiers du discours politique*. *Hermès, La Revue*, 1995/2 (n°16), p.131-146.
- GIL, A. C. *Métodos e técnicas de pesquisa social*. 5.ed. São Paulo: Atlas, 1999.
- MEJRI, Salah. *Le figement lexical: descriptions linguistiques et structuration sémantique*. Tunis: Publications de la faculté des lettres de la Manouba, 1997.
- MEJRI, Salah. *Congelamento e denominação*. *META*, 45, 4, Presses Universitaires de Montreal, Canadá, p. 609-621, 2000.
- MEJRI, Salah. *Congelamento e fraseologia em francês, Árvores e palavras*. Homenagem D. Blampin, p.169-186, Les Editions du Hazard, Bruxelas, 2006.
- MEJRI, Salah. *Délimitation des unités phraséologiques*. In: ALVAREZ, Maria L. O. (Org.) *Tendências atuais na pesquisa descritiva e aplicada em fraseologia e paremiologia*. Campinas, SP: Pontes, 2012.
- MEJRI, Salah. *La phraséologie française: synthèse, acquisition théoriques et descriptifs*. *Revue de linguistique française*, n. 1, 2018.
- NAVARRO, Carmen. La fraseología en el discurso político y económico de los medios de comunicación. In: AISPIA ctas del XX Congreso [AssociazioneIspanistiItaliani] pp.199-211, 2002.
- OLIVEIRA, R.U.G.S. *O léxico especializado do corte bovino: uma abordagem terminológica e terminográfica*. Tese (Tese - Doutorado em Letras - Estudos Linguísticos) - Faculdade de Letras e Artes, Universidade Federal do Pará, Belém, 2018.

- OLIVEIRA, R.U.G.S. & RAZKY, Abdelhak. *Dicionário do Corte Bovino do Pará*. Universidade Federal do Pará – Belém, Pará. 2018
- OLIVEIRA, R.U.G.S. *A terminologia do corte bovino no Pará*. Dissertação de Mestrado, Faculdade de Letras e Artes da Universidade Federal do Pará, 2013.
- ORTIZ ALVAREZ, M. L.; UNTERNBAUMEN, E. H. *Uma (re) visão da teoria e da pesquisa fraseológica*. Campinas: Pontes Editores, 2011.
- POTTIER, B. *Linguistique générale: théorie et description*. Paris, Klincksieck, 1974.
- RAZKY, A.; LIMA, A. F .de; OLIVEIRA, M. B. de. Geossociolinguística e Socioterminologia no Brasil – GeoLinTerm (projeto de pesquisa). UFPA, 2010. 19f.
- RIBOLDI, Ari. *A CPI das palavras: origem de palavras e expressões da linguagem política*. 2 ed.Caxias do Sul, RS: Lorigraf, 2008.
- SALVADOR, C.F.N. *Estudo das fraseologias do futebol brasileiro das Séries B, C e D em jornais digitais populares: construção de um dicionário eletrônico*. Tese (Doutorado em Letras - Estudos Linguísticos) – Universidade Federal do Pará, Belém, 2017.
- SCOTT, M. *WordSmith Tools*, versão 5.0. Liverpool: *Lexical Analysis Software*, 2008.
- SFAR, Inês. Les collocations dans le discours spécialisés: le cas de la terminologie ornithologique. *Language Design* 14, LDI-Université Paris 13. TIL-Université de Sousse, p. 129-38, 2013.
- SIL (*Summer Institute of Linguistics*) Internacional. *Lexique Pro* (v. 3.6), 2004 -2012.
- SOUZA, Davi Pereira de. *FRASEOLOGISMOS NO DISCURSO POLÍTICO BRASILEIRO: uma proposta de glossário*. Dissertação (Mestrado em Letras – Estudos Linguísticos) - Universidade Federal do Pará, Belém, 2018.
- 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: 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.