

UNIVERSIDADE DO VALE DO RIO DOS SINOS – UNISINOS
UNIDADE ACADÊMICA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA APLICADA
NÍVEL MESTRADO

DIEGO SPADER DE SOUZA

JOGADA DE LETRA:
UM ESTUDO SOBRE COLOCAÇÕES À LUZ DA SEMÂNTICA DE *FRAMES*

SÃO LEOPOLDO

2015

DIEGO SPADER DE SOUZA

JOGADA DE LETRA:

Um estudo sobre colocações à luz da Semântica de *Frames*

Dissertação apresentada como requisito parcial para a obtenção do título de Mestre em Linguística Aplicada, pelo Programa de Pós-Graduação em Linguística Aplicada da Universidade do Vale do Rio dos Sinos – UNISINOS

Área de concentração: Linguagem, Tecnologia e Interação

Orientadora: Prof.^a Dr.^a Rove Luiza de Oliveira Chishman

São Leopoldo

2015

S719j SOUZA, Diego Spader de.
Jogada de letra : um estudo sobre colocações à luz da
semântica de frames / por Diego Spader de Souza. –
2015.
135 f.: il. ; 30 cm.

Dissertação (mestrado) — Universidade do Vale do
Rio dos Sinos, Programa de Pós-Graduação em
Linguística Aplicada, São Leopoldo, RS, 2015.

“Orientação: Prof.^a Dr.^a Rove Luiza de Oliveira
Chishman”.

1. Colocações. 2. Semântica de Frames.
3. FrameNet. 4. Língua portuguesa - Lexicografia
Computacional. I. Título.

CDU: 801.54

Catálogo na Publicação:

Bibliotecário Alessandro Dietrich - CRB 10/2338.

DIEGO SPADER DE SOUZA

“JOGADA DE LETRA: UM ESTUDO SOBRE COLOCAÇÕES À LUZ DA SEMÂNTICA DE FRAMES”

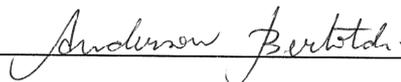
Dissertação apresentada como requisito parcial para obtenção do título de Mestre, pelo Programa de Pós-Graduação em Linguística Aplicada da Universidade do Vale do Rio dos Sinos - Unisinos

Aprovado em 03 de março de 2015

BANCA EXAMINADORA



Profa. Dra. Stella Esther Ortweiler Tagnin (USP)



Prof. Dr. Anderson Bertoldi (UNISINOS)



Profa. Dra. Rove Luiza de Oliveira Chishman (UNISINOS)

*Dedico este trabalho a meus avós
maternos e a seu legado de luta e de
perseverança.*

AGRADECIMENTOS

Ninguém vive sozinho e, ainda que este trabalho tenha sido escrito por duas mãos, ele traz consigo muitos corações. Assim, gostaria de agradecer a vocês.

Antes de mais nada, agradeço à minha família pelo incessante apoio e carinho. Em especial, à minha mãe, por absolutamente tudo que ela já fez e ainda tem feito por mim. Tenho certeza de que eu não estaria escrevendo estas palavras agora se não fosse por toda sua luta e incentivo.

Impossível não agradecer, também, à minha orientadora, Prof.^a Dr.^a Rove Chishman. Rove, há quatro anos tu vens me orientando e me inspirando. Sinto-me sempre honrado por ter a oportunidade de fazer parte do SemanTec e por ter a tua orientação, que é muito, muito importante pra mim. Obrigado.

Muito obrigado, também, a todos os membros do grupo de pesquisa SemanTec (especialmente a vocês, Aline, Ana, João e Lurdes). Mais do que colegas de pesquisa, vocês são amigos que moram permanentemente no meu coração. A amizade de vocês não tem preço. Agradeço pelas nossas reuniões, pelos sábios conselhos, por todo auxílio e por estarem sempre presentes.

Um agradecimento especial vai para o Prof. Anderson Bertoldi, por sempre me escutar, de ouvidos e coração abertos, em momentos em que tudo que eu precisava era falar sobre o trabalho. Obrigado por todas as valiosas dicas e opiniões.

Agradeço, também, à Prof.^a Alexandra Müller, que, junto com o Prof. Anderson, integrou a minha banca de qualificação, que muito contribuiu para o trabalho e para minha formação.

Demais professores e colegas da UNISINOS – tanto do Programa de Pós-Graduação quanto do Curso de Letras, muito obrigado! Cada um de vocês tem um pedaço desta história.

Não posso deixar de agradecer, também, à secretária do Centro 3, e especialmente à Valéria, por todo o carinhoso auxílio durante o Mestrado.

Agradeço, obviamente, a *todos* os meus amigos, sem exceção. Desta vez, prefiro não citar nomes, não quero esquecer de ninguém. Saibam, no entanto, que a importância de vocês na minha vida é simplesmente imensurável. Obrigado por me ampararem, por me apoiarem, por terem me ensinado tanto e, sim, por terem me mostrado os meus erros, quando eu, sozinho, não fui capaz de vê-los.

Obrigado, ainda, aos membros da banca pela leitura cuidadosa do trabalho e por terem aceitado o convite para fazer parte de um momento tão importante para mim.

Por fim, expresso gratidão à CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pelo incentivo, na forma de bolsa de estudos.

Obrigado! Amo vocês.

P.S.: Obrigado, Marina, por ter consultado, às 02:09 da manhã deste sábado, a sua edição de *O Hobbit* apenas porque eu queria a versão em português da minha epígrafe. Ah!, e lembre-se: *jamais ria de dragões vivos*.

— O que você quer dizer com isso? —
perguntou ele. — Está me desejando um
bom dia, ou quer dizer que o dia está bom,
não importa que eu queira ou não, ou quer
dizer que você se sente bem neste dia, ou
que este é um dia para se estar bem?

J. R. R. Tolkien, *O Hobbit*

RESUMO

O objetivo da presente dissertação é discutir a relação existente entre o fenômeno linguístico das colocações e os conceitos da teoria da Semântica de *Frames* (FILLMORE, 1982; 1985). O trabalho se insere no contexto de dois projetos de pesquisa desenvolvidos pelo grupo SemanTec, o Field – Dicionário de Expressões do Futebol (CHISHMAN, 2014), já disponível para consulta na *web*, e o Dicionário Eletrônico Modalidades Olímpicas (CHISHMAN, 2014), ainda em fase inicial. Os dois dicionários citados se organizam a partir da noção de *frame* semântico proposta por Fillmore (1982; 1985), de forma que a dissertação busca evidenciar de que forma esse conceito (e os conceitos que o cercam) repercutem no tratamento lexicográfico dispensado às colocações. Nesse sentido, a revisão da literatura, apresentada nos capítulos 2 e 3, discute as bases teóricas para o estudo das colocações e da Semântica de *Frames*. O método da pesquisa consiste na análise de 74 colocações da linguagem do futebol. A escolha dessas estruturas parte do estudo de 500 combinações lexicais extraídas a partir de um *corpus* em português brasileiro do discurso do futebol através do software Sketch Engine. A análise das 74 colocações selecionadas acontece em duas fases: a primeira se dedica a averiguar os aspectos quantitativos do conjunto de dados e as características estruturais das colocações da linguagem do futebol; a segunda etapa foca na relação dessas combinações com os preceitos teóricos da Semântica de *Frames* e da sua contraparte computacional, a FrameNet, a fim de perceber de que modo esse arcabouço teórico oferece subsídios para o tratamento das colocações em contextos lexicográficos. Entre os principais resultados da primeira fase de análise, destaca-se o fato de que a maior parte das colocações do futebol designa estruturas verbais, como *fazer gol* e *mandar bola*, o que demonstra que a linguagem esportiva é marcada pela dinâmica das ações e dos eventos que ocorrem durante a partida. Além disso, foi possível perceber que as colocações nominais estão fortemente ligadas aos materiais, participantes e locais do contexto futebolístico. A segunda parte demonstrou que as colocações, no âmbito de dicionários baseados em *frames*, atuam como unidades lexicais, conceito proveniente da FrameNet. Ao serem consideradas unidades lexicais, as colocações são evocadoras de *frame*, o que as caracteriza como termos que devem estar presentes na lista principal de verbetes. Foi possível notar, contudo, que a evocação de *frame* a partir das

colocações muitas vezes não segue o modelo tradicional presente na FrameNet, especialmente quando se trata das colocações nominais, que não evocam *acontecimentos*, mas entidades estáticas, como *cartão vermelho* e *tabela de classificação*. A presente dissertação evidencia a relevância da Semântica de *Frames* e da FrameNet para o estudo de unidades complexas como as colocações em contextos lexicográficos. Outro aspecto a ser mencionado é a importância dos recursos metodológicos da Linguística de Corpus para a área em que o estudo se insere.

Palavras-chave: Colocações. Semântica de *Frames*. FrameNet. Lexicografia Computacional. Linguística de Corpus.

ABSTRACT

The present thesis aims at the discussion of the relation that exists between the linguistic phenomenon of collocations and the concepts of Frame Semantics theory (FILLMORE, 1982; 1985). The study has arisen in the context of two research projects developed by the SemanTec group, Field – Football Expressions Dictionary (CHISHMAN, 2014), already available on the web, and Olympic Modalities Electronic Dictionary (CHISHMAN, 2014), still in early stage. Both dictionaries are organized around the notion of semantic frame proposed by Fillmore (1982; 1985), in such a way that the thesis seeks to demonstrate in which way this concept (and the concepts surrounding it) are related to the lexicographic treatment given to collocations. Thus, the literature review, presented in chapters 2 and 3, discusses the theoretical basis of the studies of collocations and Frame Semantics. The research method consists of the analysis of 74 collocations of football language. The choice of these structures was made after the study of 500 lexical combinations extracted from a Brazilian Portuguese corpus of football discourse through the Sketch Engine software. The analysis of the 74 collocations happens in two steps: the first one is dedicated to investigate the quantitative aspects of the data set and the structural characteristics of football language collocations; the second phase focuses on the relation between these combinations and the theoretical assumptions of Frame Semantics and its computational counterpart, FrameNet, in order to see in which way this theoretical outline treats collocations in lexicographic contexts. Among the main results of the first phase of analysis is the fact that a major part of football collocations are verbal, such as *score goal* and *send the ball*, which demonstrates that sport language is marked by the dynamics of actions and events that take place in a game. Besides, it was also possible to realize that nominal collocations are strongly connected to the materials, participants and places of football context. The second phase demonstrated that collocations in the scope of frame-based dictionaries act as lexical units, concept arising from FrameNet. Because they are considered lexical units, collocations are seen as frame evokers, thus characterizing them as terms that must be displayed in the main list of entries. However, it was also possible to note, however, that the frame evoking by collocations many times does not follow the traditional model of FrameNet, especially when it comes to nominal collocations that do not evoke events, but static entities, such as *red card* and

classification table. The present thesis evidences the relevance of Frame Semantics and FrameNet for the study of complex units such as collocations in lexicographic contexts. Another aspect to be mentioned is the importance of the methodological resources of Corpus Linguistics to the area in which this study is included.

Keywords: Collocations. Frame Semantics. FrameNet. Computational Lexicography. Corpus Linguistics.

LISTA DE FIGURAS

Figura 1 - Elementos de <i>frame</i> do <i>frame</i> Agriculture	48
Figura 2 – Relações <i>frame</i> a <i>frame</i> de Agriculture	52
Figura 3 - FrameGrapher para o <i>frame</i> Agriculture.....	53
Figura 4 - Exemplo de palavra-chave em contexto com o verbo "argue"	57
Figura 5 - Índice de ULs da FrameNet	64
Figura 6 - Mapa conceitual dos <i>frames</i> do Field.....	70
Figura 7 - Estrutura do <i>frame</i> Gol Contra.....	71
Figura 8 - Tela do <i>Word List</i> do Sketch Engine	74
Figura 9 – Seleção da busca por <i>collocations</i>	74
Figura 10 - Opções de saída da busca	75
Figura 11 - Resultado duplicado da colocação <i>abrir placar</i>	78

LISTA DE QUADROS

Quadro 1 - Relações <i>frame a frame</i>	50
Quadro 2 - Classificação das colocações	79

LISTA DE ABREVIATURAS E SIGLAS

V	Verbo
S	Substantivo
Adj.	Adjetivo
Art.	Artigo
Prep.	Preposição
UL	Unidade lexical

SUMÁRIO

1 INTRODUÇÃO	18
2 COLOCAÇÕES: REVISÃO DA LITERATURA	21
2.1 O VIÉS LEXICOLÓGICO	23
2.2 O VIÉS LEXICOGRÁFICO	34
2.3 UM APANHADO DO CAPÍTULO	37
3 A SEMÂNTICA DE <i>FRAMES</i>, SUAS APLICAÇÕES LEXICOGRÁFICAS E AS COLOCAÇÕES	39
3.1 A SEMÂNTICA DE <i>FRAMES</i> : BASES TEÓRICAS	39
3.1.1 A FrameNet	46
3.2 <i>FRAMES</i> SEMÂNTICOS E LEXICOGRAFIA: IMPLICAÇÕES PARA O TRATAMENTO DE COLOCAÇÕES	54
3.2.1 Colocações na lexicografia tradicional e na FrameNet	60
3.3 UM APANHADO DO CAPÍTULO	65
4 ETAPAS METODOLÓGICAS DE ANÁLISE	67
4.1 CONTEXTUALIZAÇÃO: OS PROJETOS DE PESQUISA	67
4.1.1 Os <i>frames</i> do Field – Dicionário de Expressões do Futebol	69
4.1.2 O <i>corpus</i> deste estudo	71
4.1.3 O Sketch Engine	72
4.2 PRIMEIRA FASE: PROCEDIMENTOS PRELIMINARES E ANÁLISE ESTRUTURAL	73
4.3 SEGUNDA FASE: ANÁLISE A PARTIR DE <i>FRAMES</i>	76
5 ANÁLISE E DISCUSSÃO DOS DADOS	77
5.1 QUESTÕES PRELIMINARES E ESTRUTURAS	77
5.1.1 Quanto à seleção das colocações	77
5.1.2 Quanto à estrutura	79
5.2 QUESTÕES RELACIONADAS A <i>FRAMES</i>	92
6 CONSIDERAÇÕES FINAIS E PERSPECTIVAS FUTURAS	101
REFERÊNCIAS BIBLIOGRÁFICAS	104
APÊNDICE A – COLOCAÇÕES SELECIONADAS A PARTIR DA BUSCA NO SKETCH ENGINE	108
APÊNDICE B – CLASSIFICAÇÃO DAS COLOCAÇÕES A PARTIR DE HAUSMANN (1989) E DE <i>FRAMES</i>	111

ANEXO A – RELAÇÕES <i>FRAME A FRAME</i>	124
ANEXO B – BUSCA REALIZADA NO SKETCH ENGINE (500 PRIMEIROS RESULTADOS)	125
ANEXO C – <i>FRAMES DO FIELD</i> – DICIONÁRIO DE EXPRESSÕES DO FUTEBOL	132

1 INTRODUÇÃO

A presente dissertação tem como objetivo refletir a respeito da relação entre as colocações e a Semântica de *Frames* (FILLMORE, 1982; 1985), levando em consideração um contexto lexicográfico. Tal contexto, no âmbito deste trabalho, diz respeito ao Dicionário Eletrônico Modalidades Olímpicas (CHISHMAN, 2014) e ao Field – Dicionário de Expressões do Futebol (CHISHMAN, 2014). Ambos são projetos do grupo de pesquisa SemanTec, do Programa de Pós-Graduação em Linguística Aplicada da Universidade do Vale do Rio dos Sinos – UNISINOS. Os dois dicionários, diferentemente de recursos lexicográficos tradicionais, são organizados a partir do conceito de *frame* semântico, da Semântica de *Frames*, o que significa que as palavras que os compõem são agrupadas e contextualizadas a partir dos diferentes cenários em que ocorrem. É nosso objetivo, assim, mostrar em que medida a relação mencionada acima contribui para o tratamento das colocações em um dicionário baseado nos *frames* semânticos propostos pelo linguista Charles J. Fillmore. O Field – Dicionário de Expressões do Futebol já está disponível para consulta, enquanto o recurso para modalidades olímpicas teve início em 2014, com previsão de lançamento para o primeiro semestre de 2016. É pertinente ressaltar, no entanto, que o trabalho com o Field continua através da inclusão de mais informações sobre a linguagem futebolística.

Nesse sentido, podemos elencar como nossos objetivos:

- Estabelecer critérios para o tratamento a ser dispensado às colocações, tendo em vista o contexto em que a pesquisa se insere;
- Promover o confronto entre a Semântica de *Frames* e – principalmente – a FrameNet e as colocações, tendo em vista a fraca abordagem que a contraparte computacional e lexicográfica da teoria de Fillmore dispensa a essa classe das unidades complexas; e
- Comprovar a importância da Semântica de *Frames* para a descrição da língua e, especialmente, no contexto deste trabalho, das colocações.

Assim sendo, iniciamos a pesquisa a partir de uma revisão da literatura acerca das colocações. A partir de um estudo bibliográfico sobre essas estruturas, é possível perceber a existência de uma grande pluralidade teórica, uma vez que

diferentes teóricos se debruçaram sobre o tema a partir de diferentes perspectivas, resultando em estudos que nem sempre abordam o fenômeno da mesma forma. É notável, contudo, o fato de que algumas características prevalecem, possibilitando entender que as colocações dizem respeito a agrupamentos lexicais que ocorrem frequentemente na linguagem. Essa frequência de uso faz com que tais construções passem a ser vistas como uma unidade. No âmbito deste estudo, nos atemos às colocações relacionadas ao contexto do futebol. A partir da necessidade de estudarmos essas combinações, iniciamos o trabalho com uma revisão teórica, levando em consideração diferentes autores que trabalharam as colocações a partir de diferentes pontos de vista. Em nosso estudo, dividimos esses pontos de vista em duas abordagens: a primeira, que chamamos de lexicológica, agrupa os teóricos que demonstraram uma preocupação com a descrição do fenômeno; a segunda abordagem, que constitui um viés lexicográfico, compreende autores que, além disso, se dedicaram também a discutir de que forma a lexicografia deve tratar as colocações em dicionários.

Levando em consideração que tratamos dessas combinações lexicais no contexto de dicionários não tradicionais, damos seguimento ao trabalho em um segundo capítulo teórico, dedicado à teoria da Semântica de *Frames*, à sua interface computacional e lexicográfica, a FrameNet, e à relação com as colocações.

A Semântica de *Frames*, desenvolvida por Charles Fillmore (1982; 1985), se estabelece como um modelo semântico da Linguística Cognitiva, cujo objetivo é estudar os significados dos itens lexicais de uma língua a partir de estruturas chamadas de *frames*. Um *frame* é como um enquadre, uma descrição esquematizada das nossas experiências. Para Fillmore, aquilo que sabemos sobre o sentido de uma palavra está diretamente relacionado com o seu *frame* e com a condição de conhecermos este *frame*, isto é, de termos uma determinada experiência. Por oferecer condições para o estudo do significado, a Semântica de *Frames* adquiriu um certo *status* na área da lexicografia. O principal resultado dessa relação é a FrameNet, que pode ser descrita como a contraparte computacional e lexicográfica da teoria de Fillmore, uma vez que se trata de uma base de dados online cujo objetivo é a descrição do léxico da língua inglesa a partir de evidência empírica e de *frames* semânticos. Importa ressaltar, no entanto, que a contribuição da FrameNet para a Semântica de *Frames* transcende o mero papel de atuar como uma contraparte aplicada. A FrameNet, podemos dizer, deu continuidade à

Semântica de *Frames* também sob o ponto de vista teórico. Em outras palavras, falar da teoria proposta por Fillmore pressupõe falar sobre a FrameNet.

Destacamos a contribuição da FrameNet para o estudo e tratamento de unidades multivocabulares, como os verbos-suporte, por exemplo, e as próprias colocações. A partir de uma perspectiva lexicográfica mais tradicional, que também apresentamos, ainda que de forma breve, as colocações são incluídas em outros verbetes do dicionário. A FrameNet não concorda com tal metodologia e apresenta subsídios para uma abordagem distinta que consideramos inovadora, uma vez que leva em consideração os *frames* evocados por estas estruturas. Em detalhes, tratamos dessas questões na última parte do terceiro capítulo.

Na parte que se segue, apresentamos a nossa metodologia, a fim de descrever os processos e etapas que compõem a realização desta pesquisa. Primeiramente, nos voltamos aos recursos utilizados no estudo, que incluem os projetos, o nosso *corpus* e a ferramenta computacional da qual extraímos nossos dados, o *software* Sketch Engine. O Sketch Engine é um programa que permite a análise e manipulação de *corpora* linguísticos, dando ao pesquisador a oportunidade de trabalhar com evidências empíricas, de uso real da língua pelos falantes. O *corpus* é exclusivo da linguagem do futebol e conta com cerca de 1 milhão de palavras. Na segunda parte do capítulo metodológico, descrevemos o processo de análise dos dados em duas etapas: na primeira, nos voltamos para os aspectos estruturais e quantitativos das colocações, de forma a estudar como as colocações da linguagem do futebol se organizam e se apresentam nesse contexto; na segunda etapa, focamos na relação dessas combinações com a Semântica de Frames e com a FrameNet, estabelecendo de que modo serão tratadas no âmbito de dicionários baseados em *frames*.

Ao fim da metodologia, iniciamos o capítulo de análise e discussão dos dados da pesquisa, que é dividido em duas partes:

- Uma dedicada a analisar as colocações a partir de uma perspectiva mais estrutural, buscando compreender o comportamento e a natureza das colocações da linguagem do futebol; e
- Uma cujo objetivo é relacionar as colocações e os *frames* semânticos, de forma a discutir as implicações dessa associação para o contexto lexicográfico da pesquisa.

2 COLOCAÇÕES: REVISÃO DA LITERATURA

You shall know a word by the company it keeps.

– J. R. Firth

Construções como *melhor amigo* e *suar frio*, tão comuns no nosso dia a dia, fazem parte de um conjunto de expressões conhecidas como *colocações*. Neste capítulo, nosso principal objetivo é a exploração desse fenômeno. Para tanto, nas páginas que seguem, nos valem de uma revisão da literatura, partindo de diferentes autores – com diferentes pontos de vista – que se dedicaram a estudar a classe das colocações.

A primeira questão que devemos levantar é: afinal, o que é uma colocação? Diferentemente do que se possa pensar, responder essa questão não é uma tarefa simples. Como mencionamos anteriormente, existe uma reconhecida pluralidade teórica acerca do fenômeno aqui explorado, o que acaba por resultar em diferentes pontos de vista no que se refere à descrição de expressões dessa categoria. Começamos pela visão de quem introduziu o termo nos estudos linguísticos.

O termo *colocação* é creditado ao linguista britânico J. R. Firth, na obra *Papers in Linguistics*, de 1957. Nas palavras do autor (1957, p. 181, tradução nossa)¹, “colocações de uma dada palavra são afirmações dos lugares habituais ou costumeiros dessa palavra”. As colocações, dessa forma, atestam o comportamento das palavras da língua, apontando para os relacionamentos que mantêm umas com as outras em um nível sintagmático. Não é à toa que, de acordo com Tagnin (2013, p. 63), Firth cunhou o termo no intuito de “designar casos de coocorrência léxico-sintática, ou seja, palavras que usualmente ‘andam juntas’”. Cruse (1986, p. 40) também contribui para o estudo das colocações com sua visão acerca do fenômeno, que conceitua como “sequências de itens lexicais que habitualmente coocorrem”.

Podemos citar, também, a definição de McEnery e Hardie, que dizem que

Em poucas palavras, o termo *colocação* denota a ideia de que importantes aspectos do significado de uma palavra (ou outra unidade linguística) não estão contidos na palavra em si, considerada em isolamento, mas sim subsiste nas associações características em que a palavra participa, ao lado de outras palavras ou estruturas com as quais frequentemente co-ocorre, no

¹ “Collocations of a given word are statements of the habitual or customary places of that word”.

que Firth (1968, p. 196) chama de “uma abstração do nível sintagmático. (McENERY e HARDIE, 2012, 122-123, tradução nossa)²

Dessa forma, para os autores, fica claro que a formação de colocações está relacionada aos sentidos de uma palavra, sentidos esses que existem não como conteúdo exclusivo *da* palavra, mas que são firmados nos relacionamentos que as palavras mantêm entre si. Relacionado a isso, a definição de Hunston (2002, p. 68) também nos interessa quando diz que “Colocação é a disposição das palavras de serem tendenciosas na forma como coocorrem”. O que isso quer dizer? A autora aponta, nesse caso, para o fato de que as palavras não se juntam a outras sem motivo, mas, ao invés disso, respeitando determinados critérios, como o contexto. Para tanto, a autora (2002) traz o exemplo da palavra “brinquedos”, que co-ocorre mais com “criança” em uma estrutura do que com “homem” ou “mulher”, tendo em vista que brinquedos fazem parte do universo infantil, não adulto (nesse caso, não se está dando um exemplo de *colocação* em si, mas de como algumas palavras naturalmente aparecem perto de determinadas outras). No entanto, é importante notar que, na contramão disso, a autora também informa que algumas colocações, como “chá forte”, parecem acontecer de forma *não-motivada* (uma afirmação problemática, uma vez que a colocação é motivada por uma característica física do chá), de acordo com Halliday (1976)³.

É pertinente notarmos que as definições parecem sempre abordar o fato de que as colocações representam estruturas recorrentes, habituais. Nesse sentido, podemos dizer que uma colocação é uma *combinação lexical* que ocorre *frequentemente*.

Tendo definido, ainda que de forma preliminar, o que é uma colocação, nos voltamos agora ao estudo de autores que abordaram as colocações a partir de uma abordagem que podemos chamar de lexicológica, isto é, trataram das colocações a partir de um viés puramente dedicado à exploração e à descrição das colocações no nível lexical.⁴ No âmbito dessa abordagem, figuram estudos que transitam,

² “In short, the term *collocation* denotes the idea that important aspects of the meaning of a word (or another linguistic unit) are not contained within the word itself, considered in isolation, but rather subsist in the characteristic associations that the word participates in, alongside other words or structures with which it frequently co-occurs, in what Firth (1968: 196) calls ‘an abstraction of the syntagmatic level’”.

³ O trabalho de Halliday com as colocações será abordado mais profundamente no decorrer do capítulo.

⁴ A distinção entre os vieses lexicológico e lexicográfico feita aqui se baseia na de Beneduzi (2008).

principalmente, na Lexicologia e na Linguística de Corpus. A parte seguinte do capítulo tratará das colocações a partir de um viés lexicográfico, esclarecendo, de antemão, que não trazemos a discussão lexicográfica, em si, para o trabalho. A divisão entre duas seções tem o objetivo de proporcionar uma melhor organização ao leitor.

2.1 O VIÉS LEXICOLÓGICO

Embora Firth tenha sido pioneiro ao tratar o fenômeno através do termo *colocação*, pouco pode ser dito sobre seu trabalho com o tema. Segundo Alonso Ramos (1994-1995, p.12), “Pensamos que o mérito de Firth está em haver despertado o interesse pela co-ocorrência lexical, mas o nível colocacional que propôs baseado na frequência de aparição resulta de um interesse limitado”. Desde que Firth (1957) cunhou o termo e o incorporou à linguística, uma série de autores se debruçaram sobre o tema, gerando a mencionada pluralidade teórica que se encontra quando o assunto é o fenômeno linguístico das colocações. No entanto, autores como o linguista francês Charles Bally já haviam tratado da questão, embora não a partir da denominação de Firth.

Bally foi um dos discípulos de Saussure e, ao trabalhar com o tema, apontou para o fato de que a língua se dá através de associações e agrupamentos entre as palavras (BALLY, 1909), os situando no eixo sintagmático de Saussure (1916). Contudo, Bally aponta também para o fato de tais associações e agrupamentos não ocorrerem de maneira uniforme, constatando nuances que permitem diferenciar e classificar essas combinações. Bally, primeiramente, cria uma distinção entre *agrupamentos momentâneos* e *unidades indecomponíveis*.⁵ Os agrupamentos momentâneos são autônomos e se desligam com facilidade, dando origem a outras combinações, enquanto as unidades indecomponíveis possuem natureza contrária: os elementos associados não são autônomos, dado que seu sentido está condicionado à forma combinada. O fato de o sentido ser limitado pela combinação demonstra falta de transparência dos elementos. Em semântica, as noções de *transparência* e *opacidade* estão estreitamente relacionadas com as expressões idiomáticas, que dividem com as colocações e com outras formas, como os verbos-

⁵ *Groupements passagers* e *unités indécomposables*.

suporte, a categoria de unidades complexas do léxico. Ser transparente denota que o significado de cada elemento de uma determinada construção pode ser identificado separadamente, como em “roupa suja”. O contrário disso é expresso pela noção de opacidade, característica principal das expressões idiomáticas, como “lavar a roupa suja”, visto que o sentido enunciado pela expressão reside não só na soma das partes (isto é, na soma de cada significado), mas em um sentido complementar. Consoante Biderman (2005, p.751), “[...] colocações são sequências semanticamente transparentes, formadas de itens lexicais que co-ocorrem.” Percebemos, neste caso, uma divergência entre a compreensão de Bally e a de Biderman acerca das colocações (lembrando que, embora o discípulo de Saussure não utilize o termo, ele foi pioneiro no estudo de agrupamentos lexicais). Não só diferentes épocas separam os dois autores, sendo relevante mencionar a preocupação de Bally com a descrição linguística dessas colocações, evidenciando um ponto de vista lexicológico, enquanto Biderman foi uma reconhecida lexicógrafa e terminóloga brasileira, isto é, especialista na elaboração de dicionários de língua geral e especializada. Trataremos da distinção entre esses dois enfoques mais detalhadamente no decorrer do capítulo.

Bally (1909), dentre as unidades indecomponíveis, identifica também o que classifica como *séries fraseológicas*⁶: combinações cujos elementos são autônomos, ao mesmo tempo em que co-ocorrem guiados por convencionalidade. Inseridas nas séries fraseológicas, estão também as *séries verbais* e as *séries de intensidade*⁷. As séries verbais designam construções que atuam como paráfrases de verbos, como “fazer uma viagem” em relação ao verbo “viajar”⁸, enquanto as séries de intensidade caracterizam agrupamentos em que uma palavra se une a um substantivo, adjetivo ou verbo no intuito de ressaltar seu significado, como “brisa agradável”, “inversamente proporcional” e “chorar copiosamente”.

A classificação de Bally (1909) nos é relevante, uma vez que, embora tenha surgido em caráter pioneiro, já traz aspectos importantes do fenômeno das colocações: noções como convencionalidade e frequência, por exemplo, sendo esta última de extrema importância para estudiosos inseridos no contexto da Linguística

⁶ *Séries phraséologiques.*

⁷ *Séries verbales e séries d'intensité.*

⁸ É valioso destacar que a paráfrase constitui uma construção com verbo-suporte, isto é, um verbo cujo conteúdo semântico é bastante esvaziado e que se une a um sintagma nominal, formando um sentido global.

de Corpus, como Halliday, linguista “neo-firthiano” (ALONSO RAMOS, 1994-1995, p.12). Trataremos de Halliday e de outros autores que seguem a mesma linha mais adiante.

Assim como Bally, os autores apresentados a seguir seguem uma visão teórica de cunho semântico, que visa à identificação das relações de sentido que existem entre as partes das colocações.

Outro teórico que tratou do problema das combinações lexicais antes de Firth estabelecer o termo colocações para o fenômeno foi Porzig (1950), que defendia a disposição das palavras de uma língua em *campos verbais* e *campos significativos*. Lembrando Firth (1957), que procurava mostrar através de seus estudos que o significado de uma palavra dependia de sua relação com outras do mesmo contexto, Porzig acreditava que cada palavra era definida pela sua relação com as demais do mesmo campo. É importante notar neste ponto que tanto Bally, quanto Porzig e Firth posicionam as combinações no eixo sintagmático, isto é, na relação que cada elemento possui com os restantes. Mais adiante, veremos que os estudos de autores como Halliday propõem um enfoque mais voltado para o léxico como entidade independente do que para a gramática. Porzig (1950) também tocou na questão referente ao limite de combinações que as palavras podem gerar, o que significa que as colocações possuem natureza restrita.⁹ Segundo o autor, “o círculo dessas possibilidades pode ser menor ou maior”¹⁰ (1950, p.133), o que indica variabilidade, ou seja, tal restrição depende das palavras que compõem as combinações. Porzig (1950) ainda afirma que, a partir do momento em que uma palavra é selecionada, então o falante já não mais tem total liberdade de escolha na formação do discurso. Isso porque, por as combinações estarem atadas ao eixo sintagmático, elas devem respeitar critérios de gramaticalidade.

Mel’čuk (1995)¹¹ também toca na questão da restrição combinatória. Segundo o autor, que desenvolve seu trabalho a partir de um viés estruturalista, as combinações de unidades lexicais podem ser classificadas em duas categorias: a primeira, combinações que ocorrem livremente; e a segunda, chamada de *frasemas*

⁹ Como veremos através de Hausmann (1989).

¹⁰ Essas possibilidades se referem às possíveis formas de combinação.

¹¹ Igor Mel’čuk é professor aposentado do Departamento de Linguística e Tradução da Universidade de Montréal. É, também, um dos desenvolvedores da *Meaning-text theory* (Teoria do texto-significado, em tradução livre).

semânticos,¹² que designam um sintagma de dois elementos lexicais, em que o significante não é composto de forma livre e o significado não é composto de forma regular” (ALONSO RAMOS, 2001), o que confirma que a formação de combinações léxicas não ocorre de forma arbitrária, portanto tornando o número de combinações possíveis menor.

Os frasemas semânticos são internamente classificados em três tipos distintos, que variam de acordo com o grau de opacidade semântica. São eles:

- (a) Frasema completo ou expressão idiomática:¹³ o sentido da construção é determinado pela estrutura e não pela soma dos sentidos dos elementos.

Exemplo: *Um dos lances mais bonitos de Neymar na partida foi um **drible da vaca**.*

- (b) Semi-frasema ou colocação:¹⁴ há a presença de dois elementos, A e B. O significado da colocação é formado pelo de A e pelo de C (um novo significado), que é expresso por B quando em contato com A.

Exemplos: *A **área de meta** delimita o local da cobrança do tiro de meta.*

*O técnico Roger já estava na **área técnica** comandando a equipe.*

*A **área penal** é a região onde o goleiro pode segurar a bola com a mão.*

Nesses exemplos, *área* é o elemento A, enquanto os complementos encontrados em cada exemplo designam os elementos B. A ação dos complementos sobre *área* delimita esse elemento A, criando um novo sentido, o elemento C.

- (c) Quase-frasema:¹⁵ o significado é construído a partir dos sentidos das unidades lexicais envolvidas acrescido de um ajuste semântico que transcende a soma dos significados das palavras da expressão. No exemplo abaixo, por exemplo, tal ajuste se expressa no fato de que o placar não está

¹² *Semantic phrasemes.*

¹³ *Full phraseme e idioms.*

¹⁴ *Semi-phraseme e collocation.*

¹⁵ *Quasi-phraseme.*

sendo, de fato, aberto, mas usado pela primeira vez em uma partida para marcar um ponto.

Exemplo: *Não demorou muito e o Timão **abriu o placar** no Pacaembu.*

Outro aspecto interessante do trabalho de Mel'čuk é que, além da restrição combinatória, o autor também demonstra preocupar-se com a questão da frequência em que as combinações ocorrem (MEL'ČUK, 2003). Corpas Pastor (1996) também toca neste ponto.

A proposta de Corpas Pastor (1996) para o estudo das combinações lexicais é de que seu tratamento aconteça dentro das pesquisas sobre fraseologia. A autora especifica seis características-chave das combinações lexicais, que podem ser aplicadas diretamente às colocações (CORPAS PASTOR, 1996): (a) frequência, (b) institucionalização, (c) estabilidade, (d) idiomaticidade, (e) variação e (f) gradação.

- a) Refere-se à frequência de coaparição dos elementos e de uso da construção;
- b) A partir da frequência de uso, a construção torna-se reconhecida pelos falantes;
- c) Refere-se ao número restrito de possíveis combinações e à especialização semântica;
- d) O sentido não é a pura soma dos sentidos dos elementos, mas sim um sentido complementar;
- e) As variações de uso a partir de diferentes registros;
- f) Denota os graus em que as combinações apresentam tais características.

Como mencionamos antes, Corpas Pastor (1996) também se preocupa com questões relacionadas à frequência e à restrição das combinações léxicas. A autora é, no entanto, criticada pela escolha da frequência como uma das características que descrevem as combinações lexicais e, mais especificamente, as colocações. Uma observação importante sobre o critério de frequência parte de Alonso Ramos (1994-1995, p.15), que defende que uma construção como “ver uma árvore” *pode* apresentar uma alta frequência de aparição e de combinação entre os elementos, sem designar, no entanto, uma colocação. Segundo a autora, a construção apenas obedece às regras da sintaxe, já que o sintagma “uma árvore” pode ser substituída

por qualquer outra que designe algo que possa também ser *visto*. Corroboramos com a visão expressa por Alonso Ramos, uma vez que prova que a identificação de colocações deve levar em consideração mais aspectos do que a frequência.

Assim, tal crítica nos é pertinente pelo fato de que, a partir de então, os autores apresentados fazem parte de uma perspectiva estatística, que leva muito em consideração a frequência como característica determinante para identificar as colocações.

Após Firth (1957) ter cunhado o termo *colocações* e estudado o fenômeno a partir de um viés guiado pela noção de frequência, Halliday (1961) deu continuidade ao seu trabalho com as colocações, se opondo ao tratamento dado por teóricos anteriores, que determinavam o estudo das combinações lexicais no eixo sintagmático. Halliday (1961), conquanto, localiza esse estudo no nível do léxico, não identificando hierarquia entre os elementos das combinações. O trabalho de Halliday com o fenômeno das colocações sofreu críticas com relação à importância dada pelo autor à questão da frequência de aparição das estruturas combinadas, como a de Corpas Pastor (1996), que aponta para o fato de que, aparentemente, a concepção de colocação de Halliday se resume à coaparição frequente.

Dentro do grupo de estudiosos que possuem uma visão estatística, Tagnin (2013) se destaca por situar as colocações (estudadas pela autora dentro do que chama de *expressões convencionais*) no nível sintático. As colocações, de acordo com a autora (2013), são caracterizadas pelo fato de que seus elementos tendem a aparecer juntos devido ao uso que essas combinações têm através dos falantes. A autora também aponta para a questão da restrição combinatória: para ela, algumas construções podem apresentar um grau de restrição maior do que outras. Seria o caso, por exemplo, de *coroca*, em que o adjetivo só consegue se agrupar com o substantivo *velha*, e de *iminente*, que pode se agrupar com um grupo bem maior de possibilidades (*azar, perigo, horror*) (TAGNIN, 2013). Dessa forma, Tagnin propõe uma classificação¹⁶ que divide as colocações em colocações adjetivas, colocações nominais, colocações verbais e colocações adverbiais.

¹⁶ É digno de nota expor que a classificação de Tagnin (2013) é baseada em Hausmann (1989), que veremos mais adiante.

1. Colocações adjetivas

No grupo das colocações adjetivas, de acordo com Tagnin (2013), estão as estruturas do tipo Adj + S, sendo que, no caso dessas colocações, tanto o adjetivo quanto o substantivo podem ser convencionalizados. Para esmiuçar essa questão, a autora recorre a exemplos em inglês e português, a fim de mostrar como diferentes línguas realizam diferentes processos de convencionalização. No caso das colocações em que o adjetivo é convencionalizado, podemos citar o par abaixo:

Feliz Natal! - Merry Christmas!

Em inglês, o adjetivo *feliz* é traduzido literalmente como *happy*, não como *merry*. É notável ainda que a expressão *Feliz Ano-Novo*, em inglês, é *Happy New Year*. Por que, então, *Feliz Natal* não se traduz como *Happy Christmas*? Isso demonstra que as colocações passam por processos culturais de convencionalização de suas partes.

Tagnin (2013) cita que pode ser o substantivo o elemento convencionalizado em uma colocação, como no par que segue:

Interesse comum – Common ground

Interesse, se traduzido literalmente, seria *interest*, não *ground*, que significa chão. Ressaltamos que os exemplos apresentados aqui expressam relevância também para a área da tradução.

2. Colocações nominais

A estrutura das colocações nominais é formada por dois substantivos, sendo que um deles será sempre convencionalizado – e, às vezes, os dois podem ser. Colocações desse tipo, segundo Tagnin (2013) são comuns em linguagens técnicas, designando ferramentas, peças etc. (como *trava do cilindro*, por exemplo). É preciso, no entanto, apontar para o fato de que a representação da estrutura das colocações nominais apresenta variações. Para defender esse ponto, a linguista (2013) mostra que, enquanto, em inglês, existe uma predominância de estruturas S + S, em português ocorre a estrutura S + Prep + S. É o caso de *bay leaf – folha de louro*.

Entretanto, isso não é regra. Em alguns casos, tanto em inglês quanto em português, nos deparamos com a estrutura S + S: *stem cell – célula tronco*. Além disso, é possível notarmos também a presença da estrutura S + Prep + (Art) + S, como em *curvature of the spine – curvatura da espinha*.

3. Colocações verbais

De acordo com Tagnin (2013), existem substantivos que co-ocorrem de forma natural com alguns verbos. No entanto, notemos que tais verbos diferem de uma língua para outra. No inglês, por exemplo, temos *make arrangements* e *make a good impression*. Em português, diferente de como ocorre na língua inglesa, utilizam-se dois verbos distintos para traduzir os dois exemplos dados: *tomar providências* e *causar uma boa impressão*, respectivamente.

As colocações verbais, portanto, seguem na maioria das vezes a estrutura V + S: *abrir processo* e *tomar cuidado* são exemplos disso. Entretanto, podemos observar também a ocorrência de outras estruturas, como V + Prep + S, que resulta em *ficar na linha* ou em *levar em consideração*, por exemplo, e como V + Adj, que resulta em colocações como *cair sentado*, *dar certo* e *errar feio*.

É pertinente chamarmos a atenção, também, para o fato de que as expressões *levar em consideração* e *dar certo* são construções com verbo-suporte, isto é, os verbos utilizados atuam em relação de dependência com o sintagma nominal por terem carga semântica reduzida (NEVES, 2000). Isso comprova que as fronteiras entre os diferentes tipos de expressões multivocabulares podem, muitas vezes, ser difusas, impedindo que se estabeleça que uma determinada construção é colocação enquanto outra caracteriza um caso de verbo-suporte sem que se tenha consciência da necessidade de analisar a estrutura representada. Em casos como os dos exemplos citados, arriscamos dizer, inclusive, que a expressão poderia ser classificada dentro dos dois fenômenos, uma vez que não negamos o enquadre feito por Tagnin (2013) nas colocações.

Por fim, apresentamos as colocações adverbiais.

4. Colocações adverbiais

Segundo Tagnin (2013), as colocações adverbiais dividem-se em dois grupos:

- a) Aquelas em que o advérbio modifica o adjetivo: *brutalmente assassinado, gravemente ferido e profundamente ofendido*;
- b) Aquelas em que o advérbio modifica o verbo: *chorar copiosamente, concordar plenamente e continuar infinitamente*.

Três aspectos interessantes das colocações adverbiais devem ser destacados: o primeiro diz respeito ao fato de que, muitas vezes, um determinado advérbio ocorre com uma série de verbos que pertencem ao mesmo campo semântico. É o caso, por exemplo, do advérbio *cegamente*, que compõe colocações relacionadas com confiança, como *acreditar cegamente, obedecer cegamente e confiar cegamente*. O segundo aspecto está relacionado à posição do advérbio na colocação, que pode vir antes ou depois da outra parte. Por exemplo: *A aprovação do projeto subiu assustadoramente* e *A atriz é popularmente conhecida por seus papéis na televisão*. Por fim, temos também a característica das colocações adverbiais que está relacionada a como algumas dessas colocações ocorrem, preferencialmente, de uma determinada forma, como com *mentir descaradamente*. Dificilmente se encontra essa colocação empregada na primeira pessoa do singular ou do plural, provavelmente porque ninguém gostaria de se acusar assim.

Stella Tagnin faz parte de um grupo de teóricos que defende o estudo das colocações a partir de provas empíricas encontradas no trabalho com *corpora* textuais. A Linguística de Corpus tem proporcionado aos estudos linguísticos a possibilidade de comprovar hipóteses a partir de evidências do uso da língua em textos reais que circulam no nosso dia a dia. De modo que a Linguística de Corpus representa grande importância para o trabalho, nos voltamos brevemente agora para os estudos de Tony McEnery e Andrew Hardie (2012) e de Susan Hunston (2002), autores cujas definições de *colocação* citamos anteriormente, no início do capítulo. As obras desses autores, *Corpus Linguistics: Method, Theory and Practice* e *Corpora in Applied Linguistics*, respectivamente, discutem as colocações a partir do olhar prático da Linguística de Corpus, oferecendo uma análise baseada em métodos empíricos de pesquisa com textos autênticos. É importante ressaltar que esta parte do capítulo é essencial para a descrição dos procedimentos metodológicos que adotamos no trabalho, detalhados no capítulo 5.

Começemos com McEnery e Hardie (2012).

Os autores estudam as colocações a partir daquilo que chamam de Linguística de Corpus Neo-Firthiana, que se desenvolve a partir do trabalho realizado por Firth e se dedica a dois temas principais: as próprias colocações e o discurso. Para os teóricos (2012), a maioria dos teóricos hoje concorda que a única forma de identificar as colocações de uma palavra ou de outro tipo de estrutura de forma confiável é através da Linguística de Corpus, o que acaba, inclusive, por contrastar com o estudo pioneiro de Firth, que, segundo Esser (1999 *apud* McENERY e HARDIE, 2012), se limitou a prover exemplos clichês. Embora ainda existam estudiosos que fazem uso da intuição para delimitar instâncias de colocações, os autores (2012) delimitam seu estudo a colocações em termos de *padrões de co-ocorrência observados em corpus*.

Tendo estabelecido o trabalho com *corpus*, nos vemos diante de uma questão de extrema importância: o que conta e o que não conta como colocação? Diferentes pesquisadores e ferramentas de software utilizam *colocação* como forma de definir uma vasta gama de padrões de co-ocorrência que podem ser extraídos de um *corpus*. A partir disso, é possível traçar dois pontos de vista: o primeiro está relacionado à Linguística Computacional e ao conceito de *n-gramas* (sequências de *n* palavras) e à Linguística de Corpus inglesa. Nesses casos, fala-se de estruturas fixas, sequências de duas ou mais palavras. O segundo ponto, que, segundo McEnery e Hardie (2012), é provavelmente mais comum, vê as colocações em termos de *proximidade*, isto é, descartando o caráter fixo da ordem dos componentes. Com isso, devemos nos perguntar qual a extensão de uma colocação; a partir de que momento uma estrutura passa a ser uma colocação e a partir de quando ela deixa de ser. A partir do momento em que assumimos que uma colocação é uma estrutura fixa, descartamos a possibilidade de construções como *ofendido profundamente*. Concordamos que esta é uma estrutura perfeitamente possível, visto que atende a uma das estruturas mencionadas por Tagnin (2013) para as colocações adverbiais: o advérbio modifica o adjetivo. No entanto, após nos depararmos tantas vezes com a noção de frequência e de sua importância para o estudo das colocações, podemos questionar até que ponto *ofendido profundamente* é uma colocação, se, através de corpus, ficasse comprovado que a estrutura ocorre muito menos dessa forma do que com o advérbio precedendo o adjetivo. Em outros casos, como o das colocações verbais, é mais fácil de defender um caráter fixo para

as colocações, visto que não é possível trocar *tomar providências* por *providências tomar*. Destacamos aqui que o critério da frequência, ainda que algumas vezes criticado, como exposto anteriormente, demonstra relevância nesse caso, permitindo que se considere qual das construções apresenta-se mais regularmente no uso feito pelos falantes.

Dentre os estudiosos Neo-Firthianos, os autores (2012) destacam Sinclair et al. (2004), que define os termos *nó* e *colocado*. O *nó* é descrito como o item cujos padrões de co-ocorrência estão sendo examinados, enquanto o *colocado* corresponde a cada item ao redor do *nó*. Seria o caso, por exemplo, de ter *café* como *nó* e *quente, preto, forte e aguado* como colocados. Assim, não se trata de uma relação hierárquica entre *nó* e *colocado*, mas apenas de uma questão de foco analítico: a partir do momento em que definimos *forte* como *nó*, *café* passa a ser um possível colocado.

A partir de uma busca por colocações em um *corpus*, se defende que o cálculo de significância dos resultados está relacionado à frequência. Os autores (2012) destacam que este é um ponto em que estudiosos divergem de forma veemente. “Significância, nesse contexto, deve claramente ser interpretada como significância *estatística* [...]” (McENERY e HARDIE, 2012, p. 125, tradução nossa), algo que é criticado, por exemplo, por Stubbs (1995), uma vez que procedimentos estatísticos complexos não seriam necessários. Por exemplo, considerando o verbo *causar* como *nó* e tendo *acidente, confusão, danos* e outras palavras como colocados, o linguista não necessita de frequência para comprovar que essas unidades são semanticamente relacionadas.

Hunston (2002), de forma similar, também utiliza os conceitos de *nó* e de *colocado*, quando tratando da forma como softwares tratam as colocações. De acordo com a autora (2002), qualquer ferramenta de *corpus* que calcule colocações centraliza uma palavra *nó*, determinando um *span* (*vão*, em inglês). Por exemplo, em um *corpus* de sentenças com a palavra *café*, podemos definir *café* como a palavra *nó*. O *span* se refere à contagem de palavras à esquerda e à direita da palavra *nó*: quatro palavras para cada lado, por exemplo. Dessa forma, a ferramenta é capaz de calcular a frequência das palavras que aparecem próximas do *nó* pesquisado. O problema de utilizar a frequência como principal critério, nesse sentido, está no fato de que as palavras com maior incidência serão as ditas *palavras gramaticais*, como artigos e pronomes, por exemplo. Assim, podemos apontar para a importância da

quantidade de dados a que se tem disposição, tendo em vista que, quanto maior o *corpus*, maior a possibilidade de encontrarmos, com boa frequência, as palavras que, junto do nó, de fato formam colocações.

Os autores que seguem na sequência do capítulo fazem parte da abordagem que podemos chamar de lexicográfica, o que significa que tais teóricos se preocuparam em descrever as colocações, tendo em mente a aplicação dessas expressões em recursos lexicográficos. Este capítulo do trabalho se dedica a refletir e estudar as colocações a partir de suas estruturas e suas características. Por isso, discussões a respeito de como e por que as colocações aparecem em dicionários (muito importantes para o nosso trabalho) são guardadas para o capítulo 3, inteiramente dedicado à Lexicografia e sua relação com as colocações. Assim, o que se apresenta aqui sobre o trabalho de pesquisadores da área lexicográfica se restringe à forma como eles descreveram o fenômeno.

2.2 O VIÉS LEXICOGRÁFICO

Podemos citar, de início, Ettinger (1982), que define as colocações como uma tendência dos níveis sintático e semântico de itens isolados de uma língua a se combinarem apenas com um número limitado de outros itens, confirmando o que vimos na seção anterior. Ettinger (1982) ainda defende que, no estudo das colocações, se deve levar em consideração também aspectos extralinguísticos. Dessa forma, o autor aponta para como as colocações estão conectadas ao seu uso, o que nos lembra do enfoque estatístico apresentado anteriormente.

Anthony Cowie (1981), lexicógrafo cujo grande interesse é o fenômeno das fraseologias, classificou as combinações lexicais tendo como critério o grau de transparência e sua variação lexical e gramatical, tendo em vista que, segundo ele, cada tipo de combinação apresenta, ao usuário de um dicionário, dificuldades diferentes. Assim, a classificação de Cowie (1981) aponta quatro tipos de combinações: colocações abertas, colocações restritas, expressões idiomáticas figuradas e expressões idiomáticas puras.

As colocações abertas correspondem às combinações livres, isto é, as combinações cujos elementos, sozinhos, são transparentes, resultando em um vasto número de combinações, como *beber chá*, *beber café*, *beber água* etc. Já nas colocações restritas, a transparência reside apenas no significado total da

construção, o que restringe o número de combinações possíveis. A estrutura das colocações restritas é formada por dois elementos: um literal e um figurado. É o caso, por exemplo, de *palavra-chave*. O elemento figurado determina a ocorrência do literal.

As expressões idiomáticas figuradas são uma categoria proposta pelo próprio Cowie (1981), intermediando as colocações e as combinações fixas. Assim como as colocações restritas, as expressões idiomáticas figuradas são formadas por um elemento literal e outro figurado. Entretanto, nessa categoria, a troca de algum elemento da estrutura é rara. Além disso, as expressões idiomáticas puras caracterizam as estruturas em que não há transparência e em que não ocorre substituição de elementos, como *chorar o leite derramado*.

A contribuição de Cowie (1981) para o estudo das combinações lexicais no âmbito da lexicografia se deve principalmente ao fato de o autor ter criado uma distinção entre as expressões idiomáticas figuradas e puras, antes tratadas como uma coisa só.

Já Benson (1986) define as colocações como grupos de palavras recorrentes na língua, cuja estrutura é fixa e identificável, situando-as entre as combinações livres e as expressões idiomáticas. Além disso, para o estudioso (1986), as combinações que formam as colocações acontecem de forma arbitrária, o que dá a elas caráter imprevisível. O autor também contribuiu para os estudos sobre as colocações com uma classificação do fenômeno, dividindo-as em *colocações gramaticais* e *colocações lexicais*. A principal diferença entre esses dois tipos está na presença ou não de uma palavra dominante na estrutura, o que se aplica somente às colocações gramaticais.

As colocações lexicais subdividem-se em:

- a) Verbo + substantivo: nessa subdivisão, o verbo carrega sentido de criação ou ativação, como em *construir uma vida*;
- b) Verbo + substantivo: nessa subdivisão, o verbo carrega sentido de anulação, como em *desfazer o contrato*;
- c) Adjetivo + substantivo: *café preto* (considerando que, em inglês, o adjetivo precede o substantivo: *black coffee*);
- d) Substantivo + verbo: *bombas explodem*;

- e) Substantivo + (prep.)¹⁷ + substantivo: *xícara de chá*;
- f) Advérbio + adjetivo: *imensamente feliz*;
- g) Verbo + advérbio: *amar incondicionalmente*.

Por questões morfossintáticas, o trabalho de Benson sofreu críticas (ALONSO RAMOS, 1994-1995) que apontam para o fato de que muito do que se encontra na classificação do autor serve apenas para explicar as colocações a partir da estrutura gramatical da língua inglesa. Nos estudos em língua espanhola, as colocações foram introduzidas por Seco (1978), inspirada pelo trabalho de Firth (1957), embora a autora siga a linha lexicográfica, que, diferente da de Firth, considera mais aspectos do que a simples questão da co-ocorrência lexical. Além disso, Seco (1978) agrega aos estudos sobre as colocações instruções para o seu posicionamento na estrutura das entradas dos dicionários, colocando-as no início do verbete, junto às informações etimológicas da palavra etc.

Por fim, dentro do enfoque lexicográfico, podemos citar novamente Hausmann. O teórico (1985) estuda as colocações a partir da gramática, justificando através do fato de que essas estruturas não acontecem livremente, visto que sua formação é restrita (como vimos anteriormente, essa restrição é explicada por outros autores como um reflexo da *convencionalidade* que permeia a formação de colocações). Hausmann (1989) identifica duas partes na estrutura de uma colocação: a *base* e o *colocado*. A base é o elemento considerado semanticamente independente, enquanto o colocado constitui o elemento que acrescenta sentido à base, porém sem alterar o sentido inicial dela. Nota-se, portanto, na classificação da estrutura de Hausmann (1989), uma hierarquia entre os elementos. O teórico (1989) oferece também uma classificação de diferentes estruturas de colocações:¹⁸

1. **Substantivo** + adjetivo
Exemplo: *Bola aérea*;
2. **Substantivo** + verbo
Exemplo: *Bola rolando*;
3. Verbo + **substantivo**;

¹⁷ Considerando que a classificação de Benson (1986) se aplica ao inglês, vale a estrutura S + S; no entanto, pensando em como isso se relaciona ao português, é válido introduzir a possibilidade de preposição entre esses elementos.

¹⁸ Em negrito, identificamos a base, e, sublinhado, o colocado.

Exemplo: *Ceder escanteio*;

4. **Verbo** + advérbio;

Exemplo: *Acordar cedo*;

5. **Adjetivo** + advérbio

Exemplo: *Expressamente proibido*

6. **Substantivo** + (preposição) + substantivo

Exemplo: *Jogada de letra*

Com isso, percebemos que o autor atribui à classe dos substantivos o papel de base na maioria das estruturas identificadas por ele.

Tendo passado por esses diferentes autores e seus trabalhos sobre colocações, encerramos o capítulo com a seção abaixo, em que recapitulamos as principais informações contidas nele.

2.3 UM APANHADO DO CAPÍTULO

Neste capítulo, exploramos as colocações a partir de dois diferentes enfoques: a abordagem **lexicológica**, preocupada especialmente com a pura descrição linguística do fenômeno, e a **lexicográfica**, que, além de explorar o fenômeno, dedica-se também a discutir o papel das colocações em dicionários. O capítulo nos mostrou que, embora não possamos afirmar que existe um consenso a respeito das colocações, algumas características se repetem nas falas de vários autores, como a frequência e a cristalização do significado. Nesse sentido, para finalizar, podemos voltar à definição que demos no início do capítulo, em que uma colocação é uma *combinação lexical* que ocorre *frequentemente*. O quanto disso se confirmou? De fato, a frequência tem participação essencial na descrição das colocações. Embora vários autores apontem para a questão de que só a frequência não dá conta de definir se uma construção é ou não é uma colocação (por motivos expostos no capítulo), admitimos que este é um critério a ser levado em consideração, visto que as colocações representam estruturas *consagradas* pelo uso. A convencionalização das colocações, mencionada no capítulo, depende do quanto essas expressões fazem parte da nossa linguagem cotidiana. Devemos citar também o caráter arbitrário das colocações, dado que essas estruturas surgem de uma determinada forma e não de outra. Por que, por exemplo, falamos “prato

principal” e não “prato central”? Por fim, citamos também a característica restrita das colocações, que está relacionada à arbitrariedade.

Dando seguimento ao trabalho, nos voltamos agora ao capítulo dedicado à Semântica de Frames e a relação das colocações com a teoria.

3 A SEMÂNTICA DE *FRAMES*, SUAS APLICAÇÕES LEXICOGRÁFICAS E AS COLOCAÇÕES

The semanticist's job is to tease out the precise nature of the relationship between the world and the category [...]

– Charles J. Fillmore

Neste capítulo, nos voltamos à descrição da teoria da Semântica de *Frames* e sua aplicação lexicográfica, a base de dados da FrameNet. De princípio, na seção 3.1, nos debruçamos sobre as bases teóricas da teoria de Fillmore, buscando também abordagens anteriores à Semântica de *Frames* que trazem conceitos próximos aos apresentados pelo linguista cognitivo, sendo elas provenientes dos trabalhos de Minsky (1974), da computação (mais especificamente da Inteligência Artificial), de Goffman (1975), da Sociologia, e do próprio Fillmore, no contexto da Gramática de Casos (1968). Uma segunda parte da seção, a 3.1.1, possui foco voltado para a FrameNet e a forma como se organiza.

Dando seguimento ao capítulo, na seção 3.2, passamos a explorar a relação existente entre a Semântica de *Frames* e a lexicografia, focando, principalmente, em como a FrameNet representa tal relação, a partir de estudos empreendidos por lexicógrafos, como Sue Atkins. Na seção 3.2.1, que a complementa, focamos em como as colocações se encaixam nesse cenário. Para tanto, primeiramente, tratamos de como essas combinatórias são tratadas em um contexto mais tradicional da lexicografia, para então nos voltarmos a um cenário que inclui a Semântica de *Frames* e a FrameNet.

Dito isso, damos início à seção 3.1, *A Semântica de Frames: bases teóricas*.

3.1 A SEMÂNTICA DE *FRAMES*: BASES TEÓRICAS

A Semântica de *Frames*, teoria proposta pelo linguista americano Charles Fillmore (1982) no contexto da Linguística Cognitiva, pode ser descrita, de acordo com Miriam Petruck (2001, p. 1, tradução nossa)¹⁹, como um “programa de pesquisa em semântica empírica que enfatiza as continuidades entre a língua e a experiência

¹⁹ “[...] research program in empirical semantics which emphasizes the continuities between language and experience [...]”

[...]”. Nesse sentido, a Semântica de Frames procura investigar as relações que se estabelecem entre os sentidos das palavras de uma língua e as experiências vivenciadas pelos falantes. Em outras palavras, a teoria entende que nossa compreensão dos itens da língua está subordinada à forma como percebemos o mundo ao nosso redor. Mais especificamente, no âmbito da teoria desenvolvida por Fillmore, essa compreensão é relativizada por *frames*.

Um *frame*, de acordo com o linguista (1982, p. 111, tradução nossa)²⁰, designa “qualquer sistema de conceitos relacionados de tal forma que, para entender qualquer um desses conceitos, é preciso entender toda a estrutura na qual ele se encaixa”. Quando um conceito é introduzido em um texto ou em uma conversa, todos os relacionados são automaticamente recuperados. Podemos pensar na palavra “garçom”: de acordo com a noção de *frame* de Fillmore, compreender o conceito de “garçom” é compreender todo o enquadre em que a palavra está inserida, o que acarreta na compreensão de itens relacionados, como “cardápio”, “conta”, “cliente” etc. Uma palavra, na Semântica de Frames, representa uma *categoria de experiência*. Um dos objetivos da pesquisa empreendida pela teoria se refere a descobrir as razões pelas quais uma comunidade de fala relaciona uma categoria a uma palavra, de modo a incluir tais razões na descrição do significado dessa palavra (PETRUCK, 2001). A partir disso, podemos perceber que o modelo teórico proposto por Fillmore se afasta de abordagens semânticas mais tradicionais ao considerar o papel de fatores como a experiência e o conhecimento de mundo na construção do significado.

Fillmore (1982, p. 111, tradução nossa, grifo nosso)²¹ propõe o termo *frame* como forma de cobrir um conjunto de conceitos já conhecidos na literatura, como “‘esquema’, ‘script’, ‘cenário’, ‘andaime ideacional’, ‘modelo cognitivo’, ou ‘teoria popular’”. Fica evidente, portanto, que a ideia por trás do conceito de *frame* proposto por Fillmore já figurava na linguística, ainda que sob outras denominações. Importa ressaltar que o próprio termo escolhido pelo linguista também já se fazia presente nos estudos de outros pesquisadores, ainda que em outras áreas de conhecimento, como Minsky (1974) e Goffman (1975).

²⁰ “[...] any system of concepts related in such a way that to understand any one of them you have to understand the whole structure in which it fits [...]”

²¹ “[...] ‘scheme’, ‘script’, ‘scenario’, ‘ideational scaffolding’, ‘cognitive model’, or ‘folk theory’”.

Petruck (2001) cita que o trabalho de Minsky com o conceito de *frame* foi, inclusive, de grande influência para a área da inteligência artificial, buscando compreender as formas pelas quais reagimos diante de situações descritas por Minsky (1974) como estereotipadas. O autor defende que, ao nos depararmos com uma nova situação, acessamos uma estrutura existente em nossa memória; essa estrutura é um *frame*, definido como “uma estrutura de dados para a representação de uma situação estereotipada, como estar em um certo tipo de sala de estar ou ir a uma festa de aniversário infantil” (MINSKY, 1974, p. 212, tradução nossa). Para Minsky (1974), atrelado a cada *frame*, existe um conjunto de informações que especificam as diferentes maneiras através das quais identificamos o cenário retratado. Notamos que a definição apresentada por Minsky (1974) se aproxima consideravelmente da utilizada por Fillmore na Semântica de Frames, de forma que parece seguro assumir uma relação entre os conceitos – especialmente se considerarmos que Fillmore (1982) cita o trabalho do informata em sua área de conhecimento.

Já Erving Goffman, no campo da Sociologia, utiliza o conceito de *frame* para definir estruturas que retratam a forma como indivíduos agem em situações diversas. Nas palavras do sociólogo:

Eu entendo que definições de uma situação são construídas de acordo com princípios de organização que governam eventos – ao menos os sociais – e nosso envolvimento subjetivo neles; *frame* é a palavra que uso para me referir a tais elementos básicos que eu sou capaz de identificar. Esta é a minha definição de *frame*. Minha expressão “análise de *frame*” é um slogan para me referir ao exame, nesses termos, da organização da experiência. (GOFFMAN, 1974, p. 11, tradução nossa, grifo nosso)²²

De acordo com Brooks (2007), Goffman, em *Frame Analysis* (1974), explora as correspondências que existem entre a organização de uma situação e a maneira como tal situação é compreendida pelos indivíduos. Desse modo, é possível perceber que o trabalho de Goffman, assim como o de Minsky, se relaciona com o de Fillmore, uma vez que o seu conceito de *frame* também diz respeito às formas

²² “I assume that definitions of a situation are built up in accordance with principles of organization which govern events – at least social ones – and our subjective involvement in them; frame is the word I use to refer to such of these basic elements as I am able to identify. That is my definition of frame. My phrase “frame analysis” is a slogan to refer to the examination in these terms of the organization of experience”.

como situações se organizam e como são percebidas, dando origem às nossas experiências.

No entanto, podemos dizer que o conceito de *frame* de Fillmore está mais relacionado com o de *case frame*, presente no trabalho do linguista com a Gramática de Casos (PETRUCK, 2001). No artigo *The case for case*, de 1968, Fillmore estabelece a noção de “casos”, que atuam nas sentenças como papéis semânticos, sendo eles: agentivo, instrumental, dativo, factivo, locativo e objetivo. Em uma oração, o verbo seleciona um determinado número de casos, formando um conjunto a que Fillmore se refere como *case frame*. Dessa forma, o termo *frame*, na Gramática de Casos, está relacionado às características dos elementos selecionados pelo verbo em uma sentença. Podemos pensar nisso como um quadro, em que figuram os casos selecionados. Por exemplo, na sentença *Eu dei o cartão para Renato*, o verbo exige um sujeito que realize a ação (agentivo), um objeto a ser dado (instrumental) e um destinatário (dativo). O conjunto formado pelos casos selecionados corresponde ao *case frame* da sentença. Fillmore (1982) aponta que o conceito de *frame* utilizado na Gramática de Casos é oriundo da psicologia cognitiva, lembrando que, embora seu desenvolvimento nessa área tenha acontecido de forma bastante independente em relação à linguística, seu uso na Gramática de Casos se deu de forma contínua com o uso aplicado pelo linguista na Semântica de Frames. Um *frame case*, nas palavras do autor (1982, p. 115, tradução nossa)²³, caracteriza “uma pequena ‘cena’ ou ‘situação’ abstrata, sendo que, para entender a estrutura semântica do verbo, era necessário entender as propriedades de tal cena esquematizada”. Uma afinidade com o conceito utilizado na Semântica de Frames se torna clara quando ligamos isso aos elementos que preenchem as estruturas dos *frames*.

A fim de demonstrar essa relação, nos valem de um exemplo dado por Fillmore (1982): o *frame* *Evento_Comercial*.²⁴ Esse *frame* em específico carrega elementos que incluem um comprador, um vendedor, produtos e dinheiro.²⁵ Verbos

²³ “[...] a small abstract ‘scene’ or ‘situation’, so that to understand the semantic structure of the verb it was necessary to understand the properties of such schematized scenes”.

²⁴ *Commercial_Event*. O uso da fonte Courier New nos nomes de *frames* respeita a metodologia da FrameNet.

²⁵ Petruck (2001) alerta que, no contexto da Semântica de Frames, os elementos são selecionados de acordo com papéis *situacionais*, isto é, pertencentes ao contexto e à forma como a situação ocorre, o que contrasta com os papéis semânticos apresentados na Gramática de Casos.

inseridos no *frame* `Evento_Comercial` – que incluem, por exemplo, *vender*, *comprar*, *pagar*, *cobrar* e *pechinchar* – criam diferentes enfoques dentro da estrutura da cena representada. De acordo com Petruck (2001), o verbo *pagar*, nesse sentido, enfoca o comprador, o vendedor e o dinheiro, enquanto *vender* enfoca o vendedor e o produto, como podemos perceber nas sentenças abaixo:

(1) [Raquel COMPRADOR] *pagou* [500 reais DINHEIRO] a [Jorge VENDEDOR].

(2) [Jorge VENDEDOR] *vendeu* o [celular PRODUTO].

No exemplo (1), o verbo exige que sejam informados um sujeito, uma quantia a ser paga e alguém que receba essa quantia, enquanto que, no exemplo (2), o verbo exige um sujeito e algo a ser vendido. Importa ressaltar que é perfeitamente possível que formemos frases em que todos os elementos aparecem juntos, como em *Raquel pagou 500 reais a Jorge pelo celular*. O que pretendemos apontar com os exemplos é que um verbo, quando utilizado em uma oração, abre espaços que devem, obrigatoriamente, ser preenchidos por elementos determinados por ele. O não preenchimento desses espaços acarreta em problemas no sentido da oração, como em *Jorge vendeu*, que, descontextualizada, não consegue expressar o sentido na sua totalidade. Assim, podemos dizer, ainda, que tanto a Semântica de Frames quanto a Gramática de Casos apresentam notável relevância para o estudo das valências verbais²⁶.

Segundo Petruck (2001), uma completa descrição desses verbos deve incluir informações a respeito das suas propriedades gramaticais e dos diferentes padrões sintáticos em que eles podem se inserir. Isso repercute diretamente na estrutura dos *frames*. Quais elementos ou aspectos de um *frame* são considerados obrigatórios e quais são considerados periféricos ou opcionais? Como vimos no parágrafo acima, é o verbo que define essa questão ao exigir um conjunto determinado de elementos. Assim sendo, a forma como o *frame* é visto está diretamente relacionada com o ponto de vista do falante, o que nos leva a outro importante aspecto da Semântica de Frames: a noção de *perspectiva*. Como vimos anteriormente, diferentes verbos irão focar em diferentes partes do *frame*, classificando os seus elementos como

²⁶ Valência, tema introduzido na linguística por Tesnière (1959), se vale da centralidade do verbo em uma oração, designando os espaços abertos por ele e que devem ser preenchidos pelos demais elementos da estrutura sentencial. Para um estudo mais aprofundado sobre o tema, recomendamos Neves (2000).

obrigatórios e opcionais de formas distintas. Nas palavras da autora (2001, p. 3, tradução nossa)²⁷, “Assim, as diferentes palavras assumem diferentes perspectivas em/ou esquematizações da mesma cena; entender a escolha de palavras para falar sobre a cena requer um apelo à história dos eventos que levaram até a cena”.

Outro aspecto importante da Semântica de Frames a ser destacado diz respeito ao conceito de *frame* enquanto responsável por retratar uma situação convencional, típica (ideia presente não só no trabalho de Fillmore, mas também nos de Minsky e de Goffman, como vimos anteriormente). Tal característica nos leva a tratar da noção de *protótipo*. Fillmore (1982) trata essa questão a partir do exemplo com a palavra *breakfast* [café-da-manhã], mostrando ao leitor que a compreensão de café-da-manhã está indissociavelmente relacionada com a cultura dos povos de terem uma refeição na primeira parte do dia, após um período de sono, cujo conteúdo faz parte de um cardápio consideravelmente único etc. (FILLMORE, 1982). É notável ressaltar, no entanto, que o conhecimento da palavra e do conceito de café-da-manhã dá margem para que falantes os usem em cenários atípicos. Uma pessoa pode dormir até às duas da tarde, comer torradas com suco e chamar isso de café-da-manhã; ao mesmo tempo, uma outra pessoa pode passar a noite em claro e comer a mesma coisa às sete da manhã, se referindo à refeição da mesma forma, nos lembra o linguista (1982).

Agora imaginemos que um turista de um determinado país, enquanto de férias em um outro país, cuja cultura pouco se assemelha com a de sua terra natal, se dirige ao café-da-manhã da pequena pousada em que ele está hospedado. Ao olhar as opções oferecidas pelo lugar, o turista se espanta ao perceber que nada ali faz parte do menu de um café-da-manhã prototípico do qual ele se lembre. Ainda assim, ele reconhece aquilo como um café-da-manhã, uma vez que a situação respeita os traços culturais impostos pela cena. No entanto, digamos que esse turista se depare com uma imagem do menu que está a sua frente, sem qualquer texto ou informação, antes do momento descrito: de que forma ele poderia reconhecer aquilo como a imagem de um café-da-manhã? Com isso, percebemos que o entendimento de conceitos e eventos convencionais (e de seus *frames*, por conseguinte) depende de mais do que a compreensão dos elementos que compõem

²⁷ “Thus, the different words assume different perspectives on or schematizations of the same scene; understanding the choice of words for talking about that scene requires appealing to the history of events leading up to it”.

a cena; depende, acima de tudo, de entender a dinâmica, o sistema que faz a roda girar. É através das relações e das associações entre os elementos de uma situação que somos capazes de compreendê-la – e de compreender o *frame* que a retrata. Questão não menos instigante diz respeito ao alcance dos *frames* enquanto representações de experiência, pois, se retratam situações prototípicas, vale nos perguntar: prototípicas para quem? O estudo do conceito adotado por Fillmore abre espaço para uma discussão a respeito da própria natureza dos *frames*. Seriam esses sistemas de conceitos representações de *experiências* ou de eventos socialmente concebidos? Se considerarmos experiência como algo pessoal de cada indivíduo, algo que, embora aconteça *em* sociedade, se constrói primariamente a partir das sensações e dos processos cognitivos individuais de cada um de nós, este é um questionamento que abre portas para diversas reflexões. Havemos de considerar, entretanto, a própria insistência de Fillmore (1982) em firmar a Semântica de Frames como um programa de pesquisa dedicado a descobrir as propriedades semânticas dos itens da língua e dos conceitos associados a eles (e sintáticas e morfológicas também, levando a consideração que o linguista fala da *formação* de novas palavras e frases).

Nesse caminho, mencionamos ainda que a Semântica de Frames apresenta relevância para estudos nos campos da semântica textual e da semântica gramatical (PETRUCK, 2001). Segundo a autora (2001, p. 3, tradução nossa), “o intérprete de um texto invoca um *frame* quando atribui uma interpretação a um pedaço de texto, colocando seu conteúdo em um padrão conhecido independentemente do texto”. Na sentença “Julia abrirá os presentes após soprar as velinhas e comer um pedaço de bolo”, por exemplo, somos capazes de compreender que Julia em sua festa de aniversário. Embora a palavra “aniversário” não seja mencionada na frase, ações desempenhadas pelo sujeito (abrir presentes e soprar velinhas) nos indicam a cena que está sendo retratada. Vale notarmos que o arranjo gramatical da sentença não só nos indica a situação de aniversário, mas também que Julia é a aniversariante. Quem sopra as velas e abre os presentes, tradicionalmente, é quem está aniversariando.

Destacamos, por fim, que pesquisas em diversas áreas foram empreendidas através dos postulados teóricos da Semântica de Frames, como a lexicologia, a lexicografia e a sintaxe, por exemplo (PETRUCK, 2001). Assim sendo, na seção

seguinte, passamos a tratar da FrameNet, aplicação computacional e lexicográfica da teoria.

3.1.1 A FrameNet

De acordo com Ruppenhofer et al. (2010, p. 5, tradução nossa, grifo nosso)²⁸, “O projeto da FrameNet de Berkeley está desenvolvendo uma ferramenta lexical online para o inglês, baseada na Semântica de Frames e apoiada por evidências de *corpus*”. A FrameNet, nesse sentido, pode ser descrita como uma base lexicográfica da língua inglesa, cujo objetivo é apresentar informação acerca dos significados dos itens lexicais do idioma a partir do enquadre em *frames* semânticos. Destacamos aqui o envolvimento direto de Fillmore na criação e no desenvolvimento da FrameNet, o que torna o projeto o maior exemplo da aplicação da Semântica de Frames na lexicografia. É notável ressaltar, também, que outras FrameNets vem sendo criadas para outros idiomas. A FrameNet trabalha, essencialmente, com a documentação das variadas possibilidades sintáticas e semânticas de combinação entre palavras, levando em consideração todos os sentidos que apresentam (RUPPENHOFER et al., 2010). A base de dados apresenta mais de 10 mil unidades lexicais (doravante ULs). O conceito de UL da FrameNet, que caracteriza “[...] um pareamento de uma palavra com um significado” (RUPPENHOFER et al., 2010, p. 5, tradução nossa)²⁹, repercute de forma a dar as bases para a apresentação de diversos fenômenos na interface do projeto. Considerando que cada pareamento palavra-significado denote uma UL, cada sentido de uma palavra polissêmica, por exemplo, representa uma entrada na lista de verbetes, tendo em vista que cada significado evoca um *frame* distinto. É de extrema importância ressaltar aqui que o conceito de UL postulado pela FrameNet contribui de forma valiosa para o tratamento das colocações. Veremos isso em detalhes na seção 3.2.1, mais adiante.

Consoante Ruppenhofer et al. (2010), a FrameNet se distingue de outras bases lexicais, como a WordNet³⁰. Primeiramente, o recurso baseado em *frames* apresenta, para cada UL, uma definição (escrita por um membro do grupo que desenvolve o recurso ou retirada do *Concise Oxford Dictionary*), além de exemplos

²⁸ “The Berkeley FrameNet project is creating an on-line lexical resource for English, based on frame semantics and supported by corpus evidence”.

²⁹ “[...] a pairing of a word with a meaning”.

³⁰ Disponível através do link: < <http://wordnet.princeton.edu/>>.

semanticamente anotados, o que dá ao usuário a possibilidade de perceber, a partir de textos reais (extraídos de *corpora* naturalísticos), as diferentes realizações sintáticas e semânticas que a FrameNet se dedica a descobrir e registrar. Além disso, diferente de dicionários tradicionais, a FrameNet não realiza análise palavra por palavra em ordem alfabética, mas *frame a frame*, de forma que as ULs aparecem sempre contextualizadas a partir dos usos que os falantes fazem.

O processo de trabalho da FrameNet, segundo Boas (2009) segue, tradicionalmente, os seguintes passos: primeiramente, faz-se as descrições dos *frames*, o que inclui a caracterização da situação ou do objeto que o *frame* está a retratar e a identificação dos elementos que os constituem. É feita também, nesta primeira fase, uma lista que elenca as palavras que podem fazer parte do *frame*. Após isso, são assinaladas frases do *British National Corpus* que se relacionam à situação do *frame*. Tais frases são, então, mecanicamente extraídos do *corpus* e manualmente anotadas, de acordo com a metodologia de anotação da FrameNet. A inserção das sentenças anotadas e dos *frames* é feita automaticamente nos arquivos da base de dados.

Quanto aos elementos de *frame* mencionados no parágrafo acima, a FrameNet identifica quatro tipos diferentes, baseando-se no grau de centralidade que apresentam: os elementos *core* [nucleares], os *peripheral* [periféricos], os *extra-thematic* [extra-temáticos] e os *core-unexpressed* [nucleares inexpressivos] (RUPPENHOFER et al., 2010).

- **Nucleares:** Os elementos de *frame* nucleares dizem respeito aos componentes que, obrigatoriamente, precisam fazer parte do *frame* para que este possa retratar uma determinada cena, tornando-o único quando em comparação com outros *frames*. Por exemplo, em um *frame* que retrata a marcação de um gol em uma partida de futebol, não é possível que não se inclua um elemento que instancie a figura do jogador que chuta a bola em direção à meta e acerta. Sem esse jogador, isto é, sem esse elemento não há *frame*.
- **Periféricos:** Já os elementos periféricos são aqueles que trazem informações adicionais, mas que não são imprescindíveis para a descrição do *frame*. Segundo Ruppenhofer et al. (2010), costumeiramente designam aspectos como tempo, lugar, modo, grau etc.

- **Extra-temáticos:** conceptualmente, não pertencem aos *frames* em que aparecem, visto que o seu entendimento não depende das informações específicas contidas em um *frame*, diferentemente dos elementos core e específicos. Nesse sentido, os extra-temáticos são tidos como elementos de *frame* propriamente ditos de *frames* mais abstratos, como é o caso, por exemplo, da sentença “Lennert, outro docinho em minha vida, **preparou** jantar para [mim], mmm mmm ótimo”, em que “mim” é identificado como elemento de um *frame* de benefício. (RUPPENHOFER et al., 2010, p. 20, tradução nossa).³¹
- **Nucleares não-expressos:** os elementos marcados como nucleares inexpressivos em alguns *frames* agem como nucleares nesses *frames*, porém não aparecem na anotação de *frames* que se desdobram deles.

Diferentemente do que está descrito acima, na interface apresentada aos usuários, a FrameNet opta por separar os elementos apenas em core e non-core, como é possível observar na imagem abaixo (Figura 1), do *frame* de agricultura:

Figura 1 - Elementos de *frame* do *frame* Agriculture

FEs:	
Core:	
Agriculturist [See] Semantic Type: Sentient	The Agriculturist is the agent cultivating Food .
Food [Foo] Semantic Type: Physical_entity	Food is the item cultivated by the Agriculturist . They CULTIVATE fungi .
Ground [Gro]	This is the the background or context for the Food to which the Agriculturist pays attention in his agricultural efforts. Romans were CULTIVATING the land in Northern Gaul.
Non-Core:	
Circumstances [I]	Circumstances describe the state of the world (at a particular time and place) which is specifically independent of the agricultural event itself and any of its participants.
Degree [Deg]	This frame element selects some gradable attribute and modifies the expected value for it. Most of lowland Britain is intensively CULTIVATED .
Duration [Dur]	For what length of time the Agriculturist attempted to obtain Food .
Event_description [Eve]	This FE appraises the event.
Frequency [Fre]	How often the Agriculturist cultivates Food .
Instrument [Ins]	The Instrument is used by the Agriculturist as a tool to help obtain Food . (This tool is not limited to typical artifacts and can be hands, arms, and other

Fonte: FrameNet... [2014?]³²

³¹ “Lennert, another sweetie in my life, **cooked** [me] dinner, mmm mmm good.”

³² FrameNet Data [2014?]. Disponível em: <<https://framenet.icsi.berkeley.edu/fndrupal/index.php?q=frameIndex>> Acesso em: 17 dez. 2014.

Além disso, podemos perceber ainda que a FrameNet cria descrições para cada um dos elementos.

Outro aspecto importante da FrameNet que podemos relacionar a sua contribuição à lexicografia é que são estabelecidas relações entre os *frames* e entre os elementos de *frame*. As relações entre os elementos de *frame* são três, de acordo com Ruppenhofer et al. (2010): *coreness sets* [conjuntos nucleares], *requires* [requer], e *excludes* [exclui]. Os conjuntos nucleares dizem respeito a alguns grupos de elementos nucleares que parem agir em conjunto, de forma que a presença de um membro do grupo é suficiente para “[...] satisfazer a valência semântica do predicado” (RUPPENHOFER et al., 2010, p. 21, tradução nossa).³³ Já o segundo tipo designa os casos os casos em que a presença de um determinado elemento de *frame* em uma sentença requer que um outro elemento ocorra também. Por fim, o último tipo caracteriza as situações em que, quando um elemento de *frame* de um grupo de elementos relacionados ocorre, nenhum outro do conjunto pode ocorrer.

Quanto as relações *frame a frame*, Ruppenhofer et al. (2010) esclarecem que todas as relações são assimétricas e entre dois *frames*, de modo que, em uma relação, o menos dependente e mais abstrato é chamado de *Super_frame*, enquanto o mais dependente e menos abstrato é referido como *Sub_frame*. Além disso, de forma geral, todos os *frames* estão relacionados a algum outro, ainda que com algumas exceções à regra. No quadro a seguir, estão especificados os diferentes tipos de relação *frame a frame*.

³³ “[...] to satisfy a semantic valence of the predicator.”

Quadro 1 - Relações *frame a frame*

Relação	<i>Sub_frame</i>	<i>Super_frame</i>
Herança	Filho	Pai
Perspectiva_em	Perspectivado	Neutro
Subframe	Componente	Complexo
Precedência	Precedente	Antecedente
Incoativo_de	Incoativo	Estado
Causativo_de	Causativo	Incoativo/estado
Uso	Filho	Pai
Veja_também	Entrada referida	Entrada principal

Fonte: RUPPENHOFER et al. (2010, p. 75, tradução nossa)³⁴

Vejam, abaixo, cada uma das relações, de acordo com os autores (2010):

- 1. Herança:** trata-se da relação *frame a frame* mais forte, correspondendo às do tipo *é-um* de muitas ontologias. Nesse sentido, o *sub_frame* Filho é um tipo do *super_frame* Pai. Nessa relação, as características do Pai devem estar presentes no Filho de forma igual ou mais específica.
- 2. Perspectiva_em:** essa relação é um refinamento de uma mais geral, a de Uso (RUPPENHOFER et al., 2010) e indica a presença de, no mínimo, dois pontos de vista que partem do Perspectivado e são empregados no Neutro.
- 3. Subframe:** a complexidade de alguns *frames* resulta em estruturas que comportam sequências de transações e de estados, de forma que cada uma pode ser descrita separadamente como um *frame*. Os *frames* derivados de um maior, o dito complexo, são os componentes. Os elementos de *frame* presentes no complexo se estendem aos componentes, embora não precisem constar em sua totalidade em cada um dos subframes.
- 4. Precedência:** relação que ocorre apenas entre dois *frames* componentes de um complexo (especificados na relação acima). A precedência designa a sequência de eventos dentro do cenário maior. Nesse sentido, o *frame* antecedente está relacionado a um acontecimento anterior ao retratado pelo *frame* precedente.

³⁴ O quadro original, com as denominações em inglês utilizadas pela FrameNet, encontra-se em anexo no fim deste trabalho.

5. **Causativo_de e Incoativo_de:** essas duas relações são apresentadas de forma combinada por Ruppenhofer et al. (2010). Relações de “não-herança”, participam delas os *frames* que atuam como estativos, incoativos e causativos quando em combinação a outros *frames*. De acordo com os autores (2010), relações desses tipos na FrameNet ainda não foram feitas (considerando a data da obra). Estranhamente, os autores (2010) afirmam que *frames* que atuam como causativos herdam do *frame* *Ação_Transitiva*, enquanto os que designam incoativos herdam de *Evento* e os que são de estado, do *frame* *Atributo_Gradual*. Tal afirmação parece causar conflito com a proposição anterior, de que essas relações não são de herança. De qualquer forma, Ruppenhofer et al. (2010) não especificam nada mais quanto a este tópico.
6. **Uso:** quando um *frame* faz referência à estrutura de outro, mais abstrato e esquemático, estabelece-se a relação de uso. Comumente, tal relação está presente quando algum aspecto no cenário retratado por um *frame* filho se refere a um *frame* pai. Além disso, é dito que um *frame* pode usar mais de um simultaneamente. (RUPPENHOFER et al., 2010).
7. **Veja_também:** quando há grupos de *frames* similares, é necessário garantir que sejam tratados de forma cuidadosa, de modo a garantir que as diferenças e contrastes entre eles estejam claros. Dessa forma, cada *frame* do grupo é ligado a um membro representativo através de uma relação de *veja_também*. Na descrição do membro representativo, é preciso que as características que separam um *frame* do outro sejam especificadas. Devemos destacar, inclusive, que esta relação é muito utilizada também em dicionários ditos tradicionais.

Podemos citar dois exemplos de como a FrameNet apresenta as relações *frame* a *frame* na interface disponibilizada aos usuários. A primeira é na forma de listagem na estrutura dos *frames*, como podemos observar abaixo. A lista disponibilizada mostra todas as possibilidades de relação com outros *frames*. Como é possível notar na figura, o *frame* *Agriculture* possui três relações: de herança, de perspectiva_em e de subframe. Consideramos que essa forma de visualização pode ser confusa ao usuário, uma vez que a FrameNet não lista apenas as relações

disponíveis para o *frame* em particular, mas todas as previstas para a base como um todo (Figura 2).

Figura 2 – Relações *frame a frame* de Agriculture

Frame-frame Relations:

Inherits from: [Attempt_obtain_food_scenario](#)

Is Inherited by:

Perspective on:

Is Perspectivized in: [Growing_food](#)

Uses:

Is Used by:

Subframe of:

Has Subframe(s): [Food_gathering](#), [Planting](#)

Precedes:

Is Preceded by:

Is Inchoative of:

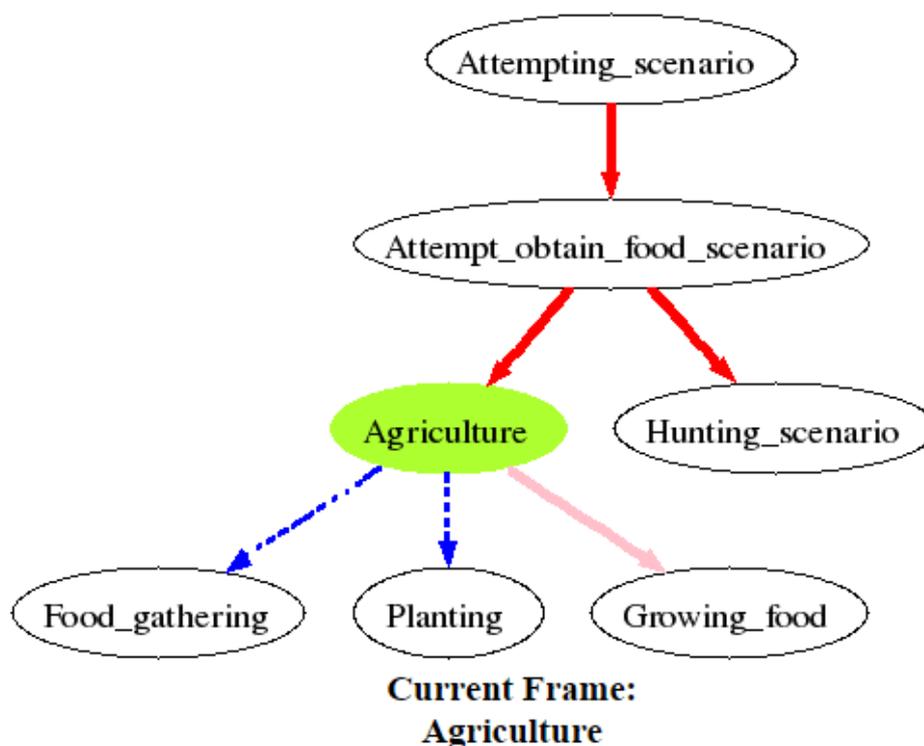
Is Causative of:

See also:

Fonte: FrameNet... [2014?]³⁵

A segunda forma de apresentação das relações entre *frames* é através da ferramenta conhecida como FrameGrapher, que pode ser acessada a partir da página inicial da base, na aba FrameNet Data. O FrameGrapher é descrito como um recurso de visualização dos relacionamentos *frame a frame* e entre os elementos de *frame*. A partir disso, é possível notar que esta forma de apresentação das relações é mais completa do que a primeira, tendo em vista que inclui as conexões estabelecidas entre elementos de *frame*. A imagem (Figura 3) abaixo apresenta o resultado para uma busca a partir do *frame* Agriculture.

³⁵ FrameNet Data [2014?]. Disponível em: <
<https://framenet.icsi.berkeley.edu/fndrupal/index.php?q=frameIndex>>. Acesso em: 17 dez. 2014.

Figura 3 - FrameGrapher para o *frame* Agriculture

Fonte: FrameGrapher [2014?]³⁶

Da mesma forma, pensamos que o tipo de visualização gerada pelo FrameGrapher, ainda que valioso, também pode resultar em uma difícil compreensão pelo usuário. A tela em que o gráfico é disponibilizado não traz indicações quanto ao significado das cores utilizadas ou dos diferentes tipos de flechas que ligam um elemento a outro.

Tendo em vista o que apresentamos nesta seção, podemos dizer seguramente que a FrameNet é mais do que uma aplicação da Semântica de Frames no âmbito computacional e lexicográfico. Não se restringindo a apenas a utilizar conceitos presentes na teoria, a FrameNet realiza uma reflexão que contribui com a Semântica de Frames em um nível que nos possibilita dizer que a base de dados, na verdade, deu continuidade ao modelo semântico proposto por Fillmore entre os anos 1970 e 1980.

Assim sendo, nesta seção 3.1, nos dedicamos a apresentar as bases que fundamental a Semântica de Frames e a FrameNet. Na seção seguinte, 3.2, nos

³⁶

FrameGrapher [2014?]. Disponível em: <<https://framenet.icsi.berkeley.edu/fndrupal/FrameGrapher>>. Acesso em: 17 dez. 2014.

voltamos, mais especificamente, à relação que a teoria de Fillmore (considerando a contribuição da FrameNet) estabelece com a área da lexicografia, explorando também as implicações disso para o estudo das colocações no contexto deste trabalho.

3.2 FRAMES SEMÂNTICOS E LEXICOGRAFIA: IMPLICAÇÕES PARA O TRATAMENTO DE COLOCAÇÕES

Nesta seção, como dito anteriormente, abordamos a relação da Semântica de Frames, a partir da FrameNet, com a lexicografia, apresentando as formas através das quais a teoria de Fillmore contribuiu com os estudos lexicográficos. A segunda parte, seção 3.2.1, foca em como essa relação contribui com o estudo das colocações.

Fillmore e Beryl T. “Sue” Atkins³⁷, em *Toward a Frame-Based Lexicon: The Semantics of RISK and its Neighbors*, de 1992, apresentam os primeiros resultados de uma pesquisa voltada à semântica do lexema inglês *RISK* [risco; arriscar]. De início, os autores argumentam que a descrição que pretendem justificar com o estudo não poderia ser facilmente apresentada e um dicionário tradicional e impresso, de forma que seria um necessário um recurso lexical online baseado em *frames* (FILLMORE; ATKINS, 1992). Em tal recurso, os significados dos itens lexicais da língua, bem como as relações entre palavras polissêmicas ou semanticamente próximas, estariam ligados a *frames*.

As descrições dos *frames*, consoante Fillmore e Atkins (1992), deveriam apresentar formas de acesso a expressões relacionadas ao contexto do item lexical visualizado pelo usuário, o que garantiria ao recurso, também, o aspecto de *enciclopédia*. Mais do que isso, um dicionário baseado em *frames* deveria ser capaz de oferecer ao consulente todo tipo de informação que os falantes têm acerca das palavras da língua. Os autores (1992) esclarecem que isso faz com que tal recurso não seja limitado por aspectos como, por exemplo, o espaço comumente disponível em dicionários tradicionais impressos. Além disso, um dicionário baseado em *frames*, nos moldes do que Fillmore e Atkins (1992) tinham em mente, tampouco poderia ser produzido a partir de informações disponíveis em dicionários comuns,

³⁷ Atkins é uma renomada lexicógrafa. Atualmente, ela atua no projeto FrameNet, em Berkeley, Califórnia, no *American National Corpus* e no *International Journal of Lexicography*.

uma vez que as conexões necessárias para mostrar os relacionamentos entre palavras e *frames* não estão, geralmente, disponíveis nesse tipo de publicação impressa. Outra razão, nas palavras dos autores, aponta que “pesquisa lexical inevitavelmente descobre muito mais informações sobre as palavras do que dicionários comuns tem espaço para acomodar” (FILLMORE; ATKINS, 1992, p. 76, tradução nossa)³⁸. Considerando o status atual da FrameNet, as predições dos autores parecem ter se confirmado, ao menos no que diz respeito ao conjunto de informações disponibilizadas na base de dados, que certamente ultrapassa a quantidade presente em dicionários tradicionais.

Uma questão fortemente defendida pelos autores (1992) no texto diz respeito a como o uso de *frames* semânticos na prática lexicográfica auxilia, por exemplo, o tratamento da polissemia. Através de exemplos reais do uso da linguagem (lembrando que a Semântica de Frames caracteriza um programa de pesquisa *empírico*), é possível enquadrar os diferentes sentidos de um mesmo item lexical em *frames* distintos, o que proporciona um melhor meio de descrever o modo como esses sentidos se relacionam. Conforme apontam Fillmore e Atkins,

A Semântica de Frames torna possível separar a noção dos fundamentos conceptuais de um conceito da maneira precisa através da qual palavras ancoradas neles são usadas. Precisamos de formas para associar uma palavra (ou grupo de palavras, ou grupo de usos de palavras) com *frames* semânticos particulares, então descrevendo as formas variáveis em que os elementos do *frame* conseguem realização sintática. (FILLMORE; ATKINS, 1992, p. 101, tradução nossa, grifo nosso)³⁹

Assim sendo, através da forma como a palavra se relaciona com outras em uma determinada porção de texto, é possível conectá-la ao *frame* ao qual pertence, estabelecendo a ela um significado. Em outra porção de texto, contudo, é possível que a mesma palavra esteja se relacionando com diferentes elementos, tornando possível perceber sentidos adicionais.

³⁸ “[...] research into the lexicon inevitably uncovers much more information about words than standard dictionaries have room for”.

³⁹ Frame Semantics makes it possible to separate the notion of the conceptual underpinnings of a concept from the precise way in which the words anchored in them get used. We need the means of associating a word (or a group of words, or a group of word uses) with particular semantic frames, and then to describe the varying ways in which the elements of the frame are given in the syntactic realization.

Dando continuidade, nos voltamos agora a outro trabalho de Atkins, o texto *The Contribution of FrameNet to Practical Lexicography*, de 2003, em parceria com Michael Rundell e Hiroaki Sato. Como podemos perceber pelo título da obra, o sonho de um recurso baseado em *frames*, visionado por Atkins e por Fillmore, já havia se realizado. Neste texto, os autores discutem de que forma a FrameNet contribui com a área da lexicografia.

Os autores (2003) argumentam que o arcabouço teórico e metodológico da FrameNet tem o potencial de **apoiar**, **acelerar** e **enriquecer** a lexicografia tradicional. A fim de justificar esse ponto de vista, os autores (2003) elencam quatro aspectos em que a base de dados de Berkeley demonstra sua relevância para a prática lexicográfica: (i) diferenciação de sentido, (ii) análise do significado, (iii) descrição valencial e (iv) verbos-suporte. Vejamos cada uma delas.

1. Diferenciação de sentido

Atualmente, com o advento da Linguística de Corpus, lexicógrafos e linguistas têm abordado a polissemia a partir de uma perspectiva mais sofisticada, segundo Atkins, Rundell e Sato (2003). Isso porque um modelo de pesquisa baseado em *corpus* oferece a possibilidade de tratar os significados das palavras como instâncias que carregam características em comum, formando grupos de sentido. Consoante os autores,

Se as fronteiras entre os sentidos são vistas como sendo menos nítidas do que estávamos habituados a acreditar, existe, não obstante, um forte argumento pragmático (bem como um imperativo comercial) para descrever o comportamento de palavras complexas de forma a atender as necessidades da maioria dos usuários de dicionários. (ATKINS; RUNDELL; SATO, 2003, p. 334, tradução nossa)⁴⁰

Nesse sentido, os autores se perguntam: de que modo a FrameNet pode oferecer suporte para o trabalho do lexicógrafo? De acordo com eles (2003), um dos primeiros passos seguidos pelo profissional da lexicografia ao dar início à descrição de uma nova entrada de dicionário é estabelecer quais são os possíveis e plausíveis sentidos que a palavra possui – em outras palavras, estabelecer quais são as ULs.

⁴⁰ If the boundaries between senses are seen as being fuzzier than we used to believe, there is nevertheless a strong pragmatic case (as well as a commercial imperative) for describing the behavior of complex words in ways that meet the needs of the majority of dictionary users.

Para o lexicógrafo no contexto da FrameNet, o trabalho inicia a partir da análise dos dados disponíveis na forma de *palavra chave em contexto*, como podemos ver na imagem abaixo (Figura 4):

Figura 4 - Exemplo de *palavra-chave em contexto* com o verbo *argue*

a of his bringing Jane to Afghanistan,	arguing that she would interfere with his
sposable income. But legal aid experts	argue that low income and capital limits
ca n't be a tourist forever.' They	argued about it for the rest of the after
ed for other types of brokerage and she	argues that it would give her business r
old man thought he was mad. Tried to	argue him out of it, even suggested -
. I forget what we had been talking and	arguing about, but probably it was some
retending. Dot wondered if they 'd been	arguing . She beckoned Gloria into the room
with the following words : 'I would n't	argue that you have to smoke crack to un

Fonte: ATKINS; RUNDELL; SATO, 2003, p. 335

Chamada na Linguística de Corpus de *palavra nó*, a palavra-chave em contexto designa o item lexical a ser analisado e descrito pelo lexicógrafo. À direita e à esquerda da palavra, figuram os *cotextos*, as porções de texto anteriores e posteriores à palavra-chave, da forma como estão presentes no *corpus* utilizado. Atkins e Rundell (2003) afirmam que programas – como o WordSketch, por exemplo – vêm sendo desenvolvidos, possibilitando ao lexicógrafo a visualizações de padrões sintáticos e de frequência no *corpus*. Vale termos em mente que o texto de Atkins e Rundell é de 2003; de lá para cá, as ferramentas tecnológicas provenientes da Linguística de Corpus avançaram consideravelmente.

A partir da anotação semântica baseada em *frames* e da análise das possibilidades sintáticas e agrupamentos da palavra, o lexicógrafo consegue estabelecer quantos significados o item lexical carrega, o que possibilita o enquadre de cada um desses significados – que passam a caracterizar, cada um, uma UL – em um *frame*.

Com isso, podemos dizer que a FrameNet contribui de forma a auxiliar a lexicografia de forma a realizar um estudo mais rico do significado.

2. Análise do significado

De acordo com os autores (2003), o objetivo da FrameNet é registrar todo e qualquer padrão valencial possível, ainda que raro, para cada uma das ULs disponíveis na base. Tendo em vista que isso envolve traços tanto semânticos quanto sintáticos, a análise do significado desempenhada pela FrameNet acaba por ser muito mais “granular” em comparação com o que comumente se apresenta em dicionários tradicionais. A partir disso, Atkins, Rundell e Sato (2003) apresentam três aspectos que demonstram de que formas a FrameNet realiza tais análises, facilitando o trabalho do lexicógrafo e abrindo caminho para a descoberta de “características distintivas significativas que poderiam ter, de outra forma, permanecido enterradas no conjunto de evidências”. (ATKINS; RUNDELL; SATO, 2003, p. 340, tradução nossa).

Os aspectos são: prosódia semântica, a desambiguação de sinônimos próximos e o desemaranhar dos vizinhos semânticos⁴¹ *say* (dizer) e *tell* (contar).

Com relação à prosódia semântica, os autores (2003) especificam que este é um tópico relativamente novo que diz respeito a como nossas escolhas lexicais comumente geram expectativas a respeito de “o que vem a seguir” (ATKINS; RUNDELL; SATO, 2003). Eles citam o caso do verbo *cause*, que geralmente demonstra carregar uma forte carga negativa, aparecendo frequentemente com substantivos como “morte”, “estrago”, “ressentimento” etc. Para o fazer lexicográfico, isso apresenta relevância no sentido de que demonstra haver uma necessidade de se ir além das definições comumente encontradas, como “fazer algo acontecer”, para o verbo em questão. Consoante os autores,

[...] embora contextos positivos ou neutros não são, de forma alguma, ‘impossíveis’, a preferência da palavra por objetos indesejáveis é tão marcada que qualquer consideração dela que falhe em mencionar isso poderia ser considerada como deficiente”. (ATKINS; RUNDELL; SATO, 2003, p. 341, tradução nossa)⁴²

Já referente à desambiguação de sinônimos próximos, uma abordagem baseada na Semântica de Frames para o tratamento dos significados das palavras

⁴¹ Semantic neighbors.

⁴² [...] though neutral or positive contexts are not in any sense “impossible”, the word’s preference for undesirable objects is so marked that any account of it that fails to mention this could be regarded as deficient.

possibilita formas de comparar e de contrastar quase-sinônimos. Através da extração de uma quantidade considerável de dados naturalísticos, a equipe da FrameNet é capaz de realizar análise de palavras sinônimas que pertencem ao mesmo *frame*, investigando e descobrindo os pontos em que essas palavras convergem ou divergem.

Por último, os autores (2003) citam o processo de *desemaranhar* os vizinhos semânticos *say* e *tell*. Segundo Atkins, Rundell e Sato (2003), esta é uma questão muito problemática para aprendizes de inglês, uma vez que, em muitas línguas, esses dois verbos do inglês podem ser traduzidos como um único verbo na língua materna do aprendiz. Fica claro, nesse caso, mais uma vez, a importância dos padrões de valência investigados e disponibilizados pela FrameNet nas entradas de cada UL, visto que é através desses padrões que o usuário tem acesso às diferentes realizações que cada uma das palavras pode gerar na língua.

3. Descrição valencial

No contexto da obra, os autores (2003) definem valência como sendo as construções gramaticais que o lexicógrafo precisa levar em consideração no momento em que analisa os dados e compila entradas. A FrameNet aborda o conjunto de dados a partir do ponto de vista dos *frames*, de modo que a organização das construções gramaticais das valências de cada UL se dá de forma bem mais complexa do que simples subcategorização. Em outras palavras, por estarem subjugadas a *frames*, as valências dos verbetes passam por uma análise que busca descobrir todos as possíveis realizações gramaticais. Dessa forma, isso também representa um auxílio ao trabalho do lexicógrafo, uma vez que contribui para um maior e mais profundo estudo dos significados dos itens lexicais da língua.

4. Verbos-suporte

Por fim, os pesquisadores (2003) se voltam para uma breve descrição do papel dos verbos-suporte na FrameNet. Tal categoria é verbal é descrita, no contexto da FrameNet, como um conjunto de verbos que transformam um “substantivo alvo” (evento ou estado) em uma espécie de predicado. É o caso, por exemplo, da frase “Andreia fez uma viagem importante”, em que a construção com

verbo-suporte *fez uma viagem* deposita no complemento nominal parte da carga predicadora, isto é, o verbo, por si só, não é suficiente para construção do sentido. Atkins, Rundell e Sato (2003) afirmam também que nenhum outro recurso apresenta os verbos-suporte de forma tão completa quanto a FrameNet. Para o trabalho do lexicógrafo, este aspecto da FrameNet importa uma vez que é possível analisar, através de um substantivo, por exemplo, as diferentes combinações que o nome possui com uma série variada de verbos, suporte ou não.

Assim, como vimos até aqui, de acordo com os autores (2003), a FrameNet representa um avanço para a prática da lexicografia tradicional, pois, além de retificar o trabalho com dados naturalísticos extraídos de *corpora*, o arcabouço da Semântica de Frames oferece novas perspectivas para o estudo do significado. Não negando o ponto de vista dos autores, lexicógrafos renomados, gostaríamos de destacar nosso entendimento de que contribuições para a prática lexicográfica não necessariamente significa que a FrameNet é o dicionário baseado em *frames* idealizado por Fillmore e por Atkins (conclusão a que consideramos possível o leitor chegar), considerando as questões que mencionamos anteriormente e que dizem respeito à forma como os dados são recebidos pelo usuário ao acessar a FrameNet. Dando seguimento ao capítulo, a subseção abaixo discute o papel das colocações nas estruturas de dicionários tradicionais e da FrameNet (o que repercute em outros recursos também baseados em *frames*, como veremos mais adiante).

3.2.1 Colocações na lexicografia tradicional e na FrameNet

Como vimos acima, a FrameNet apresenta, em seu arcabouço teórico e metodológico, formas de lidar com fenômenos linguísticos, tais como a polissemia e verbos-suporte, no contexto da prática lexicográfica. Com as colocações não é diferente. Como veremos nesta subseção, o recurso baseado em *frames* propõe, além de uma alternativa para a exposição das colocações na estrutura da base de dados, formas de identificação e de classificação de expressões multivocabulares. Começamos por este segundo aspecto, voltando à obra *FrameNet II: Extended Theory and Practice* (RUPPENHOFER et al., 2010), também conhecido como *The Book*.

De acordo com Ruppenhofer et al. (2010), em um determinado texto, por exemplo, a metodologia da FrameNet (que vimos anteriormente, ainda que de forma

breve) não pode simplesmente ser aplicada às palavras ali presentes sem um reconhecimento prévio, visto que sequências de palavras nos textos devem ser identificadas como sintagmas fixos ou como colocações duras (RUPPENHOFER et al., 2010). O texto dado como exemplo é do domínio jornalístico e traz estruturas como *White House*, *Robert Pickett* e *Secret Service*, que são nomes próprios, e *held without bond*, *assaulting a federal officer with a deadly weapon*, *preliminary hearing*, *firing shots* e *enter a plea*, que são colocações⁴³. Os autores (2010) alertam para o fato de que, embora todas essas construções sejam semanticamente transparentes, elas são “entrincheiradas” (2003). *Held without bond*, por exemplo, segundo eles (2010), é uma das frases padrão usadas para reportar uma decisão em uma audiência. O mesmo ocorre com as outras, que evocam *subframes* de um *frame* relacionado ao processo criminal americano – fiança, julgamento entre inocente ou culpado, sentença etc. Uma unidade lexical (lembrando que uma colocação é uma unidade lexical), quando evoca um *frame* dessa complexidade, evoca simultaneamente o evento menor e o evento maior.

Segundo Ruppenhofer et al (2010), a FrameNet fornece informações a respeito da identificação de colocações e expressões multivocabulares de várias formas. Vejamos abaixo (RUPPENHOFER et al, 2010):

1. **Unidades lexicais multivocabulares:** os autores (2003) esclarecem que alguns lemas (ou verbetes) da FrameNet foram inseridos como unidades multivocabulares desde o início. Tais unidades agregam compostos nome-nome (*wheel chair*, por exemplo), lemas verbo-partícula (*trip up*) e várias formas de expressões idiomáticas;
2. **Colocações envolvendo detalhes de subcategorização:** Este item está relacionado com a extração de sentenças para anotação semântica. Os lemas, na FrameNet, são selecionados em seus contextos de subcategorias (como vimos anteriormente, processo criminal-fiança, por exemplo).

⁴³ *Held without bond*: detido sem fiança;

Assaulting a federal officer with a deadly weapon: agredir um agente federal com uma arma mortal;

Preliminary hearing: audiência preliminar;

Firing shots: lançar tiros;

Enter a plea: entrar com um apelo (tradução nossa).

3. **Compostos nominais com elemento de *frame core* como modificador:** na FrameNet, uma das possibilidades para ampliação dos *frames* relacionados a um substantivo é a modificação dele através de outro substantivo (capitão – capitão da Marinha) ou de um adjetivo relacional (comandante – comandante naval);
4. **Colocações através de substantivos transparentes:** designam, por exemplo, “tipos, agregados, partes, porções, classificadores, unificadores etc.” (RUPPENHOFER et al, 2010, p. 365, tradução nossa).
5. **Colocações com substantivos transparentes:** é o caso de pares lexicais como *bando de gansos*, *enxame de abelhas*, *surto de gripe* e *caso de hepatite*. De acordo com os autores (2003), designar um status especial aos substantivos transparentes contribui de duas formas para a identificação de expressões multivocabulares: primeiro, há a possibilidade de haver pareamentos lexicalmente relevantes entre os dois substantivos em uma construção, e, segundo, a coleta de colocados linguisticamente relevantes pode ser concebida de forma que pode ser demonstrado que o segundo substantivo, não o primeiro, figura na relação colocacional.
6. **Substantivos eventivos com verbos-suporte:** verbos-suporte representam o caso em que o objeto seleciona o verbo, ao invés de o verbo selecionar o objeto (AKIMOTO, 1989; MEL’CUK, 1995, 1996, 1998 *apud* RUPPENHOFER et al, 2010). Para a anotação semântica da FrameNet, isso se reflete na identificação do *frame*, evocado pelo complemento, e não pelo verbo.

Além da classificação acima, a FrameNet também se posiciona a respeito do lugar que as colocações ocupam na estrutura do recurso lexicográfico. Anteriormente, dissemos que o conceito de UL proposto pela FrameNet representava grande importância para este estudo. Vejamos, agora, o motivo disso.

Primeiramente, nos voltamos rapidamente para a lexicografia tradicional, no intuito de descobrir de que forma as colocações são dispostas em dicionários comuns, contrastando com a alternativa encontrada pela FrameNet.

Svensén (2009), por exemplo, quando versando acerca da relação entre as estruturas das colocações e o tratamento dispensado a elas em dicionários, diz que uma consequência das propriedades da base e do colocado é que as colocações são geralmente fáceis de compreender, ao passo em que são difíceis de serem produzidas. Isso representa um fator de importância para a relação entre as colocações e a lexicografia.

Cop (1990) define duas perspectivas para as colocações em dicionários: a semasiológica e a onomasiológica.

Na perspectiva semasiológica, os usuários dos dicionários iniciam buscando por uma determinada expressão (o que significa que ele já fornece a base e o colocado). Nesse caso, o dicionário é utilizado para **recepção**. O significado da base geralmente não oferece problemas para a compreensão do usuário, enquanto o mesmo não acontece com o colocado, que precisa ser especificado. Dessa forma, quando o dicionário é usado para recepção, a informação colocacional deve ser dada na entrada do colocado, com a base sendo usada para descrever propriedades de ordem sintagmática dos colocados (Svensén, 2009), como no exemplo abaixo:

exorbitante [definição]: *preço*

Já na abordagem onomasiológica, o dicionário é usado para **produção**. O usuário inicia a busca não a partir de uma expressão já dada, mas de algo imaginado por ele. Nesse caso, comumente o usuário fornece a base. Dessa forma, na perspectiva onomasiológica, a informação colocacional deve ser apresentada na entrada da base, informando o usuário qual/quais colocado/colocados podem estar atrelados ao elemento da entrada (Svensén, 2009):

suspeitas [definição]: *levantar*

Quando um dicionário é usado para produção, o usuário pode ter o desejo de variar a colocação que ele já conhece. Assim, é importante que todas as informações colocacionais estejam dispostas juntas dentro da entrada da base. Dessa forma, percebemos que, na lexicografia tradicional, as colocações são “absorvidas” pelas entradas que designam ou a base ou o colocado.

Tendo em vista tudo que foi apresentado no capítulo anterior quanto à natureza e às características das colocações, percebemos que tais combinatórias possuem sentido pleno, muitas vezes distante do sentido que encontramos só na palavra que constitui a base ou o colocado. Assim, em um recurso não convencional baseado em *frames* semânticos, como a FrameNet, o tratamento das colocações é diferenciado, partindo do ponto de vista dos *frames*.

A FrameNet, partindo do princípio de que as colocações unidades de sentido que evocam *frames*, as caracteriza como ULs. Dessa forma, as colocações são vistas como um verbete autônomo, uma vez que o seu sentido não é, simplesmente, uma extensão da base ou do colocado, transcendendo a combinatória lexical, como vimos no capítulo 2. Conforme podemos perceber na imagem abaixo (Figura 5), as colocações *April Fool's Day*, *Armistice Day* e *a bit*, por exemplo, fazem parte da lista de ULs da FrameNet, evocando o *frame* *Comemorativo* para as duas primeiras expressões e *Massa_Quantificada* para o terceiro e último⁴⁴.

Figura 5 - Índice de ULs da FrameNet

A

- AIDS.n ([Medical_conditions](#)) **Created Lexical Entry**
- AK-47.n ([Weapon](#)) **Finished Initial Lexical Entry Annotation**
- Alzheimer's.n ([Medical_conditions](#)) **Created Lexical Entry**
- American ((N and S Am)).n ([People_by_origin](#)) **Created Lexical Entry Annotation**
- American.n ([People_by_origin](#)) **Created Lexical Entry Annotation**
- April Fool's Day.n ([Commemorative](#)) **Needs_SCs Lexical Entry Annotation**
- April.n ([Calendric_unit](#)) **Finished Initial Lexical Entry Annotation**
- Armistice Day.n ([Commemorative](#)) **Created Lexical Entry**
- August.n ([Calendric_unit](#)) **BTDT Lexical Entry Annotation**
- a bastard.n ([Difficulty](#)) **Needs_SCs Lexical Entry Annotation**
- a bit.n ([Quantified_mass](#)) **Created Lexical Entry Annotation**
- a bitch.n ([Difficulty](#)) **Insufficient Attestations Lexical Entry**
- a breeze.n ([Difficulty](#)) **Needs_SCs Lexical Entry Annotation**
- a cinch.n ([Difficulty](#)) **Finished Initial Lexical Entry Annotation**
- a doddle.n ([Difficulty](#)) **Created Lexical Entry**
- a few.art ([Quantified_mass](#)) **Created Lexical Entry Annotation**
- a good.a ([Proportional_quantity](#)) **Add Annotation Lexical Entry Annotation**
- a little bit.adv ([Degree](#)) **Created Lexical Entry**
- a little.adv ([Degree](#)) **Created Lexical Entry Annotation**

Fonte: FrameNet... [2014?]⁴⁵

⁴⁴ Commemorative e Quantified_mass.

⁴⁵ FrameNet [2014?]. Disponível em: <<https://framenet.icsi.berkeley.edu/fndrupal/index.php?q=lulIndex>>. Acesso em: 17 dez. 2014.

Assim, percebemos que a FrameNet, ao oferecer novos modelos para a descrição e análise dos significados das palavras e demais itens lexicais da língua, possibilita à lexicografia outros modos não só de conceber fenômenos linguísticos como as colocações, mas também de apresentá-los aos usuários dos dicionários.

3.3 UM APANHADO DO CAPÍTULO

Em linhas gerais, os propósitos do capítulo foram: (i) apresentar a Semântica de Frames como um modelo semântico que leva em consideração o meio, o contexto e as múltiplas vivências dos falantes na descrição do significado, e (ii) demonstrar como a teoria e o seu desdobramento eletrônico contribuem para a área da lexicografia, ainda que com as ressalvas mencionadas ao longo do capítulo e lembradas abaixo.

Na seção 3.1, nos voltamos à exposição dos conceitos e noções que fundamentam a Semântica de Frames, no intuito de abrir caminho para discussões a serem abordadas nas etapas que se seguem neste trabalho. Ainda nesta seção, demos atenção especial à FrameNet, tendo em vista a importância que a base de dados representa para o nosso estudo, como buscamos evidenciar. Na segunda parte do capítulo, seção 3.2, e em seu desdobramento, a 3.2.1, nos voltamos especificamente à relação da Semântica de Frames e da FrameNet com a lexicografia, destacando as consequências disso para o tratamento das colocações em dicionários baseados em *frames*.

Conforme esperamos ter asseverado ao longo do capítulo, a Semântica de Frames estabelece-se como uma alternativa diferenciada para o estudo do significado. Fillmore construiu um modelo teórico capaz de evidenciar o peso das nossas experiências e conhecimentos no modo como nos expressamos através da linguagem. Em nosso entendimento, o conceito de *frame* proposto pelo linguista oferece à semântica a possibilidade de estabelecer novas perspectivas e novos horizontes. É impossível desconsiderar o papel da FrameNet nesta proposição, tendo visto o trabalho empreendido pelo grupo de pesquisadores envolvidos em descrever a léxico da língua inglesa a partir do trabalho de Fillmore com a Semântica de Frames.

É importante, no entanto, termos em mente as ressalvas citadas anteriormente. Designar a FrameNet como um recurso lexicográfico pode levantar

algumas questões, principalmente relacionadas ao aspecto *dicionarístico* propriamente dito. Devemos frisar que não é objetivo do nosso trabalho travar uma discussão a respeito da legitimidade da FrameNet enquanto recurso lexicográfico. No entanto, nada nos impede de perceber que a base apresenta seus dados, muitas vezes, de uma forma que pode ser considerada confusa e pouco amigável a um usuário não-inserido no contexto da Semântica de Frames. Sendo assim, alguém poderia se perguntar: quem é o *usuário* da FrameNet? De que forma uma pessoa não inserida no contexto da teoria de Fillmore pode se beneficiar do trabalho semântico da FrameNet? Considerado a especificidade teórica da base de dados, alguém poderia argumentar que a FrameNet trata da Semântica de *Frames* servindo aos propósitos da Semântica de *Frames*, perdendo de vista o usuário. Responder a essas questões, entretanto, demanda uma reflexão muito maior do que a que propomos no âmbito deste trabalho, bem como um estudo bem mais profundo sobre a FrameNet. Por ora, concordamos com a contribuição do projeto para nossa pesquisa, no sentido de oferecer subsídios para o estudo das colocações.

Tendo em vista a revisão da literatura sobre colocações que compôs o nosso capítulo dois, defendemos a relevância do conceito de UL (e, conseqüentemente, do de *frame* semântico) para nosso estudo. A partir de agora, passamos aos procedimentos metodológicos e à análise de dados do trabalho, na qual intentamos demonstrar que as colocações – no campo da lexicografia –, levando em consideração a semântica presente em suas estruturas, são mais do que meros desdobramentos de outras entradas, como vimos na seção 3.2.1.

No intuito de nos aprofundarmos nesta premissa, partimos para o capítulo 4, *Etapas metodológicas de análise*.

4 ETAPAS METODOLÓGICAS DE ANÁLISE

Neste capítulo, nos voltamos à apresentação dos procedimentos metodológicos que permitem a execução desta pesquisa. Para tanto, primeiramente, apresentamos os projetos Dicionário Modalidades Olímpicas 2016 (CHISHMAN, 2014) e Field – Dicionário de Expressões do Futebol (CHISHMAN, 2012) e o *corpus* utilizado para extração dos dados. A partir disso, abordamos brevemente o *software* Sketch Engine e as funcionalidades usadas no âmbito deste trabalho. Tendo feito isso, passamos à exposição da metodologia propriamente dita, evidenciando os passos necessários para que consigamos desempenhar a análise proposta.

Assim sendo, iniciamos pelos projetos.

4.1 CONTEXTUALIZAÇÃO: OS PROJETOS DE PESQUISA

Para melhor poder descrever os projetos maiores que acolhem esta pesquisa, é importante apresentar o contexto em que surgem, o que diz respeito ao início da relação do grupo de pesquisa SemanTec com a linguagem esportiva.

A primeira incursão do grupo nessa relação se estabeleceu com o início do projeto Kicktictionary_Br, cujo objetivo foi a criação de uma base de dados lexical monolíngue da linguagem do futebol. Tal projeto baseou-se, principalmente, na anotação semântica de *corpus* a partir dos *frames* do The Kicktictionary (SCHMIDT, 2009), projeto pioneiro que deu origem à contraparte em português do Brasil desenvolvida pelo SemanTec. Importa ressaltar que o Kicktictionary_Br não é um dicionário, mas uma base de dados que documenta a semântica presente na linguagem do futebol a partir do arcabouço disponibilizado por Schmidt (2009), o que designa os *frames* e as cenas⁴⁶ presentes no The Kicktictionary.

Com a finalização do processo de anotação semântica do *corpus*, o Kicktictionary_Br ganhou novo fôlego com o desdobramento que deu origem ao projeto Kicktictionary_Br Copa 2014 (CHISHMAN, 2012), do qual surgiu o Field – Dicionário de Expressões do Futebol⁴⁷. Diferentemente do seu antecessor, o Copa 2014 se

⁴⁶ “Cena” é um conceito utilizado por Fillmore nos primeiros estágios de desenvolvimento da Semântica de Frames, designando estruturas *cognitivas*, enquanto os *frames* designavam estruturas *linguísticas*. Mais tarde, Fillmore abandonou tal distinção, optando por *frame* apenas (PETRUCK, 2000).

⁴⁷ Disponível em: <www.dicionariofield.com.br>.

estabeleceu como um projeto lexicográfico, uma vez que o seu produto final seria um dicionário trilingue (português – inglês – espanhol) da linguagem futebolística baseado em *frames*. Uma característica importante do Field – Dicionário de Expressões do Futebol que vale ser mencionada é que ele se destina ao público geral, isto é, não-especializado, o que o distingue de recursos anteriores, como a própria FrameNet, como vimos anteriormente. Nesse sentido, exigiu-se do grupo de pesquisa um posicionamento com relação ao próprio conceito de *frame* presente no dicionário, uma vez que o usuário previsto não estaria familiarizado com as teorias da Linguística Cognitiva. Assim, para fins de apresentação, os *frames* passaram a se chamar “cenários”. A versão final do Field – Dicionário de Expressões do Futebol conta com cerca de 700 verbetes nos três idiomas e 38 cenários ilustrados.

A partir do lançamento do dicionário, o SemanTec deu início ao desenvolvimento de um segundo recurso lexicográfico, seguindo os moldes do Field, porém voltado à linguagem dos esportes olímpicos. Assim, o projeto Dicionário Eletrônico Modalidades Olímpicas (CHISHMAN, 2014) tem por objetivo a criação de um dicionário trilingue (português – inglês – espanhol) baseado em *frames* dos léxicos dos esportes olímpicas.

O dicionário “parte da expertise adquirida pelo grupo neste primeiro trabalho e pretende estender a pesquisa de modo a incluir o vocabulário dos esportes que fazem parte do evento esportivo Olimpíadas” (CHISHMAN, 2014, p. 3). Assim, da mesma forma como ocorreu com o projeto anterior, cujo resultado foi o dicionário Field, o objetivo é disponibilizar um recurso na época dos Jogos Olímpicos 2016.

Embora o projeto Dicionário Eletrônico Modalidades Olímpicas (CHISHMAN, 2014) comprometa-se em seguir a mesma linha metodológica empregada no Field, prevemos a necessidade de adaptações e de novas funcionalidades, levando em consideração que o novo dicionário lida com as linguagens de diversos esportes olímpicos, e não apenas com a do futebol. Para tanto, esperamos as seguintes etapas: (1) estudos teórico-metodológicos, (2) compilação dos *corpora*, (3) trabalho lexicográfico e (4) desenvolvimento da interface.

Assim sendo, podemos dizer que nosso trabalho se insere tanto no Field quanto no dicionário de modalidades olímpicas, uma vez que o tema deste estudo, as colocações, está presente tanto num quando noutro. Embora o Field já esteja disponível para consulta, é importante destacarmos que o trabalho com o dicionário não está finalizado, uma vez que o SemanTec segue estudando a linguagem do

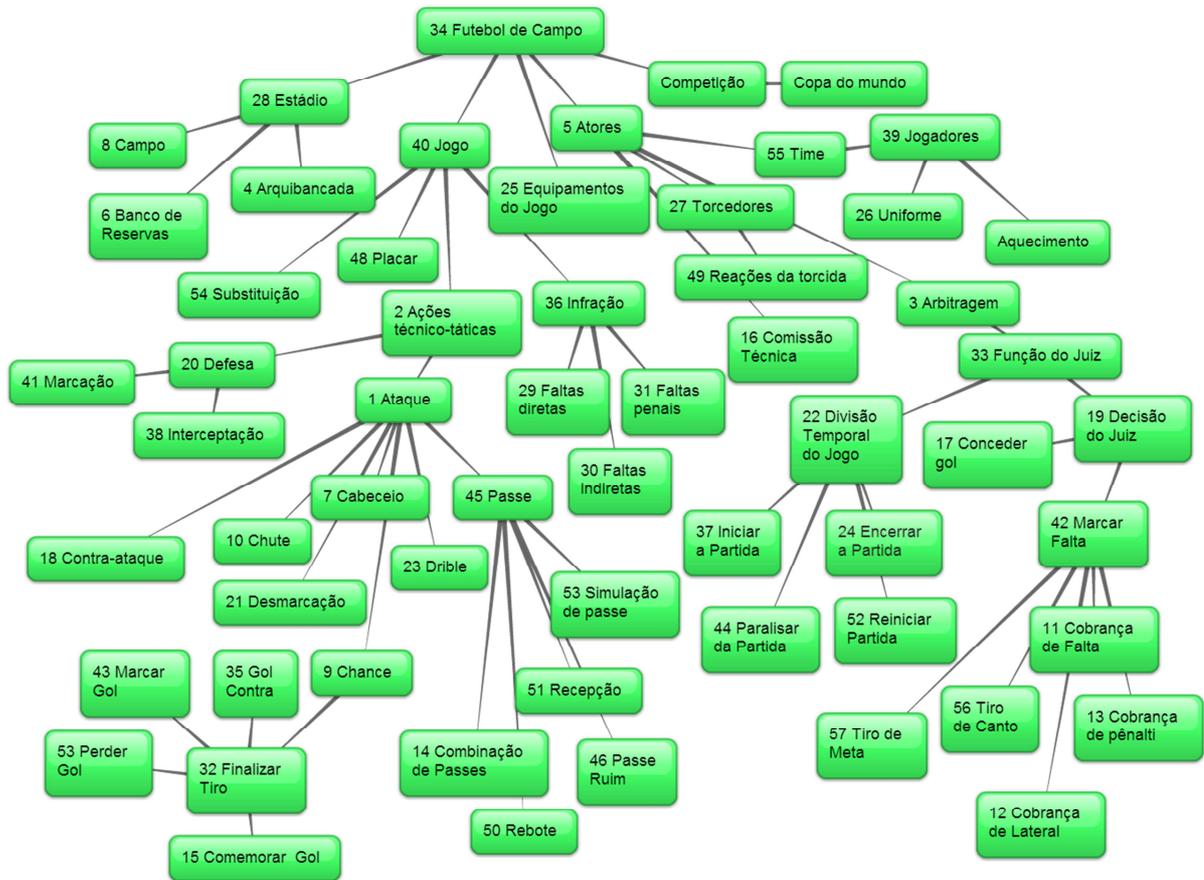
futebol para manutenção e ampliação dos dados e recursos do projeto. Nesse sentido, os resultados desta dissertação podem ainda contribuir com o Field, tanto quanto contribuirá com o Dicionário Eletrônico Modalidades Olímpicas 2016.

A seguir, tratamos brevemente do conjunto de *frames* do dicionário Field.

4.1.1 Os *frames* do Field – Dicionário de Expressões do Futebol

O Field possui, em sua estrutura, 35 *frames* semânticos da linguagem do futebol, chamamos de cenários no contexto do recurso. Uma das principais características dos *frames* do Field diz respeito à presença de dois diferentes tipos de *frames*. Primeiramente, há os *frames* mais típicos, como Comemorar Gol, Passe e Reações da Torcida, que descrevem **ações, eventos**. Além desses, o Field conta também com *frames* de caráter ontológico, como Participantes, Equipamentos e Campo, que contextualizam os **materiais** e **pessoas** envolvidos no contexto futebolístico.

O mapa conceitual abaixo (Figura 6) é uma representação de como os *frames* do Field estão dispostos:

Figura 6 - Mapa conceitual dos *frames* do Field

Fonte: o mapa foi elaborado pelo grupo de pesquisa SemanTec

Vale ressaltar a grande importância do mapa conceitual para a identificação e elaboração dos *frames* semânticos do Field, uma vez que esse tipo de representação permite uma melhor organização desse tipo de dado. Além disso, como é possível perceber pelo número de *frames* dispostos na imagem, o mapa conceitual é uma versão preliminar do conjunto de cenários do dicionário. Ainda assim, é notável a forma como os *frames* ontológicos tendem a aparecer nos primeiros níveis do mapa (como Estádio, Competição e Copa do Mundo, na segunda linha). Tal aspecto se deve ao fato de que esses cenários *alimentam* aqueles de estrutura mais prototípica, uma vez que organizam os elementos que irão, na partida, realizar todas as ações que dizem respeito ao futebol.

A estrutura de cada *frame* do Field conta com: (1) definição do cenário, (2) áudio em português brasileiro, (3) ilustração, (4) palavras que compõem o cenário e (5) cenários relacionados. O Field, buscando proporcionar uma estrutura lexicográfica amigável ao seu público alvo, não apresenta, no *layout* disponibilizado

ao usuário, informações específicas acerca de relações *frame a frame*. A imagem a seguir (Figura 7) é um exemplo da estruturação dos *frames* a partir do cenário Gol Contra.

Figura 7 - Estrutura do *frame* Gol Contra

Fonte: FIELD... (2014)⁴⁸

Ao clicar em uma das palavras que fazem parte do cenário, o usuário é levado à estrutura referente às unidades lexicais, em que ele terá acesso, entre outras informações, a exemplos que ilustram a palavra selecionada.

Na seção abaixo, fazemos uma apresentação do *corpus* do nosso estudo.

4.1.2 O *corpus* deste estudo

O *corpus*, que é em português do Brasil, conta com cerca de 1 milhão de palavras e foi coletado a partir de portais de notícias (IG, Terra, UOL e Zero Hora), de *sites* oficiais de clubes de futebol (Corinthians, Flamengo, Fluminense, Grêmio, Palmeiras e Vasco) e de perfis esportivos no Twitter. Os textos são enquadrados no gênero *match report*, caracterizando resenhas e resumos de partidas de futebol. Tal

⁴⁸ Field – Dicionário de Expressões do Futebol. 2014. Disponível em: <<http://dicionariofield.com.br/scenes/84/gol-contra>>. Acesso em: 30 dez. 2014.

gênero mostra-se relevante para a composição do *corpus* tendo em vista o fato de que os textos evidenciam os eventos que ocorrem no âmbito das partidas de futebol.

Sendo assim, podemos dizer que o *corpus* é *representativo* da linguagem do futebol, uma vez que combina as características *tamanho* e *especificidade* (SARDINHA, 2000): com 1 milhão de palavras, o corpus pode ser considerado grande, além de ser específico da linguagem do esporte – todos os textos coletados dizem respeito ao tema estudado.

Outras características do *corpus* que podemos elencar são (de acordo com Sardinha (2000)):

- Escrito: os textos foram todos coletados de material escrito;
- Contemporâneo: os textos dizem respeito ao período de tempo atual;
- Dinâmico: o *corpus* pode ser expandido com mais conteúdo;
- Especializado: diz respeito a um tema específico;
- Produzido por falantes nativos.

Os textos coletados foram armazenados no programa Microsoft Word e salvos no formato .TXT (UTF-8), possibilitando o upload dos arquivos no *parser* PALAVRAS (BICK, 2000). Feito isso, os arquivos foram processados pelo *software* Maestro, pré-requisito para a entrada do *corpus* no Sketch Engine, uma vez que converte os dados para o formato exigido pelo programa.

Sobre o Sketch Engine falamos brevemente na seção seguinte.

4.1.3 O Sketch Engine

“O Sketch Engine é para qualquer um que queira pesquisar como as palavras se comportam” (SKETCH..., [S.I.])⁴⁹. Assim, trata-se de um sistema de consulta de *corpus* que permite que os usuários tenham acesso a uma série de informações quantitativas, como, por exemplo, a frequência de uma determinada palavra ou construção dentro do *corpus* (KILGARRIFF et al., 2004). É possível ao usuário, também, a consulta às combinações possíveis (no âmbito do *corpus* usado) de uma palavra com outras. A ferramenta, assim, revela, através dos *corpora*, padrões

⁴⁹ Informação disponível no portal do Sketch Engine: <<http://www.sketchengine.co.uk/>>

gramaticais encontrados nas línguas. Tal funcionalidade do Sketch Engine se chama *Word Sketch*. Kilgarriff et. al (2004) exemplificam isso da seguinte forma: se, numa busca pelas *word sketches* do verbo *spread* (*espalhar*), o usuário clicar na palavra *toast* (*torrada*), ele terá acesso a uma lista de combinações, onde ele encontrará, por exemplo, *toast* como objeto de *spread*.

O Sketch Engine é um *software* pago. No entanto, novos usuários podem acessar gratuitamente por um período de 30 dias. De início, o usuário conta com *corpora* previamente disponibilizados, como o *British National Corpus* (não disponível no período gratuito), o ACL Anthology Reference Corpus (ARC), o *British Academic Written English Corpus* (BAWE), o *British Academic Spoken English Corpus* (BASE), o *Brown Corpus* e o *Susanne Corpus* (os quais estão disponíveis para o usuário do período de 30 dias gratuitos) (SKETCH..., [s.i.]). Além disso, como dito anteriormente, o usuário tem possibilidade de utilizar seus próprios *corpora*. No âmbito deste estudo, utilizamos a ferramenta Word List para realizar a extração dos dados.

A análise de tais dados pretende, principalmente, refletir acerca do lugar das colocações em recursos lexicográficos baseados em *frames*. Para isso, dividimos a metodologia em duas fases: na primeira, seguimos um aspecto mais quantitativo decorrente do trabalho com o *corpus*, além de realizar um estudo referente aos aspectos estruturais das colocações. Na segunda fase, nos voltamos à análise baseada na Semântica de Frames e na FrameNet.

4.2 PRIMEIRA FASE: PROCEDIMENTOS PRELIMINARES E ANÁLISE ESTRUTURAL

Nesta primeira fase, iniciamos usando o Sketch Engine para gerar listagens com as colocações presentes no *corpus*. Utilizando o *corpus* “Futebol”, acessamos a ferramenta Word List, no canto esquerdo da página, como podemos ver na imagem abaixo (Figura 8):

Figura 8 - Tela do *Word List* do Sketch Engine

Fonte: SKETCH..., [2014?]⁵⁰

Em seguida, dentre as opções apresentadas na aba *Search attribute*, selecionamos a opção *Collocations*, tendo em vista que pretendemos realizar uma busca por combinações lexicais.

Figura 9 – Seleção da busca por *collocations*

Fonte: SKETCH..., [2014?]

⁵⁰ Sketch Engine, [2014?]. Disponível em: <https://the.sketchengine.co.uk/bonito/run.cgi/wordlist_form?corpname=user/anderson/futebol>. Acesso em: 17 out. 2014.

Como estamos buscando por combinações lexicais, mantemos, em *Frequency figures*, a opção *Hit counts*, para que tenhamos acesso ao número de vezes em que cada caso ocorre no corpus. Em *Output type*, mantemos a opção *Simple*, uma vez que não estamos buscando *keywords*. Queremos que o todo do *corpus* seja analisado.

Figura 10 - Opções de saída da busca

Output options:

Frequency figures: Hit counts Document counts ARF

Output type: Simple Keywords

Reference (sub)corpus: Futebol (whole corpus)

Prefer: rare words common words 1

Fonte: SKETCH..., [2014?]

De posse da listagem pelo Sketch Engine, é necessário que seja feita uma análise preliminar, a fim de identificar quais combinações apresentadas são, de fato, colocações do domínio do futebol. É importante destacar a importância desta fase, uma vez que, a partir dela, é possível refletir acerca de questões apresentadas no capítulo dois do trabalho, que dizem respeito à frequência em que as colocações ocorrem (e o que isso significa para um estudo desse fenômeno) e à linha tênue que separa as colocações de outras categorias de expressões multivocabulares, como construções livres e verbos-suporte.

A partir do momento em que temos definido o grupo de colocações da listagem provida pelo Sketch Engine, realizamos uma análise estrutural de cada uma das construções, tendo como base a classificação de Hausmann (1989). A relevância deste aspecto da análise está diretamente conectada com a questão lexicográfica, tendo em vista que um dos critérios para o tratamento de colocações em dicionários diz respeito à estrutura de tais combinações lexicais.

No entanto, como dito anteriormente, o dicionário em que este estudo se insere é um dicionário não-tradicional, uma vez que se organiza a partir do conceito de *frame* semântico de Fillmore (1982), de forma que partimos para a segunda fase da análise, que diz respeito às implicações da FrameNet e do conceito de *frame* para as colocações em dicionários deste tipo.

4.3 SEGUNDA FASE: ANÁLISE A PARTIR DE *FRAMES*

Nesta segunda fase da análise, nosso objetivo é demonstrar de que forma os conceitos oriundos da Semântica de Frames e da FrameNet contribuem para o estudo e a apresentação das colocações em um dicionário organizado ao redor da noção de *frame* semântico. Segue-se, portanto, os seguintes passos:

- 1) Assim, primeiramente, realizamos o enquadre das colocações selecionadas nos *frames* presentes no Field – Dicionário de Expressões de Futebol;
- 2) Feito isso, refletimos acerca do *status* de UL dessas colocações, a partir do seu enquadre em *frames* semânticos e do sentido que apresentam, buscando fazer um contraponto com a forma como seriam tratadas no âmbito de um dicionário considerado tradicional;
- 3) Por fim, apresentamos uma reflexão final sobre a relação da Semântica de Frames com as colocações no contexto apresentado por este estudo.

5 ANÁLISE E DISCUSSÃO DOS DADOS

Neste capítulo, damos seguimento ao trabalho a partir dos procedimentos metodológicos expostos no capítulo anterior. Respeitando os critérios de análise previamente descritos, dividimos o capítulo em duas seções:

- Seção 5.1, em que exploramos os aspectos quantitativos e estruturais do conjunto de dados da pesquisa; e
- Seção 5.2, em que focamos na relação dos dados com os postulados da Semântica de *Frames* e da FrameNet, discutindo a forma através da qual o modelo teórico iniciado por Fillmore contribui com o tratamento lexicográfico das colocações em dicionários baseados em *frames*.

5.1 QUESTÕES PRELIMINARES E ESTRUTURAIS

Como estabelecido acima, esta seção é dividida em duas partes, cada uma dedicada a desvendar um aspecto da análise. A parte que segue designa uma análise *preliminar* que diz respeito à seleção das colocações em meio aos 500 primeiros resultados apontados pelo Sketch Engine na busca realizada.

5.1.1 Quanto à seleção das colocações

Nesta parte, descrevemos, em linhas gerais, o processo de escolha das colocações a partir dos dados obtidos no Sketch Engine.

Os 500 primeiros resultados da busca (descrita no capítulo anterior, dedicado à metodologia da pesquisa), disponíveis no Apêndice B ao fim do trabalho, constituem um grupo de combinatórias lexicais que não se enquadram, em sua totalidade, no tipo de dados que buscamos. Isso porque não apenas algumas colocações não refletem o domínio do futebol, como também algumas combinações não são, de fato, colocações, de acordo com nossa interpretação.

O primeiro aspecto sobre o qual tratamos é o da frequência. Considerando que se tratam das 500 primeiras estruturas apontadas e que o Sketch Engine organiza esse tipo de busca do resultado mais frequente para o menos frequente, temos em mãos uma lista que inclui as colocações que mais ocorrem no *corpus* utilizado, o que nos permitiu perceber, de início, que o número de combinatórias a

integrar o nosso estudo se torna mais esparso ao passo em que a frequência das expressões diminui. Dentre os primeiros 50 resultados, identificamos 16 colocações, como, por exemplo, *abrir o placar* (a mais recorrente no *corpus*) e *cartão amarelo*. Na revisão teórica do segundo capítulo do trabalho, o aspecto da frequência foi inúmeras vezes citado. Como bem lembramos, muitos autores, inclusive, criticaram abordagens que se valeram demasiadamente desse critério para identificar expressões como colocações. Corroboramos com esses autores no sentido de que a frequência, sozinha, não é suficiente para determinar se uma combinação lexical é ou não uma colocação, o que fica evidente quando olhamos para resultados como *não conseguir*, estrutura apontada pelo Sketch Engine com 698 ocorrências. Trata-se, neste caso, de um mero agrupamento lexical, gramaticalmente correto, que designa o *não*-ato de ser capaz de realizar uma determinada ação. Há sentido e há propósito em seu uso, mas não consideramos uma colocação. Percebemos mais uma vez, por conseguinte, a necessidade de que se olhe para as colocações a partir de uma perspectiva que leve em consideração o significado, a semântica dessas estruturas.

Outro aspecto percebido diz respeito a como o Sketch Engine duplica alguns resultados. Para a colocação *abrir placar*, por exemplo, são apresentados dois itens na busca: *abrir* objeto *placar* e *placar* objeto_de *abrir* (Figura 11). Os números indicando a frequência de cada uma é o mesmo, levando para o mesmo conjunto de sentenças que apresentam a colocação. Assim, estabelecemos que esses casos designam apenas uma colocação na lista de combinações selecionadas para o estudo.

Figura 11 - Resultado duplicado da colocação *abrir placar*

1		
2	Word sketch collocations	Freq
3	<i>abrir</i> <small>objeto</small> <i>placar</i>	784
4	<i>placar</i> <small>objeto_de</small> <i>abrir</i>	784

Fonte: Elaborado pelo autor.

Outro aspecto importante a ser destacado diz respeito à presença de combinações de outras naturezas, como verbos-suporte. É o caso, por exemplo, de *dar assistência*, em que o verbo se apoia no objeto direto para conseguir expressar sentido completo. É valioso mencionar, ainda, a presença de estruturas que são

colocações, mas que não fazem parte do domínio futebolístico, como *surtir efeito* e *sair na frente*. Embora elas tenham sentido dentro da linguagem do esporte, elas se aplicam também a uma série de outras situações desvinculadas desse contexto. Sem estarem empregadas em um contexto, não há como o falante determinar que são colocações do domínio em que nosso trabalho se insere. Por isso, tais expressões não foram consideradas aptas a integrar o grupo de colocações selecionadas para a segunda parte da análise.

Observando esses aspectos, temos uma lista com 74 colocações da linguagem do futebol para análise, que, de acordo com o exposto no início do capítulo, se refere a aspectos de estrutura e de enquadre em *frames* semânticos.

5.1.2 Quanto à estrutura

Como antes explicitado, nesta seção realizamos a classificação das 74 expressões selecionadas a partir dos tipos de colocação expostos por Hausmann (1989). A apresentação dos dados é inicialmente feita na forma de quadro, tomando como base as seguintes informações: (1) a colocação, (2) sua frequência no *corpus*, e (3) sua classificação⁵¹. As informações referentes às bases e aos colocados são dadas através de negrito (base) e sublinhado (colocado) nas colocações. Ao fim do quadro, discutimos a análise.

Quadro 2 - Classificação das colocações

Colocação	Freq.	Classificação
<u>Abrir</u> (o) placar	784	V + S
Cartão <u>amarelo</u>	457	S + Adj.
<u>Marcar</u> gol	441	V + S
<u>Fazer</u> gol	409	V + S
<u>Cobrança</u> de falta	394	S + (prep.) + S
<u>Cobrar</u> falta	390	V + S
Zona de <u>rebaixamento</u>	364	S + (prep.) + S
Entrar <u>em campo</u>	345	V + Adv. (Loc.)

⁵¹ Consideramos, aqui, a estrutura Verbo + locução adverbial, representada por V + Adv. (Loc.), como um desdobramento da classificação V + Adv. proposta por Hausmann (1989).

Linha de <u>fundo</u>	334	S + (prep.) + S
<u>Passar</u> bola	329	V + S
<u>Fazer</u> jogada	326	V + S
<u>Balançar</u> a rede	245	V + S
Time da <u>casa</u>	231	S + (prep.) + S
<u>Cobrança</u> de escanteio	230	S + (prep.) + S
<u>Invadir</u> área	226	V + S
<u>Levar</u> perigo	224	V + S
<u>Receber</u> passe	198	V + S
<u>Criar</u> chance	189	V + S
<u>Bater</u> bola	163	V + S
<u>Cobrar</u> escanteio	152	V + S
Canto <u>direito</u>	149	S + Adj.
Oitavas de <u>final</u>	143	S + (prep.) + S
Lado <u>esquerdo</u>	138	S + Adj.
Ficha <u>técnica</u>	136	S + Adj.
Gol da <u>vitória</u>	136	S + (prep.) + S
Canto <u>esquerdo</u>	133	S + Adj.
Tabela de <u>classificação</u>	124	S + (prep.) + S
Bola na <u>área</u>	121	S + (prep.) + S
Quartas de <u>final</u>	121	S + (prep.) + S
Cartão <u>vermelho</u>	120	S + Adj.
<u>Chutar</u> cruzado	120	V + Adv.
<u>Receber</u> lançamento	120	V + S
<u>Dar</u> passe	118	V + S
<u>Chegar</u> ao <u>gol</u>	117	V + Adv. (Loc.)
<u>Ter</u> trabalho	116	V + S
Saída de <u>bola</u>	114	S + (prep.) + S
<u>Mandar</u> bola	113	V + S
<u>Sofrer</u> falta	110	V + S
<u>Criar</u> oportunidade	109	V + S
<u>Chegar</u> com perigo	108	V + Adv. (Loc.)
Campo de <u>defesa</u>	107	S + (prep.) + S

Receber <u>cartão</u>	105	V + S
<u>Afastar</u> perigo	104	V + S
<u>Entrar</u> <u>na área</u>	102	V + Adv. (Loc.)
<u>Arriscar</u> <u>de fora da área</u>	100	V + Adv. (Loc.)
<u>Campo</u> de <u>ataque</u>	98	S + (prep.) + S
<u>Receber</u> <u>na área</u>	96	V + Adv. (Loc.)
<u>Arriscar</u> <u>de longe</u>	95	V + Adv. (Loc.)
<u>Cruzar</u> <u>na área</u>	92	V + Adv. (Loc.)
<u>Fazer</u> jogo	92	V + S
<u>Mandar</u> <u>para escanteio</u>	91	V + Adv. (Loc.)
<u>Sofrer</u> gol	90	V + S
<u>Troca</u> de <u>passse</u>	90	S + (prep.) + S
<u>Bola</u> <u>parada</u>	88	S + Adj.
<u>Acertar</u> a trave	83	V + S
<u>Livre</u> <u>na área</u>	80	Adj. + Adv. (Loc.)
<u>Ampliar</u> vantagem	78	V + S
<u>Banco</u> de <u>reserva</u>	77	S + Prep. + S
<u>Tocar</u> <u>na saída</u>	77	V + Adv. (Loc.)
<u>Jogo</u> <u>válido</u>	76	S + Adj.
<u>Receber</u> <u>na entrada</u>	73	V + Adv. (Loc.)
<u>Zona</u> de <u>classificação</u>	71	S + (prep.) + S
<u>Equipe</u> da <u>casa</u>	69	S + (prep.) + S
<u>Deixar</u> o campo	68	V + S
<u>Marcar</u> pênalti	68	V + S
<u>Bola</u> na <u>trave</u>	66	S + Prep. + S
<u>Perigo</u> ao <u>gol</u>	66	S + Prep. + S
<u>Gol</u> de <u>empate</u>	65	S + (prep.) + S
<u>Levar</u> <u>na área</u>	63	V + Adv. (Loc.)
<u>Perder</u> gol	61	V + S
<u>Vencer</u> jogo	58	V + S
<u>Garantir</u> vitória	57	V + S
<u>Meio</u> da <u>área</u>	55	S + (prep.) + S
<u>Fazer</u> <u>passse</u>	53	V + S

Fonte: Elaborado pelo autor.

A partir da classificação, podemos iniciar nossas considerações a respeito daquilo que os dados demonstram acerca das colocações do domínio do futebol. Primeiramente, percebemos uma alta incidência de colocações que possuem verbo. Das 74 coletadas, 44 expressões são verbais – 31 ocorrências da estrutura V + S e 13 da estrutura V + Adj. (Loc.) –, o que denota, aproximadamente, 59,4% do grupo selecionado. De antemão, é possível considerar esse aspecto um reflexo do futebol como esporte dinâmico, formado por ações e acontecimentos que movimentam as partidas. Focamos na relação das colocações com os *frames* na segunda parte do capítulo, mas vale adiantarmos que essa característica das colocações do futebol é bastante importante para justificar a opção por uma abordagem baseada em *frames*.

Primeiramente, nos voltamos às colocações de estrutura V + S, que constituem 30 das 43 colocações verbais.

Um traço interessante apresentado na análise dessas colocações diz respeito à presença de seis colocações diferentes para descrever o momento de gol: *marcar gol*, *fazer gol*, *abrir placar*, *balançar a rede*, *ampliar a vantagem* e *sofrer gol*.

Marcar gol e *fazer gol*, por exemplo, são expressões que, de acordo com seus usos no *corpus*, podem servir de substitutas uma da outra. Observemos as sentenças:

- (1) A pressão surtiu efeito e aos 33 minutos da etapa final, Juanfran **marcou gol** contra e deixou tudo igual no estádio El madrigal.
- (2) Logo aos 2 min, o zagueiro colombiano Jeison Murillo **marcou gol** contra e colocou o Getafe em vantagem.
- (3) Aos 28 min, Alderweireld tentou cortar cruzamento e acabou **marcando gol** contra, ao desviar de cabeça e tirar o goleiro Courtois da jogada.
- (4) Embora não tenha **feito gol**, o camisa 22 teve uma atuação excelente, ajudando muito a equipe rubro-negra na partida.
- (5) A gente tem prazer em ver-lo (sic) **fazendo gols**, elogiou o treinador.

(6) Zagueiro volta e **faz gol**.

Como podemos perceber pelos exemplos, em qualquer uma das sentenças a colocação *marcar gol* poderia ser substituída por *fazer gol* e vice-versa. Alguém poderia argumentar que existe, na realidade, uma diferença semântica, ainda que sutil, entre as duas expressões: *fazer gol* denota o puro ato de conseguir fazer com que a bola atravesse a meta do time adversário, enquanto *marcar gol* carrega um traço a mais, que corresponde ao gol ter sido validado e *marcado* pelo árbitro. Entretanto, contrariando essa característica, os exemplos demonstram que, no uso, esse traço se perde: marcar um gol é fazer um gol. Além disso, observamos também que, dentre as duas, a expressão mais utilizada é *marcar gol*, que conta com 441 ocorrências, ao passo em que a segunda possui 401. Ainda que não haja uma discrepância muito grande entre a frequência de uma e de outra, este é mais um aspecto que comprova algo que poderíamos chamar de *apagamento* deste traço relacionado ao verbo *marcar* e às ações do árbitro, tendo em vista que, nos exemplos com a colocação mais recorrente, sequer se cita a presença do juiz.

É interessante notarmos, ainda, a presença da colocação *marcar pênalti* em nosso conjunto de dados. Diferentemente do que acontece em *marcar gol*, em que o verbo diz respeito a uma ação desempenhada pelo jogador, na colocação *marcar pênalti*, o verbo está relacionado ao árbitro, que é quem define que a penalização para uma determinada infração é o pênalti. Podemos perceber isso através das sentenças que seguem:

(7) O árbitro **marcou pênalti** e o veterano Blanco converteu a cobrança com categoria, ampliando o marcador.

(8) O juiz não **marcou pênalti**, o que gerou muitas reclamações do técnico Claudinei Oliveira, expulso do jogo posteriormente.

Em ambos os casos, o sujeito do verbo é o árbitro e não um jogador. Podemos dizer que isso representa um contraste curioso e digno de ser pesquisado mais profundamente, tendo em vista que, de acordo com as regras do esporte, mesmo em *marcar gol* o verbo estaria relacionado a uma ação do árbitro. Ao mesmo tempo em que o *corpus* não evidencia nenhum caso em que *marcar pênalti* esteja

relacionado a algo realizado por um jogador, existe o fato de que, em *marcar gol*, sempre quem marca é um atleta e não o juiz.

A terceira colocação relacionada ao ato de fazer gol é *abrir placar*, com 784 ocorrências. Notamos aqui, entretanto, um traço que restringe o uso dessa colocação, que indica apenas o primeiro gol de uma partida, como em

- (9) Com o jogo bastante concentrado no meio-campo, o Grêmio conseguiu **abrir o placar** aos 31 minutos.

É pertinente ressaltarmos, também, um forte aspecto metafórico no uso do verbo *abrir*. Ainda que se possa pensar que o verbo *marcar* apresenta um aspecto figurado – ainda que bastante sutil – que deriva, principalmente, do fato de que o uso determina uma ação realizada pelo jogador, tal aspecto se apresenta de forma muito mais sólida e evidente no caso da colocação *abrir placar*. O placar é um objeto, um instrumento parte do evento partida de futebol. Ao dizer que o time do Grêmio, usando o exemplo da sentença (9), abriu o placar, muito dificilmente alguém interpretará como se time tivesse, de fato, o aberto fisicamente. Nesse sentido, reforçamos aqui algo bastante presente na nossa revisão teórica: o significado de uma colocação possui conexões com o contexto, com o uso. Além disso, podemos nos lembrar de outras colocações que carregam esse sentido metafórico de *abrir*, como *abrir o jogo*. Na oração *Carina abriu o jogo com Pedro*, entendemos que Carina está falando a verdade sobre uma determinada situação ou assunto (e que tal verdade pode causar algum tipo de atrito). No caso dessa colocação, o traço metafórico de *abrir* está relacionado a um sentido de *exposição*, como quando abrimos um armário e tiramos algo de dentro; o que antes não podia ser visto, agora pode. Já com a colocação futebolística *abrir placar*, notamos um sentido de *inauguração*, como em abrir uma loja, abrir um negócio.

A quarta colocação, *balançar a rede*, que tem frequência de 245 ocorrências, apresenta uma certa mudança no foco, na perspectiva através da qual o momento do gol é percebido.

- (10) Quando Vagner Love **balançou a rede** do Palmeiras, neste domingo (18.11), em Volta Redonda, não foram só os torcedores do Flamengo que vibraram muito.

- (11) Chegaram a **balançar a rede** paulista, mas o gol foi bem anulado.

Nos parece que, enquanto as colocações anteriores possuem um enfoque maior na bola e no jogador, a colocação *balançar a rede* foca no resultado, em uma consequência física do chute do jogador em direção ao gol. Ressaltamos mais uma vez como essas colocações sugerem uma necessidade de se olhar para o contexto e para a situação em que a expressão se insere e o que ela retrata. Em uma partida de futebol, o que mais poderia balançar a rede senão a bola? Talvez alguma outra coisa, um objeto ou mesmo um jogador a se machucar, conseguisse tal feito. No entanto, ainda assim, a colocação *balançar a rede* não seria utilizada, pois ela designa um dos maiores acontecimentos de uma partida de futebol, que é a marcação de um gol.

A quinta colocação é *ampliar a vantagem*, com 78 ocorrências. Assim como acontece em *abrir o placar*, essa colocação apresenta um caráter figurado no uso do verbo. Percebemos também uma relação com placar, tendo em vista que a vantagem diz respeito à diferença no número de gols do time que está na frente em relação ao que está para atrás. Olhemos para o seguinte excerto:

- (12) O **primeiro gol veio** com Ortega, aos 31 min. Sete minutos mais tarde, Antunes **ampliou a vantagem** para os donos da casa e, no início da segunda etapa, Juanmi Jiménez **balançou as redes** pela terceira vez.

No texto, o narrador, por três vezes, faz menção à marcação de gols. Na primeira, ao optar pela forma *o primeiro gol veio*, ele não usou uma colocação. No entanto, nas duas menções seguintes, ele utiliza as combinações *ampliar a vantagem* e *balançar a rede* como um recurso textual, que garante não só a não-repetição de palavras no texto, mas também um caráter informal e próximo ao leitor, que ressoa a popularidade do futebol e a sua forte presença na vida das pessoas.

A sexta e última colocação a retratar o momento do gol é *sofrer gol*, com 90 casos no *corpus*. É possível dizer que essa colocação está no mesmo nível de *marcar gol* ou *fazer gol*, se referindo apenas ao ponto de vista do time adversário àquele que marcou um ponto, como em

- (13) Mas não conseguiu acabar com a sina de **sofrer gols** logo no início.

O número elevado de colocações que retratam a realização de um gol pode estar relacionado ao fato de que, como dito acima, fazer gols é o grande objetivo da partida de futebol, ainda mais ao percebermos que as três colocações mais recorrentes no *corpus* dizem respeito a essa cena. E a importância desse momento é suficiente para que os fãs do esporte criem diversificadas maneiras de se referir a ele, levando em consideração diferentes perspectivas. Nesse sentido, notamos que *sofrer gol* retrata o mesmo cenário, a mesma situação, mas a partir do ponto de vista de outro elemento que compõe tal cena. Como vimos anteriormente na revisão teórica, um *frame* compreende uma quantidade de elementos – atores e objetos – que, ao se relacionarem, desencadeiam a situação retratada no cenário. No caso dessa colocação, a situação apenas está sendo descrita sob outro prisma.

Outro caso em que colocações verbais designam dois pontos de vista contrários de um mesmo ato é o das expressões *dar passe/fazer passe* e *receber passe*. Dar passe (118 ocorrências) e *fazer passe* (53 ocorrências) representam o momento em que um jogador transfere a posse da bola para um colega de equipe. Vejamos exemplos do *corpus*:

- (14) Ivanildo desceu pela meia direita, **fez o passe** para Magique, que apareceu na cara do goleiro e tocou para o fundo das redes.
- (15) Leandro **faz o passe** para o centroavante.
- (16) O jovem fez um gol e **deu passe** para outro.
- (17) Responsável pela armação, Ganso **deu passes** precisos que criaram boas chances.

Verificamos, a partir das sentenças, uma tendência de a colocação vir seguida de um objeto indireto (utilizando a preposição *para*) que indica para quem foi o passe. No exemplo (17), em que o substantivo se apresenta no plural, no entanto, isso não ocorre, aspecto que se confirma quando analisamos outros casos:

- (18) Buscou fugir da marcação, conseguiu trabalhar com Luís Fabiano e Aloísio e **deu passes** e dribles.
- (19) O argentino **deu passes** longos, de calcanhar, se movimentou pelas pontas, driblou no meio, fez a parede, armou, desarmou, deu carrinho e muito mais.

Acreditamos que isso possa estar relacionado a como o uso do substantivo no singular faz referência a um passe específico, enquanto as ocorrências no plural tendem a descrever, de forma mais geral, as ações de um jogador.

Já a colocação *receber passe* (198 ocorrências), como o próprio verbo sugere, refere-se ao ponto de vista do jogador escolhido pelo colega de equipe para receber a posse da bola, como é possível perceber pelas sentenças extraídas do *corpus*:

- (20) Com 25 minutos, Edenílson **recebeu passe**, invadiu a área e quando se preparava para tentar fazer o gol, foi derrubado por Élton.
- (21) Primeiro, ele **recebeu passe** de Dadá e tocou sem chances para o camisa 1 do Coxa.
- (22) Aos 42, ele **recebeu passe** de Paulinho e chutou com extrema categoria sem chances de defesa para Fabio um golaço!
- (23) O castigo veio aos 25 minutos. Após **receber passe** de Vidal, Giovinco deixou Zapata no chão e bateu no canto de Abbiati para marcar um belo gol e virar o placar em Turim.

Dos quatro exemplos acima, três possuem objeto indireto do verbo *receber* (com a preposição *de*) que especifica a origem do passe, isto é, o jogador que passa a bola.

Seguindo em frente com a análise das colocações verbais, nos voltamos agora àquelas de estrutura V + Adj. Das 43 estruturas com verbo identificadas, 13 correspondem a essa combinação com advérbio. Uma primeira característica a ser

observada diz respeito a como, dessas 13 colocações, 12 apresentam locuções adverbiais. Apenas uma, *chutar cruzado*, possui advérbio. Além disso, dentre as 12, 10 possuem locuções adverbiais que indicam lugar. São elas: *entrar em campo* (345 ocorrências), *chegar ao gol* (117), *entrar na área* (102), *arriscar de fora da área* (100), *receber na área* (96), *arriscar de longe* (95), *cruzar na área* (92), *tocar na saída* (77), *receber na entrada* (73) e *levantar na área* (63).

Como podemos perceber, grande parte das locuções adverbiais de lugar dizem respeito à área do gol e ao próprio gol. Observemos os exemplos do *corpus*:

- (24) O Criciúma respondeu e quase **chegou ao gol**, quando Daniel Carvalho, aos 17, cobrou uma falta da entrada da área e a bola passou raspando a trave do goleiro Aranha.
- (25) Ele **entrou na área**, driblou o goleiro Jefferson e, com o gol vazio, tocou por cima do travessão.
- (26) Aos 13, Lodeiro **arriscou de fora da área**, para a defesa do goleiro Alessandro.
- (27) Aos 29 minutos, Jadson **recebeu na área** e tentou a finalização, mas a bola passou longe do gol de Cássio.
- (28) Aos dois minutos, Vinícius Bovi **cruzou na área** alvinegra e Danilo Sacramento dominou sozinho.
- (29) O lateral **levantou na área**, mas a zaga afastou o perigo.

No exemplo (24), percebemos que a combinação do verbo *chegar* com a locução adverbial apresenta um sentido relacionado a um *alvo*, o local ao qual o sujeito se dirigiu. O mesmo ocorre no exemplo (25), com o verbo *entrar*. Já na sentença 26, notamos que a preposição *de* denota uma origem, o local do qual foi realizada a ação descrita na colocação. O caso (27), ainda que apresente a mesma locução adverbial de (25), não parece carregar o mesmo sentido de atingir, chegar a um alvo, algo referente à presença do verbo *receber*. Diferentemente do exemplo

anterior, em que o sujeito realiza uma ação *de fora para dentro* da área, no exemplo (27) o sujeito está dentro do local. Tais nuances só podem ser percebidas a partir de um estudo que leve em consideração tanto a colocação como um todo quanto os elementos que a compõem, uma vez que cada uma dessas partes constitui um item lexical preenchido de informação semântica. O mesmo ocorre nos exemplos (28) e (29). Nesse sentido, o que demonstramos aqui é que classificar a colocação como sendo de um determinado tipo não nos é suficiente (pensando na interface com os *frames*), pois, ainda que as colocações dos exemplos acima sejam do mesmo grupo, elas também apresentam diferenças (diferenças que devem ser do interesse do semanticista e do lexicógrafo).

A partir do exposto acima, concluímos a discussão das colocações verbais e passamos ao grupo das colocações de estrutura S + Adj., S + (prep.) + S e Adj. + Adv.

Dos 74 resultados que compõem o grupo de colocações, 29 correspondem aos três tipos especificados, sendo que:

- S + (prep.) + S = 18 ocorrências (24,32%);
- S + Adj. = 8 ocorrências (10,8%); e
- Adj. + Adv. = 1 ocorrência (1,35%).

Começando pelo primeiro tipo, notamos, de início, que as colocações S + (prep.) + S estão relacionadas a termos que descrevem partes do campo de futebol, como *linha de fundo*, *campo de defesa*, *campo de ataque* e *meio da área*. A primeira delas, com 334 ocorrências no *corpus*, designa as linhas que ficam nos dois extremos do campo, no centro das quais fica a grande área. Os exemplos abaixo ilustram o uso da colocação:

(30) Éderson vai até a **linha de fundo** e cruza para trás.

(31) Matsui investiu nas jogadas de **linha de fundo** pela direita.

É relevante ressaltar como, na sentença (31), a colocação serve, também, de nome para um tipo de jogada.

Campo de defesa e campo de ataque, com 107 e 98 ocorrências respectivamente, não são locais inteiramente fixos no campo. Para um time A, digamos, o lado esquerdo é o campo de defesa, onde fica o goleiro desse time, enquanto o lado direito é o campo de ataque, lado para o qual os jogadores devem tentar levar a bola e fazer gol. Para o time B, no entanto, tal organização é ao contrário. As sentenças abaixo demonstram as colocações no *corpus*:

- (32) O relógio marcava 14 minutos quando o zagueiro Betão estava com a bola dominada no ***campo de defesa***, pela direita, e resolveu recuar para seu companheiro.
- (33) Na sequência, o atacante Rômulo roubou a bola no ***campo de ataque***, mas chutou fraco, para a fácil defesa do goleiro Michel Alves.

Por fim, a última colocação a designar uma parte do campo é *meio da área*, que apresenta 55 ocorrências. O meio da área indica o espaço central de dentro da área do gol, como podemos ver no exemplo abaixo:

- (34) Com um chute do ***meio da área*** após uma bola enfiada de Jorge Henrique para Kleber, o garoto de 18 anos, melhor em campo, deixou o marcador novamente igual.

Seguindo com a análise, percebemos que, assim como ocorreu com as colocações verbais do tipo V + S, aqui também há a presença de colocações que se referem a um mesmo objeto: *time da casa* e *equipe da casa*, com 231 e 69 ocorrências no *corpus* respectivamente. Podemos notar que existe uma larga diferença no número de vezes que as duas colocações aparecem, indicando uma preferência dos falantes pela primeira forma, *time da casa*. A palavra *time*, por exemplo, ocorre 4.204 vezes no *corpus*, enquanto *equipe* ocorre pouco mais da metade disso, 2.455, o que parece confirmar essa hipótese. Vejamos as sentenças:

- (35) O ***time da casa*** virou aos 18 minutos, quando Leandro recebeu, cortou a marcação e bateu firme.

- (36) Aos 34 minutos, a **equipe da casa** voltou ao ataque com Gilsinho, que recebeu cruzamento e cabeceou com perigo.
- (37) A **equipe da casa** continuou tentando o ataque, mas dando pouco trabalho para o goleiro do Náutico.

Percebemos que a substituição de uma colocação pela outra pode ser feita sem acarretar problemas quanto ao sentido. Nesse caso, é possível afirmar que as duas se equivalem, e que o fato de *time da casa* ocorrer 162 vezes do que *equipe da casa* é, de fato, apenas uma questão de preferência dos falantes. No entanto, é importante observarmos que, fora do contexto futebolístico, uma substituição desse tipo não pode ser feita de forma assim tão segura. Tomemos como exemplo a colocação *trabalho de equipe*, que podemos utilizar em diversos domínios diferentes. Não nos parece que a troca por *trabalho de time* seja bem-sucedida. Ainda que seja necessário um estudo muito mais profundo, já podemos imaginar que *time* e *equipe* apresentam, no âmbito do futebol, um grau de sinonímia que, talvez, não seja tão marcante fora dele.

Outro aspecto interessante das colocações S + (prep.) + S de nosso estudo é a presença de combinatórias que carregam elementos em comum e que, ainda que não signifiquem o mesmo acontecimento (como *marcar gol* e *balançar e rede*, no caso das V + S), denotam eventos similares, como: *cobrança de falta* e *cobrança de escanteio*, *zona de rebaixamento* e *zona de classificação* e *oitavas de final* e *quartas de final*. O substantivo *cobrança*, presente nas duas colocações do primeiro par citado, permite compreendermos que as colocações estão relacionadas não a um mesmo acontecimento da partida de futebol, mas a um mesmo contexto em que o evento ocorre. O mesmo ocorre no segundo par, *zona de rebaixamento* e *zona de classificação*. Já no terceiro, é o elemento (prep.) + S que se repete, porém ainda representando uma relação das duas colocações a algo em comum – etapas finais de um campeonato, nesse caso.

Passamos agora à categoria das colocações de estrutura S + Adj., que diz respeito a 8 colocações do conjunto de 74 expressões usadas neste trabalho, ou 10,8%. A mais recorrente deste grupo é a colocação *cartão amarelo*, com 457 ocorrências, o que faz dela a segunda colocação mais frequente deste estudo. É interessante notar, inclusive, que *cartão vermelho* possui recorrência bastante

menor, com 120 ocorrências apenas. Logicamente, compreendemos que isso pode ser devido ao fato de que é mais comum os árbitros levantarem cartões amarelos do que vermelhos, tendo em vista que os últimos são destinados a infrações mais graves. As sentenças abaixo ilustram o uso das colocações no *corpus*:

- (38) Aos 9 minutos, Pereira faz falta em Tevez e toma **cartão amarelo**.
- (39) O árbitro deu o **cartão vermelho** direto e anotou pênalti, convertido por André Ayew.

Colocações do tipo aparecem também como termos para partes do campo, o que é o caso de *canto direito*, *lado esquerdo* e *canto esquerdo*, com 149, 138 e 133 ocorrências respectivamente. O grupo de colocações S + Adj. incluem também *ficha técnica* (136), *bola parada* (88) e *jogo válido* (76).

A categoria Adj. + Adv. apresenta apenas uma colocação: *livre na área*, que possui apenas 80 ocorrências.

5.2 QUESTÕES RELACIONADAS A *FRAMES*

Nesta segunda parte da análise, partimos para a relação das colocações com a Semântica de *Frames* e a FrameNet, discutindo as repercussões de tal relacionamento para o contexto deste estudo.

Em primeiro lugar, esclarecemos aqui porque o conceito de UL da FrameNet nos é de tanta importância, como dissemos anteriormente. A FrameNet, ao conferir status de UL a uma unidade, não só especifica que essa unidade está relacionada a um *frame*, como também garante sua posição na lista principal de verbetes da base de dados. Nesse sentido, defendemos, em nosso trabalho, que as colocações devem possuir status de UL, de verbete, ao contrário do que vimos a partir de Svensén (2009), no segundo capítulo. Assim, é preciso que analisemos de que forma as colocações estão para os *frames* semânticos. Para esta análise, utilizamos os *frames* futebolísticos do dicionário Field.

Seguindo a ordem de análise das colocações iniciada na primeira parte do capítulo, iniciamos pelas colocações de estrutura V + S. De início, como pudemos perceber através do exemplo do *frame* `Evento_Comercial`, no Capítulo 2, é

comum (e até prototípico) que verbos atuem como evocadores de *frame* em sentenças. Nesse sentido, as colocações verbais V + S apresentam uma estrutura a partir da qual a identificação do *frame* é consideravelmente fácil. No entanto, é preciso frisar a importância de olhar para a colocação como um todo, uma combinação lexical cujo sentido reside em cada uma das partes. Se analisarmos, por exemplo, a colocação *abrir placar* apenas a partir do verbo, não é possível realizar a evocação de um *frame* da linguagem do futebol, é necessário levar em consideração o peso semântico do objeto direto *placar* para que seja possível identificar a relação da colocação com o *frame* Marcar Gol. Algo parecido ocorre com construções com verbo-suporte que atuam como evocadoras de *frame*, questão estudada por Souza (2012) em Trabalho de Conclusão de Curso. Assim como ocorre com as colocações V + S, estruturas com verbo-suporte dependem da participação do complemento do verbo na evocação do *frame*.

Vejamos uma análise de algumas colocações V + S:

Marcar gol

(40) Os franceses ainda não **marcaram gols** no Mundial.

O caso da colocação *marcar gol* é especial. Diferentemente de como ocorre com a colocação *abrir placar*, aqui o verbo é forte o suficiente para evocar o *frame*, tendo em vista que é possível pensarmos numa sentença como *O jogador marcou nos minutos finais do segundo tempo*. No entanto, é importante destacarmos que isso só acontece devido à presença do elemento de *frame* que designa um jogador ou um time (*os franceses*, no caso da sentença acima). Se o elemento que realiza a ação fosse, por exemplo, o árbitro, o verbo sozinho não teria a mesma força, ele precisaria de um complemento que especificasse o que está sendo *marcado*. Assim, tal característica se apresenta apenas no caso do *frame* Marcar Gol.

Fazer gol

(41) Aos 45 do segundo, Rivellino **fez o gol** corinthiano.

A colocação *fazer gol*, ainda que se enquadre no mesmo *frame* da colocação acima, não apresenta o mesmo caráter independente para o verbo, uma vez que *fazer* não é capaz de evocar *Marcar Gol* sem a presença do complemento.

Cobrar falta

- (42) Emerson ***cobrou falta*** com categoria e achou Romarinho livre na área, que cabeceou para o fundo das redes.

Ainda que o verbo dê uma ideia acerca do tipo de evento, necessita-se do complemento para compreender a cena e enquadrá-la no *frame* *Cobrança de Falta*.

Fazer jogada

- (43) No último minuto, Renan Oliveira ***fez jogada*** individual na entrada da área, conseguiu passar pela marcação, invadiu a área pela esquerda, mas sofreu a marcação no carrinho do zagueiro Gabriel na hora do chute.

No caso da colocação *fazer jogada*, percebemos o mesmo que ocorre com *fazer gol* e com construções com verbo-suporte. De forma alguma é possível dizer que o verbo, desconectado do complemento, é capaz de evocar o *frame* retratado na sentença, que é o *frame* *Táticas do Field*.

Com a análise das colocações V + S, percebemos que tais estruturas, ainda que próximas do modelo mais prototípico de evocador de *frame* presente no arcabouço da FrameNet, mantêm a necessidade de que se aborde as colocações como as combinatórias lexicais que são. É preciso levar em consideração cada elemento da expressão. Devemos lembrar, ainda, que, de acordo com Hausmann (1989), nas colocações do tipo V + S, o substantivo caracteriza a base da colocação, isto é, a parte semanticamente independente. A partir de uma abordagem baseada em *frames* semânticos, notamos, portanto, que nem sempre tal característica se aplica, se considerarmos, por exemplo, a colocação *cobrar falta*. Necessitamos do

verbo para identificar, com precisão, qual a ação e quem a está desempenhando, pensando que poderíamos ter, por exemplo, *sufrer falta* ou *cometer falta*.

Seguindo, nos detemos um pouco sobre as colocações verbais do tipo V + Adj. e V + Adj. (Loc.).

Ainda que, como no tipo anterior, V + S, essas colocações apresentam a necessidade de considerar o elemento nominal na evocação do *frame*, percebemos alguns casos em que o verbo apresenta força suficiente para atuar, sozinho, nesse processo. O caso mais vidente disso é o da colocação *chutar cruzado*, que evoca o *frame* Chute. Como podemos ver nos exemplos retirados do *corpus* abaixo, o elemento *cruzado* atua apenas como uma especificação do tipo de chute realizado pelo jogador. Nesse caso, o verbo é, sim, capaz de evocar o *frame* sozinho.

- (44) Aos 44, Nixon recebeu de Vagner Love e **chutou cruzado** para ótima defesa de Fernando Prass.
- (45) Lucas Fernández **chuta cruzado**.

Outro caso em que percebemos forte participação do verbo é *receber na área*, enquadrada no *frame* Passe. Ainda que não seja um verbo intrinsecamente ligado à linguagem do futebol, como *chutar*, o verbo *receber* está ligado ao ato de um jogador transferir a posse da bola para um colega do time. Nesse sentido, *na área* apenas especifica o local no qual o passe foi feito e recebido.

- (46) Aos 29 minutos, Jadson **recebeu na área** e tentou a finalização, mas a bola passou longe do gol de Cássio.
- (47) Aos 22, Novakovic **recebeu na área** e chutou, mas foi prensado pela zaga.

Como podemos notar nas sentenças, há sempre a presença do elemento que corresponde ao jogador que recebe a bola. Em uma ação de passe, há também de haver um jogador que realize a ação, que passe a bola. No entanto, no caso da colocação *receber na área*, as sentenças não evidenciam obrigatoriedade de que esse elemento se faça presente.

Em outros casos, em que o verbo é mais *geral* e menos específico da linguagem do futebol, voltamos a perceber a participação ativa do sintagma nominal, como em *entrar em campo*, *chegar ao gol* e *chegar com perigo*. Vejamos os exemplos extraídos do *corpus*:

- (48) O Tricolor ***entra em campo*** na quarta-feira às 21h50 para enfrentar o Cruzeiro.
- (49) Na noite desta quarta-feira (03), o Corinthians ***entra em campo*** pela quinta rodada da fase de grupos da Libertadores.
- (50) Se não estava tão fácil construir jogadas, o melhor jeito de ***chegar ao gol*** foi destruindo.
- (51) O Flamengo tentava ***chegar ao gol*** de qualquer maneira, e abria espaços para os contra-ataques.
- (52) O Vasco ***chegou com perigo*** aos 24 minutos: após cruzamento de Wendel da esquerda, André se antecipou no primeiro pau e desviou de cabeça.
- (53) O Corinthians ***chegava com perigo*** e mais uma vez balançou o barbante, dessa vez com Romarinho, aos 29.

A colocação *entrar em campo*, presente nos exemplos (48) e (49), evoca o *frame* *Competição*. O que percebemos é que *em campo* participa da evocação, pois faz parte da ação desempenhada pelas equipes. Já as colocações *chegar ao gol* e *chegar com perigo* apresentam um aspecto interessante. As duas possuem o mesmo verbo e evocam dois *frames* diferentes: *Chance* e *Ataque*, respectivamente. A distinção entre um e outro só é permitida pelos sentidos presentes nos elementos nominais de cada colocação.

A partir de agora, passamos a tratar das colocações nominais, que representam um desafio para a relação do fenômeno com a Semântica de *Frames* no contexto deste trabalho. De início, percebemos que, ao não possuírem verbo em

sua estrutura, as colocações nominais não estão relacionadas à evocação de um evento ou de uma ação, o que faz com que as estruturas desse tipo não se enquadrem no modelo típico de evocador da FrameNet. Começamos a análise das colocações nominais pelas do tipo S + Adj.

Colocações como *canto direito* e *lado esquerdo* são termos da linguagem futebolística, uma vez que designam partes do campo de futebol, aparecendo em sentenças como

- (54) O atacante boliviano quase deixou o seu aos 32 minutos, quando recebeu belo passe de Rafinha e chutou forte no **canto direito** para grande defesa do goleiro santista.
- (55) Quando todos imaginavam que o zero não sairia do placar, Danilo recebeu no **lado esquerdo** e foi derrubado pelo zagueiro Lucas Claro.

É possível notar que as sentenças acima estão, na verdade, evocando outros *frames*, como *Chute (chutou forte)* no primeiro caso e *Passe (recebeu)* no segundo. No entanto, as colocações marcadas fazem parte do vocabulário esportivo e não se pode pensar em um dicionário da linguagem do futebol sem termos que descrevam os lugares e materiais utilizados no âmbito da partida. Seguindo a lógica da FrameNet, tais colocações merecem lugar na lista de verbetes, algo com o qual concordamos. No entanto, estar na lista principal requer estar relacionada a um *frame*. Para isso, a saída é relacionar colocações desse tipo a *frames* chamados *ontológicos*, isto é, que não designam ações e momentos, mas descrevem objetos e participantes. Dentre tais *frames*, as colocações *canto direito* e *lado esquerdo* evocam o *frame* *Campo*. Assim, o que acaba por acontecer é uma relação entre esses *frames* de caráter mais ontológico e outros *frames* mais prototípicos, uma vez que os do primeiro tipo *alimentam* os segundos, fornecem as ferramentas.

O mesmo ocorre, por exemplo, com as colocações *cartão amarelo* e *cartão vermelho*, que designam materiais utilizados pelo árbitro, sendo, portanto, enquadradas no *frame* ontológico *Equipamentos*. Vejamos exemplos:

- (56) Digão recebeu o terceiro **cartão amarelo** e não joga.

- (57) Para aumentar o prejuízo do Verdão do Oeste, Athos dividiu pelo alto, usou a mão para desviar e foi flagrado pelo árbitro Marcelo de Lima Henrique, que aplicou o **cartão vermelho**.

É importante salientar que os verbos que antecedem as colocações (*receber* e *aplicar*, respectivamente) dizem respeito a eventos que incluem cartões amarelos e vermelhos, mas como possíveis elementos de *frame*. No caso da colocação *cartão amarelo*, temos o jogador que o recebe. Nesse caso, o cartão amarelo foi dado pelo participante da partida responsável por esse tipo de atividade, o árbitro. O mesmo acontece na segunda sentença, com *cartão vermelho*, embora o sujeito seja o árbitro, que aplica. Em ambos os casos o *frame* evocado pela sentença (e não pela colocação) é o *Decisões do Juiz*. Assim, os *frames* ontológicos servem a dois propósitos: (1) permitir a descrição dos objetos e participantes que constituem uma partida de futebol e (2) mostrar de que forma tais elementos se distribuem nas diversas ações desempenhadas dentro do campo. É pertinente ressaltar, também, que algumas colocações do tipo S + (prep.) + S também se enquadram nas características discutidas até aqui, como *zona de rebaixamento*, *meio da área* etc., de forma que, quanto a esse tipo de colocação, nos voltaremos a outro aspecto.

Colocações como *troca de passe*, *cobrança de falta* e *cobrança de escanteio* são nominais e, no entanto, evocam *frames* de ações. Notemos os exemplos que seguem:

- (58) Pressão! Após boa **troca de passes**, Damião reclama de pênalti.
- (59) O São Paulo ainda teve chance em **cobrança de falta** de Carleto, aos 46 minutos, mas Cássio defendeu com facilidade.
- (60) Após **cobrança de escanteio** da esquerda, Dorval ajeitou para Everton Costa, na direita, que chutou cruzado.
- (61) Logo aos 2 minutos, outra boa **cobrança de escanteio** de Alex Telles, pela direita.

No exemplo (58), ao invés de utilizar uma construção verbal, como *Após terem trocado passes*, por exemplo, o narrador opta pela forma nominal *troca de passes*, que, ainda que não apresente verbo, serve de evocadora do *frame* *Passes*. Talvez isso ocorra devido ao fato de que se trata de um elemento muito típico do futebol, o passe. No entanto, é interessante notarmos ainda que a opção por uma forma verbal não permitiria o uso do adjetivo *boa* acoplado à colocação, o que demonstra, ainda, de que não se trata de uma escolha arbitrária – a opção pela colocação está baseada em conseguir uma versatilidade sintática que não seria possível alcançar através da forma verbal (o que nos lembra das construções com verbo-suporte, nesse sentido).

Notamos que o mesmo acontece nos exemplos (59) e (60). Os *frames*, *Cobrança de Falta* e *Cobrança de Escanteio*, respectivamente, são evocados pela colocação. Algo curioso a ser apontado é que, nesses casos, era possível ter utilizado construções com verbo (*após ter cobrado escanteio* e *teve chance de cobrar falta*) sem ter modificado o sentido da sentença, algo que não ocorre em (61); novamente há a presença do adjetivo *boa*, que não poderia ser utilizado caso o narrador tivesse optado por *Logo aos 2 minutos, Alex Telles novamente cobrou escanteio pela direita*, por exemplo.

O último caso analisado é o da colocação *livre na área*, a única a apresentar estrutura do tipo Adj. + Adv. Foi possível perceber que essas colocações aparecem, geralmente, acompanhadas de verbos que evocam *frames* diferentes, de forma que *livre na área* serve para informar ao leitor (ou ouvinte, pensando em narrações de partidas) o modo como o jogador se encontrava e onde ele estava, como podemos ver nas sentenças abaixo:

- (62) Aos 36, Luciano recebeu ***livre na área*** e empatou a partida novamente.
- (63) O lampejo de bom futebol veio de Nilmar que, após um passe de Luís Fabiano, chutou ***livre na área***.
- (64) O volante Marino chegou ***livre na área***, mas chutou por cima.

Na sentença (62), a colocação serve para identificar a posição e o estado do jogador Luciano no momento em que ele recebeu a bola. Se o evocador marcado na

sentença fosse o verbo *receber*, o *frame* seria *Passé*. Já no caso (63), Nilmar chuta a bola (*frame* *Chute*) enquanto livre na área. O último exemplo, de número (64), também se enquadra no *frame* *Chute*. Ressaltamos aqui, como já foi dito anteriormente, que colocações como essa não representam ações, acontecimentos no âmbito da partida. No entanto, tomando como base o *corpus* e a própria linguagem do futebol, estar *livre na área* faz parte do vocabulário desse universo, os falantes usam a colocação. Nesse sentido, defendemos que colocações sejam incluídas na lista de unidades lexicais. A necessidade de relação a um *frame*, nesse caso, é resolvida a partir do *frame* *Competição*, que abarca “Conjunto de ações, elementos e etapas relacionadas aos eventos em que ocorrem as disputas de futebol.” (FIELD..., 2014).

Para concluir esta análise, reiteramos, mais uma vez, que o fato de os nossos dicionários serem baseados nos postulados da Semântica de *Frames* e da FrameNet garante o lugar das colocações nas listas de unidades lexicais de tais recursos. Apoiamos esse aspecto, que é lexicográfico, tendo em vista a forma como as colocações futebolísticas designam elementos importantes da linguagem desse esporte. A popularidade do futebol em nosso país resulta em uma pluralidade de expressões de diferentes naturezas, que refletem a proximidade que existe entre o esporte e a nossa cultura, o nosso dia a dia. Um dicionário dedicado a descrever a linguagem do futebol não pode dar as costas a fenômenos como as colocações. Não obstante, o processo de criação de um dicionário desse tipo baseado em *frames* precisa refletir a respeito do papel e do lugar de expressões como essas aqui analisadas, criando subsídios para um tratamento adequado de tais estruturas, ainda que isso signifique apontar eventuais problemas no arcabouço teórico e metodológico utilizado (como fizemos de forma muito singela no capítulo 2).

Para fechar este estudo, nos voltamos agora às nossas considerações finais e perspectivas futuras.

6 CONSIDERAÇÕES FINAIS E PERSPECTIVAS FUTURAS

O principal objetivo desde trabalho foi realizar um estudo acerca do fenômeno das colocações a partir do contexto de um dicionário da linguagem do futebol baseado em *frames* semânticos. Chegando ao fim desta pesquisa, acreditamos ser possível dizer que cumprimos com nossos objetivos, respondendo a todas as questões levantadas no início do trabalho e que guiaram a execução da pesquisa. Sendo assim, aproveitamos os parágrafos que seguem para apresentar nossas considerações finais e perspectivas futuras.

No início do trabalho, com a revisão da literatura sobre colocações, nos deparamos com o conceito de *frequência*, sempre presente nos estudos dos teóricos que se dedicaram ao estudo do fenômeno, às vezes como aspecto importante da descrição das colocações, outras vezes como parte de críticas feitas por alguns estudiosos. Ainda que concordemos que, de fato, a noção de frequência não é suficiente para estudar a fundo o fenômeno das combinatórias lexicais, apontamos para as valiosas informações que a frequência nos proporcionou a respeito da natureza das colocações do contexto futebolístico, das quais destacamos a alta incidência de estruturas verbais, o que demonstra uma característica não só das colocações do esporte, mas da própria linguagem do futebol: o fato de ela ser voltada para ações, para acontecimentos. O discurso futebolístico é um discurso dinâmico, recheado de expressões que se relacionam às diversas formas de atuação no campo de futebol.

A possibilidade de usar a frequência como um dos critérios de estudo está totalmente relacionada com as ferramentas de análise da Linguística de Corpus, abordagem que utilizamos em nosso trabalho a partir de uma perspectiva mais metodológica. Isso porque a área da Linguística de Corpus, em amplo desenvolvimento no Brasil, oferece a oportunidade de nos aliarmos à rapidez e à precisão das ferramentas computacionais. Destacamos que o Sketch Engine e em especial seu maquinário dedicado à extração das colocações se mostraram altamente eficientes, permitindo que a extração das colocações para análise ocorresse de forma parcialmente automática (considerando a análise manual dos 500 primeiros resultados, como especificamos anteriormente).

Outro aspecto das colocações do futebol pertinente de ser mencionado é a forma como diversas colocações nominais são usadas para descrever partes do

campo e equipamentos dos participantes, o que, em nossa visão, reitera a importância de se ter as colocações na lista de verbetes do dicionário. Essas colocações são *termos*⁵² da linguagem futebolística, e não se pode conceber um dicionário do discurso desse esporte sem levar em consideração esses elementos.

A parte seguinte da revisão teórica deu conta de apresentar a teoria da Semântica de *Frames* e a FrameNet. Quando relacionado às colocações, o arcabouço da teoria de Fillmore (1982; 1985) evidenciou duas questões principais: (1) a FrameNet, ainda que não traga uma reflexão aprofundada acerca das colocações (e de outros fenômenos lexicais), apresenta subsídios para o tratamento dessas expressões em contexto lexicográfico, e (2) a Semântica de *Frames* demonstra sua relevância para a conceptualização da linguagem do futebol, o que averiguamos através da estruturação dos *frames* do Field e de como esses cenários se associaram às colocações. Os *frames* permitem que se organizem os diferentes momentos de uma partida de futebol e os seus elementos de forma a relacionar uma coisa à outra, criando uma rede de eventos e de significados.

Foi possível evidenciar, como visto na análise, que as colocações nominais, ainda que um tanto atípicas em relação ao conceito de unidade lexical evocadora de *frame*, podem ser enquadradas em cenários de caráter mais ontológico, mais estático (em oposição ao aspecto dinâmico dos *frames* mais tradicionais). As colocações verbais, por sua vez, ainda que apresentem um elemento habitualmente conectado à evocação de *frames* semânticos, em muitos casos evidenciaram a necessidade de levar em consideração o sintagma nominal.

Outro ponto a ser destacado aqui diz respeito a um contexto um pouco mais amplo: o das expressões multivocabulares. Acreditamos ter sido possível evidenciar que as fronteiras que separam um fenômeno do outro são, muitas vezes, difusas. Em mais de uma ocasião, ao longo do trabalho, destacamos a proximidade entre as colocações e as estruturas com verbo-suporte.

Ainda que não esteja relacionado às colocações em si, é importante salientar, também, nossos questionamentos acerca do caráter lexicográfico da FrameNet. Como asseveramos ainda no segundo capítulo, não foi nosso objeto neste trabalho desenvolver um estudo sobre a FrameNet ser ou não ser um dicionário, ser ou não

⁵² Vale ressaltar que utilizamos *termo*, aqui, de forma a designar unidades linguísticas intrínsecas a um discurso de contexto particular. Em outras palavras, não foi nosso objetivo realizar uma reflexão acerca de o que um termo é e quais suas características.

ser um recurso lexicográfico. Questionamos, sim, contudo, a estrutura da FrameNet e a forma como as informações estão ali dispostas. Quem é o usuário da FrameNet? *Para quem a FrameNet foi idealizada?* Tais questões nos interessam, tendo em vista que o recurso baseado em *frames* faz parte do conjunto teórico-metodológico das pesquisas desenvolvidas pelo SemanTec.

Quanto ao tratamento lexicográfico dispensado às colocações no âmbito de dicionários organizados em torno do conceito da teoria de Fillmore (1982; 1985), foi possível comprovar que as colocações não devem ser inseridas dentro de outros verbetes, como proposto por Svensén (2009), mas disponibilizadas como verbetes – ou unidades lexicais – na lista a que o usuário tem acesso. Acreditamos que essa discussão é importante não só para os projetos desenvolvidos pelo grupo de pesquisa, mas para qualquer recurso lexicográfico baseado em *frames* semânticos. As colocações são parte de nosso uso diário da língua, possuem sentido. Defendemos, assim, que não devem ser tratadas como meros *desdobramentos* de outras palavras.

Com o fim deste trabalho, destacamos que a equipe do grupo de pesquisa SemanTec dedicada aos dicionários esportivos baseados em *frames* segue atuando em duas frentes: (1) a manutenção e expansão do Field – Dicionário de Expressões do Futebol, com a inclusão de novos dados obtidos através do contínuo estudo da linguagem do futebol e (2) o desenvolvimento do dicionário voltado às modalidades olímpicas, tendo em vista os Jogos Olímpicos de 2016, a serem realizados no Brasil.

Além disso, destacamos também a pesquisa a nível de doutorado que iniciamos neste ano de 2015, a qual trata, especificamente, da relevância do conceito de *frame* semântico para a lexicografia, de forma que buscaremos responder às questões levantadas anteriormente sobre a FrameNet.

Por fim, esperamos que este trabalho contribua não só com as pesquisas desenvolvidas pelo SemanTec, mas também com outros estudos que, por ventura, sejam realizados.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALONSO RAMOS, M. Aproximación a un nuevo modelo lexicográfico: El Dictionnaire explicatif et combinatoire du français contemporain. **Recherches lexico-sémantiques de Igor Mel'cuk**. Verba. Santiago de Compostela, v. 16, 421-450, 1989.
- ATKINS, S.; RUNDELL, M.; SATO, H.: **The contribution of FrameNet to Practical Lexicography**. International Journal of Lexicography, Oxford, v. 16, n. 3, p. 333-357, 2003.
- BALLY, C. **Traité de stylistique française**. Paris: Klincksieck, 1951.
- BENEDUZI, R. **Colocações substantivo + adjetivo**: propostas para sua identificação e tratamento lexicográfico em dicionários ativos português-espanhol. Dissertação (Mestrado em Letras) 212f. Universidade Federal do Rio Grande do Sul: Porto Alegre, 2008.
- BENSON, M. Lexical Combinability. **Papers in Linguistics**. Edmonton, v. 18, n. 1, p. 3-15, 1986.
- BICK, E. **The Parsing System Palavras** - Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework. Aarhus: Aarhus University Press, 2000.
- BIDERMAN, M. T. 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.
- CHISHMAN, R. L. O. **Dicionário Eletrônico Modalidades Olímpicas 2016**. Projeto apresentado à Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), no âmbito do Edital FAPERGS 02/2014. São Leopoldo, 2014.
- CHISHMAN, R. L. O. (Org.) **Field** – Dicionário de Expressões do Futebol. São Leopoldo: UNISINOS, 2014.
- CHISHMAN, R. L. O.; KRIEGER, M. G.; RIGO, S. **Kicktionary-Br – Copa 2014**. Projeto apresentado ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), no âmbito do Edital Universal nº 14/2011. São Leopoldo, 2011.
- COWIE, A. P. The Treatment of Collocations and Idioms in Learners' Dictionaries. **Applied Linguistics**, 1977. p. 143-161.
- CORPAS PASTOR, G. **Manual de Fraseología Española**. Madrid: Gredos, 1996.
- COP, M. **The function of collocations in dictionaries**. In: MAGAY, T.; ZIGÁNY, J. (Eds.), 1990, p. 35-46.

CRUSE, D. A. **Lexical Semantics**. Cambridge Textbooks in Linguistics. Cambridge University Press: Cambridge, 1986.

ESSER, J. Collocation, Colligation, Semantic Preference and Semantic Prosody: New Developments in the Study of Syntagmatic Word Relations. In: FALKNER, W.; SCHMID, H. (Eds.) **Words, Lexemes, Concepts** – Approaches to the Lexicon. Studies in Honor of Leonhard Lipka. Tübingen: Gunter Narr, 1999. p. 155-165.

ETTINGER, S. Formación de palabras y fraseología en la lexicografía. In: HAENSCH et al. **La lexicografía**. De la lingüística teórica a la lexicografía práctica. Madrid: Gredos, 1982, p. 233-258.

FILLMORE, C. J. The case for case. In: BACH, E.; HARMS, R. T. (Org.). **Universals in linguistic theory**. Nova York: Holt, Rinehart and Winston, 1968. p. 1-88.

_____. Frame Semantics. In: **Linguistics in the Morning Calm**. Seoul, Hansinh Publishing Co., 1982. p. 111-137.

_____. Frames and the semantics of understanding. In: **Quaderni di Semantica**, vol. 6, n. 2, 1985. p. 222-254.

FILLMORE, C. J.; ATKINS, B. T. Towards a Frame-based organization of the lexicon: the semantics of RISK and its neighbors. In: LEHRER, A.; KITTAY, E. (Orgs.) **Frames, Fields, and Contrasts: New Essays in Semantics and Lexical Organization**. Hillsdale: Lawrence Erlbaum, 1992, p. 75-102.

FIRTH, J. R. Modes of meaning. In: _____. **Papers in linguistics 1934-1951**. London: Oxford University, 1957. p. 190-215.

FONTENELLE, T. A Bilingual Lexical Database for Frame Semantics. In: **International Journal of Lexicography**, vol. 13, n. 4, 2000. p. 232 - 248.

FRAMENET. Berkeley, [2013?]. Disponível em: <<https://framenet.icsi.berkeley.edu>>. Acesso em: 10 abr. 2014.

GOFFMAN, E. **Frame Analysis: An Essay on the Organization of Experience**. Cambridge, MA (US): Harvard University Press, 1975.

HALLIDAY, M. A. K. Categories of the Theory of Grammar. **Word**. New York, v. 17, n. 3, p.241-292, 1961.

HAUSMANN, F.J. Le dictionnaire de collocations. In: HAUSMANN, F.J.; REICHMANN, O.; WIEGAND, H. E.; ZGUSTA, L. (Orgs.). **Wörterbücher, Dictionaries, Dictionnaires**. Ein Internationales Handbuch zur Lexikographie, v. 1. Berlin: Walter de Gruyter, 1989, p. 1010-1019.

HUNSTON, S. **Corpora in Applied Linguistics**. Cambridge: Cambridge University Press, 2002.

KILGARRIFF, A. et al. The Sketch Engine. In WILLIAMS, G.; VESSIER, S. (Eds.) **Proceedings of the Eleventh Euralex International Congress**. Lorient: Université de Bretagne-Sud, 2004. p. 105-116.

McENERY, T.; HARDIE, A. **Corpus Linguistics: Method, Theory and Practice**. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press, 2012.

MEL'ČUK, I. Phrasemes in Language and Phraseology in Linguistics. In: EVERAERT, M. et al. (Eds.). **Idioms: Structural and Psychological Perspectives**. New Jersey: Lawrence Erlbaum, 1995. P. 167-232.

_____. Les collocations: définition, rôle et utilité. In: GROSSMAN, F.; TUTIN, A. (Eds.). **Les Collocations: analyse et traitement**. Amsterdam: De Werelt, 2003, p. 23-31.

MINSKY, M. A framework for representing knowledge. In: **Artificial Intelligence Memo**, n. 306. Cambridge, MA: Massachusetts Institute of Technology, 1974.

NEVES, M. H. M. **Gramática de usos**. São Paulo: Unesp, 2000.

PETRUCK, M. R. L. **Frame semantics**. Berkeley: University of California, 2001.

PORZIG, W. **El mundo maravilloso del lenguaje**. Madrid: Gredos, 1970.

RUPPENHOFER, J. et al. **FrameNet II: Extended Theory and Practice**. Berkeley, California: International Computer Science Institute, 2010. Disponível em: <<https://framenet2.icsi.berkeley.edu/docs/r1.5/book.pdf>>. Acesso em: 09 abr. 2014.

SARDINHA, T. B. **Linguística de Corpus: histórico e problemática**. DELTA, vol.16, n.2, 2000. p. 323-367.

SAUSSURE, F. **Curso de linguística geral**. São Paulo: Cultrix, 2004.

SCHMIDT, T. The Kicktionary – A multilingual lexical resource of football language. In: BOAS, H. C. (Ed.) **Multilingual FrameNets in computational lexicography: Methods and applications**. Berlin/New York: Mouton de Gruyter, 2009, p. 102-132.

SECO, M. Problemas formales de la definición lexicográfica. In: **Estudios de lexicografía española**. Madrid: Paraninfo, 1987a. p. 15-34.

SINCLAIR, J. M. et al. **English Collocation Studies: The OSTI Report**. Londres: Continuum, 2004.

SKETCH Engine. East Sussex, [2013?]. Disponível em: <<http://www.sketchengine.co.uk/>>. Acesso em: 03 jul. 2014.

SOUZA, D. S. **Os verbos-suporte e a temática do futebol: desafios para anotação semântica**. 2012. 95 f. Trabalho de Conclusão de Curso (Licenciatura em Letras) – Curso de Letras. Universidade do Vale do Rio dos Sinos: São Leopoldo, 2012.

STUBBS, M. **Collocations and Cultural Connotations of Common Words.** Linguistics and Education 7, 1995. p. 379-390.

SVENSÉN, B. **A Handbook of Lexicography:** the theory and practice of dictionary-making. Cambridge: Cambridge University Press, 2009.

TAGNIN, S. E. O. **O jeito que a gente diz:** combinações consagradas em inglês e português. São Paulo: Disal Editora, 2013.

TOLKIEN, J. R. R. **O Hobbit.** 6ª edição. São Paulo: Editora WMF Martins Fontes, 2012.

WANNER, L. **Lexical functions in lexicography and natural language processing.** Amsterdam: John Benjamins, 1996.

**APÊNDICE A – COLOCAÇÕES SELECIONADAS A PARTIR DA BUSCA NO
SKETCH ENGINE**

	Colocação	Frequência
1	Abrir placar	784
2	Cartão amarelo	457
3	Marcar gol	441
4	Fazer gol	409
5	Cobrança de falta	394
6	Cobrar falta	390
7	Zona de rebaixamento	364
8	Entrar em campo	345
9	Linha de fundo	334
10	Passar bola	329
11	Fazer jogada	326
12	Balançar a rede	245
13	Time da casa	231
14	Cobrança de escanteio	230
15	Invadir a área	226
16	Levar perigo	224
17	Receber passe	198
18	Criar chance	189
19	Bater bola	163
20	Cobrar escanteio	152
21	Canto direito	149
22	Oitavas de final	143
23	Lado esquerdo	138
24	Ficha técnica	136
25	Gol da vitória	136
26	Canto esquerdo	133
27	Tabela de classificação	124
28	Bola na área	121
29	Quartas de final	121

30	Cartão vermelho	120
31	Chutar cruzado	120
32	Receber lançamento	120
33	Dar passe	118
34	Chegar ao gol	117
35	Ter trabalho	116
36	Saída de bola	114
37	Mandar bola	113
38	Sofrer falta	110
39	Criar oportunidade	109
40	Chegar com perigo	108
41	Campo de defesa	107
42	Receber cartão	105
43	Afastar perigo	104
44	Entrar na área	102
45	Arriscar de fora da área	100
46	Campo de ataque	98
47	Receber na área	96
48	Arriscar de longe	95
49	Cruzar na área	92
50	Fazer jogo	92
51	Mandar para escanteio	91
52	Sofrer gol	90
53	Troca de passe	90
54	Bola parada	88
55	Acertar a trave	83
56	Livre na área	80
57	Ampliar vantagem	78
58	Banco de reserva	77
59	Tocar na saída	77
60	Jogo válido	76
61	Receber na entrada	73
62	Zona de classificação	71

63	Equipe da casa	69
64	Deixar o campo	68
65	Marcar pênalti	68
65	Bola na trave	66
67	Perigo ao gol	66
68	Gol de empate	65
69	Levar na área	63
70	Perder gol	61
71	Vencer jogo	58
72	Garantir vitória	57
73	Meio da área	55
74	Fazer passe	53

APÊNDICE B – CLASSIFICAÇÃO DAS COLOCAÇÕES A PARTIR DE HAUSMANN (1989) E DE *FRAMES*

1. Abrir placar

Frame: Marcar Gol

Classificação: V + S

Base: placar / Colocado: abrir

2. Cartão amarelo

Frame: Equipamentos

Classificação: S + Adj.

Base: cartão / Colocado: amarelo

3. Marcar gol

Frame: Marcar gol

Classificação: V + S

Base: gol / colocado: marcar

4. Fazer gol

Frame: Marcar gol

Classificação: V + S

Base: gol / colocado: fazer

5. Cobrança de falta

Frame: Cobrança de Falta

Classificação: S + Prep. + S

Base: cobrança / Colocado: falta

6. Cobrar falta

Frame: Cobrança de Falta

Classificação: V + S

Base: falta / Colocado: cobrar

7. Zona de rebaixamento

Frame: Competição

Classificação: S + Prep. + S

Base: zona / Colocação: rebaixamento

8. Entrar em campo

Frame: Tempo

Classificação: V + Adv. (Loc.)

Base: entrar / Colocado: em campo

9. Linha de fundo

Frame: Campo

Classificação: S + prep. + S

Base: linha / Colocado: fundo

10. Passar bola

Frame: Passe

Classificação: V + S

Base: bola / Colocado: passar

11. Fazer jogada

Frame: Táticas

Classificação: V + S

Base: jogada / Colocado: fazer

12. Balançar rede

Frame: Marcar Gol

Classificação: V + S

Base: rede / Colocado: balançar

13. Time da casa

Frame: Participantes

Classificação: S + Prep. + S

Base: time / Colocado: casa

14. Cobrança de escanteio

Frame: Escanteio

Classificação: S + Prep. + S

Base: cobrança / Colocado: escanteio

15. Invadir área

Frame: Ataque

Classificação: V + S

Base: área / Colocado: invadir

16. Levar perigo

Frame: Chance

Classificação: V + S

Base: perigo / Colocado: levar

17. Receber passe

Frame: Passe

Classificação: V + S

Base: passe / Colocado: receber

18. Criar chance

Frame: Chance

Classificação: V + S

Base: chance / Colocado: criar

19. Bater bola

Frame: Competição

Classificação: V + S

Base: bola / Colocado: bater

20. Cobrar escanteio

Frame: Escanteio

Classificação: V + S

Base: escanteio / Colocado: cobrar

21. Canto direito

Frame: Campo

Classificação: S + Adj.

Base: canto / Colocado: direito

22. Oitavas de final

Frame: Competição

Classificação: S + Prep. + S

Base: oitavas / Colocado: final

23. Lado esquerdo

Frame: Campo

Classificação: S + Adj.

Base: lado / Colocado: esquerdo

24. Ficha técnica

Frame: Equipamentos

Classificação: S + Adj.

Base: ficha / Colocado: técnica

25. Gol da vitória

Frame: Placar

Classificação: S + Prep. + S

Base: gol / Colocado: vitória

26. Canto esquerdo

Frame: Campo

Classificação: S + Adj.

Base: canto / Colocado: esquerdo

27. Tabela de classificação

Frame: Competição

Classificação: S + Prep. + S

Base: tabela / Colocado: classificação

28. Bola na área

Frame: Movimento da Bola

Classificação: S + Prep. + S

Base: bola / Colocado: área

29. Quartas de final

Frame: Competição

Classificação: S + Prep. + S

Base: quartas / Colocado: final

30. Cartão vermelho

Frame: Equipamentos

Classificação: S + Adj.

Base: cartão / Colocado: vermelho

31. Chutar cruzado

Frame: Chute

Classificação: V + Adv.

Base: chutar / Colocado: cruzado

32. Receber lançamento

Frame: Passe

Classificação: V + S

Base: lançamento / Colocado: receber

33. Dar passe

Frame: Passe

Classificação: V + S

Base: passe / Colocado: dar

34. Chegar ao gol

Frame: Ataque

Classificação: V + Adv. (Loc.)

Base: chegar / Colocado: ao gol

35. Ter trabalho

Frame: Ataque

Classificação: V + S

Base: trabalho / Colocado: ter

36. Saída de bola

Frame: Movimento da Bola

Classificação: S + Prep. + S

Base: saída / Colocado: bola

37. Mandar bola

Frame: Chute

Classificação: V + S

Base: bola / Colocado: mandar

38. Sofrer falta

Frame: Infração

Classificação: V + S

Base: falta / Colocado: sofrer

39. Criar oportunidade

Frame: Chance

Classificação: V + S

Base: oportunidade / Colocado: criar

40. Chegar com perigo

Frame: Ataque

Classificação: V + Adv. (Loc.)

Base: chegar / Colocado: com perigo

41. Campo de defesa

Frame: Campo

Classificação: S + Prep. + S

Base: campo / Colocado: defesa

42. Receber cartão

Frame: Infração

Classificação: V + S

Base: cartão / Colocado: receber

43. Afastar perigo

Frame: Defesa

Classificação: V + S

Base: perigo / Colocado: afastar

44. Entrar na área

Frame: Chance

Classificação: V + Adv. (Loc.)

Base: entrar / Colocado: na área

45. Arriscar de fora da área

Frame: Chute

Classificação: V + Adv. (Loc.)

Base: arriscar / Colocado: de fora da área

46. Campo de ataque

Frame: Campo

Classificação: S + Prep. + S

Base: campo / Colocado: ataque

47. Receber na área

Frame: Passe

Classificação: V + Adv. (Loc.)

Base: Receber / Colocado: na área

48. Arriscar de longe

Frame: Chute

Classificação: V + Adv. (Loc.)

Base: arriscar / Colocado: de longe

49. Cruzar na área

Frame: Passe

Classificação: V + Adv. (Loc.)

Base: cruzar / Colocado: na área

50. Fazer jogo

Frame: Competição

Classificação: V + S

Base: jogo / Colocado: fazer

51. Mandar para escanteio

Frame: Escanteio

Classificação: V + Adv. (Loc.)

Base: mandar / Colocado: para escanteio

52. Sofrer gol

Frame: Marcar Gol

Classificação: V + S

Base: gol / Colocado: sofrer

53. Troca de passe

Frame: Passe

Classificação: S + Prep. + S

Base: troca / Colocado: passe

54. Bola parada

Frame: Movimento da Bola

Classificação: S + Adj.

Base: bola / Colocado: parada

55. Acertar trave

Frame: Perder Gol

Classificação: V + S

Base: trave / Colocado: acertar

56. Livre na área

Frame: Competição

Classificação: Adj. + Adv. (Loc.)

Base: livre / Colocado: na área

57. Ampliar vantagem

Frame: Chance

Classificação: V + S

Base: vantagem / Colocado: ampliar

58. Banco de reserva

Frame: Banco de Reservas

Classificação: S + Prep. + S

Base: banco / Colocado: reserva

59. Tocar na saída

Frame: Chute

Classificação: V + Adv. (Loc.)

Base: tocar / Colocado: na saída

60. Jogo válido

Frame: Competição

Classificação: S + Adj.

Base: jogo / Colocado: válido

61. Receber na entrada

Frame: Passe

Classificação: V + Adv. (Loc.)

Base: receber / Colocado: na entrada

62. Zona de classificação

Frame: Competição

Classificação: S + Prep. + S

Base: zona / Colocado: classificação

63. Equipe da casa

Frame: Equipe

Classificação: S + Prep. + S

Base: equipe / Colocado: casa

64. Deixar o campo

Frame: Substituição

Classificação: V + S

Base: campo / Colocado: deixar

65. Marcar pênalti

Frame: Funções do Juiz

Classificação: V + S

Base: pênalti / Colocado: marcar

66. Bola na trave

Frame: Perder Gol

Classificação: S + Prep. + S

Base: bola / Colocado: trave

67. Perigo ao gol

Frame: Chance

Classificação: S + Prep. + S

Base: perigo / Colocado: ao gol

68. Gol de empate

Frame: Placar

Classificação: S + Prep. + S

Base: gol / Colocado: empate

69. Levar na área

Frame: Chance

Classificação: V + Adv. (Loc.)

Base: levar / Colocado: na área

70. Perder gol

Frame: Perder gol

Classificação: V + S

Base: gol / Colocado: perder

71. Vencer jogo

Frame: Competição

Classificação: V + S

Base: jogo / Colocado: vencer

72. Garantir vitória

Frame: Competição

Classificação: V + S

Base: vitória / Colocado: garantir

73. Meio da área

Frame: Campo

Classificação: S + Prep. + S

Base: meio / Colocado: área

74. Fazer passe

Frame: Passe

Classificação: V + S

Base: passe / Colocado: fazer

ANEXO A – RELAÇÕES *FRAME A FRAME*

6.1. *FRAME-TO-FRAME RELATIONS*

Relation	Sub	Super
Inheritance	Child	Parent
Perspective_on	Perspectivized	Neutral
Subframe	Component	Complex
Precedes	Later	Earlier
Inchoative_of	Inchoative	State
Causative_of	Causative	Inchoative/State
Using	Child	Parent
See_also ¹	Referring Entry	Main Entry

Table 6.1: Types of Frame-frame Relations

Fonte: RUPPENHOFER et al., 2010, p. 75.

ANEXO B – BUSCA REALIZADA NO SKETCH ENGINE (500 PRIMEIROS RESULTADOS)

Word sketch collocations	Freq		
abrir objeto placar	784	posse pp_de bola	315
placar objeto_de abrir	784	chegar pp_a+o ponto	295
não modifié conseguir	698	voltar pp_a campo	246
conseguir modifier não	697	balançar objeto rede	245
entrada pp_de+o área	547	rede objeto_de balançar	245
não modifié ser	462	válido pp_por+o rodada	244
ser modifier não	462	time pp_de+o casa	231
amarelo modifié cartão	457	cobrança pp_de escanteio	230
cartão modifier amarelo	457	chutar modifier forte	228
defesa objeto_de fazer	454	forte modifié chutar	228
fazer objeto defesa	454	invadir objeto área	226
gol objeto_de marcar	441	área objeto_de invadir	226
marcar objeto gol	441	levar objeto perigo	224
não modifié ter	424	perigo objeto_de levar	224
ter modifier não	424	etapa modifier inicial	208
chance objeto_de ter	422	inicial modifié etapa	208
ter objeto chance	422	final modifié minuto	206
min pp_de+o tempo	414	minuto modifier final	206
gol objeto_de fazer	409	chance pp_de gol	200
fazer objeto gol	407	passe objeto_de receber	198
cobrança pp_de falta	394	receber objeto passe	198
minuto pp_de+o tempo	393	chance objeto_de criar	189
cobrar objeto falta	390	criar objeto chance	189
falta objeto_de cobrar	390	individual modifié jogada	185
zona pp_de rebaixamento	364	jogada modifier individual	185
etapa modifier final	355	muito modifié ser	182
final modifié etapa	355	ser modifier muito	182
entrar pp_em campo	345	bola sujeito_de sair	178
de=fundo modifié linha	334	sair sujeito bola	178
linha modifier de=fundo	334	gol sujeito_de sair	174
bola sujeito_de passar	329	sair sujeito gol	174
passar sujeito bola	329	fundo pp_de+o rede	172
fazer objeto jogada	326	ter sujeito time	169
jogada objeto_de fazer	326	time sujeito_de ter	169
		minuto pp_de jogo	167

minuto modifier seguinte	164
seguinte modifié minuto	164
bater sujeito bola	163
bola sujeito_de bater	163
complementar modifié etapa	163
etapa modifier complementar	163
não modifié poder	162
poder modifier não	162
tiro pp_de meta	162
chance objeto_de perder	158
perder objeto chance	158
bola sujeito_de sobrar	154
mais modifié ser	154
sobrar sujeito bola	154
ser modifier mais	153
cobrar objeto escanteio	152
escanteio objeto_de cobrar	152
ainda modifié ter	151
bola sujeito_de desviar	151
desviar sujeito bola	151
gaúcho modifié time	151
ter modifier ainda	151
time modifier gaúcho	151
canto modifier direito	149
direito modifié canto	149
ponto objeto_de ter	148
ter objeto ponto	148
afastar sujeito zaga	146
zaga sujeito_de afastar	146
mais modifié ter	143
oitava pp_de final	143
ter modifier mais	143
mineiro modifié time	141
time modifier mineiro	141
esquerdo modifié lado	138
lado modifier esquerdo	138
marcar modifier quase	137
min pp_de+o etapa	137
quase modifié marcar	137

entrar pp_em+o lugar	136
ficha modifier técnico	136
gol pp_de+o vitória	136
técnico modifié ficha	136
bater modifier forte	134
forte modifié bater	134
canto modifier esquerdo	133
esquerdo modifié canto	133
agosto pp_de 2013	128
dificuldade objeto_de ter	125
ter objeto dificuldade	125
tabela pp_de classificação	124
minuto pp_de+o etapa	123
oportunidade objeto_de ter	122
ter objeto oportunidade	122
bola pp_em+o área	121
quarta pp_de final	121
cartão modifier vermelho	120
chutar objeto cruzado	120
cruzado objeto_de chutar	120
lançamento objeto_de receber	120
ponto objeto_de somar	120
receber objeto lançamento	120
somar objeto ponto	120
vermelho modifié cartão	120
dar objeto passe	118
passe objeto_de dar	118
chegar pp_a+o gol	117
direito modifié lado	117
lado modifier direito	117
ter objeto trabalho	116
trabalho objeto_de ter	116
dar modifier não	115
haver modifier já	115
já modifié haver	115
não modifié dar	115
saída pp_de bola	114
bola objeto_de mandar	113
mandar objeto bola	113

carioca <small>modifié</small> time	112
time <small>modifieur</small> carioca	112
falta <small>objeto_de</small> fazer	111
fazer <small>objeto</small> falta	111
competição <small>modifieur</small> nacional	110
falta <small>objeto_de</small> sofrer	110
nacional <small>modifié</small> competição	110
sofrer <small>objeto</small> falta	110
criar <small>objeto</small> oportunidade	109
desviar <small>pp_de</small> cabeça	109
oportunidade <small>objeto_de</small> criar	109
chegar <small>pp_com</small> perigo	108
ficar <small>modifieur</small> mais	108
mais <small>modifié</small> ficar	108
campo <small>pp_de</small> defesa	107
jogo <small>sujeito_de</small> ser	107
já <small>modifié</small> ter	107
ser <small>sujeito</small> jogo	107
ter <small>modifieur</small> já	107
cartão <small>objeto_de</small> receber	105
receber <small>objeto</small> cartão	105
ser <small>modifieur</small> também	105
afastar <small>objeto</small> perigo	104
perigo <small>objeto_de</small> afastar	104
também <small>modifié</small> ser	104
agora <small>modifié</small> ser	103
ser <small>modifieur</small> agora	103
demorar <small>modifieur</small> não	102
entrar <small>pp_em+o</small> área	102
jogo <small>sujeito_de</small> começar	102
não <small>modifié</small> demorar	102
tarde <small>pp_de+este</small> domingo	101
arriscar <small>pp_de=fora=de+o</small> área	100
começar <small>sujeito</small> jogo	100
campo <small>pp_de</small> ataque	98
estar <small>modifieur</small> não	98
não <small>modifié</small> estar	98
paulista <small>modifié</small> time	98
time <small>modifieur</small> paulista	98

apito <small>modifieur</small> final	97
final <small>modifié</small> apito	97
final <small>modifié</small> número	96
gol <small>pp_de+o</small> partida	96
número <small>modifieur</small> final	96
receber <small>pp_em+o</small> área	96
arriscar <small>modifieur</small> de=longe	95
chegar <small>pp_a</small> ponto	95
de=longe <small>modifié</small> arriscar	95
aproveitar <small>objeto</small> cruzamento	94
bem <small>modifié</small> ser	94
cruzamento <small>objeto_de</small> aproveitar	94
posse <small>objeto_de</small> ter	94
ser <small>modifieur</small> bem	94
ter <small>objeto</small> posse	94
chance <small>objeto_de</small> desperdiçar	92
cruzar <small>pp_em+o</small> área	92
desperdiçar <small>objeto</small> chance	92
fazer <small>modifieur</small> não	92
fazer <small>objeto</small> jogo	92
jogo <small>objeto_de</small> fazer	92
não <small>modifié</small> fazer	92
bem <small>modifié</small> jogar	91
direito <small>modifié</small> ponta	91
ficar <small>pp_com</small> bola	91
jogar <small>modifieur</small> bem	91
mandar <small>pp_para</small> escanteio	91
noite <small>pp_de+este</small> quarta-feira	91
ponta <small>modifieur</small> direito	91
time <small>modifieur</small> tricolor	91
tricolor <small>modifié</small> time	91
errar <small>objeto</small> passe	90
gol <small>objeto_de</small> sofrer	90
passe <small>objeto_de</small> errar	90
sofrer <small>objeto</small> gol	90
troca <small>pp_de</small> passe	90
bater <small>objeto</small> cruzado	89
cruzado <small>objeto_de</small> bater	89
estar <small>modifieur</small> já	89

evitar objeto gol	89
gol objeto_de evitar	89
já modifié estar	89
bola modifier parado	88
cara pp_de+o gol	88
parado modifié bola	88
passe objeto_de trocar	88
trocar objeto passe	88
marcar sujeito árbitro	87
árbitro sujeito_de marcar	87
acertar objeto chute	86
arriscar objeto chute	86
chute objeto_de acertar	86
chute objeto_de arriscar	86
não modifié sair	86
sair modifier não	86
sair pp_em tiro	86
esquerdo modifié ponta	85
ponta modifier esquerdo	85
bola objeto_de roubar	84
chegar pp_a+o empate	84
chute pp_de=fora=de+o área	84
conseguir sujeito time	84
defesa pp_de+o goleiro	84
para=cima modifié partir	84
partir modifier para=cima	84
roubar objeto bola	84
time sujeito_de conseguir	84
acertar objeto trave	83
em=casa modifié jogar	83
equipe sujeito_de ter	83
jogar modifier em=casa	83
sair pp_em+o frente	83
ter sujeito equipe	83
trave objeto_de acertar	83
haver modifier não	82
jogo objeto_de ter	82
não modifié haver	82
ter objeto jogo	82

ir pp_a ponto	81
partida modifier válido	81
válido modifié partida	81
bola objeto_de receber	80
livre pp_em+o área	80
marcar modifier não	80
não modifié marcar	80
receber objeto bola	80
ainda modifié ser	79
catarinense modifié time	79
ser modifier ainda	79
só modifié ter	79
ter modifier só	79
time modifier catarinense	79
ampliar objeto vantagem	78
cruzamento objeto_de receber	78
receber objeto cruzamento	78
vantagem objeto_de ampliar	78
alvinegro modifié time	77
banco pp_de reserva	77
bola objeto_de perder	77
fazer sujeito goleiro	77
goleiro sujeito_de fazer	77
perder objeto bola	77
time modifier alvinegro	77
tocar pp_em+o saída	77
bola objeto_de colocar	76
colocar objeto bola	76
dificuldade objeto_de encontrar	76
encontrar objeto dificuldade	76
jogo modifier válido	76
válido modifié jogo	76
bola objeto_de tocar	75
tocar objeto bola	75
mais modifié subir	74
rodada pp_de+o competição	74
ser modifier já	74
subir modifier mais	74
chance modifier claro	73

claro <small>modifié</small> chance	73
fase <small>pp_de</small> grupo	73
inicial <small>modifié</small> minuto	73
minuto <small>modifié</small> inicial	73
receber <small>pp_em+o</small> entrada	73
time <small>sujeito_de</small> vencer	73
vencer <small>sujeito</small> time	73
gol <small>sujeito_de</small> ser	72
partida <small>sujeito_de</small> ser	72
ser <small>sujeito</small> gol	72
ser <small>sujeito</small> partida	72
ficar <small>sujeito</small> jogo	71
final <small>pp_de</small> semana	71
final <small>pp_de+o</small> partida	71
ir <small>modifié</small> não	71
jogo <small>objeto_de</small> empatar	71
jogo <small>sujeito_de</small> ficar	71
não <small>modifié</small> saber	71
resultado <small>sujeito_de</small> ser	71
saber <small>modifié</small> não	71
ser <small>sujeito</small> resultado	71
zona <small>pp_de</small> classificação	71
empatar <small>objeto</small> jogo	70
não <small>modifié</small> ir	70
bola <small>sujeito_de</small> ir	69
bomba <small>objeto_de</small> soltar	69
conquistar <small>objeto</small> vitória	69
encerrar <small>sujeito</small> árbitro	69
equipe <small>pp_de+o</small> casa	69
final <small>pp_de+o</small> tempo	69
ir <small>sujeito</small> bola	69
saída <small>pp_de+o</small> goleiro	69
ser <small>modifié</small> só	69
soltar <small>objeto</small> bomba	69
só <small>modifié</small> ser	69
vitória <small>objeto_de</small> conquistar	69
árbitro <small>sujeito_de</small> encerrar	69
a=menos <small>modifié</small> um	68
campo <small>objeto_de</small> deixar	68

cruzar <small>pp_de+o</small> esquerda	68
deixar <small>objeto</small> campo	68
marcar <small>objeto</small> pênalti	68
minuto <small>pp_de</small> acréscimo	68
pênalti <small>objeto_de</small> marcar	68
setembro <small>pp_de</small> 2013	68
um <small>modifié</small> a=menos	68
buscar <small>objeto</small> empate	67
começar <small>sujeito</small> partida	67
empate <small>objeto_de</small> buscar	67
estar <small>modifié</small> muito	67
final <small>pp_de+o</small> jogo	67
muito <small>modifié</small> estar	67
partida <small>sujeito_de</small> começar	67
bola <small>pp_em+o</small> trave	66
chegar <small>sujeito</small> equipe	66
dar <small>objeto</small> número	66
equipe <small>modifié</small> mineiro	66
equipe <small>sujeito_de</small> chegar	66
estar <small>sujeito</small> time	66
mineiro <small>modifié</small> equipe	66
número <small>objeto_de</small> dar	66
perigo <small>pp_a+o</small> gol	66
time <small>sujeito_de</small> estar	66
atuação <small>objeto_de</small> ter	65
carioca <small>modifié</small> equipe	65
chegar <small>sujeito</small> time	65
conseguir <small>objeto</small> vitória	65
equipe <small>modifié</small> carioca	65
estacionar <small>pp_em+o</small> ponto	65
final <small>pp_de+o</small> etapa	65
gol <small>pp_de</small> empate	65
ter <small>objeto</small> atuação	65
time <small>sujeito_de</small> chegar	65
vitória <small>objeto_de</small> conseguir	65
acertar <small>objeto</small> travessão	64
fazer <small>modifié</small> mais	64
jogada <small>pp_por+o</small> direita	64
mais <small>modifié</small> fazer	64

travessão objeto_de acertar	64
% pp_de aproveitamento	63
derrota e_ou empate	63
empate e_ou derrota	63
empate pp_sem gol	63
fim pp_de+o jogo	63
levantar pp_em+o área	63
campo pp_em+o sábado	62
efeito objeto_de surtir	62
esquerdo modifié perna	62
garantir objeto vaga	62
gol pp_de+o jogo	62
paranaense modifié time	62
perna modifier esquerdo	62
surtir objeto efeito	62
ter pp_por+o frente	62
time modifier paranaense	62
vaga objeto_de garantir	62
autor pp_de+o gol	61
começar objeto jogo	61
domingo modifier horário	61
equipe sujeito_de voltar	61
gol objeto_de perder	61
horário modifié domingo	61
jogo objeto_de começar	61
julho pp_de 2013	61
lance modifier seguinte	61
perder objeto gol	61
positivo modifié resultado	61
resultado modifier positivo	61
seguinte modifié lance	61
voltar sujeito equipe	61
assistência objeto_de dar	60
consecutivo modifié vitória	60
dar objeto assistência	60
noite pp_de+este sábado	60
ser sujeito time	60
tempo objeto_de ter	60
ter objeto tempo	60

time sujeito_de ser	60
vitória modifier consecutivo	60
adversário modifié goleiro	59
chance sujeito_de ser	59
cruzar pp_de+o direita	59
goleiro modifier adversário	59
ir modifier para=cima	59
para=cima modifié ir	59
placar pp_a+o min	59
sair pp_por+o linha	59
ser sujeito chance	59
abrir objeto marcador	58
chute modifier forte	58
compromisso sujeito_de ser	58
forte modifié chute	58
jogo objeto_de vencer	58
marcador objeto_de abrir	58
ser sujeito compromisso	58
vencer objeto jogo	58
alcançar modifier não	57
bomba objeto_de mandar	57
contra-ataque objeto_de puxar	57
fazer objeto partida	57
garantir objeto vitória	57
jogo sujeito_de seguir	57
mandar objeto bomba	57
não modifié alcançar	57
partida objeto_de fazer	57
puxar objeto contra-ataque	57
seguir sujeito jogo	57
vitória objeto_de garantir	57
cometer objeto falta	56
falta objeto_de cometer	56
fazer modifier quase	56
quase modifié fazer	56
bola sujeito_de explodir	55
chance objeto_de dar	55
conseguir modifier só	55
conseguir sujeito equipe	55

cruzar pp_para área	55
dar objeto chance	55
dar objeto lugar	55
equipe modifier paulista	55
equipe sujeito_de conseguir	55
explodir sujeito bola	55
grêmio sujeito_de ter	55
horário modifié quarta-feira	55
horário modifié sábado	55
jogo pp_de ida	55
lugar objeto_de dar	55
meio pp_de+o área	55
paulista modifié equipe	55
placar pp_a+o minuto	55
quarta-feira modifier horário	55
saldo pp_de gol	55
sábado modifier horário	55
só modifié conseguir	55
ter sujeito grêmio	55
time sujeito_de voltar	55

voltar sujeito time	55
bola sujeito_de ficar	54
cruzar pp_em+o medida	54
empatar objeto partida	54
empurrar pp_para rede	54
ficar sujeito bola	54
início pp_de+o etapa	54
jogar sujeito time	54
oportunidade pp_de gol	54
partida objeto_de empatar	54
time sujeito_de jogar	54
volta pp_de+o intervalo	54
baiano modifié time	53
como modifié ter	53
cruzamento pp_em+o área	53
defender sujeito goleiro	53
equipe sujeito_de vencer	53
goleiro sujeito_de defender	53
passé objeto_de fazer	53

ANEXO C – FRAMES DO FIELD – DICIONÁRIO DE EXPRESSÕES DO FUTEBOL

Frame	Definição
Ataque	Conjunto de ações ofensivas realizadas por uma das equipes, no intuito chegar à meta do time adversário.
Arquibancada	Uma das partes do estádio em que ocorre uma partida de futebol, onde ficam os torcedores dos jogos.
Banco de reservas	Designa o local onde os jogadores reservas de cada time permanecem durante o jogo, entrando na partida conforme a decisão do técnico.
Cabeceio	O ato de um jogador deslocar a bola com a cabeça para outro lugar no campo.
Campo	Campo é a área sobre a qual os times jogam futebol e acontecem todas as ações relacionadas à partida. O campo é dividido em linhas e regiões diversas, de acordo com as regras oficiais do jogo.
Chance	A chance que um jogador ou time tem a oportunidade de marcar um gol.
Chute	Evento em que um jogador desloca a bola com o pé para determinado local no campo.
Cobrança de Falta	Momento em que, após um jogador ou seu time serem penalizados por uma infração, um jogador do time adversário cobra essa penalização, podendo ser um tiro livre direto ou tiro livre indireto.

Cobrança de Lateral	O ato de um jogador de um dos times reiniciar a partida ao lançar a bola com as duas mãos da linha lateral para dentro do campo.
Cobrança de Pênalti	Um jogador do time adversário ao que cometeu falta chuta a bola em direção ao gol, tendo o goleiro como único obstáculo.
Comemorar Gol	Celebração de gol convertido pelo time que marca ponto no placar, juntamente com sua torcida.
Competição	Competição que dura em torno de 90 minutos, da qual um time termina vencedor. A partida é disputada por 11 jogadores em cada time, em um campo que possui duas metas, sendo que o objetivo é fazer com que a bola entre no gol adversário.
Contra-ataque	Tática efetuada pelo time defensor ao tomar a bola do adversário e partir em velocidade para tentar marcar um gol.
Defesa	Conjunto de ações defensivas realizadas por uma das equipes, no intuito de evitar o gol do time adversário.
Drible	Ato de enganar o adversário através de movimentos corporais, de modo a ultrapassá-lo com a bola.
Escanteio	O momento em que um jogador do time atacante chuta a bola de um dos cantos do campo para reiniciar a partida.
Equipamentos	Os elementos e objetos utilizados pelos jogadores em campo. Entre tais elementos e objetos, estão os equipamentos do jogo, como a bola, e o uniforme dos jogadores.

Estádio	Designa o local onde o jogo de futebol acontece, o qual é formado por arquibancadas, além do banco de reservas e do vestiário. O estádio geralmente pertence a um time.
Funções do Juiz	Designa as ações do juiz durante o jogo, as quais consistem em fazer a partida ocorrer conforme as regras oficiais. O juiz toma decisões pertinentes ao regulamento e estabelece os momentos da partida.
Gol Contra	O ato de marcar gol contra sua própria equipe.
Infração	Momento em que uma equipe, através de um ou mais jogadores, comete uma irregularidade marcada pelo árbitro, que concede vantagem ao time adversário.
Interceptação	Evento em que um jogador intercepta a passagem da bola, impedindo que o time adversário possa dar continuidade à jogada.
Marcação	Ato de um ou mais jogadores de um time vigiarem e procurarem controlar os movimentos de jogadores do time adversário.
Marcar Falta	O momento em que o árbitro, diante de uma infração, marca uma falta, penalizando o jogador que a cometeu.
Marcar Gol	Evento em que um jogador acerta a bola entre as traves e abaixo do travessão da meta do time adversário. O gol resulta em ponto para o time do jogador que marcou.
Participantes	As pessoas envolvidas na partida de futebol. Entre esses participantes, figuram os jogadores e a comissão técnica de cada time, a arbitragem e os torcedores.

Passe	Ato de um jogador transferir a posse da bola para um colega de equipe.
Perder Gol	O ato de um jogador falhar em acertar a bola dentro da meta do time adversário.
Placar	O resultado em números da partida de futebol.
Reações da Torcida	O modo como a torcida reage às ações realizadas na partida.
Rebote	O momento em que a bola, quando chutada na direção do gol, se choca com a trave ou com a defesa do time adversário e retorna.
Substituição	O momento em que o técnico, com autorização do árbitro, substitui um jogador em campo por um jogador do banco de reservas durante uma partida.
Táticas	As ações desempenhadas pelos jogadores dos times na defesa e no ataque.
Tempo	As ações do árbitro relacionadas à duração da partida, quando determina seu início ou final, intervalos, recuperação de tempo e prorrogações.
Tiro de Meta	O momento em que um jogador do time defensor chuta a bola de um ponto da área da sua meta para reiniciar a partida, após ter sido tocada por jogador do time adversário.