

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

CASSIANE OGLIARI

***FRAMES E MULTILINGUALIDADE: EQUIVALENTES DE TRADUÇÃO  
NO DICIONÁRIO BILÍNGUE DO FUTEBOL DA COPA DE 2014***

SÃO LEOPOLDO,

2013

Cassiane Ogliari

***FRAMES E MULTILINGUALIDADE:***

**Equivalentes de tradução no dicionário bilíngue do futebol da Copa de 2014**

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

Orientadora: Profa. Dra. Rove Luiza de Oliveira Chishman

SÃO LEOPOLDO,

2013

Ficha catalográfica

O35f Ogliari, Cassiane

*Frames* e multilinguagem : equivalentes de tradução no dicionário bilingue do futebol da Copa de 2014 / por Cassiane Ogliari. – 2013.

128 f. : il., 30cm.

Dissertação (mestrado) — Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Linguística Aplicada, 2013.

Orientação: Prof<sup>ª</sup>. Dr<sup>ª</sup>. Rove Luiza de Oliveira Chishman.

1. Semântica de frames. 2. Multilinguagem.  
3. Equivalência de tradução. 4. Futebol. 5. Dicionário da Copa.  
I. Título.

CDU 82.035

Catálogo na Fonte:  
Bibliotecária Vanessa Borges Nunes - CRB 10/1556

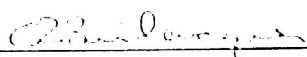
CASSIANE OGLIARI

FRAMES E MULTILINGUALIDADE: EQUIVALENTES DE TRADUÇÃO NO DICIONÁRIO  
BILÍNGUE DO FUTEBOL DA COPA DE 2014

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

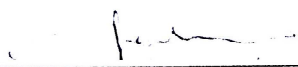
Aprovada em 13 de setembro de 2013

BANCA EXAMINADORA



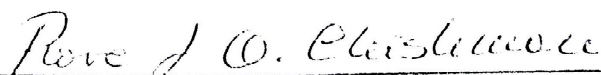
---

Profa. Dra. Cleci Regina Bevilacqua (UFRGS)



---

Profa. Dra. Silvia Matturro Panzardi Foschiera (UNISINOS)



---

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

## AGRADECIMENTOS

Neste momento não posso deixar de agradecer algumas pessoas que caminharam ao meu lado e que fazem parte, portanto, desta conquista. A conclusão desta pesquisa representa para mim muito mais do que a finalização de um trabalho acadêmico: ela representa a concretização de um sonho, a conclusão de uma etapa de crescimento profissional, a abertura de portas para novos desafios e também a certeza de que tudo é possível quando realmente queremos algo.

Nada disso seria possível se eu não tivesse contado com a ajuda de algumas pessoas que, de diferentes formas, me apoiaram, me inspiraram e me orientaram neste caminho.

Agradeço à minha orientadora, Prof. Dra. Rove Luiza de Oliveira Chishman, por ter me recebido em seu grupo de pesquisa, me apresentado a Semântica de Frames e pelo suporte prestado em todos os momentos desta caminhada.

Aos colegas do Grupo de Pesquisa Kicktionary\_Br, que têm desenvolvido um ótimo trabalho graças a esforços coletivos. Sinto-me muito orgulhosa de fazer parte deste time!

Aos professores do Programa de Pós-Graduação em Linguística Aplicada da Unisinos, exemplos de dedicação e modelos a serem seguidos, pela orientação e incentivo durante o curso.

Aos meus pais, Antônio e Lisete Ogliari, que me incentivaram e apoiaram em todas as minhas escolhas e que não mediram esforços para que eu chegasse até aqui. Sem vocês, nada disso seria possível.

Às minhas irmãs, Carina e Evandisse, meus cunhados, Luciano e Ricardo, e meus sobrinhos, Lucas e Beatrice, pelas palavras de motivação sempre e por acreditarem que eu seria capaz.

Ao “comitê técnico”, Antônio Ogliari e Lucas Bavaresco, pela ajuda com encenações e explicações sobre o até então pouco conhecido futebolês.

Às minhas amigas e colegas de trabalho, pelo apoio, companheirismo e paciência.

Muito obrigada por tudo!

*Futebol é simples: quem tem a bola, ataca. Quem não tem, se defende.*

*(Neném Prancha)*

## RESUMO

O objetivo desta pesquisa é refletir sobre os equivalentes de tradução do domínio do futebol no âmbito de um projeto multilíngue baseado nos preceitos da Semântica de *Frames* (Fillmore, 1982; 1985). Tal proposta fundamenta-se na hipótese de que, ao tratar-se de um domínio específico, como o do futebol, no qual o plano conceitual de dois idiomas seja comparável, as equivalências lexicais podem ser abordadas de forma a privilegiar a relação da unidade lexical com o *frame* ao qual ela pertence, que seria relacionável em ambos os idiomas. Para verificar a procedência desta hipótese, o trabalho está subdividido em três capítulos principais: 1) *Tradução, Língua e Cultura*, que aborda o tema das equivalências de tradução, dos diferentes tipos de equivalência e das estratégias para definição de equivalências; 2) *Semântica de Frames e Linguística Cognitiva*, que aborda os preceitos da Semântica de *Frames* e exemplifica seu emprego no desenvolvimento de bases de dados lexicais; e 3) *Frames e multilinguagem*, que aborda a forma como o conceito de *frames* pode influenciar em trabalhos em mais de um idioma e apresenta os Dicionários Kicktionary e Kicktionary\_Br Copa 2014 (projeto ao qual a presente pesquisa está vinculada). A parte aplicada da pesquisa prevê dois diferentes momentos: 1) Análise de uma amostra de 100 unidades lexicais em inglês e português com o intuito de verificar os tipos de equivalências predominantes no *corpus*; e 2) Reflexão específica sobre cada um dos três tipos de equivalência encontrados no *corpus* da pesquisa. Finaliza-se apresentando casos de equivalência exata de conteúdos, intersecção de conteúdos e de equivalência por superioridade e faz-se uma reflexão sobre as entradas lexicográficas destas equivalências no dicionário do futebol, Kicktionary\_Br Copa 2014.

**PALAVRAS-CHAVE:** Semântica de Frames; Multilinguagem; Equivalência de Tradução; Futebol; Dicionário da Copa.

## ABSTRACT

This research aims at reflecting about the translation equivalents on the soccer domain regarding to a multilingual project based on the fundamentals of Frame Semantics (Fillmore, 1982; 1985). This proposition underlies on the hypothesis that, when dealing with a specific domain, as soccer, in which the conceptual plan of both language is comparable, the lexical equivalences can be observed in a way to privilege the relation between the lexical unit and its *frame*, which would also be comparable in both languages. To be able to test this hypothesis, this study is subdivided in three chapters: 1) *Translation, Language and Culture*, that deals with translation equivalence, the different kinds of equivalence and the strategies to define equivalences; 2) *Frame Semantics and Cognitive Linguistics*, that talks about the principles of Frame Semantics and exemplifies their use when developing a lexical database; and 3) *Frames and Multilingualism*, that approaches the way in which the concept of *frame* can influence in multilingual studies and presents the dictionaries Kicktionary and Kicktionary\_Br 2014 World Cup (project which this dissertation is linked to). The applied part of this research mentions two moments: 1) The analysis of a sample composed by 100 lexical units in English and Portuguese with the objective of verifying the dominant kinds of equivalence with more frequency in our *corpus*; and 2) The specific reflection about each of the three kinds of equivalence found in *corpus*. In its last part, this study presents cases of exact equivalence of content, intersection of content and superiority equivalents and ponders about the lexical entries of such equivalences in the soccer dictionary, Kicktionary\_Br 2014 World Cup.

**KEYWORDS:** Frame Semantics; Multilingualism; Translation; Soccer; World Cup Dictionary.



## LISTA DE ILUSTRAÇÕES

FIGURA 1 - APRESENTAÇÃO DO FRAME <i>COOKING_CREATION</i> .....	44
FIGURA 2 - ELEMENTOS DE FRAME DO VERBO <i>COZINHAR (COOK)</i> – IMAGEM PARCIAL .....	44
FIGURA 3 - EXEMPLO DE RELAÇÃO ENTRE <i>FRAMES</i> NO FRAMEGRAPHER .....	45
FIGURA 4 - ABORDAGEM DE FUSÃO SEGUNDO LÖNNEKER-RODMAN (2007, p. 4).....	56
FIGURA 5 - ABORDAGEM DE EXPANSÃO SEGUNDO LÖNNEKER-RODMAN (2007, p. 5).....	56
FIGURA 6 - EXEMPLO DE <i>SYNSET</i> NO FRAME <i>AWARD_GOAL</i> .....	59
FIGURA 7 - IMAGEM PARCIAL DA CONSULTA POR CENAS NO KICKTIONARY .....	61
FIGURA 8 - APRESENTAÇÃO DO FRAME <i>REFEREE_DECISION</i> NO KICKTIONARY .....	62
FIGURA 9 - APRESENTAÇÃO DA UL <i>AWARD</i> NO KICKTIONARY .....	62
FIGURA 10 – REMODELAGEM DE <i>FRAMES</i> EM SUA PRIMEIRA VERSÃO .....	65
FIGURA 11 - EXEMPLO DE LISTA GERADA PELO CONCORD NA BUSCA PELA UL <i>GOLEIRO</i> .....	74
FIGURA 12 – EXEMPLO DE <i>CLUSTERS</i> , OBTIDO NA CONSULTA POR <i>GOLEIRO</i> NO CONCORD.....	75
FIGURA 13 - EXEMPLO DE BUSCA POR UNIDADES COMPLEXAS, COMO <i>DRIBL* O GOLEIRO</i> , NO CONCORD .....	76
FIGURA 14 - TELA DE FREQUÊNCIAS NO WORDLIST DO WORDSMITH TOOLS .....	79
FIGURA 15 - TIPOS DE EQUIVALÊNCIA NA AMOSTRA DE 100 ULS DO DOMÍNIO DO FUTEBOL .....	80
FIGURA 16 - RESULTADO DA BUSCA POR <i>BALL</i> NO CONCORD DO WORDSMITH TOOLS .....	85
FIGURE 17 - RESULTADO DOS <i>CLUSTERS</i> PARA <i>BALL</i> NO CONCORD DO WORDSMITH TOOLS.....	85
FIGURA 18 - RESULTADO DA CONSULTA POR <i>BOLA</i> NO CONCORD DO WORDSMITH TOOLS .....	86
FIGURE 19 - RESULTADO DOS <i>CLUSTERS</i> PARA <i>BOLA</i> NO CONCORD DO WORDSMITH TOOLS .....	87
FIGURA 20 - POSIÇÃO DO FRAME <i>JOGADORES</i> NA REMODELAGEM DE FRAMES PROPOSTA PELO KICKTIONARY_BR .....	90
FIGURA 21 - RESULTADO DA BUSCA POR <i>GOLEIRO</i> NO CONCORD DO WORDSMITH TOOLS.....	91
FIGURA 22 - RESULTADO DOS <i>CLUSTERS</i> PARA <i>GOLEIRO</i> NO CONCORD DO WORDSMITH TOOLS.....	92
FIGURA 23 - RESULTADO DA CONSULTA POR <i>KEEPER</i> NO CONCORD DO WORDSMITH TOOLS ...	93
FIGURA 24 - RESULTADO DOS <i>CLUSTERS</i> PARA <i>GOALKEEPER</i> NO CONCORD DO WORDSMITH TOOLS.....	94
FIGURA 25 - RESULTADO DOS <i>CLUSTERS</i> PARA <i>KEEPER</i> E 'KEEPER' NO CONCORD DO WORDSMITH TOOLS.....	95
FIGURA 26 - POSIÇÃO DO FRAME <i>TIRO</i> NA REMODELAGEM DE <i>FRAMES</i> PROPOSTA PELO KICKTIONARY_BR .....	99

FIGURA 27 - RESULTADO DA BUSCA POR <i>SHOT</i> NO CONCORD DO WORDSMITH TOOLS .....	101
FIGURA 28 - RESULTADO DA BUSCA POR <i>TIRO</i> NO CONCORD DO WORDSMITH TOOLS .....	102
FIGURA 29 - RESULTADO DA BUSCA POR <i>JOGADA</i> NO CONCORD DO WORDSMITH TOOLS .....	102
FIGURA 30 - RESULTADO DA BUSCA POR <i>LANCE</i> NO CONCORD DO WORDSMITH TOOLS .....	103
FIGURE 31 - POSIÇÃO DA UL <i>MARCAR FALTA</i> NA REMODELAGEM DE <i>FRAMES</i> PROPOSTA PELO KICKTIONARY_BR .....	106
FIGURA 32 - <i>CLUSTERS</i> RESULTANTES DA CONSULTA POR <i>FOUL*</i> NO CONCORD DO WORDSMITH TOOLS.....	108
FIGURA 33 - RESULTADO DA CONSULTA POR <i>BOOK* FOR A FOUL</i> NO CONCORD DO WORDSMITH TOOLS.....	109
FIGURA 34 - POSIÇÃO DO FRAME TIME NA REMODELAGEM DE FRAMES PROPOSTA PELO KICKTIONARY_BR .....	111
FIGURA 35- RESULTADO DA BUSCA POR <i>TEAM</i> NO CONCORD DO WORDSMITH TOOLS.....	113
FIGURA 36 - RESULTADO DA BUSCA POR <i>EQUIPE</i> NO CONCORD DO WORDSMITH TOOLS.....	114
FIGURA 37 - RESULTADO DOS <i>CLUSTERS</i> PARA <i>EQUIPE</i> NO CONCORD DO WORDSMITH TOOLS	114
FIGURA 38 - RESULTADO DA BUSCA POR <i>TIME</i> NO CONCORD DO WORDSMITH TOOLS.....	115
FIGURA 39 - RESULTADO DA CONSULTA POR <i>SELEÇÃO</i> NO CONCORD DO WORDSMITH TOOLS	115

## LISTA DE QUADROS

QUADRO 1 - RESULTADO DA BUSCA POR <i>BALL</i> EM DICIONÁRIOS ONLINE .....	84
QUADRO 2 - RESULTADO DA BUSCA POR <i>GOLEIRO</i> EM DICIONÁRIOS ONLINE .....	90
QUADRO 3- RESULTADO DA BUSCA POR <i>SHOT</i> EM DICIONÁRIOS ONLINE .....	100
QUADRO 4 - RESULTADO DA BUSCA POR <i>MARCAR</i> , <i>MARCAR FALTA</i> E <i>FALTA</i> EM DICIONÁRIOS ONLINE.....	107
QUADRO 5 - RESULTADO DA BUSCA POR <i>TEAM</i> EM DICIONÁRIOS ONLINE .....	112

## LISTA DE ENTRADAS LEXICOGRÁFICAS

ENTRADA LEXICOGRÁFICA 1 - BALL.....	88
ENTRADA LEXICOGRÁFICA 2 – BOLA.....	89
ENTRADA LEXICOGRÁFICA 3 – GOLEIRO .....	96
ENTRADA LEXICOGRÁFICA 4 - GOALKEEPER .....	97
ENTRADA LEXICOGRÁFICA 5 - KEEPER .....	97
ENTRADA LEXICOGRÁFICA 6 – 'KEEPER .....	98
ENTRADA LEXICOGRÁFICA 7 - SHOT .....	104
ENTRADA LEXICOGRÁFICA 8 - TIRO .....	105
ENTRADA LEXICOGRÁFICA 9 - MARCAR FALTA .....	110
ENTRADA LEXICOGRÁFICA 10 – BE BOOKED FOR A FOUL.....	110
ENTRADA LEXICOGRÁFICA 11- TEAM.....	117
ENTRADA LEXICOGRÁFICA 12 – SELEÇÃO .....	117
ENTRADA LEXICOGRÁFICA 13 - TIME .....	118
ENTRADA LEXICOGRÁFICA 14 - EQUIPE.....	118

## LISTA DE TABELAS

TABELA 1 - NÚMERO DE PALAVRAS POR TIPO DE TEXTO NO <i>CORPUS</i> EM INGLÊS .....	71
TABELA 2 - NÚMERO DE PALAVRAS POR TIPO DE TEXTO NO <i>CORPUS</i> EM PORTUGUÊS .....	73

## SUMÁRIO

<b>1 INTRODUÇÃO .....</b>	<b>15</b>
<b>2 PRESSUPOSTOS TEÓRICOS .....</b>	<b>19</b>
2.1 <i>TRADUÇÃO, LÍNGUA E CULTURA .....</i>	<i>20</i>
2.1.1 <b>Equivalentes de Tradução.....</b>	<b>23</b>
2.1.2 <b>Estratégias para definição de equivalências.....</b>	<b>27</b>
2.2 <i>SEMÂNTICA DE FRAMES E LINGUÍSTICA COGNITIVA.....</i>	<i>33</i>
2.2.1 <b>Linguística Cognitiva.....</b>	<b>33</b>
2.2.2 <b>A Semântica de <i>Frames</i> .....</b>	<b>37</b>
2.2.3 <b>A FrameNet .....</b>	<b>41</b>
2.3 <i>FRAMES E MULTILINGUALIDADE.....</i>	<i>48</i>
2.3.1 <b>Recursos lexicais multilíngues baseados em <i>frames</i> .....</b>	<b>50</b>
2.3.2 <b>Projeto Kicktionary .....</b>	<b>58</b>
2.3.3 <b>Projeto Kicktionary_Br Copa 2014.....</b>	<b>64</b>
<b>3 MATERIAIS E MÉTODOS .....</b>	<b>67</b>
3.1 <i>MATERIAIS.....</i>	<i>67</i>
3.1.1 <b><i>Corpora</i> da pesquisa.....</b>	<b>68</b>
3.1.1.1 <i>Corpus</i> em inglês .....	70
3.1.1.2 <i>Corpus</i> em português .....	72
3.1.2 <b>Ferramenta de análise do <i>corpus</i>: WordSmith Tools.....</b>	<b>73</b>
3.1.2.1 Ferramenta <i>Concord</i> do WordSmith Tools.....	74
3.2 <i>MÉTODO .....</i>	<i>77</i>
<b>4 ANÁLISE E DISCUSSÃO DOS DADOS .....</b>	<b>78</b>
4.1 <i>ANÁLISE AMPLA DOS FENÔMENOS DE EQUIVALÊNCIA NO DOMÍNIO DO FUTEBOL .....</i>	<i>78</i>
4.2 <i>ANÁLISE DE CASOS ESPECÍFICOS .....</i>	<i>82</i>
4.2.1 <b>Equivalência exata de conteúdos.....</b>	<b>82</b>
4.2.1.1 <i>Ball .....</i>	<i>83</i>

4.2.1.2 Goleiro .....	89
<b>4.2.2 Intersecção de conteúdo .....</b>	<b>99</b>
4.2.2.1 Shot .....	99
4.2.2.2 Falta.....	105
<b>4.2.3 Superioridade.....</b>	<b>111</b>
4.2.3.1 Team.....	111
<b>5 CONCLUSÃO .....</b>	<b>120</b>
<b>REFERÊNCIAS .....</b>	<b>122</b>
<b>ANEXO A - 50 PALAVRAS MAIS FREQUENTES NOS CORPORA EM INGLÊS E EM PORTUGUÊS.....</b>	<b>126</b>
<b>ANEXO B - ANÁLISE INICIAL DAS 50 PALAVRAS MAIS FREQUENTES NO CORPUS EM PORTUGUÊS.....</b>	<b>127</b>
<b>ANEXO C - ANÁLISE INICIAL DAS 50 PALAVRAS MAIS FREQUENTES NO CORPUS EM INGLÊS.....</b>	<b>128</b>

## 1 INTRODUÇÃO

Esta dissertação tem como objetivo refletir sobre os equivalentes de tradução do domínio do futebol no âmbito de um projeto multilíngue baseado nos preceitos da Semântica de *Frames* (Fillmore, 1982; 1985). Tal proposta fundamenta-se na hipótese de que, ao tratar-se de um domínio específico, como o do futebol, no qual o plano conceitual de dois idiomas seja comparável, as equivalências lexicais possam ser abordadas de uma forma que privilegie esta relação da unidade lexical com o *frame* ao qual ela pertence, sendo esta relacionável em ambos os idiomas da pesquisa (português e inglês).

Atualmente, ainda são poucas as pesquisas que abordam equivalências de tradução em recursos multilíngues baseados em *frame*. Porém, sabe-se que a Semântica de *Frames* tem sido discutida por lexicógrafos (Fontenelle, 2009) nos últimos anos, por seus preceitos terem se mostrado úteis no desenvolvimento de recursos multilíngues como dicionários (Schmidt, 2009) e bases de dados (FrameNet, 2012).

Portanto, o presente estudo espera contribuir, em termos teóricos, com a área de desenvolvimento de recursos lexicais multilíngues, assim como especificamente com os estudos em Semântica de *Frames* e multilinguagem. Além disso, em termos aplicados, acredita-se que, com o aumento na demanda por recursos lexicais multilíngues nos idiomas português/inglês devido à realização dos jogos da Copa do Mundo FIFA no Brasil em 2014, esta pesquisa possa auxiliar outros linguistas que enfrentam desafios impostos pela multilinguagem no desenvolvimento de recursos lexicais como dicionários.

Para tanto, esta pesquisa deve enfrentar os seguintes questionamentos: 1) considerando que os equivalentes de tradução assumem papel crucial em um dicionário multilíngue, como pensar nos equivalentes de tradução haja vista a sua inclusão em um recurso lexical multilíngue baseado nos preceitos da Semântica de *Frames*?; e 2) Quais são os tipos de equivalência mais comuns no domínio do futebol e como apresentar as entradas destes equivalentes de tradução em um recurso lexical baseado em *frames*?

Partindo-se destas questões norteadoras, o presente estudo aborda, portanto, a teoria de Semântica de *Frames* e sua interface com a Tradução. E para que estes questionamentos sejam respondidos, o trabalho traz inicialmente uma discussão a respeito da noção de *equivalência de tradução*, conceito central ao trabalho. Esta parte, representada no formato de pesquisa teórica, busca retomar contribuições dos estudos da Tradução e apresentá-las de forma a esclarecer os diferentes significados relacionados ao termo e a concepção defendida



pela presente pesquisa. Além disso, nesta primeira etapa, fala-se sobre as estratégias para definição de equivalências de tradução (Baker, 2006), o que serve como suporte para: 1) embasar a reflexão sobre possíveis equivalentes de tradução a serem disponibilizadas no Kicktionary\_Br Copa 2014<sup>1</sup>; e 2) verificar a hipótese de que, por tratar-se de um domínio específico que é comparável nas culturas dos idiomas em questão (português e inglês), os tipos de equivalências encontrados sejam: equivalência exata de conteúdos, interseção de conteúdos (Duval, 2008) e superioridade (Felber, 1984).

Após, apresenta-se a teoria da Semântica de *Frames* e exemplifica-se sua aplicação no desenvolvimento de bases de dados. Para isso, são trazidos exemplos de projetos online, com destaque ao projeto pioneiro na área, denominado FrameNet, desenvolvido pela equipe do professor Charles J. Fillmore<sup>2</sup>, no Instituto Internacional de Ciências da Computação (ICSI), em Berkeley (Califórnia, EUA). Trata-se de um banco de dados lexicais que pode ser lido por humanos e também por máquinas e que se baseou em textos reais para a realização de anotações semânticas e sintáticas (FrameNet, 2012). A partir deste, outros projetos passaram a ser desenvolvidos mundialmente para análise de diferentes idiomas, como a FrameNet para o chinês, alemão, espanhol, japonês, francês ou português (Boas, 2009). Mas não foram apenas FrameNets em outros idiomas que surgiram, pois outros projetos de bancos de dados em domínios específicos<sup>3</sup> também foram criados, como, por exemplo, nas áreas do futebol (Kicktionary), da biomedicina (BioOntoFN) e do turismo (parte do projeto Copa 2014, em andamento).

Tendo-se compreendido os temas abordados previamente, parte-se então para uma reflexão especificamente a respeito de equivalentes de tradução no âmbito das FrameNets multilíngues. Espera-se, com isso, verificar o que já foi dito sobre o tema e o que tem sido feito para se definir e indicar equivalências em trabalhos que levam em consideração a noção

---

<sup>1</sup> O Kicktionary\_Br Copa 2014 é um projeto acadêmico que faz parte da linha de pesquisa *Texto, Léxico e Tecnologia*, do Programa de Pós-Graduação em Linguística Aplicada da UNISINOS (Universidade do Vale do Rio dos Sinos). Trata-se de um projeto que busca desenvolver o dicionário bilíngue da Copa 2014, nos idiomas inglês e português, assim como, de uma forma mais ampla, colaborar com os estudos na área da criação de recursos lexicais multilíngues baseados em *frames*. Ele é desenvolvido pelo grupo de pesquisa SemanTec, Kicktionary\_Br Copa 2014, liderado pela Profa. Rove Chishman, em parceria com o grupo da FrameNet Brasil Brasil, liderado pela Profa. Margarida Salomão e pelo Prof. Tiago Torrent, na UFJF (Universidade Federal de Juiz de Fora).

<sup>2</sup> Charles J. Fillmore iniciou suas atividades no Departamento de Linguística na Universidade da Califórnia, em Berkeley (EUA), em 1971. Suas pesquisas trataram principalmente de questões de sintaxe e semântica lexical, com ênfase na relação entre as propriedades da forma linguística com seu significado e uso. Atualmente, Fillmore é diretor do Projeto FrameNet, desenvolvido no ICSI, onde desenvolveu a teoria da Semântica de *Frame*. Disponível em: <<http://www.icsi.berkeley.edu/icsi/gazette/2010/09/charles-fillmore>>. Acesso em 17 de março de 2013.

<sup>3</sup> Estes projetos estão descritos no site da FrameNet (2013), onde o link para cada um deles é disponibilizado. Disponível em: <[https://framenet.icsi.berkeley.edu/fndrupal/related\\_projects](https://framenet.icsi.berkeley.edu/fndrupal/related_projects)> Acesso em 17 mar. 2013.

de *frames*. No caso, como a presente dissertação aborda a linguagem utilizada no domínio do futebol, o projeto Kicktionary é, então, detalhado. Trata-se de uma ferramenta desenvolvida pelo pesquisador Thomas Schmidt<sup>4</sup>, disponibilizada na forma de um dicionário eletrônico do futebol, nos idiomas inglês, alemão e francês, que segue os preceitos da Semântica de *Frames*.

Um recurso lexicográfico baseado em *frames*, além de apresentar a descrição de uma palavra ou uma expressão propriamente dita, também indica a relação entre a unidade lexical e o *frame* ao qual ela pertence. Além disso, disponibiliza também sentenças de uso real nas quais as palavras são utilizadas e apresentam a etiquetagem dos elementos, ou participantes, deste *frame*. Isso justifica a razão pela qual as pesquisas neste campo têm crescido significativamente nos últimos anos e a sua aplicabilidade na criação de recursos lexicais tem se mostrado promissora.

Como exemplo disso fala-se dos projetos FrameNet e Kicktionary, que motivaram o grupo de pesquisa SemanTec (Semântica e Tecnologia) da Universidade do Vale do Rio dos Sinos a investigar como a noção de *frame* podem ser aplicada ao domínio do futebol no caso do português brasileiro. Assim, a primeira parte de pesquisa, desenvolvida pelo grupo como um todo, compreendeu a construção de corpus com anotação semântica, na modalidade texto corrido, para o português brasileiro (projeto Kicktionary\_Br) e, atualmente, trabalha-se em sua segunda fase, relacionada ao desenvolvimento de um recurso lexical bilíngue, nos idiomas português e inglês, para disponibilização no evento da Copa do Mundo FIFA 2014: o projeto Kicktionary\_Br Copa 2014. Espera-se que este recurso lexical colabore com a indicação de equivalentes linguísticos no par português-inglês, além de fornecer informações importantes sobre as unidades lexicais disponíveis para consulta. Para tanto, investigações sobre equivalências no âmbito de um projeto multilíngue baseado em *frames* tornam-se necessárias como uma das etapas da pesquisa a ser realizada, foco do presente estudo.

A estrutura da pesquisa teórica desta dissertação está subdividida em três capítulos principais. O capítulo sobre *Tradução, Língua e Cultura* parte do conceito de *tradução*, para refletir sobre a forma como a equivalência de tradução é compreendida em trabalhos da área. Nele, discorre-se também sobre os tipos de equivalência e as estratégias para definição de

---

<sup>4</sup> Linguista alemão que iniciou seus estudos na área da ciência da computação, Thomas Schmidt possui mestrado em Linguística Geral, estudou a relação entre Linguística e Inteligência Artificial, e possui doutorado em Linguística. Visitou o ICSI como pesquisador visitante em seu pós-doutorado, quando desenvolveu o projeto Kicktionary. Disponível em: < <http://www.kicktionary.de/cv.html> > Acesso em 10 de maio de 2013.

equivalências, para que seja possível utilizar este arcabouço na parte aplicada da pesquisa, que prevê a reflexão sobre as ULs e suas equivalências no par português e inglês no Kicktionary\_Br Copa 2014. Em seguida, o capítulo sobre *Semântica de Frames e Linguística Cognitiva* é apresentado, com o intuito de situar tal teoria nos estudos linguísticos cognitivistas. Para exemplificar como os conceitos da Semântica de *Frames* são empregados no desenvolvimento de bases de dados lexicais, o projeto FrameNet é, então, detalhado. Para concluir, o capítulo sobre *Frames e multilinguagem* aborda a forma como o conceito de *frames* pode influenciar em trabalhos em mais de um idioma. Para exemplificar isso, apresenta-se o Dicionário Kicktionary e, posteriormente, fala-se sobre o Kicktionary\_Br Copa 2014, projeto ao qual a presente pesquisa está relacionada.

Com isso, segue-se para a parte aplicada da pesquisa, que começa a ser descrita no capítulo seguinte: *materiais e métodos*. Tal segmento traz detalhes sobre o processo realizado pela pesquisa, além de disponibilizar informações sobre o *corpus* utilizado. Após esta seção, apresenta-se a *análise e discussão dos dados*, com uma reflexão sobre as equivalências no âmbito de um projeto multilíngue baseado em *frames* no domínio do futebol.

## 2 PRESSUPOSTOS TEÓRICOS

Como a proposta desta dissertação é refletir sobre a abordagem dos equivalentes de tradução no âmbito de um projeto multilíngue baseado em *frames*, procura-se, nesta etapa de investigação teórica, verificar de que forma os estudos nas áreas de Tradução e Semântica de *Frames* podem colaborar com esta pesquisa. Este trabalho será guiado pelas perguntas de pesquisa previamente mencionadas e que se preocupam em pensar sobre os equivalentes de tradução haja vista a sua inclusão em um recurso lexical multilíngue baseado em *frames* e sobre a validade deste tipo de trabalho.

Para isso, este capítulo do trabalho está dividido em três subcapítulos. O primeiro deles (subcapítulo 2.1), denominado *Tradução, Língua e Cultura*, busca conceituar o termo tradução e refletir sobre as equivalências de tradução, quais são seus tipos e como elas podem ser estabelecidas. Por tratar-se, nesta pesquisa, de um domínio específico, que é o do futebol, acredita-se que as equivalências restrinjam-se a equivalências exatas de conteúdo, intersecção de conteúdos e que haja, também, relações de superioridade na comparação das unidades lexicais em inglês e português. Esta discussão é apresentada na reflexão sobre estratégias de tradução.

Em seguida (subcapítulo 2.2) fala-se sobre a *Semântica de Frames e Linguística Cognitiva* e reflete-se sobre a forma como a noção de *frames* surgiu e o que se propõe a partir deste conceito. Com isso espera-se refletir sobre a base linguística que sustenta a presente dissertação. Também, neste subcapítulo, a FrameNet é apresentada, com o intuito de verificar como a noção de *frames* já está sendo empregada em projetos de criação de bases de dados baseadas em *frames*.

Por fim (subcapítulo 2.3), fala-se especificamente sobre *Frames e multilinguagem*. Neste subcapítulo, o projeto pioneiro desenvolvido na área, no domínio do futebol, é apresentado em detalhes: o Dicionário Kicktionary. Após, discorre-se então sobre o estudo nacional que prevê o desenvolvimento de um recurso lexical bilíngue baseado em *frames*, na área do futebol: o Kicktionary\_Br Copa 2014 (projeto ao qual a presente pesquisa está associada).

## 2.1 TRADUÇÃO, LÍNGUA E CULTURA

Como esta pesquisa discute os equivalentes de tradução no âmbito de FrameNets multilíngues, esta seção inicia discutindo-se o que os teóricos falam sobre o ato de traduzir, relacionado às questões culturais e linguísticas que envolvem a tarefa. Após, a subseção divide-se em mais dois tópicos específicos: 1) equivalentes de tradução (subseção na qual o conceito de equivalência é detalhado e os tipos de equivalência são descritos); e 2) estratégias para definição de equivalências. Espera-se, portanto, nesta etapa, esclarecer quais são as forças que direcionam o estabelecimento de equivalências e refletir sobre a forma como possíveis equivalências são definidas.

Para iniciar, salienta-se que, ao se falar em *tradução*, deve-se considerar que este termo pode remeter, segundo Jakobson (1991, p. 64-65), a três diferentes processos, considerando-se que, para ele, “o significado de um signo linguístico não é mais do que sua tradução por outro signo que lhe pode ser substituído”. Em outras palavras, a tradução existe de diferentes formas, envolvendo ou não a transferência de signos de uma língua para outra. Para o linguista, a tradução pode ser intralingual, interlingual ou intersemiótica, cujas caracterizações seguem abaixo:

- a) Tradução intralingual (ou *reformulação*): neste caso, a tradução acontece dentro da própria língua e, por este motivo, também é chamada de reformulação. No caso da palavra *mãe*, por exemplo, para atribuir-lhe um significado, pode-se pensar na utilização de outras palavras sinônimas como *genitora*. Porém, o uso de ambas as palavras (*mãe* e *genitora*) não é completamente equivalente, o que direciona a uma discussão sobre a não existência de equivalência completa, segundo o autor.
- b) Tradução interlingual (ou *tradução propriamente dita*): este é o caso da tradução partindo-se de uma língua de partida (língua fonte) para uma língua de chegada (língua alvo). Para Jakobson, este processo é “uma forma de discurso indireto: o tradutor recodifica e transmite uma mensagem recebida de outra fonte” (JAKOBSON, 1991, p. 65). Desta forma, deve existir uma equivalência nas mensagens na língua fonte e língua alvo, mesmo que isso seja apresentado através da utilização de códigos diferentes.
- c) Tradução intersemiótica (ou *transmutação*): referente à interpretação dos signos verbais através do uso de signos não-verbais, como da arte verbal para a música ou para a dança, por exemplo.

Nesta pesquisa, ao discorrer-se sobre tradução, discute-se exclusivamente a tradução interlingual, pois se fala dos equivalentes no domínio do futebol nos idiomas português e inglês. Porém, mais do que falar sobre as línguas fonte e alvo, analisa-se também, neste trabalho, a interlíngua que existe em trabalhos que contam com a etiquetagem semântica na comparação de estruturas e busca de equivalentes. Esta discussão é aprofundada no capítulo que trata sobre multilinguagem.

Hatim e Munday (2004) explicam que o termo *tradução* é polissêmico e esta afirmação tem como base o fato de que, conforme descrições apresentadas em dicionários, a tradução denomina o processo realizado pelo tradutor, ou por uma máquina, em transformar um texto fonte em um texto alvo na outra língua, e também o produto desta atividade do tradutor ou da máquina. Os autores optam, então, por falar no “âmbito da tradução” como

1. O processo de transferir um texto escrito de uma língua fonte para uma língua alvo, conduzido por um tradutor, ou tradutores, em um contexto sociocultural específico.
2. O produto escrito, ou texto alvo, que resulta de tal processo e que funciona no contexto sociocultural da língua alvo.
3. Os fenômenos cognitivos, linguísticos, visuais, culturais e ideológicos que são parte integral dos itens 1 e 2. (HATIM, MUNDAY, 2004, p. 6)

Com a aproximação da Copa do Mundo FIFA em 2014, no Brasil, os olhos do mundo estarão voltados para este país. Mas para que todos possam acompanhar suas seleções durante os jogos, um grande número de profissionais tradutores, intérpretes e jornalistas estará empenhado em garantir que as informações sobre os jogos sejam constantemente atualizadas e divulgadas nos meios de comunicação em diferentes idiomas. Enquanto as imagens dos jogos representarão de uma forma próxima a realidade e, de certa forma, imparcial, tudo que ocorre nos campos, as notícias e as manchetes utilizarão narrativas para relatar os eventos acontecidos. Estas narrativas, dependendo de variáveis como locutor, jornalista, local de publicação, público alvo, entre outras, serão representadas de diferentes formas. Porém, mais do que isso, dependendo também da cultura do narrador, outras variáveis como níveis de (in)formalidade, utilização de gírias, traços de oralidade, entre outros aspectos, também serão diferenciados. Como exemplo disso, tem-se a popular linguagem informal utilizada por brasileiros em narrações de jogos de futebol. Tanto em relatos falados como escritos, é comum notar-se traços de informalidade e de oralidade nos discursos futebolísticos na mídia brasileira, algo que não é tão comum em outros países. Além disso, outra situação do domínio do futebol é quando as línguas apresentam diferentes formas de falar sobre algo, como no

caso da expressão em português “fazer o gol **onde dorme a coruja**”, que não seria utilizada com tanta naturalidade se fosse traduzida literalmente para o inglês, pois as expressões “*score an upper-V goal*” ou “*score an upper 90 goal*” são mais usuais para descrever este tipo de gol. Casos como estes acabam sendo um desafio para tradutores e intérpretes, que têm a tarefa de traduzir textos de uma língua para outra e enfrentar esta barreira cultural.

Sapir (1949) fala sobre os desafios impostos na tradução devidos a questões culturais e destaca que

não existem duas línguas que sejam suficientemente similares para que sejam consideradas representantes de uma mesma realidade social. Os mundos nos quais as diferentes sociedades vivem são mundos distintos, não se trata simplesmente do mesmo mundo com etiquetas diferentes. (SAPIR, 1949, p. 69, *tradução nossa*<sup>5</sup>)

Além desta questão, determinados conceitos utilizados em uma língua podem não existir em outra. Um exemplo é o termo em inglês *brunch*, que, embora seja compreendido por alguns brasileiros graças à globalização, não existe como parte da cultura do país. Isso acontece porque também existem questões culturais que envolvem a tradução. Ao falar sobre isso, Baker e Saldanha (2009, p. 67) explicam que a língua e a cultura influenciam e filtram nossas experiências. Por este motivo, “torna-se difícil compreender e transmitir experiências que acontecem em diferentes sistemas de filtros, fora de nosso próprio sistema de referência”. Desta forma, a tradução cultural é compreendida pelos autores como uma perspectiva na qual os impactos neste movimento de saída de uma língua e chegada em outra são levados em consideração.

Neste sentido, surge uma hipótese diretamente aplicada à presente pesquisa: por se tratar, neste estudo, da linguagem específica do domínio do futebol, espera-se que os sistemas de referência sejam de certa forma comparáveis entre os idiomas inglês e português. Isso porque o *script*<sup>6</sup> de um jogo de futebol existe internacionalmente. Nele os participantes do evento são os times e seus integrantes, juízes e torcidas e pode-se dizer, resumidamente, que o evento é representado da seguinte forma: onze jogadores de dois times entram em um campo de futebol. Cada um dos times faz passes de bola, sem poder tocar nela com os membros superiores, com o intuito de chegar ao gol adversário. Sempre que a bola entrar no

---

<sup>5</sup> Todas as traduções de trechos de obras em inglês neste trabalho foram realizadas pela pesquisadora.

<sup>6</sup> Schank e Abelson (1977) falam sobre *script* e o definem como “uma estrutura que descreve uma sequência apropriada de eventos em um contexto particular”, na qual o todo é interconectado e cada parte pode influenciar a sequência seguinte de eventos. Assim, o *script* representa uma sequência pré-determinada e estereotipada de ações que definem uma situação socialmente conhecida. Graças a ele, é possível que as pessoas compreendam quem participa da situação sem que elas sejam ao menos mencionadas, por exemplo.

gol adversário marca-se um gol. Como os jogos da Copa do Mundo de futebol seguem este mesmo *script*, independente do país no qual ocorrem, pode-se dizer que, embora as formas de se narrar um jogo e o sistema linguístico utilizado sejam distintos, o *script* continua sendo comparável. Portanto, acredita-se que seja possível beneficiar-se desta estruturação pré-estabelecida, no momento de se localizarem equivalentes de tradução, já que se nota a existência de uma equivalência conceitual no domínio em questão, mesmo que esta equivalência não ocorra também no sistema lexical das línguas envolvidas. E, mais do que isso, assim como a noção de *script*, acredita-se que também a noção de *frames*<sup>7</sup> (a ser detalhada no subcapítulo 2.2) possa auxiliar na prestação de unidades lexicais utilizadas em momentos prototípicos de um jogo de futebol. Isso porque, ao valer-se da noção de *frames*, um pareamento entre sistemas linguísticos diferentes pode surgir, imaginando-se uma situação comum existente nos sistemas conceituais dos diferentes idiomas.

Mas como a equivalência é estabelecida após considerar-se a relação entre os sistemas conceituais de dois idiomas? E, afinal, pode-se falar em *equivalências de tradução*, sabendo-se que as línguas não apresentam simetria perfeita de conceitos e significados? Para esclarecer estas questões, a próxima seção aborda este conceito, situando-o nos estudos de tradução. Com isso, será possível, posteriormente, discutir-se a equivalência inserida especificamente no âmbito das FrameNets multilíngues.

### 2.1.1 Equivalentes de Tradução

Esta seção tem como objetivo apresentar a definição de *equivalentes de tradução* e situar este termo, central à presente pesquisa, nos Estudos de Tradução<sup>8</sup>.

Inicialmente, deve-se mencionar que um dos motivos que torna a equivalência um tema muito discutido em trabalhos da área da tradução se deve à forma como sua concepção mudou ao longo do tempo. Outro motivo é a própria polissemia do termo. De acordo com Alves, Magalhães e Pagano (2000), a equivalência de tradução é vista por alguns teóricos em analogia ao conceito de equivalência matemática, ou seja, relacionado à simetria entre as

<sup>7</sup> Neste primeiro momento, consideram-se similares as noções de *script* e *frame*. No próximo subcapítulo, através da reflexão sobre *frames*, as diferenças na essência destes termos passa a ser considerada.

<sup>8</sup> Segundo Munday (2009) o termo *Estudos da Tradução (Translation Studies)*, proposto por James S. Holmes, começou a ser empregado a partir da segunda metade do século XX, quando a tradução passou a ser vista como área de pesquisa emergente de outras áreas como: línguas modernas, literatura comparada e linguística.



duas línguas, enquanto que, para outros, refere-se apenas à correspondência possível entre termos de diferentes idiomas. E, por último, por ser a equivalência um tema central, ou uma “questão chave” (HATIM; MUNDAY, 2004, p. 7) nos Estudos de Tradução.

Oliveira (2007) faz uma interessante retomada sobre a mudança da concepção de *equivalente* e afirma que, em linhas gerais, as distintas formas de compreender o conceito de equivalência partem de três linhas teóricas: 1) a abordagem de orientação linguística, 2) a abordagem histórico-descritiva, e 3) a corrente desconstrucionista. A primeira corrente, também conhecida como abordagem científica, vê na equivalência a busca por um significado estável e que pode ser compartilhado entre as línguas de chegada e partida. Assim, os equivalentes têm o papel de “funcionar” do mesmo modo em ambas as línguas e pressupõe-se que a leitura de um texto seja única, ou seja, tenha apenas uma interpretação. Por esse motivo, a abordagem defende que “os significados estão todos no texto e o papel do tradutor é reiterá-los de forma honesta” (OLIVEIRA, 2007, p. 101-102). Catford (1980, p. 56) defende esta concepção e afirma que a tradução é a substituição de material de uma língua para a outra e que, em geral, “ocorre equivalência de tradução quando textos ou itens da LF (língua fonte) e da LM (língua meta) podem relacionar-se com os mesmos traços de substância (ou ao menos com alguns deles)”.

Com o tempo, esta abordagem passa a ser criticada por uma corrente que leva em consideração também as variáveis culturais e pessoais. Tal abordagem chama-se histórico-descritiva e representa uma nova fase na visão do conceito de equivalência. Para os autores que defendem esta perspectiva, como André Lefevere, James Holmes, José Lambert, Theo Hermans, Susan Bassnett, Gideon Toury e Itamar Even-Zohar, a tradução consiste em uma forma de reescrita do texto original e, segundo Oliveira (2007), as escolhas do tradutor passam a ser analisadas. A partir deste período, a noção de equivalência torna-se mais discutida, pois acredita-se que cada tradutor pode interpretar o texto de uma forma distinta, o que levaria a inúmeras possibilidades de textos finais na língua alvo.

A linha teórica desconstrutivista, ligada ao pós-estruturalismo, surge, então, para defender a posição do tradutor como transformador e responsável em comunicar os sentidos do texto original em outro idioma. Neste contexto, Oliveira (2007) relata que a equivalência simétrica deixa de existir, por tornar-se algo impossível de ser realizado, uma vez que não existe simetria entre as línguas. O que passa a ser buscado é “a transformação de um texto em outro, ambos com características próprias e ligados por cadeias de significação que se suplementam” (OLIVEIRA, 2007, p. 108), pois as línguas, independentemente de quão próximas possam ser, não possuem formas exatamente iguais de representar o mundo. Neste

contexto, as cadeias de significação funcionam como plano de fundo para que o texto fonte e o texto alvo relacionem-se de forma que o segundo possa ser chamado de tradução.

A equivalência, da forma que foi explicitada acima, está vinculada ao todo do processo tradutório, mas este termo também é amplamente utilizado por profissionais e estudantes que, mesmo que compartilhem de uma visão desconstrutivista de tradução, veem a equivalência como uma possível relação entre termos ou expressões em diferentes línguas. Oliveira (2007) fala sobre isso e afirma que esta ilusão da existência de equivalências, como relação entre as línguas, conforta. Mais do que isso, ela mostra como, no fundo, tantos outros nomes (como: simetria, adequação e aceitabilidade) levam todos a uma mesma discussão relacionada à prática da tradução.

Por ser um termo amplamente utilizado entre os tradutores, Baker (2006) defende sua utilização e justifica que seu uso está condicionado à consciência que deve existir por parte dos tradutores de que, quando se pensa em tradução, a equivalência é relativa, ou seja, ela depende de fatores linguísticos e culturais. Neste trabalho, assume-se a proposição de Baker e, portanto, sempre que se trata sobre equivalência, fala-se sobre uma possível tradução de determinado termo ou expressão que tem um sentido aproximado do significado existente na língua fonte, inserido em seu contexto.

Ressalta-se que, em seu livro *In other words*, Baker (2006, p. 5) fala sobre quatro tipos de equivalência, que são: 1) equivalência a nível lexical; 2) equivalência acima do nível lexical; 3) equivalência gramatical; e 4) equivalência textual. Esta diferenciação é tomada simplesmente porque a autora opta por iniciar tratando da equivalência em seu estado mais simples, ou seja, no formato de significados de palavras simples ou expressões e, ao longo do livro, variáveis como combinações de palavras e frases, categorias gramaticais, contexto cultural, escritor e leitor também são considerados.

É neste nível que se acredita estar a valia da noção de *frames* como sistema cognitivo capaz de ativar um conjunto de categorias relacionadas a um contexto motivador (Fillmore, 2006, p. 381). Esta proposição será aprofundada nos próximos subcapítulos, ao analisar-se a equivalência no âmbito das FrameNets multilíngues.

Ao falar sobre equivalência, Hatim e Munday (2004) também falam que ela existe de uma forma relativa. Por isso, segundo os autores, as decisões sobre uma ou outra equivalência de tradução mais adequada é crucial em qualquer discussão sobre equivalência e sobre o processo tradutório como um todo. Esta proposição nos aproxima do proposto por Neuber (*apud* Bassnett, 2003, p. 56-57), que vê a equivalência composta por componentes sintáticos, semânticos e pragmáticos, ou seja, ela “resulta da relação entre os próprios signos,

da relação entre os signos e aquilo que eles representam e da relação entre os signos, o que eles representam e seus utilizadores”. Tal noção não busca, portanto, uma simetria exata nas duas línguas, mas uma possibilidade de relação entre as palavras e estruturas inseridas em enunciados em diferentes idiomas, com funções similares.

Neste contexto, Duval (2008, p. 274) aborda os diferentes tipos possíveis de equivalências, descritos abaixo, exemplificados por nós no par português (PT) e inglês (EN):

- a) Quando o significante de uma língua aponta diretamente a um mesmo item na língua alvo e seu significado é um item lexical reconhecido no léxico de ambas as línguas. Exemplo: PT = computador / EN = *computer*. Felber (1984) denomina este tipo de situação como *equivalência exata de conteúdos*.
- b) Quando existe um item presente no léxico de ambas as línguas que não remete a um mesmo universo cultural. Exemplo: PT = Dia das Bruxas / EN = *Halloween*. Neste caso, a existência de traços semânticos em comum entre as línguas é que torna a equivalência possível. Felber (1984) denomina estes casos de *intersecção de conteúdos*.
- c) Quando o significante refere a algo real apenas na língua fonte. EN = *brunch*. Neste caso, Felber (1984) sugere a definição de *não equivalência de conteúdos*.

Além dos tipos mencionados acima, Felber (1984) menciona em seu trabalho que existe outro tipo de situação denominada *superioridade*. Neste caso, uma palavra que é utilizada na língua fonte tem uma extensão maior na língua alvo, ou vice-versa. É o caso de um dos verbos polissêmicos no futebol: o verbo *marcar*. Em português este verbo contempla ações como: um jogador que *marca* um adversário, um juiz que *marca* uma penalidade ou mesmo um jogador que *marca* um gol. Já em inglês, cada uma destas sentenças seria traduzida com o emprego de um verbo diferente, pois neste idioma não existe um verbo único que contemple todas estas acepções.

Esta situação exemplifica o que Piotrowski (1994, p. 112) chama de falta de estabilidade nas unidades de tradução. Segundo o linguista, nenhuma unidade de tradução é sempre estável, pois, dependendo do contexto no qual está inserida, ela pode assumir diferentes facetas e possuir, portanto, diferentes equivalências. Por este motivo retoma-se a importância da disponibilização de todas as opções possíveis de equivalência em um recurso lexical, inseridas em um contexto, para que o consulente seja capaz de precisar qual o uso adequado para a situação em questão.

Acredita-se que, por tratar-se, neste trabalho, de um domínio específico, que é o futebol, as equivalências restrinjam-se aos tipos “a” e “b” (equivalência exata de conteúdo e

intersecção de conteúdo) e que também haja relações de superioridade. Porém, somente através de uma pesquisa e da análise de um *corpus* é que esta hipótese pode ser corroborada.

Para concluir esta seção sobre equivalentes de tradução, deve-se considerar que não existe uma fórmula única a ser seguida como modelo para resolver todos os desafios enfrentados pelo tradutor na busca por equivalências. Isso é levado em consideração na presente pesquisa, que pretende apresentar, no subcapítulo a seguir, algumas estratégias propostas por estudiosos na área da tradução que têm se mostrado úteis para auxiliar os tradutores a determinarem o equivalente adequado em um contexto determinado. Espera-se, com isso, refletir sobre as possíveis estratégias utilizadas na parte aplicada da pesquisa, que examina os tipos de equivalências encontradas no Kicktionary\_Br Copa 2014 nos idiomas português e inglês e analisa as estratégias de definição de equivalências (Baker, 2006), detalhadas a seguir.

### 2.1.2 Estratégias para definição de equivalências

O processo de traduzir envolve uma preocupação constante por parte do profissional em tomar decisões sobre as equivalências que irá considerar no trabalho. Isso acontece porque, assim como citam Hatim e Munday,

ao fazer uma tradução, sempre haverá um “problema”, e um número de “soluções” possíveis. Em cada etapa do processo de tradução, escolhas são feitas, e estas, obviamente, influenciam escolhas subsequentes. (...) Todavia, como tudo que relaciona texto em contexto, decisões de tradução são raramente, se é que podem ser alguma vez, diretas e “sequenciais”. Elas tendem a ser altamente complexas e (...) “hierárquicas”. (HATIM E MUNDAY, 2004, p. 52)

Estes processos de buscar equivalentes, avaliar se são os ideais no contexto apresentado, validar termos, retomar conceitos, entre outros, fazem parte da tarefa do tradutor, que muitas vezes utiliza estratégias de tradução de forma automática devido à sua experiência. Mas ter conhecimento destas estratégias é uma das formas de validar se a equivalência encontrada é a mais adequada ao texto ou não.

Neste trabalho, ao falar-se na busca de equivalências e sua disponibilização em FrameNets multilíngues, deve-se considerar que existe um contexto mais amplo que engloba a discussão de equivalências, relacionado ao conceito de *frames*. Porém esta discussão apenas

pode ser realizada se forem consideradas, primeiramente, as estratégias de definição de equivalências em estudos na área da tradução, para que, posteriormente, seja possível realizar uma reflexão sobre o assunto no âmbito de recursos lexicais baseados em *frames*.

Para iniciar a discussão sobre estratégias para definição de equivalências, retoma-se, primeiramente, que a equivalência é relativa, dependendo da unidade de tradução<sup>9</sup> que está sendo considerada. Em alguns casos, haverá palavras equivalentes na língua fonte e na língua alvo que apontarão a um mesmo referente e poderá falar-se em equivalência exata de conteúdos, mas, em outros casos, não haverá um equivalente direto para a palavra ou expressão a ser traduzida, em outros, o referente na língua alvo será inexistente, e em outros ainda, o equivalente mais próximo não representará a mesma extensão de significados do esperado. Com isso, o tradutor se defronta com um grande desafio que é o de buscar por possíveis equivalentes para que a tradução seja capaz de realizar o objetivo da comunicação.

A falta de equivalência pode existir por diversos fatores, sendo eles linguísticos ou extralinguísticos. Baker (2006, p. 21-26) destaca os problemas mais comuns que causam a não equivalência lexical, são eles:

- a) Conceitos culturalmente específicos: este tipo de não equivalência acontece quando a língua fonte possui um conceito completamente desconhecido na língua alvo, sendo ele abstrato ou concreto. Exemplo: EN= Black Friday;
- b) Não lexicalização do conceito da língua fonte na língua alvo: quando o conceito da palavra expressada na língua fonte é conhecido na língua alvo, porém, nesta, não existe uma palavra lexicalizada para expressá-lo. Exemplo: PT = levar um **frango** / EN= to concede a **howler**;
- c) Palavra da língua fonte é semanticamente complexa: quando uma palavra expressa um conceito complexo na língua fonte e, na língua alvo, é necessário utilizar mais do que uma palavra para explicá-lo. Exemplo: EN= score / PT = marcar um gol;
- d) As línguas fonte e alvo fazem diferentes distinções em significado: quando uma língua alvo faz mais ou menos distinções em significado do que a língua fonte, mostrando que nesta língua o que é relevante não é tão relevante na outra. Exemplo: EN = slush / PT = neve semiderretida;
- e) A língua alvo não tem o hiperônimo: nos casos em que a língua alvo tem palavras específicas, mas não apresenta uma mais geral para caracterizar o conjunto de uma

---

<sup>9</sup> Shuttleworth e Cowie (1997, p. 192 *apud* Hatim e Munday, 2004, p. 17) definem a unidade de tradução como “o nível linguístico no qual o texto fonte é recodificado em texto alvo”, ou seja, os menores elementos de trabalho dos tradutores no texto fonte. Segundo os autores, “pode ser uma palavra, um grupo de palavras, uma oração, uma frase ou mesmo o texto como um todo”.

- forma menos detalhada. Exemplo: Exemplo: EN = nut (fruta de apenas uma semente revestida por casca resistente, como a noz, ou também semente revestida por casca resistente, como amendoim ou amêndoa);
- f) A língua alvo não tem o hipônimo: mais comum do que o caso anterior, a língua alvo pode não ter um termo específico para descrever algo da língua fonte. Exemplo: EN = scorer / PT = jogador pontuador, jogador que faz o gol;
- g) Diferenças de perspectiva física ou interpessoal: acontece quando a importância dada à perspectiva física varia de uma língua para outra e também quando a forma de expressar relações entre coisas e pessoa não acontece da mesma forma em duas línguas. Exemplo: EN = Here we **come!** / PT = Aqui **vamos** nós!;
- h) Diferenças no significado expressivo: quando uma palavra na língua alvo tem o mesmo sentido proposicional da língua fonte, mas apresenta um significado expressivo distinto. Isso acontece quando um equivalente é neutro se comparado à palavra da outra língua, que traz consigo outros elementos implícitos. Exemplo: EN = batter / PT = bater, mas de forma repetitiva;
- i) Diferenças na forma: neste caso, não existe um equivalente na língua alvo para uma forma da língua fonte, como no caso de alguns prefixos e sufixos. Exemplo: EN = **workaholic** / PT = **viciado em** trabalho;
- j) Diferenças em frequência e propósito na utilização de formas específicas: Exemplo: EN = I wish I **had** a car. / PT = Queria **ter** um carro;
- k) Utilização de empréstimos linguísticos no texto fonte. Exemplo: EN = Are you **au fait** with the rules of the game? / PT = Você está **bem informado** sobre as regras do jogo?.

Assim, em casos como estes de não equivalência, Baker (2006) afirma que a escolha de um possível equivalente a outro é realizada por fatores linguísticos ou extralinguísticos, ou seja, não existe uma forma única de resolver impasses na busca destes equivalentes de tradução, mas realiza-se um processo de análise e definição do equivalente mais adequado. Segundo ela, a escolha vai depender não apenas nos sistemas linguísticos das línguas trabalhadas, mas também em como tanto o escritor do texto fonte como o autor do texto alvo, ou seja, o tradutor, manipulam o sistema linguístico com o qual estão trabalhando.

Algumas estratégias utilizadas por tradutores profissionais são descritas abaixo. Baker (2006) salienta que estas não são as únicas estratégias existentes, pois concorda com o mencionado por Hatim e Munday (2004) ao citar que os desafios encontrados durante uma

tradução remetem a diferentes soluções, mas ela afirma que estas são as estratégias mais comuns utilizadas no dia a dia dos profissionais:

- a) Tradução por uma palavra mais ampla: esta é uma das estratégias mais comuns para lidar com a não equivalência e consiste em utilizar uma palavra mais geral para suprir a falta de uma palavra específica na língua alvo.
- b) Tradução de uma palavra por outra mais neutra ou menos expressiva: utiliza-se esta estratégia principalmente em caso de palavras que tenham uma conotação pejorativa na língua alvo. Neste caso, pode-se escolher um equivalente que seja mais neutro para não causar um impacto que o texto fonte não teria.
- c) Tradução por substituição cultural: esta estratégia normalmente depende da liberdade dada ao tradutor pelo solicitante da tradução e também do propósito da tradução. Isso porque envolve a alteração de algum item culturalmente específico na língua fonte, para outro item que tenha o mesmo sentido proposicional e impacto (dentro do possível) na língua alvo. Segundo a autora, “a maior vantagem na utilização desta estratégia é que dá ao leitor um conceito com o qual ele pode identificar algo familiar e atrativo” (BAKER, 2006, p. 31).
- d) Tradução utilizando uma palavra emprestada com ou sem uma explicação: a utilização desta estratégia depende da liberdade que o tradutor tem de utilizar estrangeirismos no seu trabalho. Ela é comumente utilizada principalmente quando se trata de itens culturalmente específicos, conceitos modernos e jargões. Pode ser muito útil utilizar uma explicação da palavra neste caso.
- e) Tradução por paráfrase utilizando-se palavras relacionadas: esta estratégia costuma ser utilizada quando o conceito expresso na língua fonte está lexicalizado ou ocorre com mais frequência de uma forma diferente na língua alvo. Utiliza-se, então, esta estratégia para aumentar a naturalidade da palavra no texto.
- f) Tradução por paráfrase utilizando-se palavras não-relacionadas: similar à estratégia apresentada anteriormente, mas utilizada em casos que a língua alvo não possua itens lexicalizados para relacionar com a palavra da língua fonte. Neste caso, a paráfrase pode utilizar um hipônimo ou simplesmente explicar o significado do item apresentado na língua fonte se for o caso de algo semanticamente complexo.
- g) Tradução por omissão: embora pareça uma opção drástica, segundo a autora em alguns casos é possível omitir palavras ou expressões que não afetem o desenvolvimento do texto.

h) Tradução por ilustração: utilizada em casos que o equivalente na língua alvo refere-se a algo que pode ser ilustrado.

Além de pontuar e esclarecer cada uma das estratégias mencionadas acima, a autora fala também sobre a forma como a noção de *campos semânticos* pode ser útil como estratégia de busca de equivalente de tradução. Segundo Baker (2006, p. 18), “às vezes é útil ver o vocabulário de uma língua como um conjunto de palavras que se referem a uma série de campos semânticos”, os quais são explicados posteriormente como sendo conceitos abstratos comuns a uma comunidade. Embora a autora defenda que a noção de campos semânticos seja limitada e também não aplicável em algumas situações, por ser uma “simplificação da forma na qual a linguagem realmente funciona” (Baker, 2006, p.19), ela defende que esta concepção pode ser útil ao tradutor por duas razões:

- a) por considerar o “valor” que as palavras têm como parte de um sistema; e
- b) por possibilitar o desenvolvimento de estratégias para lidar com a falta de equivalência.

O conceito de *campos semânticos* (*semantic fields*) foi sugerido pelo linguista Jost Trier em 1931 (Krifka, 2001). Para ele, o significado das palavras é determinado apenas se contrastado com outras palavras que denotem expressões relacionadas à mesma área. Os princípios básicos de sua teoria são: a) o significado de uma palavra depende do significado de outras palavras que denotem situações do mesmo domínio; b) as palavras que fazem parte de um mesmo campo semântico englobam toda a série que forma os fenômenos, sem brechas ou sobreposições, como um mosaico; e c) a mudança de significado de uma palavra pertencente a um campo semântico envolve alterações de significados também em outras palavras do campo.

Assim, quando Baker (2006) cita a forma como os campos semânticos podem ser úteis na tradução, compreende-se que se trata aqui da relação entre agrupamentos de palavras que existem considerando-se uma área comum em ambos os idiomas. No caso do futebol, por exemplo, o conjunto de termos utilizados para relatar a “marcação de um gol” em português, não terá um conjunto de termos equivalentes diretos em inglês. Porém, ao ter-se a relação de termos que fariam parte do conjunto em inglês e em português, torna-se possível uma análise e definição de equivalentes que podem ser adequados em diferentes casos. Isso porque, ao compreender as diferenças entre os candidatos a termos nas línguas, torna-se mais fácil a definição de um deles como possível equivalente. Além disso, a relação hierárquica observável nos campos semânticos pode ser considerada na hora de definir um termo mais



geral como possível equivalente de um termo específico que não tem um equivalente direto na língua alvo, por exemplo.

Todavia, sabe-se que a noção de campos semânticos apresenta uma concepção um tanto tímida no que concerne a questões culturais e pragmáticas, pois se trata de um conceito relacionado aos estudos do estruturalismo. Por este motivo, Fillmore (1985), ao apresentar seus estudos sobre *frames semânticos* (a ser detalhado no próximo subcapítulo), apresenta uma nova forma de perspectiva teórica em semântica empírica, que pode ser, de certa forma, comparável ao conceito de campos semânticos, mas que difere deste ao enfatizar a relação entre língua e experiência.

Quando Baker (2006) fala da utilização de conceitos abstratos comuns a duas comunidades linguísticas distintas como estratégia de tradução, a autora está falando especificamente sobre os campos semânticos. Porém, ao invés de tratar-se destes, sugere-se, nesta pesquisa, a observação de um conceito culturalmente situado, que é o de *frames semânticos* para a tarefa. Assim sendo, os conceitos abstratos comuns que podem servir como estratégia de tradução são os *frames* e eles passam considerar língua e cognição como indissociáveis.

Para refletir sobre esta proposta, o próximo subcapítulo deste trabalho apresenta a teoria da Semântica de *Frames*. Esta teoria, que reforça a relação entre linguagem e experiência cognitiva, defende a existência de “estruturas de expectativas” (Tannen, 1979, p. 144) capazes de disponibilizar conjuntos de unidades lexicais agrupadas por pertencer a um mesmo *frame* conceitual, o que pode interferir na forma como os conjuntos de unidades lexicais evocadoras deste item em diferentes línguas podem ser expostos em recursos lexicais.

Para que esta reflexão possa ser aprofundada, fala-se sobre a Semântica de *Frames* e seu compromisso com a Linguística Cognitiva, além de apresentarem-se os conceitos centrais desta teoria e também o projeto pioneiro de desenvolvimento de uma base de dados baseada em *frames*.

## 2.2 SEMÂNTICA DE *FRAMES* E LINGUÍSTICA COGNITIVA

Neste subcapítulo pretende-se apresentar a teoria da Semântica de *Frames* (Fillmore, 1982; 1985), situando-a no âmbito da Linguística Cognitiva, além de refletir sobre suas características e aplicação e, ao final, apresentar o projeto pioneiro de desenvolvimento de um banco de dados lexical baseado na noção de *frames*, a FrameNet. Com isso, espera-se que seja possível compreender o que é e no que se baseia a teoria da Semântica de *Frames*, para que seja iniciada, então, uma discussão, no subcapítulo seguinte, sobre a pertinência desta teoria como suporte para estudos acerca de questões multilíngues.

Vale lembrar que, assim como mencionado na seção em que se abordam as estratégias de definição de equivalências de tradução, a concepção de significado é uma questão central ao falar-se sobre equivalências em diferentes idiomas. Isso porque, não seria possível pensar-se em equivalências se não houvesse níveis conceituais comparáveis entre os idiomas, nos quais os significados de palavras pudessem ser comparados. Assim, ao discorrer, neste subcapítulo, sobre a forma como a Semântica de *Frames* organiza as unidades lexicais e as relaciona em um nível conceitual, torna-se possível compreender como esta concepção pode sustentar uma discussão acerca de equivalentes de tradução no âmbito desta pesquisa.

Para iniciar, portanto, esta reflexão sobre a Semântica de *Frames*, que teve seu desenvolvimento na linguística valendo-se de conceitos da Linguística Cognitiva, torna-se necessário também compreender tal perspectiva teórica nos estudos sobre a linguagem e a forma como as pesquisas na área se desenvolveram ao longo dos anos. A partir disso, tornar-se-á mais clara a compreensão da proposta teórica defendida pela Semântica de *Frames*, além das principais contribuições desta nas pesquisas em linguística aplicada. E este é o objetivo da próxima seção, que contextualizará a Semântica de *Frames* nos estudos do cognitivismo.

### 2.2.1 Linguística Cognitiva

Para explicar a relação entre a Semântica de *Frames* e a Linguística Cognitiva, Geeraerts (2006) utiliza uma metáfora. Segundo o linguista, cada abordagem linguística que se baseia no cognitivismo pode ser vista como uma ilha e constitui, por si só, uma entidade conceitual e sociológica. Às vezes, algumas ilhas são ligadas através de pontes, mas em

outros casos apenas os campos de batalha é que são compartilhados. Neste contexto, a Linguística Cognitiva passa a ser vista, então, como um arquipélago, pois é descrita como

(...) um conglomerado de centros de pesquisa linguística, mais ou menos extensivos, mais ou menos ativos, que estão intimamente ligados por uma perspectiva comum, mas que não estão (ainda) relacionados por uma regra comum de uma teoria bem definida. (GEERAERTS, 2006, p. 2).

A Semântica de *Frames*, neste caso, é vista como uma das ilhas que forma o arquipélago, com seus traços e perspectivas próprias, mas pertencente a uma teoria maior da qual faz parte, que é a do cognitivismo.

Embora não seja possível dispor em uma listagem as características que os estudos embasados na Linguística Cognitiva possuem, por serem, estes, autônomos o suficiente para possuírem suas próprias características centrais, Evans (2012, p. 130-131) destaca os compromissos básicos considerados por teóricos que têm seus estudos embasados no cognitivismo. São eles: o *compromisso cognitivo* e o *compromisso de generalização*. O compromisso cognitivo trata do aspecto de caracterização da linguagem levando-se em consideração o que é conhecido sobre a mente e o cérebro de forma interdisciplinar. É ele que garante o acordo entre estudos de todas as ciências cognitivistas, como psicologia, neurologia, etc. Já o compromisso de generalização está relacionado à busca da generalização de forma mais ampla para explicar os fenômenos linguísticos, ou seja, combate-se a proposta gerativista de existência de módulos distintos nas áreas de estudos linguísticos (a fonologia estuda os sons; a semântica, os significados; a morfologia, a estrutura das palavras; a sintaxe, a estrutura das frases; etc.). Com isso, este compromisso visa à investigação de como os aspectos de conhecimento linguístico se dão em relação ao conjunto comum de habilidades cognitivas humanas que as pessoas possuem.

O surgimento da Linguística Cognitiva foi marcado como um movimento que foi de encontro às ideias propostas pelo Gerativismo. Uma das críticas aos teóricos desta corrente é que, embora eles já considerassem o papel da “mente” nos estudos da linguagem, o abordavam de uma forma limitada, por tratarem apenas da maturação e do desenvolvimento da linguagem na mente devido a uma capacidade biológica inata presente em todos os seres humanos. Logo, os linguistas que não concordavam com tal limitação, como George Lakoff, Ronald Langacker, Leonard Talmy, Charles Fillmore, passaram a questionar as propostas do Gerativismo (Evans, 2012). Estes linguistas passaram a integrar, portanto, uma corrente que defendia, conforme explicado por Ferrari (2011, p. 14), que “a relação entre palavras e mundo é mediada pela cognição. Assim, o significado (...) passa a ser visto como uma

construção cognitiva através da qual o mundo é apreendido e experienciado” e a visão de que a mente não é influenciada pelo corpo, como se fosse uma parte independente do corpo, é deixada para trás (Lakoff, 1987).

Outras críticas foram feitas ao modelo gerativista, que defendia também a modularidade da mente, a racionalização de aspectos da linguagem e a autonomia da sintaxe. Isso porque, para os cognitivistas, a linguagem não é independente de outras faculdades mentais (Evans e Green, 2006). Um dos teóricos a posicionar-se desta forma foi Langacker (1987), que defendeu a integração entre as várias facetas que as estruturas linguísticas possuem e propôs a inseparabilidade da semântica e da sintaxe ao desenvolver os fundamentos da gramática cognitiva.

Teóricos cognitivistas defendem a adoção da perspectiva empirista e acreditam que não é possível ignorar a influência do corpo na mente humana, ou seja, a forma como as pessoas compreendem suas experiências e sua realidade está completamente vinculada à sua própria estrutura corpórea. Geeraerts (2006) esclarece este conceito dizendo que

(...) somos seres personificados, não simplesmente mentes. Nossa natureza orgânica influencia nossa experiência do mundo, e esta experiência se reflete na língua que utilizamos. O exemplo *atrás/em frente de* novamente fornece uma ilustração clara e simples disso: as perspectivas que usamos para conceituar a cena vêm do fato de que nossos corpos e nosso olhar sobre as coisas têm uma orientação natural (...) (GEERAERTS, 2006, p. 5)

Assim, a partir dessa premissa, os conceitos aos quais as pessoas têm acesso e a noção de realidade que elas têm relacionam-se diretamente com a estrutura de seu corpo e em sua interação com o ambiente no qual estão inseridas. É por isso que as experiências vivenciadas pelas pessoas influenciam na concepção dos conceitos básicos que elas têm das coisas. Desta forma, a partir dos estudos sob a perspectiva da Linguística Cognitiva, a linguagem passa a ser vista como algo que não é estável, mas influenciável pelo contexto no qual o falante está inserido e seu posicionamento no mundo.

Outra importante contribuição desta teoria é no que diz respeito aos estudos sobre categorização por protótipos. Lakoff (1987) destaca que Eleanor Rosch foi uma das principais responsáveis por tornar o tema reconhecido na área da psicologia cognitiva e que sua *teoria do protótipo* surgiu como um desafio à teoria clássica<sup>10</sup>. Isso porque ela

---

<sup>10</sup> A teoria clássica defende que categorias são definidas em termos de características em comum dos membros que fazem parte do grupo. Mais do que isso, este modelo sugere um conjunto de condições essenciais para definir o pertencimento a um grupo (LAKOFF, 1987). Por exemplo, no caso da categoria “pássaro”, as características poderiam ser: animal que tem asas, animal que voa e animal que bota ovos. Todos os membros

demonstrou que as características de membros de uma mesma categoria não são as mesmas e que, por isso, eles não devem ter o mesmo status dentro dela. Através de um estudo feito sobre a categorização de cores, a pesquisadora provou que assimetrias são encontradas na relação entre membros de uma mesma categoria e que alguns membros podem ser considerados melhores exemplares dentro dela. Isso porque para a Linguística Cognitiva a categorização por protótipos resulta de “modelos de natureza cognitiva” (LAKOFF, 1987, p. 45), ou seja, categorias são formadas e seus membros são compreendidos e interpretados baseados em modelos cognitivos que a pessoa formou durante sua vida.

Segundo Evans (2012), a Linguística Cognitiva se diferencia das teorias de estudo da língua porque seu alcance abrange investigações sobre as relações entre linguagem humana, a mente e as experiências sócio-físicas. Além disso, a autora destaca também que a corrente cognitivista chegou a um patamar de maturidade atualmente tanto em termos de teorias, metodologias e alcance, devido ao grande número de pesquisas sendo realizadas na área. Além disso, esta corrente conta atualmente com diversos desdobramentos. Trata-se de áreas de estudos mais focadas em determinados aspectos, mas que levam em consideração uma base comum a nortear seus estudos.

Outro aspecto importante sobre esta teoria é a forma como a mudança na posição da semântica nos estudos da linguagem acontece. Isso porque a semântica passa a ter um papel de maior destaque nestes estudos linguísticos, ou seja, embora já estivesse sendo estudada por alguns funcionalistas, assim como também pela semântica formal, ela passa a ser vista com diferentes olhos na perspectiva cognitivista. Na Linguística Cognitiva, a língua passa a ser vista “como um instrumento para organizar, processar e transmitir informações – em outras palavras, algo primordialmente semântico” (GEERAERTS, 2006, p. 3), o que garante uma importância maior a esta área do que ela vinha recebendo até então.

Sobre isso, destaca-se que esta semântica denominada cognitiva propõe uma visão diferenciada do significado linguístico, a visão enciclopédica do significado. Este conceito se opõe à chamada visão de dicionário. Em outras palavras, a concepção de que os significados linguísticos que temos em mente possam ser caracterizados da forma como os vemos em dicionários é criticada, por ser muito limitada, como seria o caso de relacionar a palavra *cão* a *animal mamífero de quatro patas, normalmente domesticado*. Para os linguistas cognitivos, mais do que isso, existe um significado mais amplo que está armazenado em nossa mente e

---

que fazem parte desta categoria devem ter cada uma das características mencionadas e, como todos os membros são considerados em base a um mesmo grupo de características, nenhum deles é destacado como exemplo prototípico, pois todos são representativos dentro da categoria.

que considera variáveis pragmáticas, relacionadas a contextos e situações específicas do uso da língua. Desta forma, as palavras não existem como oposições a outras, mas como “pontos de acesso para sistemas de conhecimento” (FERRARI, 2011, p. 20). A concepção de significado explicada acima vai ao encontro do que Fillmore (1985, p. 223) introduz ao explicar que, “para compreender cada membro de um grupo de palavras, é necessário, de certa forma, compreender o que elas são como um todo”, ou seja, o significado de *cão* também está relacionado à noção de *doméstico*, *latir*, *ração*, formando uma cena na qual diversas outras palavras surgem para complementar o conceito buscado.

Para explicar seu ponto de vista, alguns teóricos cognitivistas passaram a falar, então, de *estruturas de conhecimento esquemáticas*, que também são conhecidas como *frames* ou *domínios de experiência*. Na linguística, as investigações sobre a noção de *frames* foram inicialmente desenvolvidas por Charles J. Fillmore (1982), principal teórico da área e pesquisador no Instituto Internacional de Ciências da Computação (ICSI), em Berkeley, Califórnia, desde 1971 e diretor do Projeto FrameNet. Fillmore tratou principalmente de questões de sintaxe e semântica lexical, com ênfase na relação entre as propriedades da forma linguística com seu significado e uso e foi o principal responsável pelo desenvolvimento da teoria da Semântica de *Frame*. Atualmente esta abordagem já é conhecida mundialmente e diversos projetos pelo mundo têm utilizado a Semântica de *Frames* em estudos primordialmente relacionados ao desenvolvimento de léxicos computacionais e aplicações computacionais com vistas ao processamento de linguagem natural.

Para passar a caracterizar a Semântica de *Frames*, é importante levar em consideração os dois conceitos aos quais o termo está relacionado. Segundo Fillmore (1982, p. 111), a Semântica de *Frames* pode ser vista como: 1) um programa de pesquisa em semântica empírica; e 2) um sistema descritivo para representar os resultados de tal pesquisa. Nessa etapa do trabalho, trata-se a Semântica de *Frames* como um programa de pesquisa e, para compreendê-lo, falar-se-á, inicialmente, sobre os conceitos básicos ao qual ela se vincula para que, em seguida, o projeto FrameNet seja apresentado.

### 2.2.2 A Semântica de *Frames*

Ao detalhar a proposta da Semântica de *Frames*, destaca-se que um dos conceitos centrais desta teoria é o de *frame semântico*. A noção de *frame* não é exclusiva da área da

linguística e, inclusive, o próprio termo foi emprestado da área da inteligência artificial, onde era utilizado pelo pesquisador Marvin Minsky. Fillmore (1985) destaca os termos utilizados em diversas áreas e que também representam, de diferentes formas, a noção das estruturas cognitivas propostas por ele:

*frame*: Minsky (1975), Winograd (1975), Charniak (1975); *schema*: Bartlett (1932), Rumelhart (1975); *script*: Schank e Abelson (1977); *global pattern*: de Beaugrande e Dressler (1981); *pseudo-text*: Wilks (1980); *cognitive model*: Lakoff (1983); *experiential gestalt*: Lakoff e Johnson (1980); *base* (em oposição à *profile*): Langacker (1974); *scene*: Fillmore (1977) etc. (FILLMORE, 1985, p. 223)

Para a Semântica de *Frames*, o *frame* semântico se refere a um “plano de fundo comum de conhecimento, no qual o significado das palavras passa a ser interpretado” (BOAS, 2009, p. 68). Desta forma, é possível compreender significados a partir de estruturas de conhecimento que são socialmente compartilhadas, graças às situações vividas e compreendidas pelos envolvidos. Fillmore (1982) caracteriza-o como um sistema de conceitos relacionados no qual, para compreender um deles, é necessário compreender o funcionamento do todo, ou seja, “quando uma das coisas que faz parte da estrutura é apresentada em um texto, ou em uma conversa, todas as demais envolvidas são automaticamente relacionadas” (FILLMORE, 1982, p. 111). Desta forma, o pesquisador defende que as palavras estão relacionadas às nossas experiências e conhecimento prévios e não são compreendidas numa visão de dicionário, mas numa visão enciclopédica, como defende a Linguística Cognitiva. Inspirado em Fillmore, Gawron (2008, p. 4) afirma que o papel principal dos *frames* é o de “explicar como a nossa interpretação de textos pode (com razão) ir tão longe em relação ao que o texto diz literalmente”.

Para Lakoff (1987), a noção de *frames*, assim como também de *scripts* e *schemas* (empregadas por outros autores), relacionam a cognição a modelos computacionais da mente. Ele explica que, através de proposições obtidas graças a modelos prototípicos, torna-se possível compreender brechas e completar espaços até então abertos para interpretação. Assim, segundo ele estes conceitos

são tentativas de fornecer um formato para representar o conhecimento humano em modelos computacionais da mente. Tenta-se fazer isso fornecendo estruturas proposicionais em termos nos quais as situações podem ser compreendidas. Estas estruturas apresentam espaços em branco, que podem ser preenchidos pelos indivíduos que ocorrem em uma dada situação que é estruturada. (LAKOFF, 1987, p. 116)

Em 1975, Fillmore publica seu artigo *An Alternative to Checklist Theories of Meaning* e fala, então, sobre dois conceitos importantes para a Semântica de *Frames: cenas e frames*. Para o autor, o termo *cena* era compreendido em um sentido genérico e abrangia tipos familiares de trocas interpessoais, cenários padrão definidos culturalmente, estruturas institucionais, imagem corporal, entre outros. Já o termo *frame* era utilizado como um sistema de escolhas linguísticas, ou seja, “nos casos mais simples sendo grupos de palavras, mas também incluindo escolhas de regras gramaticais ou categorias linguísticas – que poderiam ser associadas a instâncias prototípicas de cenas” (FILLMORE, 1975, p. 124). Através desta distinção proposta por Fillmore, destacam-se duas considerações: 1) tinha-se o termo *cena* relacionado à estrutura cognitiva do evento, e o termo *frame* relacionado à estrutura linguística do mesmo, e 2) ao falar sobre a relação entre a escolha linguística baseada em cenas prototípicas e a cognição, abria-se espaço para algo importante na definição de equivalências de tradução, que é o trabalho de considerar contextos que podem existir de forma diferenciada em lugares e culturas distintos. Porém, Petruck (1996), ao retomar os conceitos propostos por Fillmore ao diferenciar “cena” e “frame” nos seus primeiros trabalhos, explica que alguns anos depois da publicação original, o conceito de “cena” deixou de ser utilizado, pois o *frame* passou a ser considerado também um mecanismo de estruturação cognitiva. Com isso, o *frame* passa a caracterizar uma “situação” abstrata, constrói os significados das palavras e as palavras o evocam, ou seja, trata-se de um “sistema de categorias estruturadas de acordo com algum contexto motivador” (FILLMORE, 2006, p. 381).

Para exemplificar este conceito, Fontenelle (2009, p. 39) fala sobre a cena de uma *transação comercial* no qual um dos *frames* é *Comércio\_comprar*<sup>11</sup>. Este *frame* envolve quatro participantes principais: um VENDEDOR, a MERCADORIA, um COMPRADOR e DINHEIRO. Cada um destes é considerado um *elemento de frame* (EF), ou seja, um participante da situação descrita pelo *frame* (Gawron, 2008), e a compreensão de um deles está relacionada também à compreensão dos demais elementos envolvidos, o que faz com que o *frame* seja integralmente interpretado. Cada *frame* tem seu próprio conjunto de *elementos de frame* e, segundo Gawron (2008, p. 8), o elemento de *frame* é “simplesmente um participante, uma característica ou um atributo habitual do tipo de situação descrita pelo *frame*”.

---

<sup>11</sup> Neste trabalho os *frames* são citados em fonte Courier New, as Unidades Lexicais, em *itálico* e os Elementos de Frame em LETRAS\_MAIÚSCULAS para facilitar sua visualização e distinção.



Em geral, a importância da noção de *frames* é justificada por Fillmore (2006, p. 380) ao exemplificar como critérios de classificação muitas vezes não são suficientes para compreendermos o significado das palavras. É o caso da expressão “café da manhã”, que é compreendida tendo-se em mente o pressuposto de que em determinadas culturas se tem três refeições por dia, em horários mais ou menos similares e que, uma delas, é feita logo de manhã, depois de um período de sono e com algumas características de cardápio semelhantes (embora existam diferenças culturais envolvidas neste caso). Mas algo a ser notado é que, mesmo que uma destas três condições amplamente associadas ao café da manhã não esteja presente, ainda pode-se utilizar a expressão. Por exemplo: no caso de alguém que trabalhe à noite e não durma neste período, assim que o sol levantar e ele for fazer uma refeição, poderá denominá-la “café da manhã”; outra situação seria alguém que dormiu toda a manhã e, ao acordar, às três horas da tarde, faça uma refeição e a chame de “café da manhã”. Ou seja, critérios como “ser após um período de sono” ou “ter um cardápio mais ou menos específico” não são suficientes para definir a expressão. Desta forma, *frames* auxiliam ao fornecer um contexto ao enunciado, situando-o dentro de um esquema conceitual, considerando quem são os participantes (elementos) envolvidos, e, com isso, a língua passa a ser observada em conjunto com este contexto.

Outro aspecto relevante considerado pela Semântica de *Frames* está associado à forma com que trata da valência verbal, pois nesta teoria consideram-se essenciais os critérios semânticos na identificação do pertencimento das unidades lexicais (doravante ULs) a uma determinada classe semântica, em sua relação com o *frame* que evocam (Boas, 2011). Já a análise sintática aparece de forma secundária. Isso porque nesta perspectiva ela “se baseia na presença e nas combinações de elementos de *frame*, que são definidos semanticamente”. Esta questão de valências será retomada no decorrer do trabalho, pois, considerando-se uma pesquisa que se compromete em refletir com aspectos de multilinguagem em recursos baseados em *frames*, a verificação dos equivalentes necessariamente se depara com a comparação de valências entre as ULs de diferentes idiomas.

No próximo subcapítulo apresenta-se o projeto pioneiro que teve seu desenvolvimento embasado no conceito de *frames*: a FrameNet. Com isso espera-se verificar como os conceitos da Semântica de *Frames* são aplicados no desenvolvimento de um recurso lexical e, mais do que isso, quais as contribuições deste tipo de procedimento. Além disso, torna-se possível compreender mais detalhadamente a aplicabilidade da Semântica de *Frames* para fins de criação de bases de dados e, conseqüentemente, os pressupostos a serem considerados também no desenvolvimento de trabalhos multilíngues (foco do próximo subcapítulo).

### 2.2.3 A FrameNet

Nesta seção fala-se sobre o projeto pioneiro de aplicação da teoria de significados da Semântica de *Frames*: a FrameNet<sup>12</sup>. Espera-se, com isso, verificar como acontece a aplicação dos conceitos trazidos no subcapítulo anterior e, também, aprofundar a reflexão a respeito da forma como as ULs são apresentadas em um recurso como este. Através disso, tornar-se-á possível aprofundar o estudo sobre o estabelecimento de equivalências via ULs (proposta a ser discutida no subcapítulo seguinte, sobre *frames* e multilinguagem).

O projeto FrameNet está sendo desenvolvido no Instituto Internacional de Ciências da Computação (ISCI), em Berkeley, nos Estados Unidos, pelo professor Charles J. Fillmore e colaboradores, e tem como objetivo construir uma base de dados lexicais de língua inglesa que pode ser lida “tanto por humanos quanto por máquinas, com exemplos de como as palavras são utilizadas em textos verdadeiros” (FRAMENET, 2012). Fontenelle (2009, p. 38) afirma que o que torna a FrameNet diferente de qualquer outra base de dados é que ela tem o objetivo de disponibilizar não apenas informações morfossintáticas das palavras consultadas, mas também descrições de todas as relações possíveis entre os elementos de *frame* e, nela, cada palavra está vinculada a um conjunto de sentenças retiradas de um *corpus* previamente anotado com informações semânticas. E é por isso que Boas (2005, p. 454) descreve a FrameNet como “um léxico online de unidades lexicais (ULs) em inglês descritas sob os termos da semântica de *frames*”.

Atualmente, segundo o site do projeto, a FrameNet conta com uma base de dados de mais de 10 mil unidades lexicais e mais de 170 mil sentenças anotadas, obtidas no *corpus* da pesquisa. O trabalho no *corpus* acontece, de uma forma geral, na seguinte ordem: (1) as descrições de *frame* para as palavras ou famílias de palavras para análise são planejadas (inicialmente caracteriza-se o tipo de entidade ou situação representada pelo *frame*, depois é feita a escolha mnemônica para a etiquetagem das entidades ou componentes do *frame*, para, finalmente, ser construída uma lista de trabalho de palavras que parecem pertencer a um mesmo *frame*); (2) são identificadas frases do *corpus* no British National Corpus que mostrem usos comuns das palavras-alvo em *frames* específicos; (3) as frases são extraídas mecanicamente e anotadas manualmente através da etiquetagem dos elementos de *frame* que apareceram no exemplo; e (4) as inserções lexicais são automaticamente preparadas e arquivadas no banco de dados (Boas, 2009, p. 70).

---

<sup>12</sup> Disponível para acesso através do link <<https://framenet.icsi.berkeley.edu>>.

Sobre os elementos de *frame*, faz-se um detalhamento maior do que simplesmente indicar quais os participantes da cena. Na FrameNet, são desmembrados três tipos de elementos de *frame*, e sua denominação vai depender da sua representatividade no *frame* em questão. Ruppenhofer et al. (2006, p. 19) os diferenciam como: elementos de *frame* centrais (*core*), periféricos (*peripheral*) e extratemáticos (*extra-thematic*) e os caracterizam da forma apresentada abaixo.

a) **Centrais:** componentes necessários do *frame*, que tornam o *frame* único e o diferenciam dos demais. Estes EFs não precisam estar presentes linguisticamente na frase, mas, mesmo quando omitidos, eles são inferidos devido à sua representatividade no *frame* ao qual pertencem.

b) **Periféricos:** não diferenciam um *frame* dos demais, pois marcam *tempo*, *lugar*, *modo* etc., e estão presentes em diferentes *frames*. Eles não introduzem eventos adicionais, independentes e distintos que acontecem no *frame* e podem, inclusive, estar presentes em diferentes *frames*.

c) **Extratêmáticos:** são utilizados por um grupo maior de *host frames*, ou seja, não se relacionam especificamente à compreensão de apenas um *frame*. Eles podem estar relacionados a estruturas temporais, lugares, participantes adicionais, entre outros.

Além da anotação dos elementos de *frame*, a FrameNet também destaca a *unidade lexical* (UL) da sentença, que é a palavra que, seja ela verbo, substantivo, adjetivo, ou mesmo pertencente a outra classes de palavras (com menos frequência), evoca um *frame* (Ruppenhofer et. al., 2006). O caso mais simples para anotação é quando a UL é o verbo e os elementos de *frame* (doravante EFs) são seus dependentes sintáticos. Notamos esta situação em sentenças como a seguinte:

O Atlético Paranaense foi atrás do prejuízo e em uma cobrança de falta perfeita de Netinho, o [Furacão TIME\_PONTUADOR] **empatou** TARGET [o jogo JOGO] aos 45 minutos.

Porém, nos casos em que os verbos atuam como verbo-suporte, verbo de ligação ou verbos de controle, o substantivo ligado a eles é que faz o papel de evocador. No caso dos verbos-suporte, ou seja, os verbos que se combinam com substantivos para criar um predicativo verbal (e que, para a Semântica de *Frames*, permitem argumentos para completar as brechas dos EFs nos *frames* evocados pelo nome), a anotação dos EFs passa a ser relacionada ao substantivo definido como UL. (RUPPENHOFER et al., 2006). Isso acontece porque, segundo Ruppenhofer et al. (2006, p. 37), os verbos-suporte não apresentam sozinhos

“nenhum significado semântico” e “as construções de substantivos com seus verbos-suporte indicam a mesma situação que o substantivo indica quando utilizado isoladamente”. Assim, o *frame* é compreendido bastando apenas a presença do substantivo, independente do tipo de verbos-suporte que esteja presente.

Ainda sobre as ULs, embora não seja tão comum, os adjetivos, advérbios e preposições também podem ser marcados com esta etiqueta, pois também podem evocar *frames*. Sobre os adjetivos, a FrameNet opta por anotá-los em duas situações: quando eles são utilizados atributivamente e quando são usados predicativamente. Como exemplo trazido por Ruppenhofer et al. (2006, p. 43) , tem-se a sentença: Phil está **feliz** [por ter passado]<sup>13</sup>. Já a anotação dos advérbios como evocadores, segundo os pesquisadores, embora pouco recorrente, justifica-se porque esta classe de palavras pode demonstrar a atitude um falante, como na seguinte sentença: “**Francamente**, minha querida, eu não dou a mínima.<sup>14</sup>”, na qual “francamente” evoca o *frame* Probabilidade. Por último, as preposições também são mencionadas como ULs, no caso em que carreguem a noção de espaço, tempo ou movimento. Um dos exemplos trazidos é: Ela colocou [sua bolsa] **sobre** [a mesa]<sup>15</sup>.

Um dos exemplos trazidos aqui para visualização do resultado deste trabalho, disponível na FrameNet, é o verbete *cook* (cozinhar), que envolve a pessoa que cozinha (*COOK*), o que vai ser preparado (*PRODUCED\_FOOD*), através do uso de ingredientes (*INGREDIENTS*). Podem ser mencionados também o utensílio no qual a comida será colocada (*CONTAINER*) e a fonte de aquecimento (*HEATING\_INSTRUMENT*). Aqui é onde entra a noção de elementos de *frame*, previamente explicitada, pois cada um destes elementos é um EF que é ativado pelo verbete *cook*. Na FrameNet, estas informações, assim como outras características importantes relacionadas ao *frame*, são trazidas de forma a mostrar “como os elementos do *frame* se encaixam sintaticamente com a palavra que evoca o *frame*” (FrameNet, 2012). Como exemplo, ao buscar o verbo *cook* na FrameNet, temos a informação de que ele pode pertencer a diferentes *frames*, sendo eles: *Cooking\_creation*, *Apply\_heat* e *Absorb\_heat*. Ao clicar em um deles, obtém-se a definição do *frame* com todos os seus elementos principais sendo destacados, conforme figura abaixo:

---

<sup>13</sup> Phil is **happy** [that he passed].

<sup>14</sup> **Frankly**, my dear, I don't give a damn.

<sup>15</sup> She dropped [her handbag] **on** [the table].

<b>Cooking_creation</b>	<a href="#">Lexical Unit Index</a>
<b>Definition:</b>	
This frame describes food and meal preparation. A <b>Cook</b> creates a <b>Produced food</b> from (raw) <b>Ingredients</b> . The <b>Heating Instrument</b> and/or the <b>Container</b> may also be specified. <b>Caitlin BAKED some cookies from the pre-packaged dough.</b>	

**Figura 1** - Apresentação do frame *Cooking\_creation*

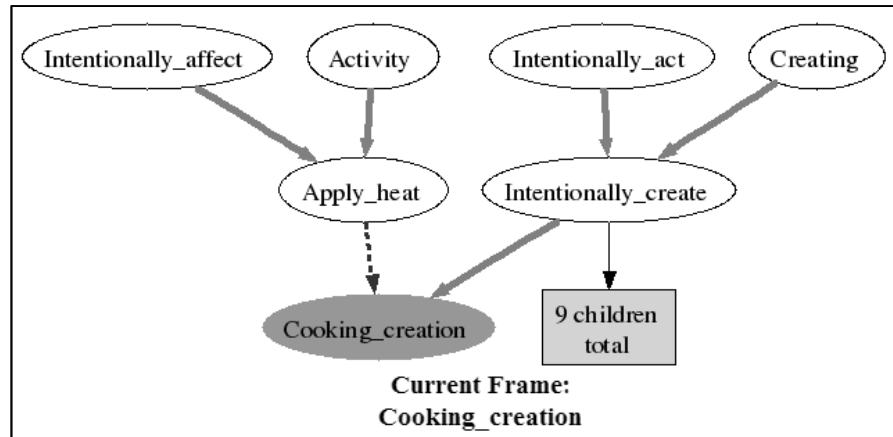
Além de mencionar quais os EFs da unidade lexical em questão, o banco de dados também lista os EFs considerados centrais (*cook, produced\_food*) e os que não são centrais (*container, degree, heating\_instruments, ingredients, manner, means* etc.). Com isso, é possível que o usuário saiba quais as informações comumente vinculadas ao uso da unidade lexical em questão e visualize exemplos de sua utilização. Segue abaixo quadro que mostra a apresentação parcial dos dados disponibilizados sobre os EFs na FrameNet:

<b>FEs:</b>	
<b>Core:</b>	
<b>Cook [Cook]</b> Semantic Type: Sentient	The <b>Cook</b> prepares the <b>Produced food</b> . Drew <b>COOKED</b> dinner for his friends. Drew <b>BAKED</b> an apple pie.
<b>Produced food [Food]</b>	The <b>Produced food</b> is the result of a <b>Cook's</b> efforts. Drew <b>PREPARED</b> dinner for his friends. Drew <b>BAKED</b> an apple pie for dessert.
<b>Non-Core:</b>	
<b>Container [Container]</b> Semantic Type: Container	This FE identifies the <b>Container</b> that holds the food being produced. <b>BAKE</b> the quiche <b>in a pie tin</b> . Things that apply the heat directly are <b>Heating Instruments</b> , e.g. crock-pot, electric griddle.
<b>Degree [Degr]</b> Semantic Type: Degree	This FE identifies the <b>Degree</b> to which an event occurs.
<b>Heating Instrument [Heat instr]</b> Semantic Type: Physical_entity	This FE identifies the <b>Heating Instrument</b> with which the <b>Cook</b> prepares the <b>Food</b> . Drew <b>PREPARED</b> vegetable soup <b>in a crock pot</b> .
<b>Ingredients [Ingr]</b>	The <b>Ingredients</b> which are altered by The <b>Cook</b> alters <b>Ingredients</b> to create the <b>Produced food</b> . Drew <b>BAKED</b> a pie <b>with apples, nuts, and raisins</b> .
<b>Manner [Manr]</b> Semantic Type: Manner	This FE identifies the <b>Manner</b> in which the cooking creation performed. Drew <b>PREPARED</b> dinner <b>hastily</b> .
<b>Means [Mns]</b> Semantic Type: State_of_affairs	This FE identifies the <b>Means</b> by which the cooking creation is performed. Drew <b>PREPARED</b> the soup <b>by throwing some vegetables into a pot of water</b> .

**Figura 2** - Elementos de frame do verbo “cozinhar” (*cook*) – imagem parcial

Outro aspecto importante da FrameNet é que ela apresenta a relação entre os *frames* através do recurso FrameGrapher, conforme figura a seguir. No caso da unidade lexical *cook*,

tem-se o acesso à informação de que o *frame* *Cooking\_creation* herda traços do *frame* *Intentionally\_create* (linha não tracejada) e utiliza-se de EFs do *Apply\_heat* (linha tracejada). Através destas classificações, torna-se possível compreender mais profundamente a ligação entre *frames* e a forma como um deles pode influenciar o outro.



**Figura 3** - Exemplo de relação entre *frames* no FrameGrapher

A FrameNet lista também a relação de unidades lexicais que também são utilizadas no *frame* em questão. A unidade lexical *cook*, por exemplo, faz parte de um conjunto de evocadores que remete a um mesmo *frame*, juntamente com *bake.v* (assar.v), *cook\_up.v* (inventar.v), *cook.n* (cozinheiro.n), *make.v* (fazer.v), *prepare.v* (preparar.v), entre outros. O usuário tem acesso a esta listagem e também pode consultar individualmente cada uma das unidades lexicais em questão para verificar sua definição, relação com outros *frames* e exemplos de sentenças anotadas. Com isso, o banco de dados disponibilizado pela FrameNet fornece informações importantes tanto semânticas quanto sintáticas sobre milhares de verbos, substantivos e adjetivos em língua inglesa.

Por se tratar de um projeto monolíngue, a FrameNet não tem o compromisso de tentar relacionar ULs, ou mesmo *frames*, em diferentes idiomas. Porém, pode-se pensar que, tendo-se *frames* conceituais similares em mais de um idioma, esta listagem de ULs evocadoras de um *frame* pode ser um caminho para pensar na indicação de equivalências. Isso porque supondo-se que, por exemplo, o *frame* *Cooking\_Creation* em inglês é comparável ao *frame* *Cozinhar* em português, em um nível conceitual, tornar-se-ia possível verificar de que forma suas ULs também podem ser comparáveis e as equivalências, estabelecidas. Além disso, como a FrameNet disponibiliza detalhes sobre as valências das ULs listadas, haveria

também a viabilidade de compará-las em um projeto multilíngue. Esta reflexão é aprofundada no próximo subcapítulo, que trata de FrameNets multilíngues.

Para concluir esta seção, pode-se dizer, resumidamente, que cada busca realizada na FrameNet remete a: (1) uma definição do *frame* ao qual a unidade lexical pertence, incluindo a definição dos elementos de *frame*; (2) frases exemplificando o uso prototípico dos elementos de *frame*; (3) outras informações referentes às relações entre *frames*; e (4) uma lista de unidades lexicais que evocam o *frame* (FrameNet, 2012).

A FrameNet contribui com diferentes áreas de pesquisa e pode ser utilizada com diversos propósitos, como: a) alunos podem utilizá-lo como um dicionário que conta com um número representativo de palavras juntamente com seus significados e usos; b) pesquisadores na área de processamento de linguagem natural podem utilizar os dados em programas como os de tradução automática, análise de sentimentos, entre outros; c) linguistas têm a possibilidade de consultá-lo e verificar propriedades de combinação de vocábulos; entre outros (FrameNet, 2012). Segundo Ruppenhofer et al. (2006, p. 5), atualmente sua base de dados possui mais de 10.000 unidades lexicais e pesquisadores de todo o mundo têm acesso a esta ferramenta. Estes pesquisadores passaram a desenvolver projetos irmãos a este em diversas partes do mundo e, por este motivo, atualmente estão em desenvolvimento FrameNets em línguas como: português<sup>16</sup>, chinês, alemão, espanhol, japonês e francês, sendo que nenhuma delas está diretamente relacionada ao projeto original, no sentido de estabelecimento de equivalentes de tradução. Além destes, outros projetos foram desenvolvidos graças ao pioneiro FrameNet, porém focando na análise linguística em domínios específicos, com bases de dados com terminologia do futebol, do direito, do turismo, da biomedicina etc.

Com a conclusão de projetos como estes, segundo Salomão (2009, p. 174), “estaríamos nos aproximando do sonho do ‘dicionário ideal’”, isso porque em bancos de dados baseados em *frames* poderíamos buscar por uma palavra qualquer e mais do que uma simples definição descontextualizada, teríamos acesso ao *frame* que evoca o determinado lexema, juntamente com os demais elementos de *frame* que têm relação com o buscado, além de uma listagem de valências da palavra, suas possibilidades combinatórias e exemplos de fácil visualização de uso real. E, no caso dos recursos multilíngues, todas estas informações

---

<sup>16</sup> Nacionalmente a FrameNet Brasil tem sido desenvolvida pela Universidade Federal de Juiz de Fora como contrapartida ao original em inglês. Trata-se de um trabalho em andamento que vem sendo discutido academicamente em dissertações e teses de alunos de mestrado e doutorado da instituição, e que tem colaborado com o avanço nos estudos na área de Semântica de *Frames* e linguística computacional. (FrameNet Brasil, 2012).

seriam, portanto, acessíveis em ambos os idiomas, tendo-se acesso a muito mais do que a simples indicação de equivalências.

Espera-se com este capítulo de revisão bibliográfica sobre a Linguística Cognitiva e a Semântica de *Frames* ter-se feito um apanhado capaz de situá-las e contextualizá-las nos estudos sobre a linguagem. Com isso, torna-se mais fácil compreender a amplitude e o alcance de trabalhos como o da FrameNet e todos os seus desdobramentos. Partindo-se destas premissas, e também do exposto no primeiro capítulo de revisão bibliográfica sobre equivalências de tradução, pretende-se, no próximo capítulo, abordar a multilinguagem em recursos lexicais baseados em *frames*. Isso porque o desenvolvimento de trabalhos deste tipo leva a um dos grandes desafios enfrentados em trabalhos em mais de um idioma: a necessidade de falar-se sobre equivalentes de tradução, foco da presente pesquisa.



### 2.3 FRAMES E MULTILINGUALIDADE

No mundo globalizado em que vivemos, cada vez mais nota-se a presença de recursos em mais de um idioma auxiliando na comunicação entre povos de diferentes países e culturas (Kay, 1995) de diferentes formas. Trata-se, por exemplo, de dicionários multilíngues, tradutores automáticos, aplicativos que traduzem a fala instantaneamente, enfim, recursos que esclarecem, de formas distintas, as dúvidas de consulentes que buscam a compreensão de textos, expressões ou termos em outro idioma. Assim, a multilingualidade passou a ter um status mais central nos estudos linguísticos e, atualmente, diversas são as pesquisas que colaboram com o desenvolvimento da área.

Este subcapítulo tem como objetivo refletir, inicialmente, sobre o que é a multilingualidade e como as pesquisas na área têm se desenvolvido atualmente no cenário mundial. Em seguida, parte-se para uma discussão mais específica que trata da multilingualidade vista a partir da perspectiva do desenvolvimento de bases de dados lexicais baseados em *frames*, com seus desafios e vantagens de aplicação, para, ao final, apresentarem-se dois projetos que utilizaram a Semântica de *Frames* no desenvolvimento de bases de dados multilíngues no domínio do futebol: Kicktictionary e Kicktictionary\_Br Copa 2014.

Destaca-se nesta pesquisa, principalmente, a preocupação na observação com os desafios impostos no estabelecimento de equivalências de tradução e a forma como os recursos multilíngues apresentados têm tratado deste assunto. Tendo-se isso claro, tornar-se-á possível refletir sobre as equivalências no âmbito de um projeto baseado em *frames* (FILLMORE, 1982; 1985), objetivo desta dissertação.

Ao definir *multilingualidade*, Kay (1995, p.1) afirma que se trata de “uma característica de tarefas que envolvam a utilização de mais do que uma linguagem natural” e afirma que houve, nos últimos anos, um aumento na demanda de pesquisas na área. Boas (2009) também fala sobre este aumento, porém reforça que tarefas que envolvem multilingualidade podem impor alguns desafios, como: lidar com estruturas polissêmicas divergentes, diferentes valências, diferenças em padrões de lexicalização e equivalências de tradução, temas que serão abordados em detalhes na próxima seção.

Estes desafios emergem porque, como mencionado na primeira parte do trabalho, cada língua representa a realidade social a que está vinculada de uma forma particular (Sapir, 1949) e tarefas que envolvam mais de um idioma devem, portanto, refletir a respeito de como

relacionar cada uma destas representações. Esta tarefa pode ser realizada de uma única forma: especificando-se correspondências de tradução.

Segundo Kay (1995), nota-se, atualmente, um aumento tanto no número de traduções propriamente ditas, como também em alternativas que possam auxiliar na tarefa. Para exemplificar, o autor cita que os estudos na área têm se relacionado principalmente à tradução automática, ferramentas de auxílio aos tradutores e produção completa ou parcial de textos originais paralelos em diferentes idiomas.

Ao falar sobre equivalência nos estudos multilíngues, um aspecto a ser considerado é que diferentes trabalhos podem observá-la de perspectivas distintas. No caso de lexicógrafos, por exemplo, conforme destacado por Piotrowski (1994), ao desenvolverem recursos lexicais em mais de um idioma, a classe gramatical das palavras listadas, assim como das possibilidades de equivalências para estas, são de primordial importância. Por outro lado, no caso de estudos similares que consideram o ponto de vista do tradutor, a classe de palavra de uma unidade de tradução tem menos importância, porque seu foco não está no trabalho segmentado de unidades lexicais. Já no caso dos recursos baseados em *frames* é comum que as classes de palavras sejam disponibilizadas para consulta, assim como as valências semânticas, porém, apenas em segundo plano, sendo o foco a relação das unidades lexicais com o *frame* ao qual pertencem.

Conforme destacado no subcapítulo anterior, projetos que utilizam *frames* na construção de bases de dados, mesmo que monolíngues, têm se mostrado úteis para diversas atividades. Porém, também são muitas as vantagens da utilização da Semântica de *Frames* no desenvolvimento de recursos lexicais em mais de um idioma e alguns dos motivos para que estudos nesta área tenham aumentado é suprir a demanda de ferramentas que auxiliem a tradução, a recuperação de informação e o ensino de idiomas (Boas, 2009). Neste contexto, a próxima seção tem como objetivo discorrer sobre o emprego da teoria da Semântica de *Frames* em trabalhos multilíngues, com destaque a forma como a equivalência de tradução é abordada neles.

A pergunta que direciona a discussão no presente subcapítulo, como um todo, é: levando-se em consideração os níveis conceituais (entre *frames*) e lexicais (entre ULs) nos diferentes idiomas, em que medida a correspondência conceitual repercute no nível lexical em termos de equivalência de tradução? Destaca-se que esta pergunta está vinculada também ao fato de observar-se que mesmo já havendo FrameNets para diversas línguas, como o japonês, o sueco e o espanhol, estes projetos não exploram de fato as questões sobre

multilinguagem, já que não têm um compromisso em abordar a forma como as FrameNets de línguas distintas interagem.

### 2.3.1 Recursos lexicais multilíngues baseados em *frames*

Diversas são as funções de um recurso lexical multilíngue: ele pode funcionar como auxílio a um usuário bilíngue que não tenha conhecimento de alguns aspectos da língua de consulta, ou a um tradutor que busca a palavra adequada para comunicar algo de um texto fonte para um texto alvo, ou mesmo um aprendiz que quer ser capaz de se comunicar com falantes de uma segunda língua e precisa, para isso, do vocabulário adequado. Porém, independentemente do objetivo da utilização de um recurso lexical multilíngue, assim como afirma Piotrowski (1994, p. 19), “a consulta a um dicionário é um processo relacionado a um texto”, ou seja, envolve muito mais do que a decodificação ou “absorção” de um símbolo, mas também a construção de um significado junto ao texto no qual ele foi e será utilizado. Desta forma, dicionários que não apresentam as palavras inseridas em um contexto funcionam como disseminadores de significados, porém, não direcionam o consulente à compreensão de qual é a palavra adequada para ele utilizar em um determinado contexto.

Neste contexto, a Semântica de *Frames* se diferencia de outras concepções de desenvolvimento de recursos lexicais multilíngues de várias formas. A principal delas é por ter como central a noção de *frame* como capaz de ativar uma estrutura conceitual que faz com que a compreensão de palavras ou expressões se relacione a um conhecimento cognitivo do todo representado na situação, conforme detalhado no capítulo sobre Semântica de *Frames*.

Assim, nesta etapa do trabalho, a multilinguagem é discutida a partir da perspectiva de sua interdisciplinaridade com a teoria da Semântica de *Frames*. Para falar sobre isso, um importante trabalho desenvolvido por Boas (2009) será revisado. Este linguista, ao falar sobre a criação de recursos lexicais em mais de um idioma, lista alguns dos desafios impostos por trabalhos deste tipo, que são:

- a) **Polissemia:** a possibilidade de uma determinada palavra ter mais de um significado pode representar, por si só, um desafio na construção de um recurso lexical. Ao adicionar a isso a problemática de que duas palavras em idiomas

distintos podem ser também polissêmicas, tem-se um desafio ainda maior. Boas (2009) fala sobre diferentes graus de correspondência em diferentes idiomas e exemplifica isso ao indicar que a palavra *cold*, em inglês, apresenta padrões de polissemia bastante similares aos de *frío*, em espanhol e a *fred*, em catalão. Porém, por outro lado, o verbo *to crawl* em inglês diverge do verbo francês *ramper* em diversos contextos e, portanto, casos como estes devem ser tratados com cuidado no desenvolvimento de bases de dados, pois seus padrões de polissemia não podem ser ignorados.

- b) **Padrões de valência sintática e semântica:** para Boas (2009), além de fornecer os diferentes significados das palavras, uma base de dados lexical multilíngue deve disponibilizar, também, informações sintáticas detalhadas sobre as diferentes formas em que os significados podem ser apresentados. Isso porque, segundo o linguista, somente tendo-se informações sobre a valência de uma palavra torna-se possível identificar seus diferentes sentidos em diferentes contextos. Além disso, ele destaca que, devido à multilinguagem, deve-se mapear os sentidos que as palavras tenham em relação às disponíveis no outro idioma, valendo-se dos padrões de valência observados.
- c) **Diferenças em padrões de lexicalização:** diferentes línguas podem apresentar diferentes padrões de lexicalização e, no desenvolvimento de uma base de dados multilíngue, é necessário apontar os padrões utilizados por cada língua para que a compreensão da utilização de determinada palavra seja possível após a consulta. Boas (2009) cita o exemplo de alguns verbos em japonês que exigem complementos distintos, diferentes do que aconteceria na língua inglesa. Neste contexto, a vantagem da aplicação da Semântica de *Frames* reside em disponibilizar os diferentes padrões de lexicalização do idioma na forma de elementos de *frame* e categoria gramatical a estes vinculados. Assim, utilizando-se mecanismos de mapeamento e de indexação numérica, pode-se “relacionar fragmentos lexicais de acordo com as configurações de elementos de *frame* equivalentes” (BOAS, 2009, p. 90).
- d) **Medição de relações de paráfrases e equivalentes de tradução:** a paráfrase pode ser utilizada para substituir um determinado sintagma por outro de sentido

similar e, quando se trata de um trabalho em mais de um idioma, Boas (2009) salienta que é importante que a base de dados inclua informações sobre como mapear tais sintagmas entre as línguas e como determinar as equivalências de tradução. Além disso, o linguista destaca que, ao se falar em paráfrases, deve-se levar em conta que algumas palavras ou expressões podem ser substituídas por outras sem alterar o significado de uma sentença. Como exemplo, Boas (2009, p. 66) cita as sentenças: (a) Jana *discutiu* com Inge sobre a teoria; e (b) Jana *teve uma discussão* com Inge sobre a teoria<sup>17</sup>. Neste caso nota-se que ambas as sentenças expressam a mesma situação, porém a estrutura para tal é diferenciada. Assim, o linguista explica que “qualquer base de dados utilizada para tradução não deve apenas levar em consideração relações de paráfrase em apenas um idioma, mas deve incluir também uma descrição de como mapear tais paráfrases na relação entre as línguas”.

Através do desenvolvimento de bases de dados multilíngues que levam em consideração a noção de *frames*, nota-se uma preocupação em se destacar mais do que simplesmente a unidade lexical, mas também considerar um conjunto de variáveis que reflete o uso da língua inserida em um contexto. E ao trazer informações como padrões de valência, elementos de *frame*, listagem de ULs evocadoras do *frame*, exemplos de sentenças anotadas, entre outros, as FrameNets abrem um importante caminho para evidenciar as equivalências em nível lexical.

Tendo-se isso em mente, os questionamentos que surgem agora são: em que medida a Semântica de *Frames* e, conseqüentemente, as FrameNets enfrentam os desafios impostos na construção de um recurso lexical em mais de um idioma? E em que medida a correspondência conceitual entre *frames* repercute no nível lexical em termos de equivalência de tradução em um recurso lexical baseado em *frames* no domínio do futebol?

Para responder estas perguntas, inicialmente deve-se considerar que, assim como previamente proposto por Jakobson (1991, p. 69), as “informações requeridas” são diferentes no sistema de cada língua. O fato de frases como “Abriu a porta e saiu” serem gramaticalmente corretas no português não implica necessariamente que a mesma estrutura seja utilizada na língua inglesa. Neste exemplo, notamos, inclusive, que a falta do sujeito expresso na sentença pode causar uma dificuldade de tradução, pois em inglês é necessário

---

<sup>17</sup> (a) Jana *argued* with Inge about the theory.

(b) Jana *had an argument* with Inge about the theory.

que o sujeito esteja presente, como em: “He opened the door and left”. A existência de um contexto pode ser extremamente importante em situações como esta e, por este motivo, é possível afirmar que “quanto mais rico for o contexto de uma mensagem, mais limitada será a perda de informação” (Jakobson, 1991, p. 69). Neste sentido, a Semântica de *Frames* pode auxiliar no papel de fornecer este contexto<sup>18</sup> de uma forma única, pois, durante o alinhamento, ela funcionará como uma *interlíngua* (Boas, 2009).

A proposta de lidar com a Semântica de *Frames* como interlíngua em estudos multilíngues foi sugerida por Boas (2009), tendo como base um estudo comparativo baseado em *frames* entre as unidades lexicais *answer*, do inglês, e *responder*, do espanhol, pertencentes ao *frame* *Communication\_Response*. Ao estudar a tabela de representação das unidades lexicais em questão, Boas (2009, p. 83) chegou à conclusão de que “a Semântica de *Frames* funciona como uma representação interlingual entre as tabelas de valência e realização das unidades lexicais em inglês e espanhol, estabelecendo, assim vínculos entre equivalentes de tradução”. Mais do isso, ele observa também que, embora esta análise demande um trabalho extenso para obtenção de equivalentes, é possível superar alguns tipos de problemas que são comumente encontrados enquanto se cria uma base de dados lexical, como, por exemplo, palavras polissêmicas, padrões de valência divergentes entre as línguas, etc.

Para verificar esta proposição, observam-se as duas frases abaixo que apresentam formas sintaticamente distintas para representar as unidades lexicais *want.v*, em inglês, em comparação com seu equivalente *querer.v*, em português. Ambas as sentenças foram retiradas das bases de dados da FrameNet e da FrameNet Brasil, respectivamente, de exemplos utilizados para apresentar os *frames* *Desiring* e *Desejar*. Este *frame* representa uma situação na qual “um EXPERIENCIADOR deseja que um EVENTO ocorra. Em alguns casos, o EXPERIENCIADOR é um participativo ativo (...) e, em tais casos, o EVENTO por si só geralmente não é mencionado, mas sim, algum PARTICIPANTE\_FOCAL, o qual estará subordinadamente envolvido no EVENTO” (Framenet Brasil, 2012). Neste *frame*, o *Focal\_Participant* (Participante\_Focal) “*them*” atua como objeto do sintagma nominal (NP Obj.) na sentença em inglês, porém, um detalhe é que em português, este mesmo EF acontece sempre precedido de preposição.

---

<sup>18</sup> Segundo Fillmore (1985), para compreender o significado de uma palavra é necessário levar em consideração os contextos linguístico e extralinguístico ligados a ela. O linguista explica também que “a pesquisa em Semântica de *Frames* pode ser vista como um esforço para compreender o processo pelo qual os *frames* são introduzidos em um texto para criar e desenvolver um contexto textual crescente”, ou seja, o contexto leva ao *frame* e este mesmo *frame* ajuda a desenvolver o contexto que envolve o texto. (FILLMORE, 1985, p. 234)

- (1) [*I*<sub>EXPERIENCER</sub>]'ve **WANTED**<sub>TARGET</sub> [*them*<sub>PARTICIPANT\_FOCAL</sub>] [*to fix the sidewalk*<sub>EVENT</sub>] [*for years*<sub>DURATION</sub>].
- (2) [Eu<sub>EXPERIENCIADOR</sub>] **QUERIA**<sub>TARGET</sub> [que eles<sub>PARTICIPANTE\_FOCAL</sub>] [consertassem a calçada<sub>EVENTO</sub>] [há anos<sub>DURAÇÃO</sub>].

Assim, com esta estrutura que é apresentada em uma anotação que segue os preceitos da Semântica de *Frames*, torna-se possível observar como os elementos de *frame* se realizam sintaticamente e semanticamente em cada uma das línguas analisadas, em sua relação com a unidade lexical *target* da sentença. Com isso, a análise das informações disponíveis entre as línguas comparadas pode ser feita de uma forma mais simples e dará acesso a mais dados do que um consulta a um dicionário que não seja baseado em *frames*.

Snell-Hornby (1984), linguista com diversos trabalhos reconhecidos na área de tradução e lexicografia contrastiva, fala sobre a polêmica da menção à interlíngua quando se fala em dicionários bilíngues e menciona sua insatisfação com o fato de estes recursos apresentarem os lexemas isoladamente. Desta forma, os dicionários bilíngues funcionam apenas como um “repositório de lexemas isolados e equivalentes estáticos” (SNELL-HORNBY, 1984, p. 279). Diferente deste modelo, a pesquisadora propõe que, para solucionar este problema, é necessário que os recursos lexicográficos em mais de um idioma apresentem também o sistema dinâmico de relações existentes entre as línguas, a função das palavras inseridas dentro de um contexto e a interdependência entre a língua, a cultura e a interação social que são fixos às palavras.

Anos mais tarde, Snell-Hornby (2005) publica um trabalho que fala sobre *frames* e multilinguagem, com foco na tradução. Este texto destaca a forma como a equivalência de tradução acontece levando-se em conta os *frames* conceituais existentes na mente do autor e do tradutor. Destaca-se aqui como os *frames* é que passam a atuar, através desta descrição, como a interlíngua que relacionará a língua fonte à língua alvo. Isso porque, segundo a autora,

o tradutor parte de um *frame* apresentado (o texto fonte e seus componentes linguísticos), que foi produzido por um autor que o criou a partir do seu próprio repertório de cenas parcialmente prototípicas. Com base no(s) *frame(s)* do texto fonte, o tradutor-leitor cria suas próprias cenas que são ativadas pela sua experiência pessoal e pelo conhecimento internalizado do assunto em questão. (...) Com base nas cenas que ele tenha ativado, o tradutor deve, então, encontrar *frames* adequados na língua alvo; isto envolve um processo constante de tomada de decisões, dependendo de sua proficiência e conhecimento da língua e da cultura alvo. (SNELL-HORNBY, 2005, p. 195)

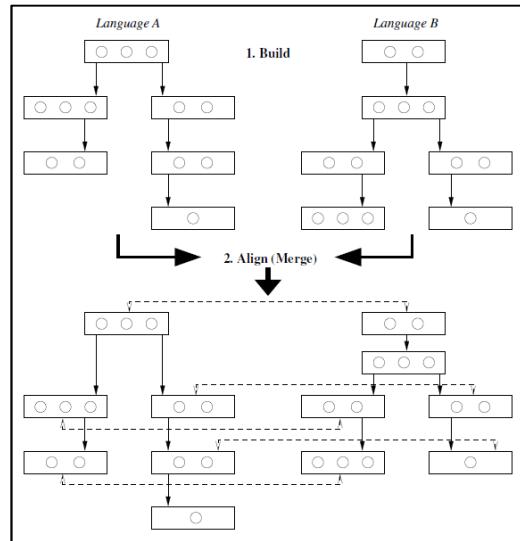
A pesquisadora utiliza a noção de *cena* para destacar o cenário conceitual prototípico em questão, em oposição ao *frame*, que refletiria as escolhas linguísticas utilizadas no mesmo. Chama a atenção neste trecho a relação entre *frame* fonte e alvo trazida pela autora. Nota-se, por exemplo, a importância dada pela autora ao fato de o tradutor conhecer bem a língua e a cultura fonte para que ele consiga encontrar um *frame* adequado na língua alvo para fazer a comparação e definir os equivalentes.

Outra pesquisadora que aborda o assunto de multilinguagem em recursos lexicais, tratando especificamente sobre aqueles baseados em *frames*, é Lönneker-Rodman (2007). Através de seu trabalho a linguista revisita diferentes modelos de construção de recursos lexicais multilíngues e mostra que o que permite a comparação entre as línguas são os blocos independentes que elas têm, mas que apresentam funções similares quando contrastadas entre os idiomas.

Outro aspecto detalhado pela linguista é a forma como as duas línguas são aproximadas, ou seja, a forma como a inter-relação entre elas acontece. Lönneker-Rodman (2007) propõe a existência de dois métodos possíveis para relacionar as línguas quando se trata de um trabalho de construção de bancos de dados baseado em *frames*. O primeiro é o que ele chama de *merge approach* (abordagem de fusão) e o segundo de *expand approach* (abordagem de expansão).

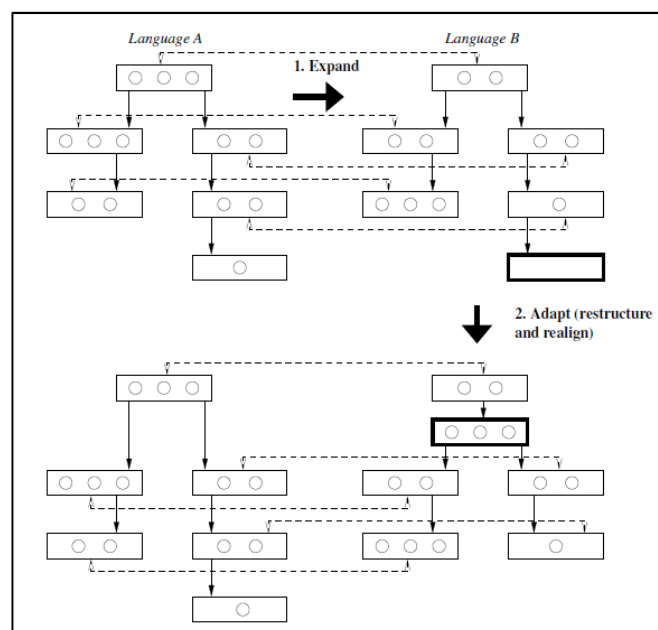
No primeiro caso, na abordagem de fusão, consideram-se inicialmente os recursos provenientes de diferentes línguas independentemente, de modo que, apenas posteriormente, os links e relações entre os componentes sejam definidos. Ou seja, tem-se uma base de dados em um idioma *x* (*Language A*) e outra base de dados no idioma *y* (*Language B*) e relacionam-se ambas, através de equivalências, para que a base de dados multilíngue seja desenvolvida. Enquanto uma das vantagens desta opção seja a liberdade que se tem ao construir o banco de dados em cada uma das línguas, a autora destaca que o estabelecimento de relações entre línguas não é uma tarefa simples, devido à divergência nos sistemas de cada uma. Para exemplificar este tipo de abordagem, segue esquema apresentado pela pesquisadora:





**Figura 4** - Abordagem de fusão segundo Lönneker-Rodman (2007, p. 4)

Já na abordagem de expansão, tem-se uma língua como original e acontece uma transferência do material já existente nela para outro idioma. Neste caso, a dificuldade existe na localização de equivalências para conceitos que são apresentados de diferentes formas em idiomas distintos, pois conta-se inicialmente apenas com uma base de dados em um idioma  $x$  (*Language A*) e, a partir deste, é que as equivalências serão buscadas para o desenvolvimento da base de dados. Optando-se por esta abordagem, é necessário, ao final, verificar-se a estrutura que se tem no idioma  $y$  (*Language B*) e fazer as adaptações necessárias para que ele fique adequadamente apresentado. Abaixo segue esquema para exemplificar como é feita a criação do banco de dados:



**Figura 5** - Abordagem de expansão segundo Lönneker-Rodman (2007, p. 5)

No Kicktionary\_Br Copa 2014, projeto ao qual a presente dissertação está vinculada, utiliza-se como referência a abordagem de fusão, porque se pretende valorizar os traços linguísticos característicos de cada um dos idiomas pesquisados e porque se tem o intuito de não negligenciar questões lexicais e estruturais tipicamente utilizadas nas línguas em questão, português e inglês. Assim, a reflexão sobre as equivalências de tradução será baseada em um *corpus* de textos originais escritos em inglês e outro comparável a este escrito originalmente em português. Maiores detalhes sobre estes *corpora* serão trazidos no capítulo sobre materiais e métodos.

Para concluir esta seção, vale lembrar que, mesmo levando-se em conta os desafios impostos pela multilinguagem, diversos motivos tornam vantajosa a aplicação da Semântica de *Frames* em atividades que envolvem mais de um idioma, como na tradução e na criação de bancos de dados multilíngues. Na área da tradução, especificamente, Snell-Hornby (2005) afirma que os *frames* têm se mostrado um modelo efetivo e inspirador de tratar a multilinguagem ao mostrar como *cenários* fonte e alvo auxiliam no trabalho do tradutor. Já na criação de recursos lexicais, Boas (2009) lista algumas das vantagens em considerar-se a noção de *frames*, que são:

- a forma como a polissemia pode ser evitada em anotações por *frames*, devido ao maior detalhamento disponível na apresentação do *frame*, que torna acessíveis informações sobre o contexto no qual a unidade lexical em questão é utilizada e os elementos de *frame* que participam deste *frame*;

- a desambiguação do significado de palavras através dos padrões de valência entre elas, disponíveis na base de dados;

- a possibilidade de apresentar os diferentes padrões de lexicalização entre as línguas, de forma a destacar os EFs presentes em cada um dos idiomas.

Neste contexto, acredita-se que, mesmo não havendo um grande número de pesquisas que abordam FrameNets e multilinguagem, mais especificamente no que se trata de equivalências de tradução, trata-se de uma área que carece de estudos mais aprofundados. Para Boas (2009, p. 16), esta ligação entre Semântica de *Frames* e multilinguagem ainda é um tema pouco explorado. E, para o linguista, o sucesso obtido até o momento na criação de diversas FrameNets em línguas que não sejam o inglês é um forte indicativo de que este tipo de trabalho deve ser priorizado.

No domínio do futebol, um importante trabalho relacionado a esta pesquisa é o projeto multilíngue Kicktionary, desenvolvido nos idiomas inglês, francês e alemão (Schmidt, 2009),

que será detalhado na subseção seguinte. Este projeto serviu de inspiração para o projeto Kicktionary\_Br, a contraparte em português da base de dados original e que, posteriormente, impulsionou o desenvolvimento do projeto Kicktionary Copa 2014. Deste último faz parte o presente estudo e, portanto, tal recurso também será detalhado.

### 2.3.2 Projeto Kicktionary

O projeto Kicktionary foi desenvolvido pelo pesquisador Thomas Schmidt, ao visitar o ICSI, na Califórnia, entre 2005 e 2006, e disponibiliza um dicionário eletrônico do futebol nos idiomas inglês, alemão e francês, com base nos preceitos da Semântica de *Frames* (Kicktionary, 2012). Segundo descrição no próprio site, o Kicktionary tem como principal objetivo

verificar como teorias linguísticas sobre semântica lexical, métodos da linguística de corpus, tecnologias para hipertexto e hipermedia e técnicas de processamento de linguagem computacional podem ajudar a fornecer recursos lexicais que são melhores do que (ou: bons, de uma forma diferente) os tradicionais dicionários de papel. (Kicktionary, 2012)

E o resultado prático deste trabalho é o Dicionário Kicktionary, disponível online, com acesso irrestrito a qualquer usuário com acesso a internet. Este recurso lexical é similar à FrameNet no sentido que também seguiu os preceitos da Semântica de *Frames*, mas trata-se, segundo Schmidt (2009, p. 101), de uma FrameNet multilíngue e em um domínio específico: o do futebol. Ele recebe esta caracterização porque, diferentemente de um dicionário tradicional, onde a busca seria por verbetes em ordem alfabética, as buscas no Kicktionary podem ser feitas por cenas<sup>19</sup>, *frames* ou unidades lexicais.

Este trabalho motivou a criação do Kicktionary\_Br, que inicialmente propunha a contrapartida deste dicionário em português. Porém, durante sua elaboração diversos aspectos do Kicktionary passaram a ser questionados (a serem discutidos nesta e na próxima seção).

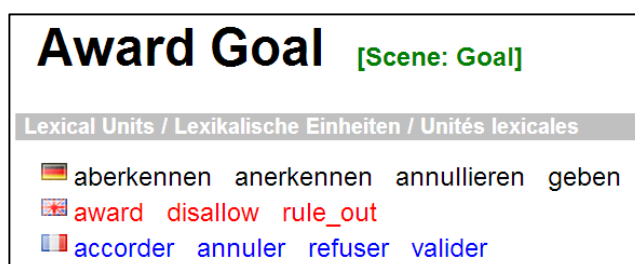
Para desenvolver a estrutura do Kicktionary, Schmidt (2009, p. 103) destaca que aproveitou-se das regras universais existentes no futebol e dos sub-eventos que compõem a partida, comuns mesmo em diferentes culturas, como por exemplo: a falta, a substituição, o gol, o passe, etc. Ao ter-se acesso a estes *frames* disponíveis no Dicionário Kicktionary, tem-

---

<sup>19</sup> Embora Fillmore (2006) tenha abandonado o uso do termo *cenar*, Schmidt (2009) opta por manter a distinção entre *cenar* (entidades conceituais) e *frames* (entidades linguísticas) e, portanto, ambos os termos são utilizados quando se fala sobre o Dicionário Kicktionary e seus desdobramentos.

se, então, acesso também aos *synsets* de cada idioma, diferentemente do que acontece na FrameNet, que não se vale deste conceito.

Os *synsets* (termo utilizados emprestados da terminologia das WordNets) são empregados no Kicktionary como forma de “não apenas agrupar unidades lexicais sinônimas de uma língua (...), mas também, juntar em pares uma dada unidade lexical em uma língua com um possível equivalente de tradução em outra língua” (Kicktionary, 2012). Desta forma, a apresentação dos *synsets* é feita da seguinte forma no dicionário:



**Figura 6** - Exemplo de *synset* no *frame* Award\_Goal

Como se pode notar na figura acima, no caso do inglês, os verbos *award*, *disallow* e *rule\_out* são considerados evocadores do *frame* Award\_Goal. Da mesma forma, os evocadores deste *frame* em outros idiomas também são mencionados. Porém, a relação entre as ULs de cada idioma, ou seja, a relação direta de equivalências, não é disponibilizada neste tipo de apresentação. Isso porque o Kicktionary, embora tenha um forte compromisso com a multilinguagem, não tem compromisso diretamente com a indicação de equivalentes de tradução. Resumidamente, pode-se dizer que o Kicktionary relaciona os *frames* em diferentes idiomas em seu nível conceitual, porém, não tem compromisso com as relações em nível lexical, indicando equivalências entre as ULs. Diferentemente deste, o Kicktionary\_Br Copa 2014 tem compromisso com a indicação de possíveis equivalências e este aspecto é um dos que diferencia estes projetos. Além disso, o Kicktionary\_Br Copa 2014 também opta por trazer a relação entre os *frames*, diferentemente do Kicktionary.

Ainda sobre o não compromisso do Kicktionary com a indicação de equivalentes de tradução, de acordo com informação obtida no site do projeto, existem casos de *synsets* que não apresentam, necessariamente, uma unidade lexical em cada idioma, resultando em uma UL poder ser disponibilizada em apenas um ou dois idiomas. Schmidt exemplifica estes casos com as seguintes ULs:

- 1) {köpfen.v; head.v} (“to play the ball with the head” - no French equivalent)
- 2) {Abspielfehler.n; Fehlpass.n} (“a bad pass” - no English and no French equivalent)

Retoma-se aqui o proposto por Baker (2006) ao afirmar que às vezes as equivalências podem não ser diretas e que, neste caso, é necessário buscar-se por possíveis equivalentes que deem conta de realizar o objetivo da comunicação. No caso das ULs *head* e *köpfen*, mencionadas acima, por exemplo, nota-se que as palavras em inglês e alemão expressam conceitos complexos e que, em francês, seria necessário utilizar mais do que uma palavra para explica-lo. Portanto, o Kicktionary apresenta apenas as palavras em inglês e alemão como partes do *synset*. Este é um dos aspectos que não foi seguido na elaboração do Kicktionary\_Br Copa 2014, pois, por se tratar de um recurso lexical multilíngue, espera-se que ele seja utilizado por usuários na busca de equivalências e, portanto, a indicação de possíveis equivalentes será apresentada. Neste sentido, desafios como o de ULs complexas como as apresentadas deverão ser enfrentados.

Atualmente o Kicktionary conta com aproximadamente mil e novecentas unidades lexicais, sendo elas substantivos, verbos, adjetivos e expressões idiomáticas, em alemão, francês e inglês. Cada uma destas ULs tem entre uma e dez sentenças anotadas como exemplos, todas retiradas do *corpus* composto por *match reports* de partidas de futebol. Aproximadamente cem *frames* foram encontrados pelo pesquisador, e eles estão inseridos em dezesseis cenas, cada uma delas correspondendo a um evento prototípico do jogo de futebol (Kicktionary, 2012).

Este tipo de recurso lexical se diferencia dos tradicionais porque as entradas são apresentadas sempre em sua relação com os *frames* aos quais pertencem, e por não ter como base de consulta apenas a típica listagem de palavras em ordem alfabética. No domínio do futebol, o Kicktionary apresenta o grande plano de fundo do jogo subdividido em três blocos: *on the pitch* (em campo), *a match in a competition* (a partida em uma competição) e *actors and objects* (atores e objetos), que são os diferentes pontos de vista dos quais as cenas e *frames* podem ser observados e, posteriormente, analisados. Segue imagem parcial que mostra algumas das cenas disponíveis na situação “em campo”, com o número de *frames* e de unidades lexicais encontrados por cena, à esquerda, e os *frames* discriminados à direita, juntamente com o número de ULs do *frame*:

<b>Scenes</b>	
<b>On the pitch</b>	
Shot 12 frames 269 lexical units	Ball_And_Goal (11 LUs) Feign (6 LUs) Finish (5 LUs) Follow_Up (7 LUs) Goal_Kickoff (7 LUs) Hit (4 LUs) Intervene (23 LUs) Miss_Goal (7 LUs) Save (46 LUs) Shoot_At (5 LUs) Shot (123 LUs) Shot_Supports (25 LUs)
Pass 11 frames 226 lexical units	Bad_Pass (2 LUs) Being_Free (14 LUs) Connect (5 LUs) Control (37 LUs) Flick_On (7 LUs) Intercept (11 LUs) Mark (20 LUs) Pass (92 LUs) Pass_Back (5 LUs) Pass_Combination (21 LUs) Supply_Pass (12 LUs)
Goal 10 frames 148 lexical units	Award_Goal (11 LUs) Celebrate_Goal (5 LUs) Concede_Goal (7 LUs) Convert_Chance (8 LUs) Goal (85 LUs) Multiple_Goals (8 LUs) Overcome_Goalkeeper (7 LUs) Own_Goal (4 LUs) Prepare_Goal (6 LUs) Score_Goal (7 LUs)
One_On_One 7 frames 129 lexical units	Beat (49 LUs) Challenge (18 LUs) Deny (19 LUs) Lose_Ball (7 LUs) One_On_One (25 LUs) Take_On (5 LUs) Trick (6 LUs)
Foul 12 frames 211 lexical units	Advantage (4 LUs) Concede_Compensation (4 LUs) Dissent (12 LUs) Foul (54 LUs) Give_Card (8 LUs) Offside (8 LUs) Receive_Card (16 LUs) Referee_Decision (12 LUs) Sanction (36 LUs) Set_Piece (45 LUs) Simulation (6 LUs) Win_Compensation (6 LUs)

**Figura 7** - Imagem parcial da consulta por cenas no Kicktionary

Esta figura nos mostra duas das possibilidades de consulta às quais o usuário tem acesso: a cena e o *frame*. Ao clicar no nome da cena, o usuário é direcionado à descrição da cena, com imagens e detalhes sobre a relação entre os *frames* que fazem parte dela. Já ao clicar no nome do *frame*, ele é conduzido a outra tela, com detalhes sobre o *frame* escolhido, incluindo as unidades lexicais que o evocam e os elementos de *frame*. No caso da cena *Foul* (Falta), segundo o exemplo acima, nota-se a subdivisão em doze *frames*, como *Advantage* (Vantagem), *Concede\_Compensation* (Dar\_Compensação), *Dissent* (Discordar), entre outros. Além disso, pode-se observar também o número de unidades lexicais por *frame* e a quantidade de ULs na cena como um todo (neste caso, 211 ULs, dividida entre os *frames*). Diferentemente deste, o Kicktionary\_Br Copa 2014 opta por realizar uma nova estruturação dos *frames* e não adota os três blocos principais apresentados pelo Kicktionary, tendo-se em vista que, nessa pesquisa, o conceito de cena deixa de ser empregado e que esse projeto pretende trabalhar também a relação entre *frames* para disponibilização no dicionários. Uma reflexão sobre este tema será realizada na próxima seção.

Na imagem a seguir (Figura 8) tem-se a tela de consulta do *frame* *Referee\_Decision* (Decisão do juiz) no Kicktionary<sup>20</sup>. Através deste tipo de consulta por *frame*, o usuário tem acesso às unidades lexicais que evocam o *frame* em questão (*ahnden*, *entscheiden*, *award*, *accorder* etc.) nos três idiomas, ou seja, os *synsets*, e, além

<sup>20</sup> Ainda que o projeto Kicktionary\_Br Copa 2014 tenha optado por remodelar a organização de *frames*, isto não significa que algumas descrições do projeto Kicktionary não tenham sido aproveitadas e considerados válidas. Um exemplo é o *frame* *Referee\_Decision*.

disso, aos elementos de *frame* core (únicos EFs indicados pelo Kicktionary). A cor da letra nas ULs e EFs é o que diferencia as palavras pertencentes a cada idioma.

Referee Decision [Scene: Foul]							
Lexical Units / Lexikalische Einheiten / Unités lexicales							
ahnden	entscheiden	erkennen	Fehlentscheidung	geben	pfleifen	zusprechen	
award	rule						
accorder	sanctionner	siffler					
Frame elements / Frame-Elemente / Éléments de frame							
	1. REFEREE	2. OFFENSE	3. COMPENSATION	4. OFFENDED_TEAM	5. CARD	6. OFFENDER	7. OFFENDED_PLAYER
	*	1	2	3	4	5	6
ahnden.v							
entscheiden.v							
erkennen.v							
Fehlentscheidung.n							
geben.v							
pfleifen.v							
zusprechen.v							
award.v							
rule.v							
accorder.v							
sanctionner.v							
siffler.v							

Figura 8 - Apresentação do *frame* Referee\_Decision no Kicktionary

Na tela de apresentação do *frame* (Figura 9), está disponível ao usuário uma lista com as ULs que evocam o *frame* para que ele clique na de seu interesse e obtenha mais informações sobre ela. Para exemplificar, segue abaixo o resultado da busca por detalhes da UL *award*. São apresentados exemplos retirados do *corpus*, com os EFs marcados em vermelho, relações de sinônimos com ULs nos outros idiomas (em alguns casos) e a listagem de EFs encontrados no *corpus*. Esta é a forma mais próxima de observarem-se os equivalentes de tradução, pois o acesso já está restrito à consulta a uma UL. Destaca-se que as informações disponibilizadas nesta tela (lista de ULs por *frame*, exemplos, listagem de EFs,) serão também apresentadas no Kicktionary\_Br Copa 2014, com a inclusão da indicação de possíveis equivalentes.

award (verb)		Scene Foul	Frame Referee_Decision
<b>Examples</b>			
1.	Manninger dived on to a back pass and [referee Luboš Michel] <sub>REFEREE</sub> <b>awarded</b> [an indirect free-kick ten metres out] <sub>COMPENSATION</sub> . [75325 / p5]		
2.	Povilas Lukšys could have doubled the lead ten minutes later but was foiled by Joubert before [the Luxembourg outfit] <sub>OFFENDED_TEAM</sub> <b>were awarded</b> [a penalty] <sub>COMPENSATION</sub> against the run of play. [79065 / p3]		
3.	On 71 minutes [Terek] <sub>OFFENDED_TEAM</sub> <b>were awarded</b> [a penalty] <sub>COMPENSATION</sub> [after Mariusz Mowlik's handball] <sub>OFFENSE</sub> , but Khomukha's spot-kick was weak and Piatek easily parried. [79286 / p5]		
4.	But [the referee] <sub>REFEREE</sub> <b>had already awarded</b> [a free-kick] <sub>COMPENSATION</sub> after adjudging Ciprian Marica to have fouled in the build up. [1077155 / p6]		
<b>Semantic relations</b>			
Synonyms	zusprechen	geben	entscheiden
	award		erkennen
	accorder		
<b>Frame elements</b>			
	COMPENSATION [Set_Piece]	OFFENDED_TEAM [Team]	OFFENSE [Offense] REFEREE [Referee]
LU	COMPENSATION	OFFENDED TEAM	OFFENSE
<b>awarded</b>	an indirect free-kick ten...		referee Luboš Michel
<b>were awarded</b>	a penalty	the Luxembourg outfit	
<b>were awarded</b>	a penalty	Terek	after Mariusz Mowlik's ha...
<b>had awarded</b>	a free-kick		the referee

Figura 9 - Apresentação da UL *award* no Kicktionary

Cabe ressaltar ainda que, por se tratar de um recurso eletrônico composto por diversos hyperlinks, a qualquer momento o usuário pode retornar à tela com as cenas principais, consultar a descrição da cena e, também, passar da consulta de uma UL do *frame* para outra, até mesmo em outro idioma.

O *corpus* utilizado no projeto é composto por *match reports* retirados do site da UEFA (www.uefa.com), sendo que quinhentos textos foram obtidos em cada um dos idiomas, resultando num total de aproximadamente duzentas mil palavras por idioma. Uma parte destes textos é tradução direta de outros, formando, portanto, um *corpus* paralelo. No idioma alemão, o *corpus* contém também textos de outro site (www.kicker.de), além de aproximadamente uma hora de transcrições de comentários de futebol de um rádio alemão (Kicktionary, 2012).

Também a respeito do *corpus*, uma das seções do dicionário disponibiliza o acesso a três textos paralelos anotados, nos idiomas inglês, francês e alemão. Novamente não existe compromisso com a indicação de equivalentes de tradução, mas existe um link que é capaz de direcionar o consulente do texto diretamente a consulta da UL que deseja, ao lê-la na anotação realizada no texto paralelo.

Para Schmidt (2009, p. 129), diversas são as vantagens da criação do Kicktionary:

- a hierarquia de cenas e *frames* é uma forma eficiente de agrupar palavras de um domínio específico, como o futebol;
- a forma como as anotações são feitas disponibiliza uma forma eficiente de incluir linguagem empírica em um dicionário;
- a noção de cenas e *frames* se encaixa muito bem à criação de um recurso *multilíngue* que pode ser utilizado na tradução;
- uma análise por cenas e *frames* é mais fácil e eficiente nas áreas às quais o vocabulário trate de atividades dinâmicas;
- o conceito de cena informando sobre eventos prototípicos dá àqueles que desenvolvem dicionários um espaço para refletir sobre a utilização de elementos multimídia, como imagens, filmes, o que pode ajudar a compreensão de palavras em uma língua estrangeira.

Assim, tendo-se, nesta seção, refletido sobre a estrutura e as características principais do Dicionário Kicktionary e a forma como este tratou da multilinguagem, torna-se possível, a partir daqui, discutir sobre a forma com que o Kicktionary\_Br Copa 2014 a trata. Evidencia-se que o Kicktionary\_Br Copa 2014 tem um compromisso com a apresentação de equivalências de tradução, diferentemente do projeto original e, portanto, embora muitas



características das apresentadas nesta seção sejam seguidas, algumas alterações foram necessárias e serão discutidas no próximo subcapítulo.

### 2.3.3 Projeto Kicktionary\_Br Copa 2014

Conforme discutido na seção anterior, o projeto Kicktionary colaborou com estudos de criação de recursos lexicais baseados em *frames* em domínios específicos e, particularmente, motivou o grupo SemanTec (Semântica e Tecnologia) de pesquisadores da Universidade do Vale do Rio dos Sinos (UNISINOS)<sup>21</sup> a iniciar um projeto próprio. Este trabalho foi denominado Kicktionary\_Br e não teve compromisso em realizar uma revisão crítica sobre o proposto por Schmidt no Kicktionary, pois seu foco era, inicialmente, construir um corpus com anotação semântica. Neste sentido, as categorias semânticas propostas pelo projeto alemão foram adotadas na etiquetagem do corpus. Porém, um desdobramento deste projeto foi impulsionado em 2012: o Kicktionary\_Br Copa 2014, que, ao contrário do projeto anterior, assumiu o compromisso de construir um recurso lexical bilíngue (português/inglês) e em formato eletrônico, abrangendo o domínio do futebol e baseado na teoria da Semântica de *Frames*.

Assim como o Kicktionary, o Kicktionary\_Br Copa 2014 também conta com um *corpus* composto por sentenças obtidas em notícias de jogos de futebol (*match reports*). Na fase monolíngue do projeto a proposta inicial era de compilar-se um *corpus* em português a ser posteriormente utilizado na etapa bilíngue do projeto. Este *corpus* em português seria comparável<sup>22</sup> ao *corpus* disponível em inglês no Kicktionary e, desta forma, seria possível tratar-se dos equivalentes. Porém, o grupo de pesquisa observou que estes *corpora* não seriam suficientes, devido à insuficiência de material disponível para consulta e exemplificação e, portanto, estes *corpora* foram replanejados. Os *corpora* utilizados para a pesquisa de equivalências são detalhados no próximo capítulo, sobre *matérias e métodos*.

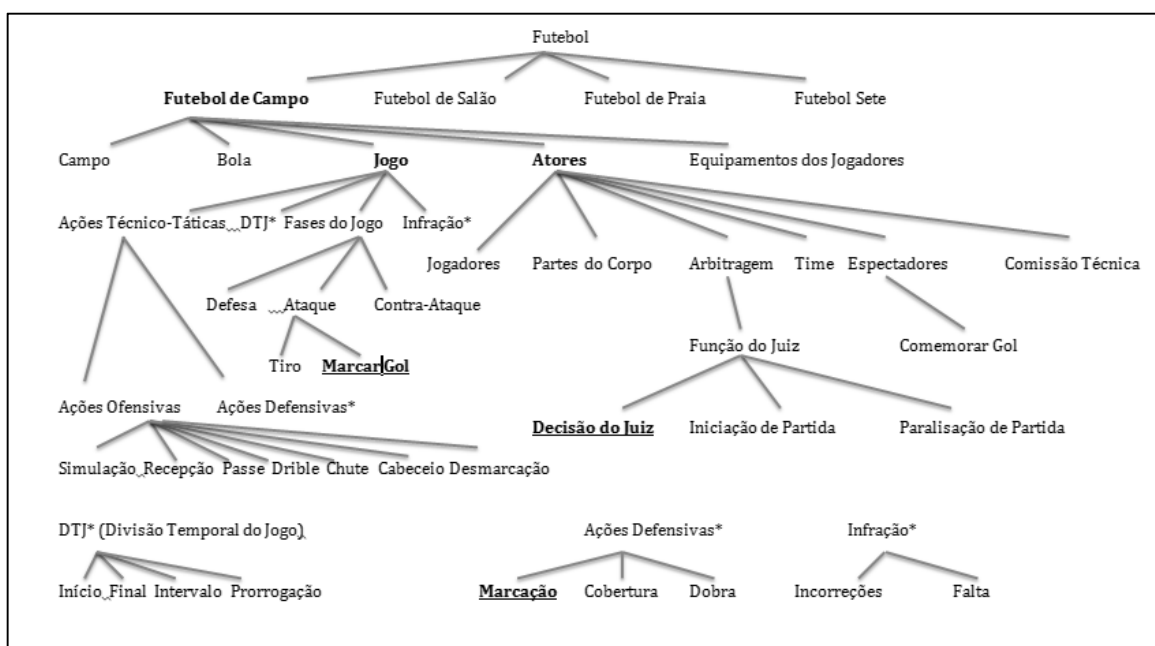
---

<sup>21</sup> A equipe do Grupo Kicktionary é liderada pela Prof. Dra. Rove Chishman, e conta com a colaboração dos professores Anderson Bertoldi, Sandro Rigo, além de um grupo de alunos de mestrado (Diego Spader e João Gabriel Padilha) e de iniciação científica (Brenda Bastos, Ana LuizaVianna, Aline Nardes, George Mileski e Lurdes Gava). Mais detalhes disponíveis na página do projeto em <http://projeto.unisinos.br/semantec/>.

<sup>22</sup> Por *corpus* comparável entende-se uma coletânea de textos originais, em idiomas distintos, sobre um mesmo assunto, pertencentes a um mesmo gênero, de tipologia similar e escritos em um período de tempo comparável. (TAGNIN, 2007).

Ainda sobre o *corpus*, chama-se atenção para a diferença na procedência dos textos no Kicktionary e no Kicktionary\_Br Copa 2014. O primeiro conta com um *corpus* principal obtido diretamente no site da União das Federações Europeias de Futebol (UEFA)<sup>23</sup>, nos idiomas Alemão, Inglês e Francês. Trata-se de traduções do texto, o que o caracteriza como *corpus* paralelo. Já o Kicktionary\_Br Copa 2014 por ter como uma das frentes de trabalho o estudo de equivalências de tradução, conta com uma coletânea de textos originais, escritos nos idiomas português e inglês, em sites a serem detalhados no próximo capítulo.

O modo com que Schmidt apresenta as cenas e *frames* no Kicktionary serve como inspiração para o Kicktionary\_Br Copa 2014, porém algumas considerações e adaptações foram realizadas. Duas importantes mudanças dizem respeito à eliminação da noção de *cenar* e à reflexão a respeito da remodelagem dos *frames* com vistas ao detalhamento da relação entre *frames*. Embora o *script* de um jogo de futebol seja mundialmente pré-estabelecido, observou-se que a subdivisão proposta pelo Kicktionary não dá conta de algumas especificidades como a relação entre *frames*, o que gerou dificuldade de análise de *frames* como o Gol, por exemplo. A figura a seguir ilustra a remodelagem de frames proposta pelo grupo de pesquisa, em sua primeira estruturação<sup>24</sup>.



**Figura 10** – Remodelagem de *frames* em sua primeira versão

<sup>23</sup> Disponível em: <www.uefa.com> Acesso em 10 jun. 2013.

<sup>24</sup> O objetivo da ilustração é apenas apresentar a proposta de mapeamento de frames, sem pretensão de indicar relações habitualmente ilustradas em um mapa conceitual ou ontologia, como é o caso de hipônimos e merônimos.

Além disso, outra questão a ser estudada pelo grupo de pesquisa, diferentemente do projeto original, é a implicação lexicográfica do recurso Kicktionary\_Br Copa 2014, ou seja, a caracterização dos *frames*, descrição e nomeação dos EFs, seleção das ULs, seleção de sentenças como exemplo e construção das entradas lexicais e glosas tendo-se em vista a disponibilização em formato de dicionário. Embora este seja um aspecto importante da pesquisa, não é o foco desta dissertação, que se propõe a refletir sobre as equivalências tradutórias.

Sobre as equivalências no Kicktionary\_Br Copa 2014, espera-se que, embora os *frames* auxiliem na relação no nível conceitual entre as línguas, o nível lexical envolva mais do que a simples tradução das unidades lexicais em questão. Piotrowsi (1994, p. 110) adverte que “o que exige maior atenção do que os termos na tradução técnica, é a ‘colocabilidade’ dos termos”, ou seja, não basta ter-se a equivalência de um termo por si só, mas também apresentá-las juntamente com as expressões às quais os termos estão associados. Neste sentido a noção de *frames* volta a ser destacada, por fornecer um contexto e também uma série de exemplos de utilização das ULs.

Para concluir, destaca-se que, devido à globalização, linguistas passaram a ver na criação de bancos de dados multilíngues uma forma de oferecer ferramentas que podem auxiliar em tarefas como a tradução humana e automática, além da recuperação de informações e até mesmo o ensino de língua estrangeira (Boas, 2009, p. 59). Porém, a criação deste tipo de recurso não é algo simples de ser realizado e, para o autor, uma das principais dificuldades é o desenvolvimento de um esquema capaz de englobar aspectos como estruturas polissêmicas, equivalentes de tradução, etc. Acredita-se que a Semântica de *Frames* possa somar neste aspecto, ao disponibilizar uma nova forma para se tratar de equivalências de tradução. Esta hipótese será testada no capítulo de análise de dados.

### 3 MATERIAIS E MÉTODOS

Como se pretende, neste trabalho, refletir sobre a forma como os equivalentes de tradução são abordados no âmbito de um projeto multilíngue baseado nos preceitos da Semântica de *Frames*, faz-se necessário observar como as equivalências podem ser explicitadas em um recurso como este, para que se possam tecer algumas considerações. Desta forma, este capítulo detalha os materiais a serem utilizados e o método que guia esta análise.

Para isso, o presente capítulo divide-se em duas subseções:

- 1) Materiais: descrição dos *corpora* utilizados na pesquisa e ferramenta de análise de corpus (WordSmith Tools);
- 2) Métodos: detalhamento sobre o passo-a-passo para verificação de possíveis equivalências e discussão sobre sua repercussão em um dicionário temático bilíngue baseado em *frames*.

Destaca-se aqui que, neste capítulo, não se pretende abranger a apresentação lexicográfica de tais equivalentes e nem discussões questões terminológicas, devido ao próprio escopo desta pesquisa e à restrição de tempo para tal aprofundamento. Porém, será apresentado um modelo de apresentação das equivalências no dicionário (sem compromisso com a apresentação lexicográfica) para demonstrar como os resultados obtidos nesta pesquisa podem ser pensados a partir daqui.

#### 3.1 MATERIAIS

Para refletir sobre equivalências, neste caso nos idiomas inglês e português em um dicionário bilíngue baseado em *frames*, pode-se contar, atualmente, com o suporte de um recurso que tem se mostrado útil em pesquisas na área da tradução: a pesquisa em *corpus*. Este subcapítulo tem como objetivo refletir sobre o uso de *corpus* em pesquisas como essa e apresentar os *corpora* comparáveis em inglês e português do domínio do futebol que serão consultados na análise e discussão dos dados desta dissertação.

### 3.1.1 *Corpora* da pesquisa

Assim como acontece na FrameNet e em projetos lexicográficos que surgiram a partir dela, o Kicktionary\_Br Copa 2014 caracteriza-se como um estudo baseado em evidência empírica, no qual os textos para análise são autênticos e disponibilizados em diferentes meios. Assim, ao aplicar-se a Semântica de *Frames* no desenvolvimento de recursos lexicais, as sentenças que exemplificam a utilização das ULs não são criadas pelos linguistas que trabalham no projeto, mas extraídas de textos reais, mesmo quando algumas adaptações necessitem ser feitas com fim lexicográfico.

Para falar sobre isso, inicialmente é necessário retomar o que Berber Sardinha (2000, p. 335-336) apresenta como *corpus*. Para ele, um *corpus* é uma coletânea de dados linguísticos naturais, ou seja, autênticos e produzidos por humanos, construída com um objetivo específico de pesquisa. O linguista cita também que o *corpus* é uma coletânea que retrata porções de linguagem que são coletadas com um objetivo e com critérios específicos para servir como uma amostra de como a língua funciona. Esta vinculação do *corpus* à seleção de uma amostra capaz de resolver alguma questão linguística faz parte da preocupação de linguistas que utilizam esta ferramenta. Até porque uma das maiores contribuições da utilização do *corpus* é que ele evidencia a linguagem em uso.

A utilização de *corpus* no desenvolvimento de dicionários iniciou no início da década de oitenta (Atkins e Rundell, 2008, p. 53) e, desde então, diversas pesquisas foram realizadas para provar como tal prática é importante. Isso porque, segundo Atkins e Rundell (2008), a utilização de dados autênticos é um pré-requisito para um dicionário ser considerado confiável. Porém, eles também citam que a quantidade de informação obtida precisa ser representativa, caso contrário, a insuficiência de dados pode limitar a apresentação real da língua.

Tendo-se isso em mente, notou-se, nesta pesquisa, a necessidade de contar-se com dois *corpora* comparáveis, sendo que ambos apresentariam características comuns, mas cada um deles formado por textos em diferentes idiomas: inglês e português, línguas de consulta no Dicionário Kicktionary\_Br Copa 2014. A opção pela utilização de *corpus* comparável nesta pesquisa parte do pressuposto de que *corpus* paralelo pode não ser a melhor opção ao pesquisar-se sobre equivalências. Isso porque cada língua possui características próprias, assim como mencionado anteriormente, e ao traduzir-se um texto, estas propriedades únicas podem se perder. Desta forma, pode-se dizer que “um *corpus* comparável, composto de textos

originalmente escritos nas duas línguas (...) fornecerá o termo efetivamente usado por aquela comunidade” (TAGNIN, 2007, p. 2), sendo mais adequado para o trabalho de observação de equivalências.

A coleta destes *corpora* foi realizada em conjunto com o grupo de pesquisa Kicktictionary\_Br Copa 2014 e levou em consideração os seguintes aspectos listados por Berber Sardinha (2000) como essenciais no planejamento de compilação e organização de um *corpus* para estudo:

- (a) A origem: os dados devem ser autênticos
  - (b) O propósito: o corpus deve ter a finalidade de ser um objeto de estudo linguístico
  - (c) A composição: o conteúdo do corpus deve ser criteriosamente escolhido
  - (d) A formatação: os dados do corpus devem ser legíveis por computador
  - (e) A representatividade: o corpus deve ser representativo de uma língua ou variedade
  - (f) A extensão: o corpus deve ser vasto para ser representativo
- (BERBER SARDINHA, 2000, p. 30)

Cada um destes tópicos se justifica pelo objetivo de pesquisa que o linguista tem ao compilá-lo. Não se pode garantir a qualidade de um *corpus* de análise que não tenha nele dados autênticos, pois, do contrário, seus dados serão duvidosos. Além disso, sua compilação deve levar em conta o objetivo da pesquisa para que sejam buscadas formas de observar como a língua se comporta. E, por último, sem se ter uma boa quantidade de material para analisar, torna-se difícil fazer uma descoberta que não seja aberta a questionamentos.

Estes aspectos serviram como guia para a compilação dos *corpora* para o Kicktictionary\_Br Copa 2014 e definiu-se, portanto, que os *corpora* utilizados no projeto seguiriam os seguintes critérios:

- Dois *corpora* comparáveis, sendo um deles em inglês e o outro em português;
- Origem: dados autênticos extraídos de sites escritos originalmente no idioma de cada *corpus*;
- Propósito: ilustrar os *frames* do futebol;
- Composição: relatórios sobre partidas de futebol disponibilizados em sites de times, notícias sobre jogos, relatos de jogos tipo minuto-a-minuto e notícias sobre jogos das seleções na última Copa do Mundo FIFA, em 2010;
- Formatação: textos salvos em formato .TXT, divididos por tipo de texto fonte e por idioma. O *corpus* em inglês é composto por quatro arquivos: TimesEN.txt, NoticiasEN.txt, Minutu-a-minutoEN.txt e CopaEN.txt. Em português o *corpus* foi salvo como: TimesPT.txt,

NotíciasPT.txt, Minutu-a-minutoPT.txt e CopaPT.txt. Esta divisão foi feita para que se pudesse verificar se, durante a análise do *corpus*, alguma estrutura aparecia apenas em um tipo de texto mais característico por exemplo e, também, devido à extensão do *corpus*.

- Representatividade: compilação realizada em diversos sites, com o objetivo de representar os frames do futebol;
- Extensão: aproximadamente 1 milhão de palavras<sup>25</sup> por idioma.

Acredita-se que, com estas características, os *corpora* sejam capazes de fornecer os dados necessários à pesquisa. Por contar-se com dados autênticos extraídos de sites escritos originalmente nos idiomas do dicionário, tem-se acesso a amostras da língua como ela realmente acontece naquele idioma. E, também, ao contar-se com aproximadamente 1 milhão de palavras em cada idioma salvas em formato .TXT, tem-se dois *corpora* de quantidade representativa, que pode ser lido facilmente pelo computador e que permite a identificação de padrões de valência, aspecto central no arcabouço das FrameNets.

A seguir seguem detalhes sobre cada um dos *corpora* da pesquisa: o *corpus* em inglês e o *corpus* em português.

### 3.1.1.1 *Corpus* em inglês

Conforme informado anteriormente, teve-se o cuidado de coletar textos escritos originalmente em inglês, para que o texto não contivesse traduções ou interferências de outras línguas em sua apresentação. Além disso, a variedade da composição dos textos também foi algo cuidadosamente planejado para que a representatividade da língua fosse o mais fiel possível aos diferentes tipos de relatos neste domínio específico. Tendo-se isso em vista, os textos foram todos obtidos em meio eletrônico e os sites utilizados foram os seguintes:

- 1) Relatórios sobre partidas de futebol: obtidos nos sites oficiais dos times Arsenal F.C. ([www.arsenal.com](http://www.arsenal.com)), Aston Villa F.C. ([www.avfc.co.uk](http://www.avfc.co.uk)), Chelsea F.C. ([www.chelseafc.com](http://www.chelseafc.com)), Everton F.C. ([www.evertonfc.com](http://www.evertonfc.com)), Fulham F.C.

---

<sup>25</sup> Atualmente, a definição sobre o tamanho do *corpus* já não é mais um assunto tão discutido porque, segundo Atkins e Rundell (2008, p. 57), a obtenção de grandes quantidades de dados tornou-se mais fácil graças à era digital. Porém, como o *corpus* tem o papel de representar a linguagem de alguma forma, sua extensão é o que garante a amplitude de seu alcance e, portanto, quanto maior o tamanho do *corpus*, mais chances de encontrar palavras que ocorram com frequências menores (BERBER SARDINHA, 2000).

(www.fulhamfc.com), Liverpool F.C. (www.liverpoolfc.com), Manchester United F.C. (www.manutd.com), Manchester City F.C. (www.mcfc.com), Newcastle United F.C. (www.nufc.com) e Tottenham Hotspur F.C. (www.tottenhamhotspur.com).

- 2) Notícias sobre jogos: obtidas nos sites: BBC (www.bbc.co.uk/sport), Daily Mail (www.dailymail.co.uk), Independent (www.independent.co.uk), The Guardian (www.guardian.co.uk), ESPN (www.espn.co.uk), Four Four Two (fourfourtwo.com), Eye Football (www.eyefootball.com), Football 365 (www.football365.com), Football (www.football.co.uk) e Mirror (www.mirror.co.uk/sport)
- 3) Relatos de jogos tipo minuto-a-minuto: BBC (www.bbc.co.uk/football) e The Guardian (www.guardian.co.uk).
- 4) Notícias sobre jogos das seleções na última Copa do Mundo FIFA em 2010: EuroSport (uk.eurosport.yahoo.com/football/world-cup) e BBC (news.bbc.co.uk/sport2/hi/football/world\_cup\_2010).

A quantidade de palavras compiladas por tipo de texto foi planejada para seguir o seguinte padrão:

<b>Composição</b>	<b>Número de palavras</b>
Relatórios sobre partidas de futebol	300 mil palavras
Notícias sobre jogos	300 mil palavras
Relatos de jogos tipo minuto-a-minuto	200 mil palavras
Notícias sobre jogos das seleções na última Copa do Mundo FIFA em 2010	200 mil palavras

**Tabela 1** - Número de palavras por tipo de texto no *corpus* em inglês

Com isso, espera-se ter um bom balanço entre os diferentes tipos de textos e, conseqüentemente, ter-se uma grande variedade de exemplos de uso de cada uma das ULs a ser analisada para tratar sobre equivalências. Outro aspecto diz respeito ao nível de formalidade e diferentes estruturas que cada um dos sites pode ter como característica, o que pode ser valorizado devido à quantidade de sites que compõem o *corpus*.



### 3.1.1.2 *Corpus* em português

O *corpus* em português tem as mesmas características do compilado em inglês, ou seja, trata-se de uma coletânea de textos autênticos, extraídos de sites escritos originalmente em português, composto por relatórios sobre partidas de futebol, notícias sobre jogos, relatos de jogos tipo minuto-a-minuto e notícias sobre jogos das seleções na última Copa do Mundo FIFA, em 2010. Os textos são salvos em .TXT e formam um *corpus* de aproximadamente 1 milhão de palavras.

Os textos do *corpus* em português foram retirados dos seguintes sites:

- 1) Relatórios sobre partidas de futebol: obtidos nos sites oficiais dos times São Paulo F.C. ([www.saopaulofc.net](http://www.saopaulofc.net)), Santos F.C. ([www.santosfc.com.br](http://www.santosfc.com.br)), Cruzeiro E.C. ([www.cruzeiro.com.br](http://www.cruzeiro.com.br)), S.C. Internacional ([www.internacional.com.br](http://www.internacional.com.br)), S.C. Corinthians Paulista ([www.corinthians.com.br](http://www.corinthians.com.br)), Grêmio ([www.gremio.net](http://www.gremio.net)), Fluminense F.C. ([www.fluminense.com.br](http://www.fluminense.com.br)), Flamengo ([www.flamengo.com.br](http://www.flamengo.com.br)), Palmeiras ([www.palmeiras.com.br](http://www.palmeiras.com.br)) e Vasco da Gama ([www.crvascodagama.com](http://www.crvascodagama.com)).
- 2) Notícias sobre jogos obtidas nos sites: O Globo ([oglobo.globo.com](http://oglobo.globo.com)), ESPN ([www.espn.com.br](http://www.espn.com.br)), Folha UOL ([www1.folha.uol.com.br/esporte](http://www1.folha.uol.com.br/esporte)), Lance Net ([www.lancenet.com.br](http://www.lancenet.com.br)), Futebol na Rede ([www.futebolnarede.com.br](http://www.futebolnarede.com.br)), Gazeta Esportiva ([www.gazetaesportiva.net](http://www.gazetaesportiva.net)), Globo Esporte ([globoesporte.globo.com](http://globoesporte.globo.com)), Placar ([placar.abril.com.br](http://placar.abril.com.br)), Esporte UOL ([esporte.uol.com.br](http://esporte.uol.com.br)) e Confederação Brasileira de Futebol ([www.cbf.com.br](http://www.cbf.com.br))
- 5) Relatos de jogos tipo minuto-a-minuto: Terra Esporte ([esportes.terra.com.br](http://esportes.terra.com.br)), Diário Catarinense ([diariocatarinense.clicrbs.com.br/sc/jogo-ao-vivo](http://diariocatarinense.clicrbs.com.br/sc/jogo-ao-vivo)) e Zero Hora ([zerohora.clicrbs.com.br/rs/esportes](http://zerohora.clicrbs.com.br/rs/esportes)).
- 6) Notícias sobre jogos das seleções na última Copa do Mundo FIFA em 2010: Terra Esportes ([esportes.terra.com.br/futebol/copa/2010](http://esportes.terra.com.br/futebol/copa/2010)), Copa do Mundo UOL ([copadomundo.uol.com.br/2010/ultimas-noticias](http://copadomundo.uol.com.br/2010/ultimas-noticias)) e Notícias UOL ([noticias.bol.uol.com.br/copa/2010/tabelas](http://noticias.bol.uol.com.br/copa/2010/tabelas)).

O *corpus* em português foi planejado para seguir o mesmo padrão de número de palavras do *corpus* em inglês, ou seja, teve como base a seguinte tabela:

Composição	Número de palavras
Relatórios sobre partidas de futebol	300 mil palavras
Notícias sobre jogos	300 mil palavras
Relatos de jogos tipo minuto-a-minuto	200 mil palavras
Notícias sobre jogos das seleções na última Copa do Mundo FIFA em 2010	200 mil palavras

**Tabela 2** - Número de palavras por tipo de texto no *corpus* em português

A definição das especificações dos *corpora* a serem utilizados nesta pesquisa foi feita com o auxílio dos integrantes do grupo de pesquisa Kicktionary\_Br Copa 2014. Isso porque o trabalho com Linguística de *Corpus* só passou a ser aprofundado a partir do momento em que a pesquisa do projeto passou para o nível lexicográfico e multilíngue. Lexicográfico porque o Dicionário da Copa prevê a seleção de ULs para as entradas no dicionário, assim como sua descrição, e multilíngue porque o trabalho com *corpus* pode ser de grande auxílio na observação de equivalências.

### 3.1.2 Ferramenta de análise do *corpus*: WordSmith Tools

A ferramenta utilizada para análise do *corpus* foi o WordSmith Tools. Este programa, criado em 1996 por Mike Scott, da Universidade de Liverpool, é destinado à análise linguística, e, mais especificamente, à realização de análises baseadas na frequência de palavras em *corpus* e em suas co-ocorrências (Berber Sardinha, 2009).

Qualquer pessoa pode baixar a versão Demo do programa gratuitamente, disponibilizada no site do linguista Mike Scott<sup>26</sup>. Esta versão oferece algumas limitações principalmente quanto à quantidade de palavras que podem ser processadas, mas dá acesso às principais ferramentas que são *WordList*, *Keywords* e *Concord*.

Segundo o site do programa, estas ferramentas funcionam, de forma geral, do seguinte modo:

<sup>26</sup> Disponível em <http://www.lexically.net/wordsmith/index.html>.

- *WordList*: mostra o quão frequentemente cada palavra aparece em cada arquivo de texto, o percentual das palavras escolhidas no texto, e em quantos arquivos de texto cada palavra foi encontrada.
- *Keywords*: lista as palavras-chave do texto, mostrando sua frequência e disponibilizando a comparação com a frequência em um corpus de referência.
- *Concord*: apresenta a concordância de todas as ocorrências de uma palavra em um determinado texto.

Nesta pesquisa, a ferramenta mais utilizada é o *Concord*, pois se tem como objetivo observar a representatividade de determinadas ULs em cada um dos *corpora* da pesquisa e também analisar sua co-ocorrência. Por este motivo, tal ferramenta é detalhada em seguida. A ferramenta *WordList* também foi utilizada, mas apenas para gerar a lista de palavras mais frequentes nos *corpora* e será mencionada apenas na primeira parte da análise.

Diversas são as vantagens da utilização do WordSmith Tools, como, por exemplo: sua gratuidade (ainda que falando-se da versão Demo); a facilidade de usá-lo, por ter uma interface simples; a rapidez em processar as buscas e apresentar os resultados.

### 3.1.2.1 Ferramenta *Concord* do WordSmith Tools

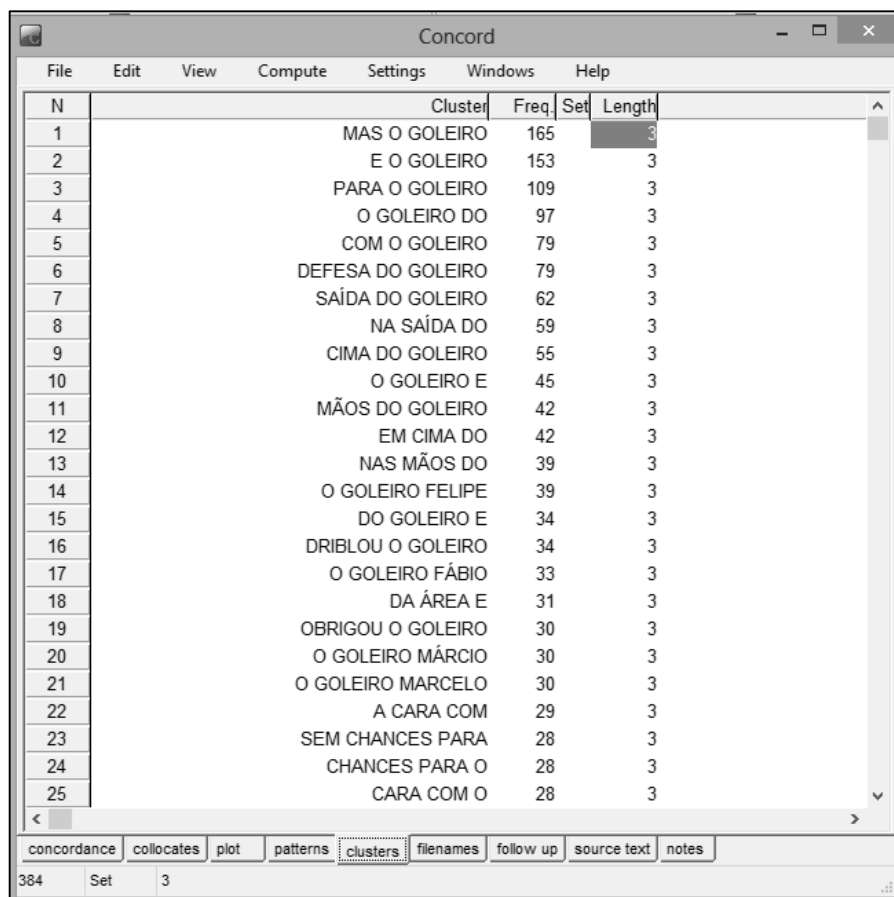
O *Concord* é o programa do WordSmith Tools que lista as ocorrências de uma determinada palavra ou expressão consultada, inserida no texto que ela tiver ao seu redor. Como exemplo, trazemos abaixo a consulta à UL *goleiro*.

Line	Text Excerpt	Word #/len	Sen
1	escanteio de Eduardo. No último lance da partida, Fred ainda cobrou falta nas mãos do goleiro Adilson Junior, que não segurou firme, porém não havia nenhum jogador de	52,950,068	035%
2	surgiu aos 15 minutos, após bom lançamento de Jean para Wellington Nem, mas o goleiro Jefferson se antecipou bem à jogada. O Botafogo ariscou chutes de fora da área,	53,066,077	035%
3	deles, Diego Cavalieri mais uma vez foi fundamental em campo. Logo aos 6 minutos, o goleiro impediu um gol de cabeça do atacante Rômulo ao se esticar e mostrar seu	52,639,043	035%
4	, respectivamente. Aos 13 minutos, depois de cobrança de escanteio de Wagner, o goleiro Adilson foi forçado a sair do gol para afastar a bola. Porém, a bola cobrou para os	52,783,054	035%
5	um bom lançamento de Jean para Nem, que chutou em cima de Jefferson. No rebote, o goleiro defendeu o chute de Jean. Aos 27 minutos, Seedorf cobrou falta e Cavalieri tocou	53,281,065	035%
6	Braga. Aos 18 minutos, o volante Fábio ariscou chute da entrada da área e assustou o goleiro Jefferson, que conseguiu defender. Poucos minutos depois, foi a vez de Ricardo	53,860,117	035%
7	chutou forte. Berna espalmou e o atacante Glauber quase pegou o rebote, mas o goleiro tricolor conseguiu se recuperar e fazer nova defesa. Após o intervalo, o time das	53,902,119	035%
8	um grande ano em sua carreira. O atacante não perdeu a oportunidade de marcar após o goleiro soltar a bola nos seus pés: 2 a 0. Mas os gols não pararam por aí. Três minutos	53,578,105	035%
9	, Jean recebeu em velocidade dentro da área e chutou cruzado no canto esquerdo do goleiro Rafael Moreno, para marcar o terceiro gol tricolor e praticamente garantir mais	53,631,103	035%
10	ariscou chute forte de fora da área e a bola passou muito próxima ao travessão do goleiro Márcio. *** Em partida válida pela segunda rodada da Copa Libertadores, o	50,535,913	033%
11	foi aos 44 minutos, com Fred. O atacante dominou na área e chutou, porém o goleiro Dida defendeu. *** Faltando uma rodada para o término da fase de grupos da	50,749,923	033%
12	Logo em seguida, foi o árbitro marcou penalidade máxima a favor do Fluminense, após o goleiro Márcio derrubar Thiago Neves. O próprio meia cobrou, mas o goleiro acertou o	50,351,893	033%
13	Fluminense, após o goleiro Márcio derrubar Thiago Neves. O próprio meia cobrou, mas o goleiro acertou o canto. O primeiro tempo chegou ao fim com empate em 1 a 1 no	50,362,907	033%
14	para o capitão, que chutou cruzado, no canto esquerdo, sem chance de defesa para o goleiro. Com a vantagem, o Tricolor passou a controlar a partida, sem dar espaços para	51,435,977	034%
15	receber de Marcos Junior e Anderson. Fred se apresentou bem e deu trabalho para o goleiro Adilson Junior. Com o resultado, o Tricolor segue na liderança do grupo B da	52,274,024	034%
16	gols, foram Marcos Junior e Anderson. Fred se apresentou bem e deu trabalho para o goleiro Adilson Junior. Com o resultado, o Tricolor segue na liderança do grupo B da	52,520,043	034%
17	depois: Fred cruzou e outra vez Marcos Junior tocou de letra para grande defesa do goleiro do Vasco. Aos 40 minutos, Pedro Ken entrou sozinho pelo meio e chutou forte	52,061,013	034%
18	. Logo aos três minutos, Thiago Neves cobrou falta com pengo para o meio da área. O goleiro Ricardo acabou afastando com os pés. No contra-ataque, Bruno Neves chutou	52,216,023	034%
19	e cruzou na medida para Wagner, que mandou de cabeça para o fundo da rede do goleiro Jefferson. Faltando menos de um minuto para o apito final, Wagner marcou	54,030,123	035%
20	fazer o corte mas a bola passou entre suas pernas caindo para Ricardinho, na frente do goleiro Johnny Herrera. O atacante gremista deu um leve toque tirando do goleiro mas a	57,549,134	038%
21	na frente do goleiro Johnny Herrera. O atacante gremista deu um leve toque tirando do goleiro mas a bola saiu um pouco desviada passando à esquerda. Inicial. Minuto 32 -	57,561,134	038%
22	passou à direita de Galatto. Minuto 33 - Eivaldo tentou surpreender chutando de longe. O goleiro Márcio espalmou. Minuto 34 - Outra vez Eivaldo chutando de longa distância.	56,538,311	037%
23	16:07 - início de jogo Minuto 06 - Bolívar na intermédiana recou para o goleiro Clemer que segurou a bola com a mão. O árbitro Carlos Simon não marcou a	56,732,323	037%
24	dentro da área, pela direita de ataque. O chute cruzado saiu desviado, sem perigo para o goleiro Marcelo Grohe. Foi a primeira real chance de gol de Corinthians. Minuto 35 -	57,594,333	038%
25	no braço do jogador do Paraná Páralho para o Grêmio. Hugo bate no canto esquerdo do goleiro e marca o 1º gol do tricolor no jogo. 2 a 1 Paraná 29 minutos: O zagueiro do	58,842,412	039%
26	da saída da bola, mas Ricardinho é mais esperto e sai com ela, mas na frente do goleiro tenta a finalização, mas Flávio leva a melhor. 31 minutos: Patrício ia invadindo a	58,880,487	039%
27	espera o final do jogo quando num vacilo da defesa, Aloísio driblou 2 jogadores e o goleiro, que para evitar o gol derrubou o atacante gremista na área. O capitão Helton	58,466,467	038%
28	, que passou longe do gol de Galatto. 9 minutos: Hugo tenta finalização de longe, e o goleiro Flávio defende com tranquilidade. 11 minutos: Alessandro divide com 2 jogadores,	58,705,474	039%
29	colocar o Tricolor em vantagem. Na segunda etapa, o Grêmio ampliou aos 16 minutos. O goleiro Nivaldo não segurou a falta cobrada por Nunes. Lucas, com oportunismo, pegou o	54,674,152	036%
30	nassa para Herrera pelo meio da zaga. O argentino tocou com perfeição no canto do goleiro do Juventude. O resultado era melhor do que os gremistas imaginavam. Aos 23	54,971,183	036%

Figura 11 - Exemplo de lista gerada pelo Concord na busca pela UL *goleiro*

Para fazer esta busca, inicialmente é necessário fazer o upload do *corpus* no programa, (clicando em File/New). Após, basta digitar a consulta que se quer fazer e o programa gera, automaticamente, listas como a apresentada na Figura 11. Estas listas podem ser organizadas de diversas formas, bastando o consulente clicar em Edit/Resort e definindo os padrões de busca. Nos trabalhos de tradução, como este, esta primeira consulta já pode apontar indícios quanto ao número de ocorrências e a forma como palavras co-ocorrem em um texto, exercício a ser aplicado no próximo capítulo desta dissertação.

Outra ferramenta importante do Concord é o acesso aos *clusters*, que são agrupamentos observados estatisticamente em torno da palavra consultada. No caso da UL goleiro, observa-se a alta frequência de *clusters* como: *mas o goleiro, e o goleiro, para o goleiro, o goleiro do, com o goleiro, defesa do goleiro*, entre outras, conforme imagem abaixo (Figura 12).



The screenshot shows the Concord software window with a menu bar (File, Edit, View, Compute, Settings, Windows, Help) and a table of clusters. The table has columns for N, Cluster, Freq, Set, and Length. The 'Set' column for all rows is highlighted in grey and contains the value '3'. Below the table, there are tabs for 'concordance', 'collocates', 'plot', 'patterns', 'clusters', 'filenames', 'follow up', 'source text', and 'notes'. The 'clusters' tab is selected. At the bottom left, it shows '384' and 'Set 3'.

N	Cluster	Freq	Set	Length
1	MAS O GOLEIRO	165	3	3
2	E O GOLEIRO	153	3	3
3	PARA O GOLEIRO	109	3	3
4	O GOLEIRO DO	97	3	3
5	COM O GOLEIRO	79	3	3
6	DEFESA DO GOLEIRO	79	3	3
7	SAÍDA DO GOLEIRO	62	3	3
8	NA SAÍDA DO	59	3	3
9	CIMA DO GOLEIRO	55	3	3
10	O GOLEIRO E	45	3	3
11	MÃOS DO GOLEIRO	42	3	3
12	EM CIMA DO	42	3	3
13	NAS MÃOS DO	39	3	3
14	O GOLEIRO FELIPE	39	3	3
15	DO GOLEIRO E	34	3	3
16	DRIBLOU O GOLEIRO	34	3	3
17	O GOLEIRO FÁBIO	33	3	3
18	DA ÁREA E	31	3	3
19	OBRIGOU O GOLEIRO	30	3	3
20	O GOLEIRO MÁRCIO	30	3	3
21	O GOLEIRO MARCELO	30	3	3
22	A CARA COM	29	3	3
23	SEM CHANCES PARA	28	3	3
24	CHANCES PARA O	28	3	3
25	CARA COM O	28	3	3

**Figura 12** – Exemplo de *clusters*, obtido na consulta por *goleiro* no Concord

O exercício de observação de *clusters* pode ser utilizado na análise de co-ocorrência de palavras e, também, para verificar, por exemplo, o comportamento de unidades

complexas. Destaca-se que esta observação pode ser o início de uma nova consulta no Concord, agora por expressões. No caso de *driblou o goleiro*, por exemplo, pode-se fazer um novo recorte de busca ao consultar-se no Concord por *dribl\* o goleiro* (esta forma de consulta abre espaço para a flexão do verbo *driblar*).

Neste exemplo, tem-se como resultado os dados abaixo (Figura 13), que podem, então, ser analisados de diversas formas: por ordem alfabética dos elementos centrais, das palavras que estão a esquerda ou direita, etc. E, além disso, é possível salvar um arquivo com um *subcorpus* como este em formato .txt diretamente em seu computador.

N	Concordance	
1	zagueiro Niell salva sobre a linha. Eduardo Cecconi: Ozil recebe na frente, dribla o goleiro, mas adianta demais. Eduardo Cecconi: Para interromper a	2
2	Schüler: Quaaase!!! Damião recebe cruzamento dentro da área, domina, dribla o goleiro, mas na hora do chute, Tiago Saletti salva a bola em cima	4
3	foi quase no minuto seguinte. Alex Afonso aproveita bobeira da zaga rival, dribla o goleiro Diego, mas bate na trave, desperdiçando uma incrível	3
4	Lopes: UUUUHHH!! Borges recebe na entrada da área, sai da marcação, dribla o goleiro e finaliza.... Glauber tira de cima da linha!!! Tati Lopes:	8
5	, Fred lançou para Wellington em velocidade. O atacante tricolor chegou a driblar o goleiro Marcelo Lomba que se recuperou e evitou a conclusão, A	29
6	, a defesa francesa ficou parada e o jovem atacante teve tempo para driblar o goleiro e tocar para as redes. A França se lançou para o ataque,	8
7	para Osvaldo. O camisa 17 achou Aloisio na área, e o atacante tentou driblar o goleiro, mas Campestrini conseguiu desarmar o são-paulino. O	14
8	catarinense teve boa ação ofensiva. Caio recebeu passe e, ao tentar driblar o goleiro Márcio, foi derrubado: pênalti para o Figueira. O mesmo	16
9	a bola para Diego Cavalieri. O centroavante do São Paulo foi veloz para driblar o goleiro, empurrar para o gol e transformar o Morumbi novamente	33
10	atacante, que havia acabado de entrar, teve a calma dos veteranos para driblar o goleiro e só empurrar para as redes. Mais. Mas o México não se	7
11	de Everton Ribeiro, partiu em velocidade para a grande área e tentou driblar o goleiro Jessé, que colocou o mão na bola fora da área. A	1
12	marcou o segundo, após passe recebido no meio dos zagueiros e de driblar o goleiro. Entretanto, o rival empatou em duas cobranças de pênalti.	13
13	fez o seu. Foi oportunista ao ganhar o presente de Osório e frio para driblar o goleiro, fazendo seu quarto gol e se isolando na artilharia do	1
14	logo com um minuto do segundo tempo, após invadir a área pela direita e driblar o goleiro. Uma das estratégias adotadas pelo Paraguai foi colocar	39
15	lbson penetrou livre na área, mas foi desarmado por Muriel quando tentava driblar o goleiro. Na volta para o segundo tempo o Internacional apresentou	10
16	para Felipe, que saiu na cara do gol e não aproveitou. O atacante tentou driblar o goleiro, mas Wilson saiu bem do gol e ficou com a bola. Quinze	29
17	esquerda de Fernando Prass. Dois minutos depois Neto Berola chegou a driblar o goleiro, mas se desequilibrou e foi dominado pela zaga vascaína.	34
18	: Borges driblou o goleiro e bateu. Mas errou a mira Tati Lopes: Borges driblou o goleiro e bateu. Mas errou a mira Tati Lopes: Borges é lançado	6
19	e fazer 2 a 0. No minuto seguinte, Torres arrancou livre do meio-campo, driblou o goleiro ainda fora da área e ampliou. David Villa ainda deixou o	6

**Figura 13** - Exemplo de busca por unidades complexas, como *dribl\* o goleiro*, no Concord

No próximo capítulo, que traz a análise e discussão dos dados, o Concord é utilizado e as funcionalidades acima são postas em prática com o intuito de analisar, especificamente, as equivalências de tradução.

### 3.2 MÉTODO

Para que seja possível discutir sobre as equivalências de tradução em recursos multilíngues baseados em *frames*, esta pesquisa prevê a realização dos seguintes procedimentos:

- 1) Observação de uma amostra de 100 ULs em inglês e português com o intuito de verificar, de uma forma ampla, o comportamento de seus equivalentes, para posterior estudo dos casos que merecem atenção especial. Tal amostra foi coletada observando-se os 50 verbos e substantivos mais frequentes em nossos *corpora* destes idiomas<sup>27</sup>, obtida no recurso WordList do programa WordSmith Tools.
- 2) Seleção de casos que destacamos no que se refere ao estabelecimento de equivalências, por apresentarem diferentes tipos de equivalência.
- 3) Análise das ULs pertencentes a cada um destes casos, com o intuito de refletir sobre como a noção de *frames* pode influenciar no estabelecimento de equivalências e em sua apresentação em um recurso lexical bilíngue baseado em *frames*. Para isso, estão previstas as seguintes etapas de trabalho:

3.1.1) Fase 1: discussão sobre as características das ULs em questão e sua valência no idioma fonte;

3.1.2) Fase 2: verificação das ULs em dicionários bilíngues, para que seja possível compreender e avaliar como as ULs são apresentadas atualmente nos recursos lexicais, juntamente com suas traduções. Nesta fase, são consultados os dicionários WordReference<sup>28</sup> e Michaelis<sup>29</sup> e tradutor automático Babylon<sup>30</sup>, por serem referências na busca por equivalências disponíveis em formato eletrônico atualmente;

3.1.3) Fase 3: busca de equivalências utilizando-se a ferramenta WordSmith Tools;

3.1.4) Fase 4: reflexão sobre o exercício de apresentação das possíveis equivalências em um recurso lexical baseado em *frames*, para que seja possível ponderar-se a valia de se considerar este conceito no desenvolvimento de recursos lexicais multilíngues.

Em seguida, a pesquisa encerra apresentando algumas considerações finais a respeito dos achados durante a análise de dados e também propostas para futuros trabalhos na área.

---

<sup>27</sup> A escolha por verbos e substantivos foi feita porque estas são as classes de palavras que evocam *frames* com mais frequência, sendo, portanto, possíveis ULs a serem disponibilizadas no Dicionário da Copa 2014.

<sup>28</sup> Disponível em: <http://www.wordreference.com/>. Acesso em 20 de julho de 2013.

<sup>29</sup> Disponível em: <http://michaelis.uol.com.br/>. Acesso em 20 de julho de 2013.

<sup>30</sup> Disponível em: <http://tradutor.babylon.com/>. Acesso em 20 de julho de 2013.

## 4 ANÁLISE E DISCUSSÃO DOS DADOS

Este capítulo tem como objetivo apresentar a reflexão sobre o exercício prático que busca analisar como a Semântica de *Frames* pode lidar com os equivalentes de tradução nos idiomas inglês/português em um recurso lexical baseado em *frames*, no domínio do futebol. Para isso, a análise será realizada em duas partes:

- 1) Análise ampla dos fenômenos de equivalências no domínio do futebol através da avaliação de uma amostra de 100 ULs;
- 2) Análise de casos específicos.

### 4.1 ANÁLISE AMPLA DOS FENÔMENOS DE EQUIVALÊNCIA NO DOMÍNIO DO FUTEBOL

Conforme mencionado anteriormente, a reflexão proposta neste estudo iniciou pela análise de uma amostra de 100 ULs<sup>31</sup> do domínio do futebol. A apreciação deste conjunto inicial de 100 ULs teve como principal propósito avaliar, em um sentido amplo, como se comportavam os equivalentes em nosso corpus e identificar, em um segundo momento, os casos de equivalência que mereciam destaque em nosso estudo.

Para isso, foram selecionadas 50 palavras em inglês e outras 50 palavras em português, definidas por serem os 50 substantivos ou verbos com maior frequência no *corpus* de cada um dos idiomas do dicionário. Esta listagem foi obtida a partir do uso da ferramenta WordList do WordSmith Tools.

A lista de palavras foi gerada para os *corpora* em inglês e português e retornou uma relação de palavras como a apresentada na figura abaixo (Figura 14).

---

<sup>31</sup> Nesta primeira análise o termo UL não se refere ainda diretamente a “evocadores de *frames*”, como o termo utilizado na Semântica de *frames*. Trata-se de possíveis ULs ou, em alguns casos, partes de unidades complexas que podem vir a ser evocadoras de *frames*.

N	Word	Freq.	%	Texts	% Lemmas	Set
1	O	38,855	5.04	4	100.00	
2	A	28,686	3.72	4	100.00	
3	DE	24,240	3.15	4	100.00	
4	E	22,437	2.91	4	100.00	
5	#	18,654	2.42	4	100.00	
6	DO	15,121	1.96	4	100.00	
7	PARA	13,976	1.81	4	100.00	
8	NO	12,379	1.61	4	100.00	
9	NA	11,905	1.55	4	100.00	
10	DA	10,901	1.42	4	100.00	
11	COM	10,794	1.40	4	100.00	
12	QUE	9,151	1.19	4	100.00	
13	EM	8,758	1.14	4	100.00	
14	AOS	7,205	0.94	4	100.00	
15	MINUTOS	5,755	0.75	4	100.00	
16	MAS	5,716	0.74	4	100.00	
17	BOLA	5,311	0.69	4	100.00	
18	GOL	5,184	0.67	4	100.00	
19	POR	4,982	0.65	4	100.00	
20	UM	4,817	0.63	4	100.00	
21	ÁREA	4,622	0.60	4	100.00	
22	OS	4,594	0.60	4	100.00	
23	FOI	4,260	0.55	4	100.00	
24	NÃO	4,133	0.54	4	100.00	
25	UMA	3,497	0.45	4	100.00	

**Figura 14** - Tela de frequências no WordList do WordSmith Tools

Esta listagem foi salva em um arquivo no formato .xls, como planilha no programa Excel, para facilitar seu manuseio e posterior análise. O mesmo procedimento foi realizado com o *corpus* em inglês. As palavras que não fossem substantivos ou verbos foram deletadas, assim como os nomes de times e jogadores. O resultado foi salvo em duas planilhas, uma delas sendo denominada WordListPT, com as 50 palavras mais frequentes do *corpus* em português, e a outra WordListEN, com as 50 palavras mais frequentes do *corpus* em inglês (Anexo A).

Neste primeiro momento foi realizada a busca por equivalentes destas palavras no dicionário WordReference, apenas para que se pudessem identificar os tipos de equivalência mais recorrentes no *corpus*. Os possíveis equivalentes apresentados nesta etapa são considerados um tipo de guia para direcionar as próximas análises, que levarão em conta sentenças do *corpus* e consulta a outros recursos lexicográficos como base para discussão sobre equivalências mais adequadas no domínio do futebol.

Para a reflexão dos tipos de equivalência, utilizaram-se as definições propostas por Felber (1984) e também por Duval (2008, p. 274), para se propor os seguintes agrupamentos:

- 1) *Equivalência exata de conteúdos*: quando o significante da língua fonte aponta a um mesmo item na língua alvo e seu significado é um item lexical reconhecido no

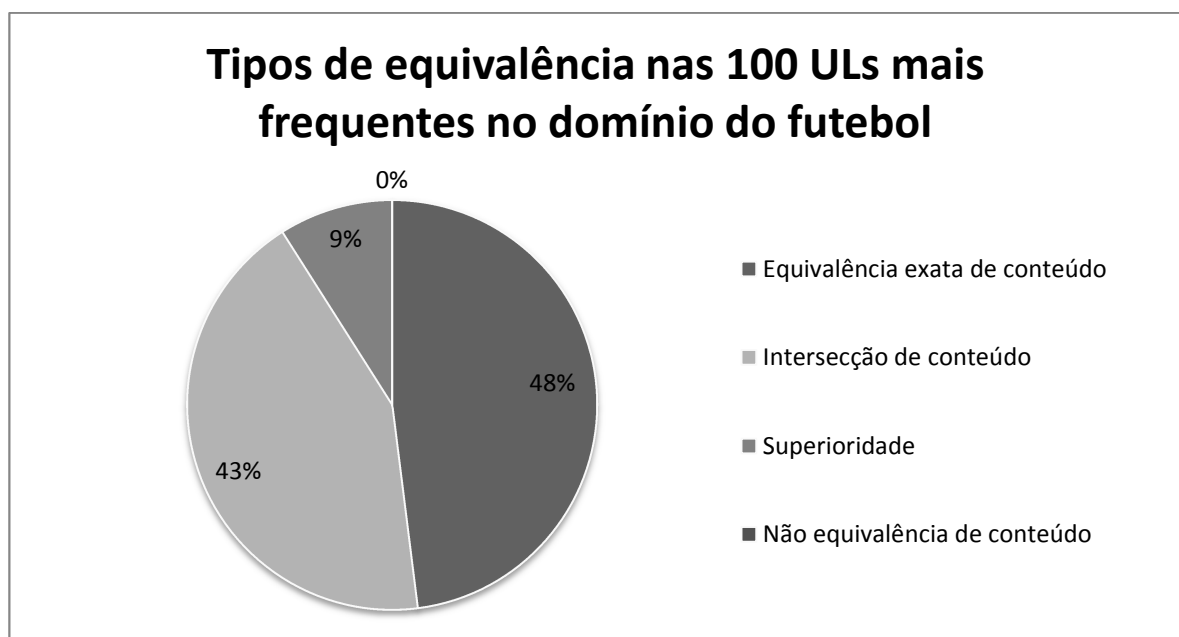


léxico de ambas as línguas. Neste caso, a unidade lexical é monossêmica e não é ambígua em nenhum dos idiomas;

- 2) *Intersecção de conteúdos*: quando existe um item presente no léxico de ambas as línguas que não remete a um mesmo universo cultural, ou não é conhecido por muitos. Porém, semanticamente existe uma relação entre o léxico dos dois idiomas, o que os torna possíveis equivalentes.
- 3) *Superioridade*: quando uma palavra que é utilizada na língua fonte tem uma extensão maior na língua alvo, ou vice-versa.
- 4) *Não equivalência de conteúdos*: quando o significante refere a algo real apenas na língua fonte;

As listagens de ULs, com seus possíveis equivalentes e o tipo de equivalência aparente, estão disponibilizadas em forma de anexos (Anexo B e Anexo C).

Ao fazer isso, observou-se o seguinte percentual de cada tipo de equivalência considerando-se o total das 100 ULs analisadas (Figura 15):



**Figura 15** - Tipos de equivalência na amostra de 100 ULs do domínio do futebol

Como se pode observar no gráfico acima (Gráfico 1), 48% das palavras analisadas possuem equivalência exata de conteúdo com ULs do idioma alvo. Alguns exemplos são: *ball* = bola, *league* = liga, *kick* = chutar, *goleiro* = *goalkeeper*, *escanteio* = *corner*, etc. São palavras monossêmicas e não ambíguas no domínio do futebol.

Já a equivalência por intersecção de conteúdo, que representa 43% dos casos encontrados, retrata casos nos quais os significados não são exatamente iguais em ambos os idiomas. Destaca-se que os casos de polissemia foram agrupados aqui, pois trata-se de palavras que não são monossêmicas e que podem ter como equivalentes expressões poliléxicas. É o caso de *chance*, em português, que embora tenha os possíveis equivalentes *chance* e *opportunity*, pode se referir a *chance de gol*, *chance de finalizar*, entre outros, evocando, inclusive, de diferente *frames*.

Os casos de superioridade encontrados (9% dos casos analisados na amostra) são os de ULs como *team*, que pode significar *time*, *equipe* ou *seleção* em português. Em casos como este, pode-se dizer que a UL em inglês tem uma amplitude de significado maior do que um equivalente disponível no léxico do português, ou seja, aparentemente não existe em português uma UL que descreva o conjunto de *times* e *seleções*, que são ULs mais específicas disponíveis no léxico do português.

Não foram encontrados exemplos de não equivalência de conteúdo. Acredita-se que isso tenha ocorrido possivelmente porque, por se tratar de um domínio específico, com regras internacionais em comum. Além disso, ainda se tratando do léxico do futebol em uma instância de campeonato internacional, favorece-se ainda mais o compartilhamento de unidades lexicais e significados comuns e a minimização de variantes regionais e níveis de linguagem. Chama-se atenção para o caso da palavra *yard*, em inglês, que, embora sua possível tradução (jarda) não seja utilizada em português, não está sendo considerado um caso de não equivalência. Isso por ser um conceito de medida conhecido também no Brasil, mesmo sabendo-se que uma tradução deste termo possivelmente transformaria a medida para o sistema métrico.

Assim sendo, ao concluir esta primeira análise, pode-se dizer que existem indícios de que a maioria das ULs no domínio do futebol possui equivalência exata de conteúdo ou equivalência por intersecção de conteúdos. Por outro lado, um percentual pouco representativo foi encontrado no caso de equivalências por superioridade (9%) e não houve casos de não equivalência de conteúdo, o que nos leva a pensar que estes casos também não ocorram com frequência em geral no léxico desta área.

No próximo capítulo serão examinados alguns casos específicos observados nesta primeira análise.

## 4.2 ANÁLISE DE CASOS ESPECÍFICOS

Nesta segunda parte da análise, foram selecionados 5 casos de equivalência, estes representando o que consideramos casos especiais em termos de equivalência de tradução. Espera-se, portanto, com a reflexão neste subcapítulo, ponderar sobre como abordar as equivalências de tradução no âmbito de uma FrameNet multilíngue em um domínio específico como o do futebol.

Destaca-se que não se tem como objetivo tratar da representação lexicográfica ou mesmo de questões terminológicas no que tange aos equivalentes, ou seja, da construção de entradas dos equivalentes no dicionário, da seleção de ULs a serem disponibilizadas ou de sua caracterização. Porém, pretende-se abordar as relações existentes entre ULs que fazem parte de um mesmo *frame* conceitual nos diferentes idiomas e a forma como a Semântica de *Frames* pode dar suporte às informações disponibilizadas em um recurso baseado em *frames*.

Os casos analisados são: 1) equivalência exata de conteúdos (ULs *ball* e *goleiro*); 2) intersecção de conteúdos (ULs *shot* e *marcar falta*); e 3) equivalências por superioridade (UL *team*). Isso será feito para que cada caso seja observado em sua particularidade.

Para a análise, as seguintes fases descritas anteriormente estão previstas: 1) discussão sobre as características das ULs em questão; 2) verificação das ULs em dicionários bilíngues com o intuito de verificar se é simples a localização de equivalentes em recursos como estes e também quais as informações disponibilizadas por eles; 3) busca de equivalências utilizando-se da análise de *corpus*; e 4) reflexão sobre o exercício de apresentação das possíveis equivalências em um recurso lexical baseado em *frames*.

Destaca-se que estas quatro fases são necessárias para que seja possível aproximar-se de uma resposta ao questionamento de pesquisa proposto no início deste trabalho e que busca investigar como pensar nos equivalentes de tradução haja vista a sua inclusão em um recurso lexical multilíngue baseado nos preceitos da Semântica de *Frames*.

### 4.2.1 Equivalência exata de conteúdos

Um dos casos encontrados na primeira análise deste trabalho foi o de ULs que apresentam equivalência exata de tradução, ou seja, casos nos quais o significante da língua fonte aponta a um mesmo item na língua alvo e seu significado é um item lexical reconhecido

no léxico de ambas as línguas (Duval, 2008, p. 274). Destaca-se que, assim como apresentado no capítulo sobre tradução, a equivalência nunca é vista como uma “simetria exata” neste estudo, pois se sabe que a equivalência perfeita não é possível tendo em vista fatores internos e externos à língua. Porém, fala-se em possíveis equivalentes tendo-se em vista características comuns às línguas em contato.

Após a realização da primeira análise deste trabalho, descrita acima, 48% das equivalências de uma amostragem de 100 ULs do domínio do futebol apresentaram equivalência exata de conteúdos. Pretende-se neste subcapítulo verificar como isso ocorre com as ULs *ball* e *goleiro*, para que se possa refletir sobre como tais ULs se comportariam, na relação com seus possíveis equivalentes, no Kicktionary\_Br Copa 2014.

#### 4.2.1.1 Ball

##### *Fase 1: Característica da UL ball*

A primeira análise, da UL *ball*, partiu do pressuposto de se tratar de um caso típico de equivalência exata de conteúdo, tendo-se em vista que descreve um referente concreto, possivelmente o objeto mais importante relacionado ao futebol, e que é conhecido também em outros esportes, o que aproxima o universo cultural entre os idiomas inglês e português ainda mais.

Esta UL é o substantivo mais frequente no *corpus* em inglês e considera-se que ela evoca o *frame* Bola. Este *frame* é usado pelos *frames* Ataque e Ações Técnico-Táticas.

Tendo-se estas considerações em vista, parte-se agora para a segunda fase da análise, que é a consulta desta UL em dicionários.

##### *Fase 2: Consulta a dicionários*

A consulta pela UL *ball* em dicionários foi realizada com o intuito de verificar o que estes recursos lexicais, mesmo não sendo específicos do domínio do futebol, trazem sobre tal UL e qual o equivalente indicado por eles em português. Esta busca gerou os seguintes resultados:

	Consulta por:
Dicionário:	<i>Ball</i>
Word Reference	(sport) bola
Michaelis	<b>1</b> bola, esfera. <b>2</b> jogo de bola. <b>3</b> tiro, arremesso da bola. <b>4</b> beisebol (jogo de bola americano).
Babylon	bola, esfera, globo.

**Quadro 1** - Resultado da busca por *ball* em dicionários online

Através desta pesquisa, notou-se que todos os dicionários apresentavam pelo menos uma equivalência possível, sendo que, no primeiro, a UL *bola* é listada como equivalente já inserida no domínio de esportes. Nos dicionários Michaelis e Babylon, notou-se que o domínio de uso de cada uma das possíveis traduções não é mencionado, mas que a palavra *bola* já aparece como primeira possibilidade de tradução em ambos os casos.

Tendo isso em mente, partiu-se então para a fase seguinte, que é a de busca de equivalências utilizando-se a linguística de corpus e a ferramenta WordSmith Tools.

### *Fase 3: Busca de equivalências utilizando-se da linguística de corpus*

Neste trabalho sugerimos como procedimento para busca de equivalências o trabalho com análise de corpus. No caso de equivalências exatas de conteúdo, acredita-se que o *corpus* pode ser útil principalmente para validar possíveis equivalentes muitas vezes já previstos, baseado na frequência e nos EFs de uma UL. Além disso, de forma mais geral, para ajudar na compreensão do *frame* do qual uma UL faz parte, fornecendo sentenças muitas vezes capazes de contextualizar o uso da UL buscada.

A primeira busca foi pela UL *ball* e resultou em uma listagem de 4.379 sentenças na ferramenta Concord do WordSmith Tools. Destas, as primeiras 25 foram lidas com o intuito de verificar se poderia haver algum tipo de polissemia na UL (o que não foi observado) e compreender melhor o contexto de utilização desta UL. A figura abaixo representa o resultado obtido na busca (Figura 16):

N	Concordance	Word #	Sen	Sen	Parz	Pa
1	but seems uncharacteristically hesitant. This time Coentrao nicks the ball from his feet as he runs towards the box. * 17' Cristiano Ronaldo	151,61	10,	449	1	2
2	hits a decent effort from 20 yards. The Real Madrid 'keeper palms the ball rather than away, the danger passes and Spain survive. * 18'	151,58	10,	479	1	2
3	Goal kick Portugal. * 24' Spain with a free kick, a chance to get the ball in the box... * 23' ESPN Stats & Info: Iker Casillas captains	151,48	10,	829	1	27
4	from Spain, Ramos' header clear only finds Almeida, who guides the ball towards the offside Cristiano Ronaldo. Spain's defence does not	151,80	10,	789	1	2
5	it. * 16' Paraguay keeper Villar, a veteran of 35, is seeing a lot of the ball as Japan hit numerous long balls towards Honda and Okubo. Too	149,16	10,	639	1	2
6	' Ramos has got the taste for it now. Running at Coentrao again, the ball finds its way to Villa who smashes it from 25-yards. Another	150,24	10,	419	1	2
7	Portugal defenders since his introduction. * 77' David Villa slots the ball past Eduardo * 76' Ramos has got the taste for it now. Running	150,22	10,	409	1	2
8	floats a glorious ball beyond the defence. Llorente can only guide the ball wide. * 85' More great work down the left from Coentrao, only	150,08	10,	949	1	2
9	gets his head to corner, Eduardo gathers comfortably. He throws the ball immediately to Ronaldo who's streaking away. Casillas is out	151,15	10,	459	1	27
10	is headed out for another corner by Carmona. * 7' Maicon gets the ball square from Kaka and wins a corner off Jean Beausejour. * 6'	146,10	10,	419	1	21
11	pull the pass down and the ball bounces out. * 5' Chile work the ball well and force Julio Cesar into a sharp piece of work. As Brazil	146,14	10,	399	1	21
12	move but their approach breaks down when Robinho fails to slip the ball to Luis Fabiano in the box. * 10' Now Kaka gets in on the act,	145,98	10,	819	1	21
13	with a sliding through-ball but Juan sticks out a leg and diverts the ball to safety. Chile continue to commit players forward though. * 11'	145,94	10,	919	1	21
14	powerful, but dragged well wide. * 29' Gerard Pique challenges for the ball in the air * 28' Villa again running at Costa, the poor full back	151,37	10,	309	1	27
15	* 32' Portugal's tactics are working so far. Spain passing the ball in front of them, Alonso getting frustrated and hitting his shot	151,30	10,	239	1	27
16	while the forward is onside, he cannot pull the pass down and the ball bounces out. * 5' Chile work the ball well and force Julio Cesar	146,14	10,	939	1	21
17	Kaka and wins a corner off Jean Beausejour. * 6' Lucio chips the ball forward to Robinho and while the forward is onside, he cannot	146,12	10,	269	1	21
18	much up front. Portugal looked poor, unless Ronaldo got on the ball. Expect better in the second half. * 45' A Mexican wave breaks	32,146	2,2	100	0	2
19	clips a nice ball into the area but though Kalou throws himself at the ball he is a yard off. Drogba is warming up. Get him on. * 48' Very	31,902	2,2	819	0	2
20	when bursting down the left and digging out a cross, but the ball had just crossed the touchline. * 61' Now left-back Michel	34,525	2,4	809	0	27
21	for a studs-upper on Demel. The Portuguese midfielder did win the ball though. * 39' Ronaldo is getting frustrated now. He has found	32,291	2,2	899	0	2
22	"Ron looks like he's in the mood - he's constantly demanding the ball." * 12' Don't forget Ronaldo hasn't scored for Portugal for 16	32,986	2,3	100	0	2
23	when he had space and time. * 17' CIV rebuild and Tiote receives the ball from Kalou before hammering a shot narrowly over the bar. Good	32,865	2,3	509	0	2
24	to Ronaldo. * 20' YELLOW CARD: Here we go. Ronaldo overruns the ball and dives when a challenge comes in, then he and Demel bump	32,800	2,3	289	0	2
25	for CIV if they don't close him down. * 10' Wow! Ronaldo gets the ball, jinks inside and launches one from 30 yards out, the 'keeper is	33,041	2,3	209	0	2

Figura 16 - Resultado da busca por *ball* no Concord do WordSmith Tools

Além disso, também foram analisadas os 25 primeiros *clusters* que utilizam a palavra *ball* (Figura 17), que também aparentemente tratam apenas do objeto esférico utilizado pelos jogadores no futebol, por mostrarem indicações de percurso desta (*ball into the, the ball in, the ball over, the ball away, ball across the, etc*).

N	Cluster	Freq	Set	Length	Related
1	BALL INTO THE	230	3		
2	THE BALL IN	202	3		
3	THE BALL INTO	189	3		
4	THE BALL BACK	128	3		
5	BALL IN THE	126	3		
6	THE BALL AND	119	3		
7	THE BALL TO	115	3		
8	AND THE BALL	110	3		
9	BALL OVER THE	105	3		
10	THE BALL AWAY	99	3		
11	THE BALL OVER	98	3		
12	OF THE BALL	97	3		
13	THE BALL ON	95	3		
14	THE LOOSE BALL	92	3		
15	THE BALL PAST	89	3		
16	THE BALL WAS	86	3		
17	THE BALL OUT	86	3		
18	INTO THE BOX	82	3		
19	BUT THE BALL	82	3		
20	A LONG BALL	78	3		
21	AS THE BALL	74	3		
22	BALL ON THE	67	3		
23	ON THE BALL	67	3		
24	BALL ACROSS THE	64	3		
25	THE BALL IS	63	3		
26	THE BALL FELL	61	3		

Figure 17 - Resultado dos *clusters* para *ball* no Concord do WordSmith Tools

Com isso observou-se que há indícios de que *ball* seja uma UL monossêmica e não seja utilizada de forma ambígua nos relatos de futebol presentes no *corpus*.

A consulta seguinte foi feita à listagem das 50 ULs mais recorrentes no português para verificar se havia algum possível equivalente disponível. Esta busca apenas confirmou o já era esperado: que *bola* seria o equivalente mais próximo à UL *ball*. Realizou-se então uma nova busca no Concord, agora no *corpus* em português pela UL *bola*. Esta busca retornou 5.311 sentenças e as primeiras 25 foram analisadas (Figura 18). Assim como no caso da UL *ball*, os resultados neste caso também apontam com exclusividade o conceito de bola como objeto esférico utilizado em jogos como o futebol.

N	Concordance	Set	Word #	Sem
1	e passando até certo sufoco quando pressionado, Carlinhos recuperou a bola na intermediária e tocou curto para Thiago Neves, que observou		260,15f	13,
2	livre dentro da área e tentou acertar o canto esquerdo de Wilson, a bola passou muito próxima da trave e não entrou. O panorama da etapa		260,83f	13,
3	que administraria até o apito final. Aos 27 minutos, o time trabalhava a bola pelo lado esquerdo do ataque quando Jean recebeu bola cruzada de		260,22f	13,
4	ângulo de Felipe, mas o goleiro rubro-negro se esticou, espalmando a bola para escanteio e evitando o primeiro gol do Peixe. Os cariocas		264,08f	13,
5	incomodar aos dez minutos. Após boa trama entre Nikão e Marcinho, a bola ficou com Roger, que, na frente de Cássio, pegou fraco e permitiu a		261,77f	13,
6	em jogada de bola parada: Rildo, que entrara no intervalo, desviou e a bola saiu à direita, com muito perigo à meta de Wilson. O abafa dos		253,25f	13,
7	tocando a bola, foi conseguindo faltas mais à frente e, em uma delas, a bola foi parar na cabeça de Felipe Azevedo. Aos 44 minutos, o atacante		252,26f	13,
8	Aos 39, mais uma vez Araújo perdeu a chance de empatar e mandou a bola por cima do gol de Renan. Mas o dia era mesmo do alvinegro		256,82f	13,
9	ao novo tiro pontepretano veio dos pés de Douglas. O meia tomou a bola na direita e bateu falta nas mãos do goleiro Edson Bastos, aos 36		261,53f	13,
10	da Macaca dominou aos 11 minutos e arriscou o chute cruzado. A bola pingou na frente de Cássio e passou perto de sua trave esquerda.		261,32f	13,
11	por sua vez, tentou o chute de fora da área aos 25 minutos, mas a bola passou por cima do gol de Victor, sem grande perigo. No final da		258,38f	13,
12	Depois de cobrança levantada na área, Bruno deu um leve soco e a bola sobrou para Bernard. O camisa 11 bateu no canto direito e a zaga		258,01f	13,
13	João Vitor, cruzou para Obina. O camisa 21 desviou desequilibrado e a bola passou devagar, mas bem próxima da meta do Atlético-MG. A		257,98f	13,
14	rasteiro. Jones Carioca precisou finalizar duas vezes para colocar a bola para a rede. Mesmo depois do gol, o Bahia seguiu melhor em		257,30f	13,
15	dentro da área e finalizou cruzado na saída de Fernando Prass, mas a bola passou por toda pequena área e ficou para a zaga cruzmaltina. A		257,14f	13,
16	puxados pelo clube carioca. Aos 26 minutos, Tenorio prendeu a bola na intermediária, driblou Wellington e deixou para Juninho		263,86f	13,
17	o Flamengo teve mais uma grande chance. Aos 22, Adryan recuperou a bola, puxou contra-ataque e abriu para Matheus. O meia rubro-negro		264,16f	13,
18	insistirem, Jones e Elias participaram de jogada na grande área e a bola ficou limpa para o volante Hélder encher o pé e bater cruzado, sem		263,08f	13,
19	que apenas observou. Na sequência, o camisa 10 do Flu trabalhou a bola com Carlinhos e obrigou Rogério a desviar para escanteio. Menos		260,10f	13,
20	cruzou para Wellington obrigar Fernando Prass a fazer grande defesa. A bola sobrou na pequena área para Luan, que aproveitou confusão de		263,58f	13,
21	a assustar Vanderlei quando Vagner Love chutou da entrada da área. A bola passou por cima do travessão, para a alegria da torcida no Couto		253,75f	13,
22	jogo. Aos 34, Eliecarlos entrou na área pela direita e cruzou rasteiro. A bola acabou nos pés de Rogério, que entrara no lugar de Dimba, mas		256,74f	13,
23	foi acionado no bico direito da área e bateu cruzado e rasteiro, vendo a bola passar rente à trave direita do arqueiro santista, que já estava		256,08f	13,
24	e cruzou. Na frente da pequena área, Souza pega de primeira, mas a bola saiu à direita do gol de Renan. Aos 39, mais uma vez Araújo		256,80f	13,
25	e achou Elkeson livre na frente da pequena área. Sem deixar a bola tocar no chão, de primeira e de letra, o atacante marcou um golaço.		256,35f	13,

**Figura 18** - Resultado da consulta por *bola* no Concord do WordSmith Tools

O mesmo aconteceu com os *clusters* gerados pelo WordSmith para esta UL, ou seja, ao observar-se os agrupamentos mais frequentes relacionados à *bola* (Figura 19), notamos ocorrências como *posse de bola*, *bola passou para*, *bola na área*, *bola desviou*, *a bola saiu*, entre outros, abordando o mesmo conceito de movimento já visto na consulta à UL *ball*.

N	Cluster	Freq	Set	Length	Related
1	E A BOLA	424	3		
2	MAS A BOLA	370	3		
3	POSSE DE BOLA	302	3		
4	A BOLA PASSOU	256	3		
5	COM A BOLA	249	3		
6	A BOLA PARA	242	3		
7	A BOLA E	194	3		
8	A BOLA NA	193	3		
9	A BOLA NO	177	3		
10	A BOLA SOBROU	167	3		
11	BOLA SOBROU PARA	141	3		
12	SAÍDA DE BOLA	137	3		
13	A BOLA FOI	128	3		
14	A POSSE DE	124	3		
15	A BOLA BATEU	121	3		
16	BOLA NA ÁREA	116	3		
17	ÁREA A BOLA	112	3		
18	DE BOLA E	105	3		
19	A BOLA DESVIOU	103	3		
20	FORA DA ÁREA	102	3		
21	A BOLA SAIU	95	3		
22	DA ÁREA E	83	3		
23	A BOLA COM	83	3		
24	DA ÁREA A	80	3		
25	LINHA DE FUNDO	73	3		

**Figure 19** - Resultado dos *clusters* para *bola* no Concord do WordSmith Tools

Tendo-se em vista os resultados de consulta dos dicionários, aliados a análise do corpus, pode-se dizer, portanto, que as ULs *ball* e *bola* são equivalentes possíveis no domínio do futebol.

Neste caso, possivelmente por se tratar de uma UL concreta e tão comumente utilizada por ambos os idiomas, a localização do equivalente é simples. Parte-se agora para a parte final da análise, que é a reflexão sobre a abordagem de *frames* levando-se em conta a equivalência proposta.

*Fase 4: Reflexão sobre o exercício de apresentação das possíveis equivalências em um recurso lexical baseado em frames.*

Nesta última fase da análise, realiza-se uma reflexão sobre a apresentação dos equivalentes propostos acima no Kicktionary\_Br Copa 2014.

Inicialmente, retoma-se que, diferentemente de um dicionário tradicional, onde a busca seria por verbetes em ordem alfabética, as buscas no Kicktionary\_Br Copa 2014 poderão ser feitas por *frames* ou unidades lexicais. Desta forma, no caso da UL *ball* e seu equivalente *bola*, ambos estarão acessíveis no *frame* *Bola/Ball* e também poderão ser consultados bastando a digitação destas ULs no campo de busca. Além disso, ambas as ULs terão um hyperlink para que, ao clicar neles, um detalhamento maior sobre as ULs seja disponibilizado, no idioma da consulta.



No Kicktionary, Schmidt utiliza-se de *synsets* como forma de juntar pares de possíveis equivalentes, porém, como destacado no capítulo sobre o projeto, esta apresentação de relação entre ULs em diferentes idiomas não é direta. Já no Kicktionary\_Br Copa 2014 a relação entre ULs e sua possível equivalência será indicada na tela de consulta. Assim sendo, prevê-se o acesso às entradas lexicográficas capazes de fornecer informações como: *frame* evocado pela UL, definição da UL, relações semânticas entre os *frames*, *equivalente(s) de tradução possível(is)* e exemplos com indicação dos EFs.

A seguir, apresentam-se sugestões de entradas lexicográficas baseadas nos resultados obtidos durante a presente análise. Destaca-se, porém, que as informações apresentadas nas entradas lexicográficas propostas não são oficiais, mas um modelo do trabalho a ser desenvolvido com foco no que concerne às relações de equivalências. As demais informações foram obtida em trabalho em conjunto com o grupo Kicktionary\_Br.

Tendo-se em vista que as ULs disponíveis no Dicionário da Copa evocam *frames*, esta relação deve ser indicada. No caso da UL *ball*, o *frame* evocado é o Bola. Portanto, as entradas lexicográficas devem apresentar informações como as que seguem:

<b>Lexic Unit:</b> ball	<b>Frame:</b> Ball ( <a href="#">hyperlink to the frame details</a> )
<p><b>Definition*:</b> a spherical object that is kicked or thrown in soccer.</p> <p><b>Frame to Frame Relations*:</b> Is used by: Offence , Technical-Tactical_Actions (<a href="#">hyperlink to the frame details</a>)</p> <p><b>Possible Translation Equivalent:</b> bola (<a href="#">hyperlink to the equivalent entry</a>)</p> <p><b>Examples*:</b></p> <ol style="list-style-type: none"> <li>1) Marko Pantelic robs the <b>ball</b> and fires an audacious 40-yard effort wide.</li> <li>2) Unsurprisingly the Seasiders sat men behind the <b>ball</b> from that point and our first attempt to draw level was...</li> <li>3) Fulham gave the <b>ball</b> away almost instantly, though, and it needed a fantastic...</li> </ol> <p style="text-align: right; font-size: small;">*Informações ainda não oficiais. Este trabalho faz parte do projeto lexicográfico em andamento</p>	

### Entrada Lexicográfica 1 - Ball

<b>Unidade Lexical:</b> bola	<b>Frame:</b> Bola ( <a href="#">hyperlink para os dados do <i>frame</i></a> )
<p><b>Definição*:</b> objeto esférico que é chutado e jogado no futebol.</p> <p><b>Relações Frame a Frame*:</b> É usado por: <a href="#">Ataque</a>, <a href="#">Ações_Técnico-Táticas</a> (<a href="#">hyperlink para os dados do <i>frame</i></a>)</p> <p><b>Possível Equivalente de Tradução:</b> <a href="#">ball</a> (<a href="#">hyperlink para a entrada do equivalente</a>)</p> <p><b>Exemplos*:</b></p> <ol style="list-style-type: none"> <li>1. Carlinhos recuperou a <b>bola</b> na intermediária e tocou curto para Thiago Neves.</li> <li>2. O meia tomou a <b>bola</b> na direita e bateu falta nas mãos do goleiro.</li> <li>3. Bruno deu um leve soco e a <b>bola</b> sobrou para Bernard.</li> </ol> <p style="text-align: right; font-size: small;">*Informações ainda não oficiais. Este trabalho faz parte do projeto lexicográfico em andamento</p>	

### Entrada Lexicográfica 2 – Bola

A vantagem da utilização dos preceitos da Semântica de *Frames* neste caso é visível na perspectiva de que o *frame* comum nos dois idiomas do dicionário é capaz de dar suporte à pesquisa realizada em análise de *corpus*. Desta forma, as unidades lexicais deixam de ser a base da consulta para localização de equivalências e passam a ser secundárias, ao dar lugar, primeiramente, à equivalência conceitual entre os *frames* comuns aos dois idiomas. A respeito do caso específico que acaba de ser analisado, pode-se dizer que as evidências disponíveis no *corpus* realmente apontam para uma equivalência exata de conteúdos e que, neste caso, a semântica de *frames* teve um papel pouco expressivo na localização da equivalência. Todavia, acredita-se que o formato de disponibilização das entradas lexicográficas pode ser útil para o consulente que tiver dúvidas sobre a utilização das ULs *ball* e *bola*.

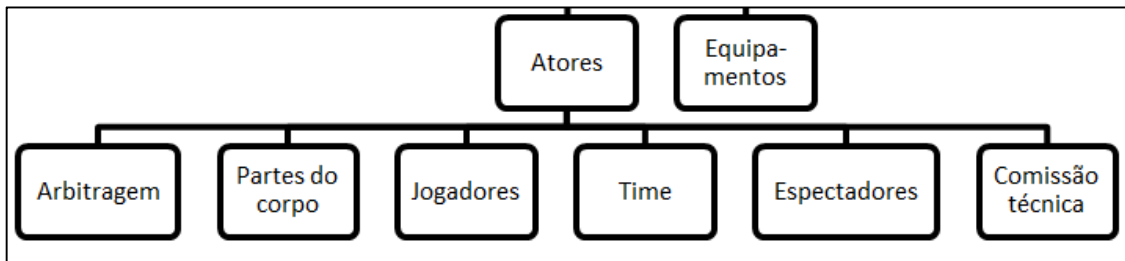
#### 4.2.1.2 Goleiro

##### *Fase 1: Característica da UL goleiro*

Também um caso de equivalência exata de conteúdo, mas com a equivalência acontecendo de forma distinta, é a UL *goleiro*, que descreve um referente concreto,

monossêmico e não ambíguo. Esta UL tem uma frequência alta no *corpus* em português (2.300 aparições) e evoca o *frame* Jogadores .

Na estruturação de *frames* proposta pelo Kicktionary\_Br este *frame* aparece da seguinte forma (Figura 20):



**Figura 20** - Posição do *frame* Jogadores na remodelagem de frames proposta pelo Kicktionary\_Br

#### *Fase 2: Consulta a dicionários*

Na consulta pela UL *goleiro* em dicionários, os seguintes resultados foram obtidos:

	Consulta por:
Dicionário:	<i>Goleiro</i>
Word Reference	Goalkeeper (informal) goalie (football, soccer) keeper
Michaelis	(sport) goalkeeper
Babylon	n. (Brazil) goalkeeper, one who guards the entrance of a goal to prevent the ball or puck from entering (Sports)

**Quadro 2** - Resultado da busca por *goleiro* em dicionários online

Observou-se que todos os dicionários apresentam pelo menos uma possível equivalência para a UL *goleiro*, sendo que a principal é *goalkeeper*, especificada quanto ao uso no domínio esportivo. No recurso Babylon, notou-se também, além da indicação da equivalência, uma breve descrição da UL, como “aquele que protege a entrada do gol para evitar que a bola ou o disco entrem neste”.

Aparentemente, valendo-se apenas desta consulta, presume-se que novamente este será o caso de uma equivalência direta na qual *goleiro* e *goalkeeper* são os equivalentes mais comuns. Para confirmar esta hipótese, parte-se então para a análise do *corpus*.

### Fase 3: Busca de equivalências utilizando-se da linguística de corpus

Assim como no caso anterior, espera-se que esta análise do *corpus* venha a validar a possível equivalência entre *goleiro* e *goalkeeper*, servindo como mais uma evidência de que esta relação é exata e direta. Como plano de fundo tem-se o *frame* Jogadores a guiar a consulta, dando suporte à proposta de que ambas as ULs remetem a uma mesma entidade em ambos os idiomas.

Inicialmente buscou-se a UL *goleiro*. Tal consulta gerou uma listagem de 2.300 sentenças na ferramenta Concord do WordSmith Tools. Ao analisar as 25 primeiras sentenças do *corpus* confirmou-se que a UL é utilizada em todos os casos descrevendo o jogador responsável por defender as tentativas do gol do time adversário. A figura abaixo representa o resultado obtido na busca (Figura 21):

N	Concordance	Set	Wo
1	a bola nos pés de Jorge Henrique, que acionou Romarinho em velocidade, porém o goleiro vascaíno saiu com precisão e evitou a conclusão do atacante corintiano. O		192
2	mesmo sem acelerar, acabou criando a primeira oportunidade aos dois minutos. O goleiro Fernando Prass errou a reposição de bola e colocou a bola nos pés de		192
3	Logo na primeira jogada, Doriva recuou de forma atrapalhada para Ricardo. O goleiro catarinense saiu de qualquer jeito de sua meta e chutou em cima de Bruno		192
4	pela direita e recuou para Auremir, que cruzou para Alecsandro na área, mas o goleiro Cássio foi mais rápido e segurou a bola antes da conclusão do atacante.		193
5	redes. Aos 40, Bruno Peres lançou para a área, Miralles ganhou a dividida com o goleiro Roberto e empurrou a bola para o gol, recolocando os santistas em		185
6	a mudar aos 12 minutos, quando Alecsandro foi lançado dentro da área e o goleiro Dida derrubou o atacante alvinegro. O árbitro goiano André Luiz de Freitas		73,
7	Cléber Santana arriscou de longe, a bola desviou em Liédson e ia no canto. Só o goleiro Victor se esticou, tocou nela e a viu tocar na trave. Nos minutos finais, o		190
8	Santana recebeu a bola na intermediária, avançou e chutou da entrada da área. O goleiro Victor voou e viu a bola bater na trave esquerda e ir para fora. Depois do		190
9	casa não deram sossego para seus adversários e ameaçaram aos 11 minutos. O goleiro Ricardo saiu jogando muito mal e Marcelo Cordeiro tentou o chute de		192
10	O argentino tabelou com William José, entrou livre na área e, ao tentar deslocar o goleiro Renan, acabou botando a bola para fora. A torcida botafoguense não		69,
11	chutar, um pouco desequilibrado, próximo à trave direita de Márcio, assustando o goleiro Márcio. Melhor no segundo tempo, o Peixe quase chegou ao empate em		206
12	bola pela esquerda, ver o levantamento se tornar finalização e quase surpreender o goleiro Marcelo Lomba. A bola terminou em cima do gol, próxima ao travessão. O		205
13	no corte para cima da marcação de Jackson e lançou Aloísio em velocidade. O goleiro Magrão leu a jogada melhor que o atacante adversário e conseguiu fazer a		204
14	habitual e colocou no ângulo esquerdo sem qualquer chance de defesa para o goleiro Magrão. Logo após sofrer o gol, o técnico Vágner Mancini colocou Hugo e		198
15	crescer na partida e exigir duas grandes intervenções de Dida. Aos 18 minutos, o goleiro pulou em passe cruzado e evitou que a bola chegasse nos pés de Loco		192
16	assustou o Atlético-GO, com um arremate de longa distância, aos 13. Atento, o goleiro Márcio espalmou para escanteio. Na sequência, Felipe Anderson bateu o		205
17	salvando o Peixe. Porém, se na primeira boa investida ofensiva do Dragão o goleiro santista salvou, no minuto seguinte Aranha não pôde fazer nada. Após troca		205
18	tempo de jogo, a Portuguesa chegou ao ataque pelas laterais e assustou o goleiro Ricardo. Aos oito minutos, Moisés tabelou com Ananias e invadiu o lado		191
19	apontou impedimento. A partir deste lance, Magrão se tornou o nome do jogo. O goleiro pegou de longe ou cara a cara chutes de Ademilson, Jadson, Cícero e de		196
20	Gabriel, que deixou para Gilberto bater no canto direito rasteiro de Rogério Ceni. O goleiro de 39 anos teve que esticar o ombro operado há seis meses para evitar o		195
21	para o gol e obrigou Titi a recuar para Marcelo Lomba. Como o passe foi curto, o goleiro teve que tirar de bico, no susto. Sem pressa, o Grêmio cadenciava o jogo		194
22	colocou Alexandre Pato na cara de Renan. O atacante de R\$ 40 milhões driblou o goleiro e incrivelmente chutou por cima. Poupado do primeiro tempo, Guerrero		71,
23	o estreante Lodeiro tentou um cruzamento pela direita e quase surpreendeu o goleiro Victor com a curva feita pela bola. A resposta atleticana veio com Leandro		220
24	desempenho bastante irregular e nas poucas vezes em que conseguiu concluir, o goleiro Felipe apareceu com destaque. Na próxima rodada, o Flamengo enfrentará		222
25	Leonardo Silva tentou o corte, mas o rebote sobrou para Andrezinho, que fuzilou o goleiro Victor abrindo os trabalhos no Independência. Aos 43, veio o troco do Galo		220

Figura 21 - Resultado da busca por *goleiro* no Concord do WordSmith Tools

Na análise dos *clusters*, (Figura 22), obtiveram-se indícios que apontam ações do *goleiro*, pois há agrupamentos falando sobre a *defesa do goleiro*, *nas mãos do (goleiro)*, *sem chances para o goleiro*, entre outros. Com isso chegou-se a duas proposições: 1) os *clusters* apontam principalmente para as ações do goleiro, o que confirma a tese da UL ser

monossêmica; e 2) os *clusters* deixam de apontar para o mesmo *frame* da UL isoladamente, como no caso de *defesa do goleiro*, que evoca o *frame* Defesa ou de *driblou o goleiro*, que evoca o *frame* Drible. Sabendo-se destas particularidades da UL *goleiro*, passou-se para a busca e análise de um possível equivalente, para verificar se tal UL também teria estas características.

N	Cluster	Freq.	Set	Length	Related
1	MAS O GOLEIRO	165	3		
2	E O GOLEIRO	153	3		
3	PARA O GOLEIRO	109	3		
4	O GOLEIRO DO	97	3		
5	COM O GOLEIRO	79	3		
6	DEFESA DO GOLEIRO	79	3		
7	SAÍDA DO GOLEIRO	62	3		
8	NA SAÍDA DO	59	3		
9	CIMA DO GOLEIRO	55	3		
10	O GOLEIRO E	45	3		
11	MÃOS DO GOLEIRO	42	3		
12	EM CIMA DO	42	3		
13	NAS MÃOS DO	39	3		
14	O GOLEIRO FELIPE	39	3		
15	DO GOLEIRO E	34	3		
16	DRIBLOU O GOLEIRO	34	3		
17	O GOLEIRO FÁBIO	33	3		
18	DA ÁREA E	31	3		
19	OBRIGOU O GOLEIRO	30	3		
20	O GOLEIRO MÁRCIO	30	3		
21	O GOLEIRO MARCELO	30	3		
22	A CARA COM	29	3		
23	SEM CHANCES PARA	28	3		
24	CHANCES PARA O	28	3		
25	CARA COM O	28	3		

**Figura 22** - Resultado dos *clusters* para *goleiro* no Concord do WordSmith Tools

Esperava-se, nesta etapa, que a localização do equivalente de *goleiro* fosse simples, assim como ocorreu no caso de *ball*. Porém, neste caso a situação foi diferente: enquanto que no *corpus* em português a UL *goleiro* teve uma frequência de 2.300 casos, a UL que era considerada um possível equivalente devido à consulta aos dicionários (*goalkeeper*) retornou apenas 533 aparições.

Supõe-se que este tipo de situação possa ter acontecer por dois motivos: 1) o idioma alvo não utiliza a UL disponível em seu léxico com tanta frequência; e 2) é possível que exista mais do que uma forma de expressar tal UL na língua alvo, dividindo a frequência entre mais de uma UL. Assim sendo, foi necessário fazer novas consultas ao *corpus* em inglês buscando-se também por sinônimos de *goalkeeper* para compreender qual o motivo de tal diferença de frequência do equivalente de *goleiro* no *corpus*.

Retomou-se, então, a consulta realizada no Dicionário WordReference na fase anterior desta análise que apontou *goalkeeper*, *goalie (inf.)* e *keeper* como traduções de *goleiro* e, um a um, estes termos foram buscados no *corpus*.

A consulta por *goalie* não retornou nenhuma sentença, porém a pesquisa por *keeper* retornou 611 entradas (aprox. 15% a mais de sentenças do que as que utilizam a UL *goalkeeper*). Neste *subcorpus* (parcialmente apresentado na Figura 23), constatou-se que a grafia de tal UL acontece de duas formas: *keeper* e *'keeper*, sendo que a segunda opção evidencia a supressão de letras de *goalkeeper*. A frequência de cada uma difere significativamente, sendo que *keeper* tem 530 menções (87% dos casos) e *'keeper* apresenta apenas 81 (13% dos casos).



**Figura 23** - Resultado da consulta por *keeper* no Concord do WordSmith Tools

Assim sendo, considerando-se o total de sentenças que citam o goleiro no *corpus* em inglês (1.144 sentenças), 47% utilizam a UL *goalkeeper*, 46% mencionam *keeper*, e apenas 7% fazem uso da UL *'keeper*. Este tipo de informação pode ser conveniente em um recurso lexical para que o consulente tenha argumentos para optar por uma ou outra como equivalente.

Para finalizar a análise do *corpus*, foram observados os *clusters* disponíveis em cada uma das ULs em inglês para verificar se existem evidências de que eles remetam ao mesmo *frame* que a UL *goleiro*. Os 25 primeiros *clusters* de *goalkeeper* (Figura 24) se mostraram pouco representativos, tendo frequência de 12 a 5 casos, diferente do que aconteceu com *goleiro*, com *clusters* de 165 a 28 repetições. O conteúdo destes agrupamentos também difere, pois não se notou uma frequência de *clusters* como *a defesa do goleiro* ou *driblar o goleiro*, ou seja, que ajudem a caracterizar as ações do mesmo. Por outro lado, reparou-se que vários destes agrupamentos apontam para o nome completo do goleiro dos times em questão.

N	Cluster	Freq.	Set	Length	Related
1	AT THE GOALKEEPER	12	3	3	
2	STRAIGHT AT THE	11	3	3	
3	THE GOALKEEPER WAS	10	3	3	
4	THE GOALKEEPER AND	9	3	3	
5	GOALKEEPER MARK SCHWARZER	9	3	3	
6	GOALKEEPER ASMIR BEGOVIC	9	3	3	
7	BUT THE GOALKEEPER	9	3	3	
8	GOALKEEPER SIMON MIGNOLET	8	3	3	
9	GOALKEEPER JUSSI JAASKELAINEN	8	3	3	
10	GOALKEEPER HAD TO	8	3	3	
11	TO THANK FOR	7	3	3	
12	GOALKEEPER WOJCIECH SZCZESNY	7	3	3	
13	GOALKEEPER TIM KRUL	7	3	3	
14	GOALKEEPER TIM HOWARD	7	3	3	
15	GOALKEEPER ROB ELLIOT	7	3	3	
16	GOALKEEPER JULIO CESAR	7	3	3	
17	GOALKEEPER BRAD GUZAN	7	3	3	
18	GOALKEEPER ADAM FEDERICI	7	3	3	
19	READING GOALKEEPER ADAM	6	3	3	
20	HIS OWN GOALKEEPER	6	3	3	
21	GOALKEEPER MARK BUNN	6	3	3	
22	DID WELL TO	6	3	3	
23	AND THE GOALKEEPER	6	3	3	
24	THE GOALKEEPER TO	5	3	3	
25	THE GOALKEEPER THE	5	3	3	

**Figura 24** - Resultado dos *clusters* para *goalkeeper* no Concord do WordSmith Tools

Pensou-se então em agrupar a consulta buscando-se, no Concord, por *goalkeeper / keeper*. Com isso, os *clusters* levariam em consideração agrupamentos de palavras de forma geral, considerando as três ULs disponíveis como possíveis equivalentes de *goleiro*. Como se pode ver abaixo (Figura 25), os agrupamentos não apontaram as ações do goleiro de forma tão direta como aconteceu no *corpus* em português, que apresentou *clusters* como: *defesa do goleiro*, *sem chance para o (goleiro)*, *driblou o goleiro*. Porém, notaram-se alguns casos que apontam a forma como a bola (possivelmente) chega ao goleiro, como *shot straight at (the keeper)*, *past the keeper*, *over the keeper*, etc.

N	Cluster	Freq	Set	Length	Related
1	STRAIGHT AT THE	29	5		
2	BUT THE KEEPER	25	5		
3	AT THE KEEPER	22	5		
4	THE KEEPER WAS	19	5		
5	AND THE KEEPER	18	5		
6	THE KEEPER BUT	13	5		
7	KEEPER SIMON MIGNOLET	13	5		
8	AT THE GOALKEEPER	12	5		
9	PAST THE KEEPER	11	5		
10	THE KEEPER AND	11	5		
11	WITH THE KEEPER	11	5		
12	KEEPER BEN FOSTER	11	5		
13	KEEPER BRAD GUZAN	11	5		
14	THE BALL OVER	10	5		
15	THE GOALKEEPER WAS	10	5		
16	OVER THE KEEPER	9	5		
17	BUT THE GOALKEEPER	9	5		
18	GOALKEEPER MARK SCHWARZER	9	5		
19	BALL OVER THE	9	5		
20	THE GOALKEEPER AND	9	5		
21	GOALKEEPER ASMIR BEGOVIC	9	5		
22	DID WELL TO	9	5		
23	THE KEEPER TO	8	5		
24	GOALKEEPER SIMON MIGNOLET	8	5		
25	PAST HIS OWN	8	5		

**Figura 25** - Resultado dos *clusters* para *keeper* e *'keeper* no Concord do WordSmith Tools

Considerando-se os resultados desta análise de *corpus*, defende-se que, embora as ULs em inglês e português não apresentem relações de frequência diretamente proporcionais quanto aos *clusters* em textos de diferentes *corpora*, o conceito concreto de jogador que defende o time de gols adversários segue sendo mantido. Além disso, devido à consulta ao *corpus*, é possível afirmar que as ULs *goalkeeper*, *keeper* e *'keeper* podem ser todas consideradas equivalentes da UL *goleiro*, porém, por ter-se acesso às frequências de cada caso, destaca-se que há indícios de que a UL *'keeper* seja utilizada com uma frequência significativamente menor que as demais.

*Fase 4: Reflexão sobre o exercício de apresentação das possíveis equivalências em um recurso lexical baseado em frames.*

No caso da UL *goleiro* e seus equivalentes, a relação a ser apresentada no Dicionário da Copa segue sendo direta e não ambígua, porém, é necessário que seja explicitado que existe mais de uma forma de expressar tal UL em inglês. Sugere-se, portanto, as seguintes entradas lexicográficas, que podem ser acessadas diretamente, uma a outra, através de *hyperlinks*.



<b>Unidade Lexical:</b> goleiro	<b>Frame:</b> Jogador ( <a href="#">hyperlink para os dados do <i>frame</i></a> )
<p><b>Definição*:</b> jogador que defende o gol de seu time, utilizando seu corpo para afastar a bola.</p> <p><b>Relações Frame a Frame*:</b></p> <p>Herda de: <i>Atores</i> (<a href="#">hyperlink para os dados do <i>frame</i></a>)</p> <p>É usado por: <i>Ações_Técnico-Táticas</i>, <i>Funções_do_Juiz</i> (<a href="#">hyperlinks para os dados dos <i>frames</i></a>)</p> <p><b>Possível Equivalente de Tradução:</b> goalkeeper, keeper, 'keeper (<a href="#">hyperlink para as entradas de cada equivalente</a>)</p> <p><b>Exemplos*:</b></p> <ol style="list-style-type: none"> <li>1) O <b>goleiro</b> Fernando Prass errou a reposição de bola e colocou a bola nos pés de Jorge Henrique, que acionou Romarinho em velocidade.</li> <li>2) O <b>goleiro</b> catarinense saiu de qualquer jeito de sua meta e chutou em cima de Bruno Mineiro para ganhar o tiro de meta.</li> <li>3) (...) mas o <b>goleiro</b> Cássio foi mais rápido e segurou a bola antes da conclusão do atacante.</li> </ol> <p style="text-align: right;">*Informações ainda não oficiais. Este trabalho faz parte do projeto lexicográfico em andamento</p>	

### Entrada Lexicográfica 3 – Goleiro

<p><b>Lexic Unit:</b> goalkeeper</p> <p>See also: <i>keeper</i>, <i>'keeper</i></p> <p>(<a href="#">hyperlink to the other LUs that evoke the same <i>frame</i></a>)</p>	<p><b>Frame:</b> Player</p> <p>(<a href="#">hyperlink to the <i>frame</i> details</a>)</p>
<p><b>Definition*:</b> player who defends its team's goal, using the body to keep the ball away.</p> <p><b>Frame to Frame Relations*:</b></p> <p>Inherits from: <i>Actors</i> (<a href="#">hyperlink to the <i>frame</i> details</a>)</p> <p>Is used by: <i>Technical-Tactical_Actions</i>, <i>Referee_Function</i> (<a href="#">hyperlinks para os dados dos <i>frames</i></a>)</p> <p><b>Possible Translation Equivalent:</b> goleiro (<a href="#">hyperlink to the <i>equivalent</i> entry</a>)</p> <p><b>Examples*:</b></p> <ol style="list-style-type: none"> <li>1) He is really a <b>goalkeeper</b> but he's really fast, so we switched him to a striker.</li> </ol>	

- 2) Cole and Suso swapped passes on the edge of the box to engineer an opening, and the former lifted the ball over the advancing **goalkeeper** to allow Shelvey to head over the line from inches out.
- 3) \* 59' ...and it's straight to Khune, the South Africa **goalkeeper**.

\*Informações ainda não oficiais. Este trabalho faz parte do projeto lexicográfico em andamento

#### Entrada Lexicográfica 4 - Goalkeeper

<p><b>Lexic Unit:</b> Keeper</p> <p>See also: <i>goalkeeper</i>, <i>'keeper</i></p> <p>(<a href="#">hyperlink to the other LUs that evoke the same <i>frame</i></a>)</p>	<p><b>Frame:</b> Player</p> <p>(<a href="#">hyperlink to the <i>frame</i> details</a>)</p>
<p><b>Definition*:</b> player who defends its team's goal, using the body to keep the ball away.</p> <p><b>Frame to Frame Relations*:</b></p> <p>Inherits from: <i>Actors</i> (<a href="#">hyperlink to the <i>frame</i> details</a>)</p> <p>Is used by: <i>Technical-Tactical_Actions</i> , <i>Referee_Function</i> (<a href="#">hyperlink to the <i>frame</i> details</a>)</p> <p><b>Possible Translation Equivalent:</b> goleiro (<a href="#">hyperlink to the <i>equivalent</i> entry</a>)</p> <p><b>Examples*:</b></p> <ol style="list-style-type: none"> <li>1) As it was, Malaysia stood firm until three minutes after half-time when a threaded ball from Aguero put in Tevez who held off his marker to hit a low left-foot drive past the <b>keeper</b>.</li> <li>2) * 80' It was a decent shot by Koren but the <b>keeper</b> will be blaming the Jbulani for that.</li> <li>3) Ramsey's shot was pushed wide by the impressive QPR <b>keeper</b>, who also denied Cazorla as Arsenal continued to press in the closing stages.</li> </ol> <p>*Informações ainda não oficiais. Este trabalho faz parte do projeto lexicográfico em andamento</p>	

#### Entrada Lexicográfica 5 - Keeper

<p><b>Lexic Unit:</b> 'Keeper See also: <i>goalkeeper, keeper</i> (<a href="#">hyperlink to the other LUs that evoke the same frame</a>)</p>	<p><b>Frame:</b> Player (<a href="#">hyperlink to the frame details</a>)</p>
<p><b>Definition*:</b> player who defends its team's goal, using the body to keep the ball away.</p> <p>① <b>Comment:</b> this lexic unit has a small usage frequency base on statistical analysis.</p> <p>① <i>Comentário:</i> esta unidade lexical apresenta uma baixa frequência de uso baseada em análise estatística</p> <p><b>Frame to Frame Relations*:</b></p> <p>Inherits from: <code>Actors</code> (<a href="#">hyperlink to the frame details</a>)</p> <p>Is used by: <code>Technical-Tactical_Actions</code> , <code>Referee_Function</code> (<a href="#">hyperlink to the frame details</a>)</p> <p><b>Possible Translation Equivalent:</b> <code>goleiro</code> (<a href="#">hyperlink to the equivalent entry</a>)</p> <p><b>Examples*:</b></p> <ol style="list-style-type: none"> <li>1) As it was, Malaysia stood firm until three minutes after half-time when a threaded ball from Aguero put in Tevez who held off his marker to hit a low left-foot drive past the <b>keeper</b>.</li> <li>2) * 80' It was a decent shot by Koren but the <b>keeper</b> will be blaming the Jbulani for that.</li> <li>3) Ramsey's shot was pushed wide by the impressive QPR <b>keeper</b>, who also denied Cazorla as Arsenal continued to press in the closing stages.</li> </ol> <p style="text-align: right;"><small>*Informações ainda não oficiais. Este trabalho faz parte do projeto lexicográfico em andamento</small></p>	

### Entrada Lexicográfica 6 – 'Keeper

Como se pode verificar, neste caso uma única UL em português remete a três possíveis equivalentes em inglês, sendo que o uso de um deles ('keeper) está condicionado ao saber de que se trata de uma UL pouco utilizada, de acordo com o observado em nosso *corpus*. Portanto, a relação que deve ser disponibilizada é de um *frame* (no caso, `Jogador`) que tem como um dos evocadores a UL *goleiro*, em português, relacionada a 3 possíveis equivalentes em inglês, sendo dois deles mais indicados devido ao número de ocorrências ser mais representativo.

Neste caso, além da vantagem de se ter os *frames* como plano de fundo para estruturar estas entradas no dicionário, foi possível notar também o auxílio fornecido pelo trabalho com *corpus*. Somente tendo-se acesso a resultados de pesquisas sobre frequências é possível definir os termos mais utilizados em determinados tipos de textos ou domínios.

## 4.2.2 Intersecção de conteúdo

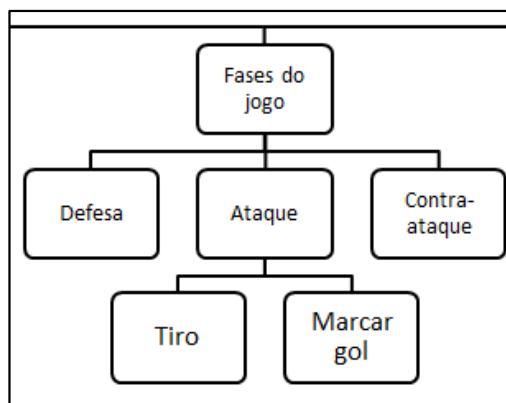
Outra situação observada na primeira parte da análise realizada nesta pesquisa foi o de equivalentes com intersecção de conteúdo. Neste caso, as equivalências não são exatas entre os dois idiomas em contato, mas ainda assim existe certo grau de equivalência capaz de associar as ULs das línguas fonte e alvo, por haver pontos em comum entre elas.

Exemplos disso são as ULs *shot* e *falta*, que tiveram alta frequência, mas que não remetem diretamente a uma única possibilidade de tradução. Cada um destes casos será analisado a seguir.

### 4.2.2.1 Shot

#### *Fase 1: Característica da UL shot*

Foram observadas 2.835 ocorrências da UL *shot* no *corpus* do Kicktictionary\_Br Copa. Na estruturação de *frames* proposta pelo Kicktictionary\_Br, a UL *shot* aparece como evocadora do *frame* Tiro, um dos desdobramentos do *frame* Ataque, conforme imagem abaixo (Figura 26):



**Figura 26** - Posição do *frame* *Tiro* na remodelagem de *frames* proposta pelo Kicktictionary\_Br

O *frame* tiro possui como EFs: o JOGADOR\_QUE\_ATIRA, o TIRO, a BOLA, a META e o GOLEIRO. Isso porque este *frame* representa as possíveis formas de o JOGADOR\_QUE\_ATIRA lançar (TIRO) a BOLA ao gol (META), o que pode ser impedido pelo GOLEIRO.

Considerando-se este detalhamento, partiu-se para a análise desta UL em dicionários.

### Fase 2: Consulta a dicionários

Ao consultar a UL *shot* obtiveram-se, de forma geral, resultados que remetiam ao substantivo *shot* e ao passado do verbo *shoot*, que também tem a mesma grafia. Vamos nos deter, nesta análise, apenas a UL *shot* na forma de substantivo. Na consulta a dicionários, obtivemos os seguintes equivalentes sendo indicados em cada um deles:

	Consulta por:
Dicionário:	<i>Shot</i>
Word Reference	tiro
Michaelis	<i>n</i> 1 tiro. 2 chumbo. 3 bala, projétil. 4 descarga de arma de fogo. 5 distância de tiro. 6 atirador. 7 lance, jogada. 8 tentativa.
Babylon	s. tiro; pontaria; jogada; foto; fotografia; pontaria; franco-atirador; chance; adivinhação, teste

**Quadro 3-** Resultado da busca por *shot* em dicionários online

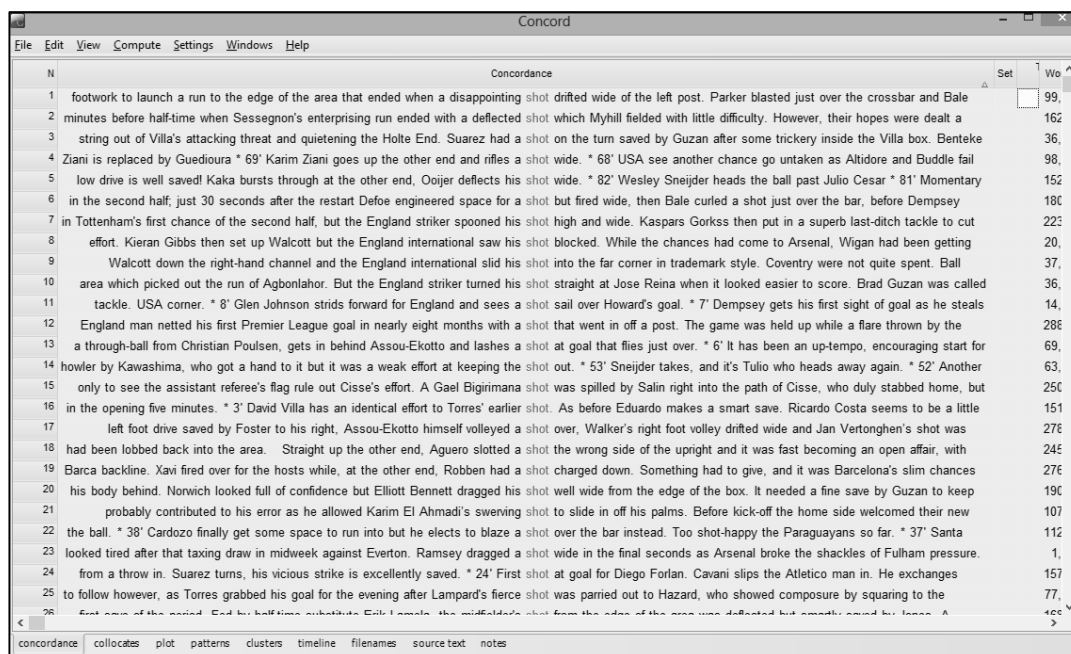
Novamente retomamos que apenas mencionam-se no quadro acima os equivalentes que não são definidos como pertencentes a outros domínios específicos. No WordReference nota-se que a equivalência principal destacada é de *tiro*, mas não há especificações sobre o uso desta UL no domínio do futebol. Já o Michaelis e o Babylon apontam uma tradução que é recorrente nesta área, que é a de *jogada*, ou, no caso do primeiro recurso, também se fala em *lance*.

Após esta consulta, algumas dúvidas surgiram a respeito destas possíveis equivalências: 1) As ULs *lance* e *jogada* são mais neutras do que a UL *tiro*, mas elas podem ser consideradas possíveis equivalentes assim mesmo?; e 2) Utiliza-se a UL *tiro* como tradução de *shot* no domínio do futebol? Buscando as respostas a estas perguntas partiu-se para a análise dos *corpora*.

### Fase 3: Busca de equivalências utilizando-se da linguística de corpus

O primeiro teste realizado foi observar o comportamento da UL *shot* no *corpus* em inglês. Das 2.835 ocorrências no *corpus*, foram selecionadas as 25 primeiras para análise (Figura 27). Observou-se que, destas 25 sentenças, todas trazem o substantivo *shot* e em praticamente metade dos casos observam-se traços de que esta jogada esteja relacionada a algo que ocorre próxima ao gol ou tendo este como alvo, como em *a disappointing shot*

*drifted wide of the left post, a deflected shot, had a shot on the turn saved by, a shot well wide from the edge of the box, first shot at goal, etc.* Com isso, aumentam-se os indícios de que tal UL está vinculada ao *frame Ataque* e, mais especificamente, evoca o *frame, Tiro*.



**Figura 27** - Resultado da busca por *shot* no Concord do WordSmith Tools

Partiu-se então para a análise da UL *tiro* em português para que se comparassem as frequências e os usos de ambas as ULs. Observou-se que a consulta por *tiro* gerou um resultado de apenas 160 entradas, sendo que 79 delas traziam o *cluster tiro de meta* (que não é um tipo de chute com o objetivo de marcar um gol, mas uma jogada de bola parada executada a partir da área do goleiro). Tendo-se isso em vista, aparentemente a UL *tiro* não seria o equivalente mais adequado para *shot*. Porém, depois de eliminarem-se os casos de entradas de *tiro de meta*, notou-se que em algumas sentenças o uso da UL *tiro* era comparável ao da UL *shot* (Figura 28), conforme exemplos disponíveis na imagem abaixo, que falam sobre *arriscar o tiro, inaugurou o marcador com um tiro, arriscou um tiro, etc.* Assim sendo, *tiro* pode ser considerado, em alguns casos, mesmo com baixa frequência, um possível equivalente de *shot*.



Figura 28 - Resultado da busca por *tiro* no Concord do WordSmith Tools

A mesma análise foi realizada com a *UL jogada*. Por *jogada* entende-se um lance do jogo, ou seja, um movimento feito por um jogador e o observado no *corpus* veio ao encontro desta definição. A listagem de 25 ULs (Figura 29), neste caso, apontou para a flexibilidade de *jogada* em expressões como *jogada ofensiva*, *jogada aérea*, *tentar a jogada individual*. Mas o mais recorrente foi o *cluster boa jogada*. Assim sendo, aparentemente este não é o equivalente mais adequado para disponibilização no Dicionário da Copa 2014.

N	Concordance	Word #	Sen S
1	Gomes. O Cruzeiro segurou o ímpeto do adversário e, logo na primeira jogada ofensiva, aos 13 minutos, abriu o marcador. O atacante Borges	12,855	713
2	pela direita e chuta muito alta, sem perigo para Marcelo Grohe. Mas a jogada do Católica foi boa Viničius Rebello: Menezes recebeu livre e o	91,088	5,6
3	vai fazer a diferença Gatti!!!! O goleiro salva o Volta Redonda em grande jogada de Roberson GENTE É INCRIVEL COMO VOCES GREMISTAS	26,622	1,8
4	jogadas para diminuir ainda no primeiro tempo, aos 29. Fred fez bela jogada, deslocou o goleiro Wilson e rolou para Fórlan, que chutou no	380,81	1,7
5	com perigo. O lance mostrou a fragilidade da defesa alvinegra na jogada aérea e aos 33 minutos, o goleiro Jefferson hesitou ao sair para	141,65	7,7
6	à trave esquerda de Berna. Alex permaneceu desfilando talento. Em jogada na lateral, ele tocou por entre as pernas de um adversário e	2,653	166
7	uma segurada nos comentários. Eduardo Cecconi: Podolski faz grande jogada e toca para o meio. Muller dá um come, deixa o zagueiro sentado	20,772	1,4
8	campo de defesa para o de ataque. Pela direita, Lucas Marques fez boa jogada e tocou para trás, encontrando Andrezinho. O camisa 17 hesitou	350,76	67
9	ganhou de Edinho e bateu para empatar: 1 a 1. Aos 18, nova grande jogada de Sobis na área, que acertou o travessão. Aos 27, Fernando	44,803	2,5
10	de Daniel Carvalho, que desviou e se perdeu pela linha de fundo. Na jogada seguinte, Juninho fez boa jogada individual pela esquerda e	139,46	7,6
11	bem, o Inter ia criando na base do abafa. Aos 28, Lucas Lima fez boa jogada, invadiu a área e chutou cruzado, mas Cassiano chegou atrasado	273,84	14
12	à direita da trave defendida por Magrão. Aos 22 minutos, Nilton tentou a jogada individual e foi puxado por Diego Ivo na entrada da área. Juninho	198,48	10
13	primeiro para ter um lance de perigo. Aos três minutos, Carlinhos fez boa jogada individual, invadiu a área pela esquerda e bateu com força com a	73,650	4,2
14	camisa 11, de frente para gol, bateu forte, mas a bola foi para fora. Boa jogada! Cícero pegou a bola no campo de defesa e arrancou, o meia	114,21	6,7
15	área e bateu rasteiro: 3 a 1. Mas a zaga tricolor voltou a falhar em uma jogada aérea. Após cobrança de escanteio da direita e o desvio no	143,45	7,8
16	na área e chuta. A bola desvia e é escanteio! Tati Lopes: UUUHHH! Boa jogada de Lúcio, driblando o marcador no lado esquerdo da área. Ele	80,038	4,9
17	o Galo voltou a sofrer aos 15 minutos. O time não emplacava nem uma jogada sequer e viu o volante Souza invadir a área pela esquerda e chutar	269,66	14
18	Austrália, Schwarzer sai e pega. Eduardo Cecconi: ESCANTEIO! Grande jogada de Muller. Para a Alemanha, pela direita. Eduardo Cecconi:	19,601	1,2
19	sai após raspar na trave. Gustavo Heldt: UUUH! Talles Cunha fez bela jogada com direito a tabela com Guto e entrada veloz pela área. O chute	35,021	2,6
20	da casa igualassem as ações rapidamente. Tanto que a primeira boa jogada foi santista. Gerson Magrão colocou a bola entre as pernas de	255,78	13
21	recebe dentro da área, tira o marcador da jogada, mas chuta fraco. Boa jogada do garoto Viničius Rebello: Uhhhh - Victor salva o Grêmio. Diego	94,319	5,6
22	para Rildo, e desta vez o atacante não desperdiçou. William fez boa jogada individual pelo canto esquerdo da área e chutou. Rildo aproveitou	358,93	471
23	coloca o time no ataque. Sai Marrone Roberson. Ricardinho faz boa jogada mas chuta muito mal. Um balão. Ricardinho leva amarelo por	59,945	4,2
24	João Paulo desvia para escanteio. Juliano Schüler: Kleber tenta combinar jogada com Andrezinho pela esquerda. Mas a marcação do VEC toma a	40,711	3,0
25	busca de um grandalhão qualquer para cabecear. Nada disso. Foi uma jogada lapidada com pés de artistas. De Ozil para Müller, de Müller para	78,958	4,5

Figura 29 - Resultado da busca por *jogada* no Concord do WordSmith Tools

E o último teste foi com a *UL lance*. Neste caso, assim como ocorreu com *jogada*, os resultados obtidos não apontaram diretamente para o mesmo significante que o da *UL shot*. Como é visível nas 25 ocorrências analisadas abaixo (Figura 30), esta *UL* acontece em diversas situações do jogo, como *um lance confuso*, *limpar o lance*, *um lance de bola parada*, *no primeiro lance do jogo*, *outro lance do primeiro tempo* e não somente em chutes próximos ao gol, o que inclusive tem uma baixa frequência, notado apenas em casos como *errou no lance do gol* ou *resultou em lance perigoso*. Tendo-se isso em vista, não se indica esta *UL* como a equivalência mais adequada para o Dicionário da Copa.

The screenshot shows the Concord window in WordSmith Tools. The window title is 'Concord' and it has a menu bar with 'File', 'Edit', 'View', 'Compute', 'Settings', 'Windows', and 'Help'. Below the menu bar is a toolbar with icons for 'Set' and 'Word'. The main area displays a concordance search for the word 'lance'. The results are listed in a table with two columns: 'Word' and 'Word'. The 'Word' column contains the text of the concordance entries, and the 'Word' column contains the frequency of each entry. The results are as follows:

Word	Word
grandes chances de perigo. Segundo tempo Aos seis minutos, em um lance confuso dentro da área, Cássio defendeu e deu rebote para o ataque para Martín Silva e a bola saiu fraca. Antes que o goleiro chegasse no lance, Rhayner invadiu a área como um raio e tocou por cima de Martín	1!
direto pela linha de fundo. Júnior Urso quase conseguiu fazer o gol em um lance estranho. Aos 15 minutos, em um cruzamento torto que saiu com o meia Zé Rafael no lugar do volante Gil. Logo em seu primeiro lance, o atleta deixou a defesa para trás e, de frente para o gol, tocou para com perigo pela primeira vez. Da entrada da área, o atacante limpou o lance e disparou uma bomba. A bola, porém, parou com violência em que havia substituído Dagoberto, senti um problema na coxa direita no lance do gol e deu lugar ao atacante Luan. O Cruzeiro voltou a levar perigo primeiros minutos. A Ponte Preta até chegou perto de marcar, mas em lance de bola parada. O zagueiro Ferron cabeceou firme bola alçada na da área, mas a bola carimbou o travessão do goleiro Gleidson. No último lance, aos 48 minutos, Gil aproveitou cruzamento e cabeceou uma bola na	1,3; 9! 1,1! 1,5! 4,9; 6,0! 2,0!
Aos 7min, Elias e Léo Moura fizeram linda tabela. Na sequência do lance, Hernane chutou em cima de Gideão. Mais Fla. Aos 13min, Luis maior posse de bola, embora não conseguisse ser envolvente. O primeiro lance mais agudo aconteceu aos 24min, em jogada individual de Vitinho. O tem de manter uma sequência boa", disse. Cássio também explicou o lance que resultou o gol da vitória cruzeirense, o pênalti cobrado por acelerando a partida e pressionando a Portuguesa. Logo no primeiro lance do jogo, Emerson arriscou chute da entrada da área, mas a bola bateu com perigo e Howard espalmou. O volante Clark, que errou no lance do gol e já tinha um amarelo, foi substituído por Maurice Edu aos de falta. Livre, só serviu Sneijder para empurrar para o gol. No último lance da partida, a Holanda vacilou e permitiu que a Eslováquia marcasse e perdeu muito, era o gol de empate. Alef recebe o vermelho pela falta. Lance era chance clara de gol. tipo Alef segura Douglas Silva na área e chances, mas o atacante Marcão fez o quarto dos atleticanos no último lance e definiu a vitória. 4 a 3. O Flamengo segue sem vencer no foi até Campinas e derrotou a Ponte Preta por 4 a 3, com gol no último lance da partida, na noite desta quarta- feira, pela 4ª rodada. Paulo Baier 8, comentou que havia percebido o goleiro adversário adiantado em outro lance do primeiro tempo e que lembrou disso para anotar o golaço. "Eu bateu e garantiu os três pontos no primeiro jogo da Arena pelo nacional. O lance ocorreu aos 24 do segundo tempo. O Grêmio pressionava e Elano e não deu certo. Mas, com certeza, se eu tivesse que escolher entre um lance e outro, escolheria o gol contra o São Paulo. Foi um jogo da bola entre o seu toque e as redes, Renato Augusto lembrou de um lance parecido, protagonizado por ele mesmo, na partida contra o Santos, primeira chance clara de gol veio com o peruano Guerrero aos 17min, num lance em que ganhou na força, cortou Titi e bateu para defesa de Marcelo fase do Campeonato Paulista de 2013. "Veio à minha cabeça aquele lance, eu me lembro de que tentei e não deu certo. Mas, com certeza, se mas errou a finalização. O auxiliar marcou impedimento da Ponte no lance. Por muito pouco o time visitante não tirou o zero do placar. Edenílson pela direita. No entanto, nenhum dos cruzamentos resultou em lance perigoso. O mais perto que o Corinthians chegou de marcar foi em	2,9; 6,1; 1,1; 1,7; 5; 1,7; 1,7; 9,4; 9,1; 2,5! 8,6; 2,6; 2,6; 3,4! 2,6! 14,4! 14,2!

At the bottom of the window, there is a toolbar with icons for 'concordance', 'collocates', 'plot', 'patterns', 'clusters', 'timeline', 'filenames', 'source text', and 'notes'. The status bar at the bottom left shows '932 entries' and 'Row 1'.

**Figura 30** - Resultado da busca por *lance* no Concord do WordSmith Tools

Assim sendo, a análise do *corpus* levou a possíveis respostas sobre os questionamentos feitos anteriormente e agora pode-se dizer que as *ULs lance* e *jogada* são mais neutras do que a *UL tiro* e, portanto, podem não ser os equivalentes mais adequados e a tradução da *UL tiro* como tradução de *shot* no domínio do futebol é possível.

Em casos como este, é provável que existam outros equivalentes, pois a *UL* em inglês tem uma frequência muito alta. Um exercício que pode ser feito para localizá-los é consultar o *corpus* por outras palavras que façam parte do *frame*. Neste caso, como o *shot* costuma estar relacionado a uma tentativa de gol, buscar-se-ia no *corpus* por “*gol*”, o que levaria a outras unidades lexicais que também poderiam ser testadas.



*Fase 4: Reflexão sobre o exercício de apresentação das possíveis equivalências em um recurso lexical baseado em frames*

A forma como a equivalência se estabelece neste exemplo mostra que, mesmo sem ser direta, os traços em comum entre as ULs podem ser suficientes para indicar possíveis equivalentes. Todavia, a busca por mais opções de equivalentes possíveis deve ser encorajada para que o consulente do dicionário tenha a possibilidade de optar pelo equivalente mais adequado em cada situação. Vamos representar nas entradas lexicais a seguir o modelo de entrada com as informações sugeridas para este equivalente, mas isso não quer dizer que ele seja o único que deva ser disponibilizado.

<b>Lexic Unit:</b> shot	<b>Frame:</b> Shot ( <a href="#">hyperlink to the <i>frame</i> details</a> )
<p><b>Definition*:</b> kicking the ball towards goal trying to score a goal.</p> <p><b>Frame to Frame Relations*:</b></p> <p>Is used by: <a href="#">Offensive_Actions</a> (<a href="#">hyperlink to the <i>frame</i> details</a>)</p> <p>Precedes: <a href="#">Score_Goal</a> (<a href="#">hyperlink to the <i>frame</i> details</a>)</p> <p>Uses: <a href="#">Players</a> , <a href="#">Ball</a> (<a href="#">hyperlink to the <i>frame</i> details</a>)</p> <p><b>Possible Translation Equivalent:</b> <a href="#">tiro</a> (<a href="#">hyperlink to the <i>equivalent</i> entries</a>)</p> <p><b>Examples*:</b></p> <ol style="list-style-type: none"> <li>1) (...) after Sommer failed to hold Lampard's <b>shot</b> from the edge of the area.</li> <li>2) Fellow replacement Jamie Mackie saw a late <b>shot</b> saved by Mignolet as the R's ended on top.</li> <li>3) Mata came close to a goal on his 25th birthday but his <b>shot</b> across goal went narrowly wide and was just out of the reach of Ba's outstretched leg.</li> </ol> <p style="text-align: right;"><small>*Informações ainda não oficiais. Este trabalho faz parte do projeto lexicográfico em andamento</small></p>	

**Entrada Lexicográfica 7** - shot

<b>Unidade Lexical:</b> tiro	<b>Frame:</b> Tiro ( <a href="#">hyperlink para os dados do frame</a> )
<p><b>Definição*:</b> chutar a bola em direção à meta tentando marcar um gol.</p> <p><b>Relações Frame a Frame*:</b>  É usado por: <a href="#">Ações_Ofensivas</a> (<a href="#">hyperlink to the frame details</a>)  Precede: <a href="#">Marcar_Gol</a> (<a href="#">hyperlink to the frame details</a>)  Usa: <a href="#">Jogadores</a> , <a href="#">Bola</a> (<a href="#">hyperlink to the frame details</a>)</p> <p><b>Possível Equivalente de Tradução:</b> <i>shot</i> (<a href="#">hyperlink para a entrada do equivalente</a>)</p> <p><b>Exemplos*:</b></p> <ol style="list-style-type: none"> <li>1) 16' Neymar domina pela meia e arrisca <b>tiro</b> para o gol, fácil para Muslera.</li> <li>2) (...) antes de a bola chegar ao gol de Wilson, tentativa de cruzamento de Hugo dois minutos depois, e <b>tiro</b> forte de Reinaldo, aos seis, pela linha de fundo.</li> <li>3) (...)o meio- campista obrigou Fernando Prass a praticar boa defesa em <b>tiro</b> forte.</li> </ol> <p style="text-align: right;"><small>*Informações ainda não oficiais. Este trabalho faz parte do projeto lexicográfico em andamento</small></p>	

### Entrada Lexicográfica 8 - tiro

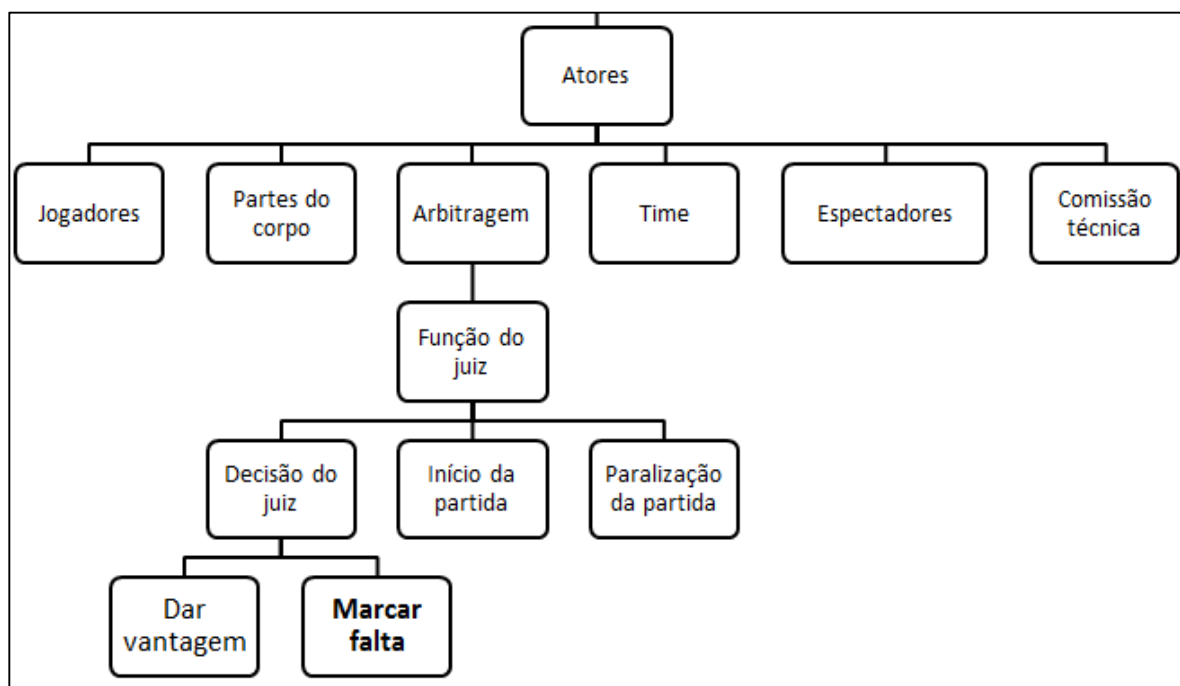
#### 4.2.2.2 Falta

##### Fase 1: Característica da UL *marcar falta*

A UL *falta* teve um alto número de ocorrências no *corpus* em português, mas a polissemia desta palavra foi o que a destacou para a realização desta análise. Neste sentido, escolheu-se uma das unidades complexas<sup>32</sup> que é formada por esta palavra, *marcar falta*, para verificar-se como a equivalência acontece.

A UL *marcar falta* faz parte do *frame* *Decisão\_do\_juiz*, que denota a atividade desempenhada pelo árbitro quando este assinala uma infração cometida por um infrator. Como se pode ver na figura abaixo (Figura 31), este *frame* herda traços de *Função\_do\_juiz*, conforme reestruturação proposta pelo grupo Kicktionary.

<sup>32</sup> Como unidade complexa consideramos o proposto por Biderman (2005, p. 747) ao destacar que a língua possui unidades muito heterogêneas e, dentre elas, algumas são monossilábicas, outras são simples e outras complexas, ou seja, “formadas de vários vocábulos e mesmo frases”. Estas, por serem utilizadas constantemente na língua, acabam se tornando fixas, algumas mais e outras menos, adquirindo significados únicos. Por isso, o termo *unidade complexa* abrange um grande número de combinações lexicais, que inclui fraseologias, colocações, expressões cristalizadas, provérbios, entre outros.



**Figure 31** - Posição da UL *marcar falta* na remodelagem de *frames* proposta pelo Kicktionary\_Br

Além da relação entre *frames*, os EFs previstos para o *frame* são: o JUIZ que marca a INFRAÇÃO de um JOGADOR\_INFRATOR contra um TIME\_QUE\_SOFRE\_A\_INFRAÇÃO, podendo, devido a isso, apresentar o CARTÃO a tal jogador.

#### *Fase 2: Consulta a dicionários*

No caso de uma unidade complexa como esta, a consulta no dicionário foi feita de três formas: 1) por *marcar*; 2) pela unidade complexa *marcar falta*; e 2) por *falta*, com o objetivo de verificar o que é disponibilizado em recursos como estes sobre tal UL.

Na tabela abaixo seguem resultados da busca. Mencionam-se aqui apenas os resultados relacionados a estas três formas de consulta, inseridos no domínio do futebol, como evocadoras do *frame* *Decisão\_do\_juiz*, ou seja, não serão apresentados todos os resultados da busca, mas apenas os que estão relacionados à UL *marcar falta*.

Dicionário:	Consulta por:		
	marcar	marcar falta	falta.n
Word Reference	mark	não disponível	não disponível
Michaelis	1 to mark, brand, seal, label. 2 to indicate, determine, designate, fix. 3 to book.	não disponível	não disponível
Babylon	mark	mark lack	n. omission, default, fail; mistake, fault; failing, forfeit; absence, need, paucity, deficiency; penury, privation

**Quadro 4** - Resultado da busca por *marcar*, *marcar falta* e *falta* em dicionários online

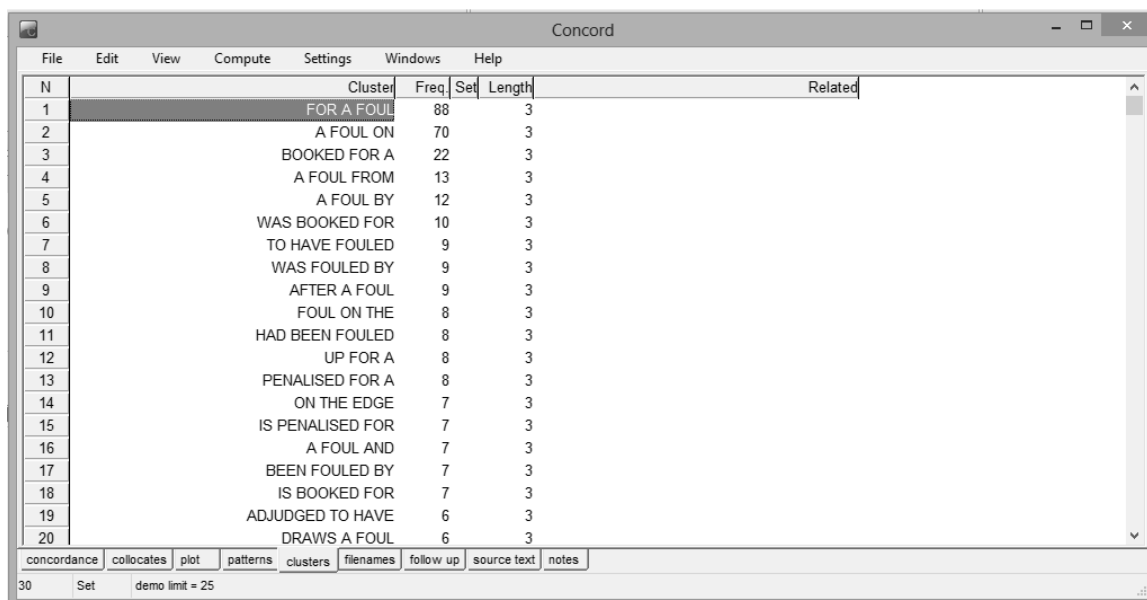
Neste primeiro momento já foi possível observar que não existem, atualmente, dicionários bilíngues (português/inglês) com um vocabulário amplo, especificamente do domínio do futebol, tanto em forma impressa quanto eletrônica. Portanto, para localizar equivalentes de tradução, um interessado precisa acessar dicionários não técnicos, nos quais alguns dos termos do futebol estão disponíveis, ou blogs e fóruns que trazem alguns termos traduzidos. Além disso, no caso do resultado remetido pelo Babylon (Quadro 1), nota-se que, ao buscar-se por *marcar falta*, a tradução aparece de forma literal, palavra por palavra, e que no caso da consulta por *falta* a equivalência disponibilizada não é apresentada diretamente, pois diversos são os equivalentes propostos para o termo, sendo necessário que o consulente descubra por conta qual seria o equivalente mais adequado no domínio de seu interesse.

O contexto, que não se encontra disponibilizado em dicionários como os trazidos acima, é extremamente importante quando se trata da consulta por equivalências de termos de um domínio específico, pois é ele que aponta o uso da unidade complexa em questão. Este é um dos aspectos que diferencia um projeto lexicográfico baseado em *frames*, proposta a ser detalhada na fase 4 desta análise.

### *Fase 3: Busca de equivalências utilizando-se da linguística de corpus*

Após o teste de busca pelo equivalente da UL *marcar falta* em dicionários eletrônicos, que não teve sucesso, parte-se agora para a busca de equivalências através da aplicação da linguística de corpus. Utilizando-se o WordSmith Tools, descrito no capítulo anterior, e o *corpus* do projeto, algumas consultas foram realizadas e serão apresentadas a seguir.

Inicialmente realizou-se uma pesquisa por *foul*, para que fosse possível observar, através da ferramenta Concord, os *clusters* mais recorrentes com esta palavra. O resultado obtido foi o seguinte:



The screenshot shows the Concord software interface with a table of clusters. The table has columns for N, Cluster, Freq, Set, Length, and Related. The data is as follows:

N	Cluster	Freq	Set	Length	Related
1	FOR A FOUL	88		3	
2	A FOUL ON	70		3	
3	BOOKED FOR A	22		3	
4	A FOUL FROM	13		3	
5	A FOUL BY	12		3	
6	WAS BOOKED FOR	10		3	
7	TO HAVE FOULED	9		3	
8	WAS FOULED BY	9		3	
9	AFTER A FOUL	9		3	
10	FOUL ON THE	8		3	
11	HAD BEEN FOULED	8		3	
12	UP FOR A	8		3	
13	PENALISED FOR A	8		3	
14	ON THE EDGE	7		3	
15	IS PENALISED FOR	7		3	
16	A FOUL AND	7		3	
17	BEEN FOULED BY	7		3	
18	IS BOOKED FOR	7		3	
19	ADJUDGED TO HAVE	6		3	
20	DRAWS A FOUL	6		3	

**Figura 32** - *Clusters* resultantes da consulta por *foul\** no Concord do WordSmith Tools

Estes resultados apontam para a alta frequência de *clusters* como: *for a foul*, *a foul on/by/from*, *booked for a (foul)*, *was booked for/by*, *to have fouled*, *after a foul*, *foul on the*, entre outros com menos recorrência. Através deste resultado, é possível analisar, como próximo passo, se algum destes *clusters* está relacionado ao *frame* *Decisão\_do\_juiz*, ou seja, se um deles, ou mais, podem ser possíveis equivalentes para a unidade complexa *marcar falta*.

Este teste é realizado consultando-se, então, pelos *clusters* que podem ser possíveis equivalentes. O primeiro teste é realizado por *book\* for a foul\**, e os resultados das 20 primeiras sentenças são os seguintes:

N	Concordance
1	occasions he didn't have answer to how to beat Pyatov. David Luiz was booked for a foul on Willian, Olexandr Kucher having earlier been
2	over wastefully. He would more than make up for that miss. Gallas was booked for a foul on Torres inside the Chelsea half (Ivanovic and
3	and Smalling caused a nervous moment with a back-pass before being booked for a foul on Kone. To be fair, the defensive substitute was
4	off for young central midfielder Barkley. Emanuelson then went into the book for a foul on Coleman, with Brede Hangeland getting in the way of
5	Nicky Shorey's curling 54th-minute free-kick after Danny Rose had been booked for a foul on McAnuff, and Mariappa headed wastefully wide after
6	made amends for a horrendous clearance before Torres earned his first booking for a foul on Tom Cleverley. But Chelsea needed only seven
7	goal at his mercy. Morgan Schneiderlin provided the cross and was then booked for a foul on Cazorla, before Puncheon fired into the side-netting
8	again came close, before another substitute, Steve de Ridder, was booked for a foul on Sagna. Szczesny parried a late Puncheon strike
9	launched himself well to tip his header over the bar. James Perch was booked for a foul on McAnuff but then provided vital intervention as
10	culminated in them taking the lead after 42 minutes. Glen Johnson was booked for a foul on Peter Odemwingie and, although Reid's free-kick
11	they had not learned their lesson and left him all alone. Ramis was then booked for a foul on Wigan old boy Mohamed Diame 25 yards out but
12	Fulham still looked the more likely team to score. Karim El Ahmadi was booked for a foul on Berbatov that checked a Fulham counter-attack.
13	the resulting corner but City were soon back on the defence. Toure was booked for a foul on substitute Adam Lallana soon after, before Fox
14	deflection. Corner kick. * 70' Yellow card Van Bronckhorst goes into the book for a foul on Aboubakar. * 69' He goes for goal but it's a really poor
15	Sub Van Persie goes off to be replaced by Huntelaar. * 58' Dirk Kuyt is booked for a foul on Gaetan Bong. * 57' ESPN Stats & Info: Should
16	at the expense of a corner. * 85' Mexico's Castro is frustrated and gets booked for a foul on the excellent Perez. * 84' Uruguay goalscorer
17	plenty of the ball but not getting too far with it. * 70' Masilela gets a booking for a foul and Mexico have a free-kick in a decent position, but
18	should have punched clear under very little pressure. * 42' Lochhead is booked for a foul on Jendrisek. But Jendrisek did go to ground rather
19	, but he's offside as he heads across the area. Meanwhile, Kuyt gets a booking for a foul on Bong. * 17' ESPN Stats & Info: The 'Indomitable
20	. Did Villar get a touch? I think so but he got away with it. * 70' Morel is booked for a foul on Cesc. * 68' Spain are regaining their composure and

**Figura 33** - Resultado da consulta por *book\* for a foul* no Concord do WordSmith Tools

No domínio do futebol, o termo *book* é compreendido como “o caderno no qual o juiz anota o nome dos jogadores que foram advertidos por falta”<sup>33</sup>. Tendo-se isso em vista, aliado à alta representatividade do *cluster* no *frame* em questão, sugere-se, nesta pesquisa, que as ULs “marcar falta” e “*be booked for a foul*” sejam possíveis equivalentes.

*Fase 4: Reflexão sobre o exercício de apresentação das possíveis equivalências em um recurso lexical baseado em frames.*

Tendo-se em vista a sugestão de equivalência apresentada acima, parte-se agora para a análise sobre a apresentação desta equivalência no Kicktionary\_Br Copa 2014. Como no Kicktionary\_Br Copa 2014 as buscas poderão ser feitas por *frames* ou unidades lexicais, no caso da UL *marcar falta* e seu equivalente *be booked for a foul*, ambos estarão acessíveis no *frame* *Decisão\_do\_juiz* e também poderão ser consultados bastando a digitação destas ULs no campo de busca correspondente. Cada um vai estar em uma base, porém as duas ULs devem aparecer como equivalentes.

Segue modelo de entrada lexical para cada uma das ULs analisadas:

<sup>33</sup> **Dicionário Oxford**. Disponível em: <<http://oxforddictionaries.com/definition/english/book>> Acesso em 27 de julho de 2013.

<b>Unidade Lexical:</b> marcar falta	<b>Frame:</b> Decisão_do_Juiz ( <a href="#">hyperlink para os dados do <i>frame</i></a> )
<p><b>Definição*:</b> frame que descreve as intervenções do juiz quando este observa uma infração cometida por um jogador infrator.</p> <p><b>Relações Frame a Frame*:</b></p> <p>Herda de: Função_do_Juiz (<a href="#">hyperlink para os dados do <i>frame</i></a>)</p> <p>Usa: Infração (<a href="#">hyperlink para os dados do <i>frame</i></a>)</p> <p><b>Possível Equivalente de Tradução:</b> be booked for a foul (<a href="#">hyperlink para a entrada do equivalente</a>)</p> <p><b>Exemplos*:</b></p> <ol style="list-style-type: none"> <li>1) Porém, o árbitro Fabrício Neves Corrêa anulou o gol e <b>marcou falta</b> do atacante colorado.</li> <li>2) O juiz anulou o lance, <b>marcando falta</b> de Neymar no zagueiro.</li> <li>3) O árbitro Wagner Tardelli <b>marcou falta</b> e expulsou o zagueiro.</li> </ol> <p style="text-align: right;"><small>*Informações ainda não oficiais. Este trabalho faz parte do projeto lexicográfico em andamento</small></p>	

#### Entrada Lexicográfica 9 - Marcar falta

<b>Lexic Unit:</b> be booked for a foul	<b>Frame:</b> Referee_Decision ( <a href="#">hyperlink to the <i>frame</i> details</a> )
<p><b>Definition*:</b> referee's intervention when he spots an offense, sanctioning an offender.</p> <p><b>Frame to Frame Relations*:</b></p> <p>Inherits from: Referee_Function (<a href="#">hyperlink to the <i>frame</i> details</a>)</p> <p>Uses: Offense (<a href="#">hyperlink to the <i>frame</i> details</a>)</p> <p><b>Possible Translation Equivalent:</b> marcar falta (<a href="#">hyperlink to the <i>equivalent</i> entry</a>)</p> <p><b>Examples*:</b></p> <ol style="list-style-type: none"> <li>1) David Luiz was <b>booked for a foul</b> on Willian, Olexandr Kucher having earlier been cautioned for the home side.</li> <li>2) Danny Rose had been <b>booked for a foul</b> on McAnuff, and Mariappa headed wastefully wide after meeting the resulting corner at the near post.</li> <li>3) Morgan Schneiderlin provided the cross and was then <b>booked for a foul</b> on Cazorla, before Puncheon fired into the side-netting from an acute angle and Ramirez struck over.</li> </ol> <p style="text-align: right;"><small>*Informações ainda não oficiais. Este trabalho faz parte do projeto lexicográfico em andamento</small></p>	

#### Entrada Lexicográfica 10 – be booked for a foul

### 4.2.3 Superioridade

Casos de superioridade, embora não sejam tão frequentes, acontecem quando uma palavra tem uma extensão no significado que difere nas línguas fonte e alvo. Isso faz com que as equivalências não sejam diretas, pois a palavra em um dos idiomas terá um alcance maior do que uma equivalência no outro idioma.

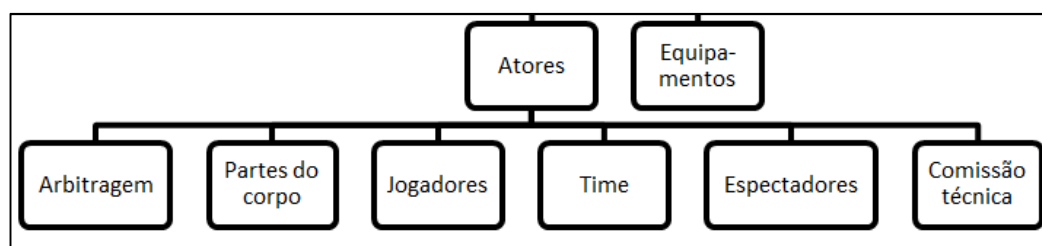
Um caso disso é o da UL *team* em inglês, que é analisada a seguir.

#### 4.2.3.1 Team

##### *Fase 1: Característica da UL team*

A palavra *team* não apareceu como uma das ULs mais frequentes no *corpus* em inglês, mas algo interessante a destacou para que fosse selecionada para esta análise: duas palavras da lista de 50 ULs da análise inicial do *corpus* em português a indicaram como possível equivalente. São elas: *time* e *equipe*.

Na estruturação de *frames* proposta pelo Kicktionary\_Br a UL *team* aparece como evocadora do *frame* Team (ou *Time*, conforme versão em português em desenvolvimento), no seguinte formato (Figura 34):



**Figura 34** - Posição do frame Time na remodelagem de frames proposta pelo Kicktionary\_Br

Este frame representa cada uma das equipes em campo, que são comumente denominadas *times*, *equipes* ou, no caso da Copa do Mundo, *seleções*. Como nosso *corpus* não traz apenas textos sobre este campeonato mundial, acredita-se que este seja o motivo da alta frequência da UL *time*.

Para compreender as especificidades sobre a UL *team*, parte-se agora para uma análise mais aprofundada desta, para posterior discussão sobre equivalências.



*Fase 2: Consulta a dicionários*

Ao consultar a UL *team* obtivemos os seguintes resultados:

	Consulta por:
Dicionário:	<i>team</i>
Word Reference	(sport) time, equipe
Michaelis	<b>n</b> 1 time, equipe, conjunto, quadro. <b>2</b> junta, parilha
Babylon	<b>s.</b> grupo, equipe; time; animal atrelado a um veículo

**Quadro 5** - Resultado da busca por *team* em dicionários online

Observa-se aqui a recorrência das palavras *time* e *equipe* como possíveis equivalentes de tradução, em todos os dicionários consultados. Isso será levado em consideração, mas análise de *corpus* nos ajudará a compreender melhor o uso de cada um destes termos. Comenta-se, porém, que a UL *seleção* não aparece nestas consultas, embora a pesquisa inversa, de português para inglês, por esta palavra remeta, nos dicionários Michaelis e WordReference, a tradução de *team* em inglês.

Assim sendo, serão testadas as hipóteses de *time*, *equipe* e *seleção* como possíveis equivalentes de *team*.

*Fase 3: Busca de equivalências utilizando-se da linguística de corpus*

A primeira consulta realizada foi pela UL *team*. Esta UL retornou 805 entradas, porém, destas, 122 eram parte da unidade composta *team-mate*. Portanto, desconsiderando-se estas unidades, tem-se que apenas 683 entradas em inglês mencionam a UL *team*.

Destes 683 casos, observaram-se os primeiros 25 mais a fundo para verificar que tipo de *team* era retratado (Figura 35). Em alguns casos, não foi possível identificar se tratava-se de clubes ou de seleções, como em *This team like each other* (exemplo 9). Em outros, notou-se que são seleções, como em *...the only African team left in the competition* (exemplo 7). E, em outros, tratava-se de clubes, como em *...a poor Spurs team in the first half* (exemplo 17).

N	Concordance	Set	W
1	14-minute period which exposed all the flaws characteristic of a newly-promoted team. Kevin Mirallas swung over a right-wing cross and although Jelavic threw		21
2	a favour against Hull after doing their own job at Leicester. Nigel Pearson's team knew nothing other than a win at a noisy King Power Stadium would		25
3	a heartening start to the festive period. Wenger made three changes from the team knocked out of the Capital One Cup at Bradford last week. Mikel Arteta		20
4	just eludes Tagoe at the back post. * 43'Prince Dep: "Serbia have a great team, Ivanovic and Vidic at the back are world class, and they have talent all		19
5	from the start by the vocal visiting support, having lost his place in the QPR team just four games after his summer move from Upton Park. The England		20
6	that had looked like a goal all the way. Harry Redknapp knew the longer his team kept City at bay, the greater his team's belief would become – and		20
7	Ghana carry the hopes of a continent tonight in Rustenburg, as the only African team left in the competition. For USA, they will feel they now stand an		13
8	go through and initially, you felt the home fans thought that was that. But their team lifted themselves again. Palacio forced a sharp save from Friedel at the		27
9	matched in spectacle by a bout of rampant man love on the sidelines. This team like each other, unlike some I could mention... * 53' Tevez's second of		13
10	out Bale and then Defoe in injury time. Match Report - Spurs 0 - 0 Lazio The team knocked and knocked but couldn't find the breakthrough against SS Lazio		29
11	part in the excitement and on Saturday afternoon Andre Villas Boas' side, a team laden with pace and purpose, were rewarded with a victory that finally		19
12	Mandzukic. The Emirates crowd chanted "come on Arsenal" with gusto as their team left the field at the break. The tie was only a quarter gone and already		8
13	day in South Africa, sunny and warm. * Honduras coach Reinaldo Rueda: "The team is not at its best right now but we will kick off he tournament with		39
14	all had shots, West Brom came close to punishing more lax defending. Fine team interplay saw the ball swept out wide to Lukaku, only for the Chelsea		10
15	opening exchanges, with Lallana seeing a shot blocked in the box after neat team interplay. The Saints captain was getting a lot of space in midfield and		24
16	Was sitting next to a bunch of Uruguay jounos today and they feel their team is up for it - France are there for the taking, if they can get one in early.		5
17	away win looked on the cards as the Blues dominated a poor Spurs team in the first half, taking the lead through a fine Gary Cahill volley.		19
18	Stats & Info: Against Chile, Switzerland had the worst passing accuracy of any team in a World Cup game since 1966 (56.7%). After the first two rounds of		12
19	were beaten at the death. Their total of 54 points is the highest by a relegated team in Championship history. Ultimately, results elsewhere meant Palace		25
20	clear. Brazil are on the ropes. * 74' What must Dunga be thinking? His team is coming apart at the seams. Robben brought down by Alves and it's a		15
21	World Cup Final tournaments for South Korea. * 63'Kofi: "The Greek team is imbalanced, they are playing strikers as wingers." * 62'Chung-Yong		18
22	that Tottenham surrendered a 10-point lead to Arsenal last year, but this Spurs team is made of stronger stuff and Wenger's side appear to lack the same		54
23	night, Paul Marshall believes France's spin doctors need to accept the fact the team is struggling. * 45' Time to point you in the direction of some of our		9
24	Brazil beat Chile 4-1. * 3' Brazil and Chile fans may be all smiles, but one team is going home tonight. * 3' Chile, as ever, are showing no fear of		14

**Figura 35-** Resultado da busca por *team* no Concord do WordSmith Tools

Sabendo-se desta característica da UL *team*, partiu-se então para a observação das possíveis equivalências que apareceram na lista de palavras com alta frequência no corpus em português, que são *time* e *equipe*.

Iniciando pela UL *equipe*, verificou-se através da análise mais aprofundada das 25 primeiras sentenças em que esta UL ocorre para se referir a clubes e também seleções, porém a quantidade de menções a clubes é muito maior. Isso pode ser visto nos casos como *equipe paulista*, *equipe mineira*, *equipe catarinense*, *equipe pernambucana*, *equipe carioca*, etc.

N	Concordance	Word #
1	para o Atlético-GO, mas também só garantiu o terceiro empate seguido da equipe paulista, desta vez por 1 a 1, no retorno ao Pacaembu após dois	199,235
2	Daniilinho. Com isso, o Botafogo procurou explorar os erros da equipe mineira e acabou premiado com um gol. Após uma falha de	220,884
3	Sua comemoração mostrou bem o sentimento do jogador e do restante da equipe: ele "lavou a alma" em uma cena de desabafo. Na frente do placar,	42,791
4	para ampliar sua vantagem. Aos 35 min, Sandro fez o terceiro da equipe catarinense, lhe dando maior tranquilidade. "Sabemos que é	31,055
5	partiu de forma desesperada para tentar descontar a vantagem da equipe pernambucana. E quase conseguiu anotar o primeiro gol aos 40	129,411
6	nesta sexta-feira, em Durban, porém, mostrou o velho problema da equipe de Dunga contra retrancas. Mesmo assim, foi suficiente para	41,371
7	do Sul de forma menos vergonhosa. Coube a Eto'o, melhor jogador da equipe, o papel de cobrar o pênalti e marcar o segundo gol dele e da sua	43,273
8	E a insistência se mostrou a novamente a melhor arma da equipe dos Estados Unidos, que já havia corrido atrás do placar nos	95,949
9	no Campeonato Brasileiro. Neste domingo, o Tricolor não teve o brilho da equipe campeã continental de 20 anos atrás, mas foi mais objetivo para	131,461
10	desde que assumiu o São Paulo foram decisivos para a reabilitação da equipe no Campeonato Brasileiro. Neste sábado, o Tricolor contou com	218,672
11	sobre um dos rivais na briga pelo acesso. A vitória tira a pressão da equipe de Kleina e a recoloca no grupo dos quatro melhores time da Série	30,485
12	Van Marwijk começou a poupar jogadores, já pensando no próximo jogo da equipe. Dessa forma, Van Persie e Kuyt foram substituídos. Se a Holanda	43,238
13	vencemos na individualidade", afirmou Márcio Rodrigues, técnico da equipe sub-17 do Palmeiras, sobre o jogo-treino desta terça-feira. * * * As	101,041
14	japoneses. Com o revés, os espanhóis terão de se inspirar na trajetória da equipe principal na Copa do Mundo de 2010. Apesar da derrota na estreia	45,141
15	André e Éder Luis provocou muita reclamação por parte dos jogadores da equipe carioca, que acusaram o zagueiro corintiano de dar uma cotovelada	193,057
16	jogo. Pablo Escobar mostrou ser um dos principais jogadores ofensivos da equipe e definiu o marcador com um gol aos 25 e outro aos 35 minutos.	401,762
17	a sobra para descontar. Mas já era tarde, e o gol não impediu a queda da equipe na Copa do Mundo. Aos 43min, Okazaki ainda fez o terceiro em	68,015
18	Alemanha, no entanto, deve dar lugar à apreensão, já que o adversário da equipe nas oitavas de final será a Inglaterra. A partida acontece neste	37,264
19	preocupar com os portugueses. Principalmente porque o principal nome da equipe resolveu desencantar. Cristiano Ronaldo participou bastante do	41,108
20	nas vagas de Gil e Kleberston, no Bahia. Aos sete minutos, o treinador da equipe mandante promoveu a volta de Hernán Barcos aos gramados,	182,851
21	Denilson, que sentiu um incômodo no joelho direito no último treino da equipe, o técnico Ney Franco escalou Wellington para compor o setor de	132,434
22	agressiva, tocando a bola com rapidez e buscando espaços na defesa da equipe carioca. Aos dois minutos, Cicinho arrancou pela direita e cruzou	144,444
23	e segue sem poder de reação. Esta foi a terceira derrota seguida da equipe no Brasileiro e acabou beneficiando o Bahia, primeiro clube fora da	234,404
24	foi derrubado ao lado da área. Na cobrança, foi bloqueado pela zaga da equipe paulista. Aos 34 minutos, o Vasco chegou a colocar a bola nas	144,664
25	boas chances com Vinicius, Leandro e Tiago Real, os mais perigosos da equipe do técnico Gilson Kleina. O América-MG, por sua vez, procurava	34,562

Figura 36 - Resultado da busca por *equipe* no Concord do WordSmith Tools

Esta observação é confirmada ao olhar-se os *clusters* disponibilizados pelo programa, que apontam a alta recorrência de estruturas como as mencionadas acima. Na figura abaixo (Figura 37) segue listagem de *clusters* com destaque aos agrupamentos que mencionam clubes nacionais.

N	Cluster	Freq	Set	Length	Related
4	A-EQUIPE-DE	105	3		
2	EQUIPE-DA-CASA	75	3		
3	A-EQUIPE-DA	68	3		
4	A-EQUIPE-DO	66	3		
5	A-EQUIPE-CARIOCA	44	3		
6	A-EQUIPE-PAULISTA	42	3		
7	QUE-A-EQUIPE	39	3		
8	A-EQUIPE-SUB	35	3		
9	DA-EQUIPE-CARIOCA	35	3		
10	COM-O-RESULTADO	31	3		
11	A-EQUIPE-GAÚCHA	30	3		
12	EQUIPE-SÃO-PAULINA	29	3		
13	RESULTADO-A-EQUIPE	28	3		
14	O-RESULTADO-A	26	3		
16	PARA-A-EQUIPE	26	3		
16	MAS-A-EQUIPE	26	3		
17	A-EQUIPE-SÃO	25	3		
18	E-A-EQUIPE	24	3		
19	DA-EQUIPE-DA	24	3		
20	A-SUA-EQUIPE	24	3		
21	A-EQUIPE-PERNAMBUCANA	24	3		
22	DA-EQUIPE-PAULISTA	22	3		
23	A-EQUIPE-FOI	21	3		
24	DA-EQUIPE-DE	21	3		
25	A-EQUIPE-VISITANTE	20	3		

Figura 37 - Resultado dos *clusters* para *equipe* no Concord do WordSmith Tools

Assim como a *UL equipe*, a *UL time* também foi analisada e apontou para um uso que compreende a menção a clubes e a seleções (Figura 38). Porém, neste caso, a frequência da palavra *time* (2.718) é quase 50% maior do que a de *equipe* (1.870).

The screenshot shows the Concord window with the following data:

N	Concordance	Set	Word#
1	zagueiro, ampliou e acendeu as esperanças sul-africanas, afinal, o time precisaria de apenas mais dois gols, caso o placar em	Set	54,37
2	Quando a Inglaterra resolveu variar a tática de "bola para Rooney", o time conseguiu marcar. Aos 23, Milner recebeu pela direita e	Set	55,67
3	do Cabo. O impressionante feito veio com tempero especial, afinal o time bateu a favorita Argentina, com gols de Muller (um dos artilheiros		49,99
4	uma grande ameaça ao gol de Joseph no primeiro tempo. Aláís, o time sequer conseguiu se acertar na marcação. Mais organizada em		54,24
5	e Diego Forlan perderam pelo menos duas chances cada um. O time ainda perdeu Lugano, machucado, e se assustou com Fucile,		48,94
6	do Norte, no Ellis Park, em Johannesburg. O resultado deixa o time verde-amarelo na liderança do Grupo G (que tem ainda Portugal		57,11
7	abalou o Corinthians. Embora ainda presente no campo de ataque, o time da casa começou a pecar ainda mais por falta de atenção.		63,52
8	o técnico Caio Júnior tirou Dinei para colocar Giancarlo. No entanto, o time gaúcho voltou melhor para o segundo tempo, e D'Alessandro		62,92
9	ao Botafogo. Apesar de ter adotado uma postura cautelosa, o time carioca não chegou a abdicar do ataque. O Botafogo estava bem		63,39
10	derrota da Grécia (que enfrentava a Argentina na mesma hora). E o time africano já foi logo fazendo a sua parte. Aos 12, Odiah ganhou a		54,74
11	no placar no jogo da Coreia já pressionava os gregos, imagine se o time oriental virasse? E foi o que aconteceu, logo no começo da		54,92
12	Elano ajudou, como no gol de Maicon. Mas foi muito pouco para o time pentacampeão. Muito pouco para quem busca o hexa. Muitos		57,33
13	entre as pernas de um coreano. Arriscou de fora da área. Viu que o time estava com dificuldade na armação e recuou ao meio-campo.		57,40
14	em uma possível partida de quartas de final. Na próxima fase, o time de Dunga pega o Chile, que perdeu a liderança da chave para a		58,18
15	das oitavas de final da Copa, nesta terça-feira, 29, em Pretória. O time se tornou a quarta seleção sul-americana a ficar entre as oito		47,60
16	Talvez até a Coreia do Norte tenha saído de campo satisfeita. Mas o time de Dunga não pode se dar esse luxo. Brasil supera marfinenses		57,49
17	ofensivo e os meias não fizeram sua função básica: armar o jogo. O time de Dunga teve duas chances ainda no primeiro tempo. Uma no		58,29
18	Depois do intervalo, o jogo ficou mais aberto. Portugal lembrou o time que goleou a Coreia do Norte por 7 a 0. E o Brasil começou a		58,57
19	também só conseguiu ir à Copa por ser anfitriã. Prova disto foi que o time não conseguiu classificação para a Copa Africana das Nações.		54,21
20	lusitanos. Com 63% de posse de bola e o dobro de finalizações, o time de Dunga dominou as ações. Nilmar mostrou gás. Mandou uma		58,49
21	logo, e bem. Para Portugal, o jogo foi bom. A retranca funcionou e o time foi perigoso, principalmente, na etapa final. Como era esperado,		58,63
22	e companhia por 3 a 0 em Johannesburg, em resultado que conduz o time nacional adiante na Copa e manda os rivais de volta a Santiago.		58,87
23	tempo, aproveitando-se de uma estratégia voltada ao contra-ataque, o time visitante contou com assistência do astro Seedorf para Rafael		63,08
24	da partida, criando as principais jogadas ofensivas. Aos poucos, o time goiano começou a adiantar as linhas de marcação e ameaçou o		66,87
25	encontrar espaços na fechada defesa visitante. Sem atacar muito, o time de Luxemburgo atuou com tranquilidade, não rifando a bola a		73,02

Figura 38 - Resultado da busca por *time* no Concord do WordSmith Tools

Por último, observou-se o comportamento da *UL seleção*, que teve 620 ocorrências (Figura 39). Neste caso todos os casos mencionados são de times que representam seus países em jogos internacionais.

The screenshot shows the Concord window with the following data:

N	Concordance	Set	Word#	Sem	Seni	Parz	Parz
1	fazer um estrago Felipe Truda: 28min - Mapemba tenta cruzamento na área da Seleção, mas sai torto para fora Felipe Truda: 3 a 0 para o Brasil. Tem um		23,606	1,6	61%	2	64%
2	rúgbi de 1995, sediado no País e que terminou com a hollywoodiana vitória da seleção local. Os torcedores deram os pésames em forma de vuvuzela. Foi um		26,460	1,3	97%	0	27%
3	no Estádio Nacional de Harare! É um caldeirão! Bom pra testar os nervos da Seleção Felipe Truda: Agora sai Luis Fabiano para a entrada de Grafite Felipe		23,982	1,6	22%	2	66%
4	sai e entra Pedro Henrique Felipe T. te liga e liga-te! Timinho ruim esse da seleção do dunga FELIPE TRUDA - favor informar o time que esta em campo.		23,518	1,6	32%	2	64%
5	na zaga e todo o restante do time pressionando. Apesar do ímpeto ofensivo da seleção europeia, poucas chances claras de gol foram criadas. Até que aos		25,490	1,3	47%	0	26%
6	de resenas e não teve a oportunidade de mostrar o seu valor com a camisa da seleção principal. As próximas partidas das duas seleções serão realizadas		404,998	2,8	97%	0	89%
7	ao vencer a embalada Colômbia por 1 a 0, nesta terça-feira, em casa. O gol da seleção vinotinto foi marcado pelo atacante Rondón, aos 13 minutos do primeiro		409,356	3,1	29%	0	96%
8	seguida, Carrilho buscou a finalização colocada e acertou a trave do camisa 1 da seleção argentina. A movimentação esboçada pelos clubes nesta etapa do		405,322	2,9	94%	0	90%
9	etapa, a pressão voltou a surtir efeito e novamente com Farfán. O camisa 10 da seleção peruana venceu o goleiro Renny Veja e sacramentou a vitória alvirubra		403,888	2,8	30%	0	87%
10	erra o alvo. Ele pode, fez um golaço... Felipe Truda: E agora, um ataque da Seleção e o juiz para o jogo alegando falta no goleiro. Já está difícil a gente		24,251	1,7	47%	2	67%
11	Henrique, Miralles, Paulo Henrique, Alison e Rafael Galhardo), dois a serviço da Seleção Brasileira (Neymar e Arouca), um por questões contratuais (Juan está		255,603	13,	52%	0	73%
12	Num momento raro de se ver, a torcida alemã gritou "olé" para os jogadores da seleção, que envolveram totalmente os argentinos a partir da metade do segundo		20,024	1,0	59%	0	21%
13	Wesley, Fernandinho e João Vitor, machucados, e Barcos, à disposição da seleção argentina, Gilson Kleina apostou em uma escalação ofensiva. E		305,398	16,	77%	0	87%
14	O são-paulino Gabriel Boschilia balançou as redes no primeiro tempo do jogo da Seleção Brasileira Sub-17 contra a Venezuela, pelo Sul-Americano da categoria,		141,168	8,2	42%	0	93%
15	Héber, que domina mal e permite a chegada de Pablo Felipe que acha da seleção???????? FELIPE PELO AMOR DE DEUS ESSA SELEÇÃO ESTA		23,491	1,6	100%	2	63%
16	ontem. Quer a real? Magão luta por vaga entre os 11. O Bolatti é titular e da Seleção. [via Twitter] blogdozini: @schmitt100 se der Inter nos pênaltis, decisão		4,145	310,	40%	0	38%
17	tempo, e o gol marcado aos 12min do segundo tempo premia a vontade da Seleção Felipe Truda: Daniel bate e o goleiro defende Felipe Truda: Depois do		24,215	1,7	49%	2	67%
18	b8PTgG [via Twitter] clicEsportesRS: #Copa2010 Robben é vetado da estreia da seleção holandesa http://bit.ly/kK7SD [via Twitter] clicEsportesRS: #Copa2010		7,817	562,	62%	0	71%
19	anulado. Foi a primeira vez desde 2007 que Messi iniciou um compromisso da seleção argentina no banco de reservas - a última aconteceu na vitória sobre o		409,862	3,1	45%	0	97%
20	3 x 0 Peñarol (Mundial - 1952) Maracanazo vingado Fluminense derruba base da Seleção do Uruguai no Maracanã Campeão do mundo em 1950 numa decisão		40,469	2,2	30%	0	27%
21	de Jorginho, teria levado o gol de empate. Aos 37, o zagueiro Wallace, titular da seleção brasileira sub-20, entrou no lugar do volante Tinga e fez a estreia dele		4,760	256	18%	0	3%
22	Cabo, e voltou a uma decisão de Copa do Mundo após 32 anos. "Carrascos" da Seleção Brasileira nas quartas, Sneijder (autor dos gols que eliminaram o Brasil)		10,063	510	13%	0	10%
23	garantir sua vaga. O clube, porém, não teria vida fácil. O Peñarol, base da Seleção do Uruguai campeão do mundo, seria um de seus adversários na primeira		40,604	2,2	32%	0	27%
24	modificado - apenas o atacante Luan não foi substituído -, o artilheiro da seleção brasileira sub-20, Vinícius Araújo, teve duas chances seguidas em uma		6,209	320	43%	0	4%
25	Maxi Velázquez é amareladol Maylon tipo me desculpe, mas tem que babar da seleção pra sair comentário aqui? Médico Paulo Rabaldo sobre Victor: "Não		50,117	3,5	73%	5	31%

Figura 39 - Resultado da consulta por *seleção* no Concord do WordSmith Tools

Tendo-se isso em vista pode-se dizer que o caso de superioridade é confirmado neste caso, sendo que a UL *team*, que pode ser considerada um possível equivalente de *time*, por seus traços semânticos similares, também assume uma posição de superioridade (no sentido de alcance de significados) quando relacionada à UL seleção. Ou seja, a UL *team* pode ter mais de um equivalente e isso acontece, porque o léxico do inglês utiliza esta palavra mais geral para falar sobre clubes e seleções. Por outro lado, o léxico do português especifica o termo *seleção* para times que representem países em competições internacionais.

Acredita-se que possivelmente a frequência do termo *seleção* esteja vinculada ao tipo de textos que compõem o *corpus*, pois os textos compilados em sites de times ou em notícias esportivas sobre competições nacionais não utilizariam o termo seleção. Apenas para confirmar esta hipótese e para compararmos a frequência dos usos dos termos *seleção* e *time* em um contexto exclusivamente de Copa do Mundo, foi realizada uma última consulta apenas aos textos que relatam as partidas da Copa de 2010 por ambos.

O resultado comprovou o que era imaginado. Neste *corpus* a frequência para cada uma das ULs foi: *Time* = 285 entradas; *Seleção* = 273 entradas. Assim sendo, pode-se dizer que ambas as ULs têm uma representatividade importante em textos sobre a Copa do Mundo.

#### *Fase 4: Reflexão sobre o exercício de apresentação das possíveis equivalências em um recurso lexical baseado em frames*

Graças ao exercício de análise delineado anteriormente, já se tem as informações necessárias para refletir sobre a apresentação das entradas lexicográficas da UL *team* e seus possíveis equivalentes. Assim como nos casos anteriores, a base desta entrada prevê a indicação do *frame* evocado pela UL, descrição, exemplificação, etc. Porém, neste caso, um diferencial será a distinção entre os possíveis equivalentes com a explicação sobre o caso de superioridade. Isso porque, por exemplo, ao traduzir uma sentença do inglês para o português que fala sobre o “*American team*”, é importante que o consulente compreenda que ele tem mais do que uma possibilidade de tradução, podendo se valer das ULs *time* e *equipe*, mas levando em conta também que, no português brasileiro, existe o termo *seleção*, que pode ser utilizado neste caso.

As ULs analisadas nesta subseção podem, portanto, ter suas entradas lexicográficas representadas com as seguintes informações em um recurso baseado em *frames*:

<b>Lexic Unit:</b> team	<b>Frame:</b> Team ( <a href="#">hyperlink to the frame details</a> )
<p><b>Definition*:</b> a group of players forming one of the sides in a soccer match.</p> <p><b>Frame to Frame Relations*:</b></p> <p>Has: <a href="#">Players</a> (<a href="#">hyperlink to the frame details</a>)</p> <p>Is used by: <a href="#">Mark</a> (<a href="#">hyperlink to the frame details</a>)</p> <p><b>Possible Translation Equivalent:</b> time, seleção, equipe (<a href="#">hyperlink to the equivalent entries</a>)</p> <p><b>Examples*:</b></p> <p>1) 64' Heinze is clearly auditioning for the Argentina high jump <b>team</b> for the next Olympics.</p> <p>2) Since 1986, no <b>team</b> has advanced to the semi-finals without scoring a goal in either of their knockout stage matches.</p> <p>3) If results hold, this would be the first World Cup under the current 32-team format in which no more than one European <b>team</b> advances from any group.</p> <p style="text-align: right;">*Informações ainda não oficiais. Este trabalho faz parte do projeto lexicográfico em andamento</p>	

### Entrada Lexicográfica 11- Team

<b>Unidade Lexical:</b> seleção	<b>Frame:</b> Time ( <a href="#">hyperlink para os dados do frame</a> )
<p><b>Definição*:</b> grupo de jogadores que formam uma das partes em um jogo de futebol.</p> <p>① <b>Comentário:</b> esta unidade lexical é utilizada exclusivamente para tratar sobre times que representam um país em campeonatos internacionais. Exemplo: Seleção <b>brasileira</b>, Seleção <b>americana</b>, etc.</p> <p>① <i>Comment: this lexic unit is exclusively used to talk about teams that represent a country in international matches. Example: "Seleção Brasileira", "Seleção Americana", etc.</i></p> <p><b>Relações Frame a Frame*:</b></p> <p>Tem: <a href="#">Jogadores</a> (<a href="#">hyperlink para os dados do frame</a>)</p> <p>É usado por: <a href="#">Marcação</a> (<a href="#">hyperlink para os dados do frame</a>)</p> <p><b>Possível Equivalente de Tradução:</b> <i>team</i> (<a href="#">hyperlink para a entrada do equivalente</a>)</p> <p><b>Exemplos*:</b></p> <p>4) Gekas, o mais perigoso da <b>seleção grega</b>, ainda quase marcou em lindo lance individual.</p> <p>5) Mas o esquema da <b>seleção portuguesa</b> não levava muito perigo, pois Raul Meirelles e Duda ficavam distantes de Cristiano Ronaldo.</p> <p>6) ...o grande pecado da <b>seleção espanhola</b>, querer caprichar demais em certas jogadas.</p> <p style="text-align: right;">*Informações ainda não oficiais. Este trabalho faz parte do projeto lexicográfico em andamento</p>	

### Entrada Lexicográfica 12 – Seleção

<b>Unidade Lexical:</b> time	<b>Frame:</b> Time ( <a href="#">hyperlink para os dados do frame</a> )
<p><b>Definição*:</b> grupo de jogadores que formam uma das partes em um jogo de futebol.</p> <p><b>Relações Frame a Frame*:</b></p> <p>Tem: <a href="#">Jogadores</a> (<a href="#">hyperlink para os dados do frame</a>)</p> <p>É usado por: <a href="#">Marcação</a> (<a href="#">hyperlink para os dados do frame</a>)</p> <p><b>Possível Equivalente de Tradução:</b> <a href="#">team</a> (<a href="#">hyperlink para a entrada do equivalente</a>)</p> <p><b>Exemplos*:</b></p> <ol style="list-style-type: none"> <li>1) O <b>time</b> tem posse de bola e, ao mesmo tempo, sabe ser vertical e objetivo.</li> <li>2) (...) a equipe europeia arrancou gritos de "olé" dos torcedores, despachou o <b>time</b> de Maradona e garantiu vaga nas (...)</li> <li>3) O jogo a esta altura acontecia debaixo de chuva forte e o <b>time</b> da casa se descontrolou.</li> </ol> <p style="text-align: right;"><small>*Informações ainda não oficiais. Este trabalho faz parte do projeto lexicográfico em andamento</small></p>	

### Entrada Lexicográfica 13 - Time

<b>Unidade Lexical:</b> equipe	<b>Frame:</b> Time ( <a href="#">hyperlink para os dados do frame</a> )
<p><b>Definição*:</b> grupo de jogadores que formam uma das partes em um jogo de futebol.</p> <p><b>Relações Frame a Frame*:</b></p> <p>Tem: <a href="#">Jogadores</a> (<a href="#">hyperlink para os dados do frame</a>)</p> <p>É usado por: <a href="#">Marcação</a> (<a href="#">hyperlink para os dados do frame</a>)</p> <p><b>Possível Equivalente de Tradução:</b> <a href="#">team</a> (<a href="#">hyperlink para a entrada do equivalente</a>)</p> <p><b>Exemplos*:</b></p> <ol style="list-style-type: none"> <li>1) O lema da torcida são-paulina para o jogo contra o Atlético-MG nesta quarta-feira, que decidiria o futuro da <b>equipe</b> na Copa Libertadores, foi "Eu vou com fé", que remete ao "Clube da Fé", título que acompanhou o São Paulo durante toda a sua história.</li> <li>2) Com o empate, o jogo ficou aberto, e cada <b>equipe</b> ainda teve uma chance clara de vencer.</li> <li>3) Os comandados de Vicente del Bosque têm agora pela frente a Alemanha, <b>equipe</b> que tem apresentado o futebol mais convincente no Mundial até então.</li> </ol> <p style="text-align: right;"><small>*Informações ainda não oficiais. Este trabalho faz parte do projeto lexicográfico em andamento</small></p>	

### Entrada Lexicográfica 14 - Equipe

Espera-se que, com entradas lexicográficas que contemplem as informações relacionadas acima, seja possível que o consulente escolha o equivalente mais adequado para a sentença que estiver traduzindo, tendo como respaldo a relação de sinonímia existente entre as ULs *time* e *equipe* e a particularidade da UL *seleção*.



## 5 CONCLUSÃO

A realização desta pesquisa, que teve como objetivo refletir sobre os equivalentes de tradução do domínio do futebol no âmbito de um projeto multilíngue baseado em *frames*, foi capaz de responder aos questionamentos de pesquisa apresentados no início do trabalho, a serem retomados a seguir.

Inicialmente, tinha-se como hipótese que, por se tratar de um domínio específico (futebol), as equivalências lexicais poderiam ser abordadas de forma a privilegiar a relação da unidade lexical com o *frame* ao qual ela pertence. Esta proposta mostrou-se válida, tendo-se em vista que, independentemente do tipo de equivalência de tradução, o *frame* serviu como plano de fundo conceitual nos dois idiomas por estes serem comparáveis nas línguas em questão.

Sobre os tipos de equivalência, um dos questionamentos da pesquisa dizia respeito aos tipos de equivalência mais comum no domínio do futebol. Neste caso, após realizar a análise de 100 ULs, observou-se que a equivalência exata de conteúdos é a mais frequente, seguida pela intersecção de conteúdos. Já a equivalência por superioridade tem um percentual pouco representativo, seguida pela não equivalência, que não retornou nenhum caso. Pode-se dizer, portanto, que se notaram indícios de que estes são os tipos de equivalência mais comuns nesta área.

Tendo-se isso em vista, pode-se por o foco no o outro questionamento, que diz respeito à forma de se pensar nos equivalentes de tradução haja vista a sua inclusão em um recurso lexical multilíngue baseado em *frames*. Esta indagação abre espaço a algumas considerações tecidas durante a parte aplicada do projeto e que serão listadas abaixo, como:

- 1) Sabendo-se que possivelmente os equivalentes de tradução no domínio do futebol caracterizam-se como do tipo equivalência exata de tradução, espera-se, no desenvolvimento do Dicionário da Copa, que grande parte das glosas seja mais facilmente definidas, prevendo-se que este se trata do caso mais simples de definição de equivalências;
- 2) Ao se ter a concepção de *frames* embasando a apresentação das entradas lexicográficas, diversas outras informações também serão disponibilizadas, como evocadores do *frame*, lista de exemplos, EFs, relação entre *frames*, etc. Estas

informações podem tornar o dicionário mais rico, capaz de esclarecer dúvidas mais específicas sobre as ULs a serem traduzidas.

- 3) Através da disponibilização de sentenças obtidas em *corpora* comparáveis, os exemplos disponibilizados no dicionário são de uso real, ou seja, modelos escritos originalmente na língua nativa da UL sendo consultada. Com isso evita-se a interferência que pode existir em *corpora* paralelos, devido à tradução realizada.
- 4) Enquanto que a linguística de *corpus* já vem sendo utilizada na definição de ULs em FrameNets multilíngues, observou-se, nesta pesquisa, que, em se tratando de FrameNets genuinamente multilíngues, esta estratégia se mostra ainda mais imprescindível haja vista o compromisso com a relação entre as duas bases de dados através da indicação de equivalências.

Espera-se, com esta pesquisa, ter contribuído com o trabalho de pesquisa em desenvolvimento pela equipe do grupo Kicktionary\_Br Copa 2014 e, de forma mais ampla, também auxiliar novas pesquisas na área da multilingualidade e desenvolvimento de recursos lexicais baseados em *frames*.

## REFERÊNCIAS

ALVES, Fábio; MAGALHÃES, Célia; PAGANO, Adriana. **Traduzir com Autonomia: estratégias para o tradutor em formação**. São Paulo: Contexto, 2000.

ATKINS, B.T. Sue; RUNDELL, Michael. **The Oxford Guide to Practical Lexicography**. Oxford: Oxford University Press, 2008.

**Aulete iDicionário**. Disponível em: <<http://aulete.uol.com.br>>. Acesso em 3 ago. 2012.

BAKER, Mona. **In other words: a coursebook on translation**. London & New York: Routledge, 2006.

\_\_\_\_\_, Mona; SALDANHA, Gabriela. **Routledge Encyclopedia of Translation Studies**. London & New York: Routledge, 2009.

BASSNETT, Susan. **Estudos de tradução**. Tradução de Vivina de Campos Figueiredo. Lisboa: Fundação Calouste Gulbenkian, 2003.

BERBER SARDINHA, Tony. Linguística de corpus: histórico e problemática. In **D.E.L.T.A.** Vol. 16, N.º 2, p. 323-367, 2000.

\_\_\_\_\_, Tony. **Pesquisa em Lingüística de Corpus com WordSmith Tools**. Campinas: Mercado de Letras, 2009.

BIDERMAN, M. T. C. Unidades complexas do léxico. In: Rio-Torto, G.; Figueiredo, O. M.; Silva, F. (Org.). **Estudos em Homenagem ao Professor Doutor Mário Vilela**. Porto, Portugal: Faculdade de Letras da Universidade do Porto, v. II, p. 747-757, 2005.

BOAS, Hans C. Semantic frames as interlingual representations for multilingual lexical databases. **International Journal of Lexicography**. 18(4), p. 445-478, 2005.

\_\_\_\_\_, Hans C. **Multilingual FrameNets in computational lexicography: methods and applications**. Berlin: Mouton de Gruyter: 2009.

\_\_\_\_\_, Hans C. "A frame-semantic approach to syntactic alternations with build-verbs". In: GUERRERO MEDINA, P. (ed.). **Morphosyntactic alternations in English**. London: Equinox. 207-234, 2011.

\_\_\_\_\_, Hans C. Constructing parallel lexicon fragments based on English FrameNet entries: Semantic and syntactic issues. In: HEDELAND, Hanna; SCHMIDT, Thomas; WÖRNER, Kai (eds.). **Multilingual Resources and Multilingual Applications**. Proceedings of the Conference of the German Society for Computational Linguistics and Language Technology (GSCL), 2011. Disponível em: <<http://www.corpora.uni-hamburg.de/gscl2011/downloads/AZM96.pdf>> Acesso em: 30 nov. 2012.

CATFORD, John Cunnison. **Uma teoria lingüística da tradução: um ensaio em lingüística aplicada**. Tradução do Centro de Especialização de Tradutores de inglês do Instituto de

Letras da Pontifícia Universidade Católica de Campinas. São Paulo: Cultrix; Campinas: Pontifícia Universidade Católica de Campinas, 1980.

CHISHMAN, Rove Luiza de Oliveira; FIGUEIRÓ, Guilherme Tiecher. **Anotação semântica de frames dos verbos de movimento na temática do futebol**. Anais do IX encontro do CELSUL. Santa Catarina: 2010.

DUVAL, Alain. Equivalence in Bilingual Dictionaries. In: FONTENELLE, Thierry (ed.). **Practical lexicography: a reader**. Oxford: Oxford University Press. 2008.

EVANS, Vyvyan; GREEN, Melanie. **Cognitive Linguistics: An Introduction**. Edinburgh University Press, 2006.

EVANS, Vyvyan. Cognitive Science. In: **Wiley Interdisciplinary Reviews: Cognitive Science**. Vol. 3, Issue 2, P. 129-141, 2012.

FELBER, H. **Manuel de terminologie**. Paris: UNESCO, INFOTERM, 1984.

FERRARI, Lilian. O que é Linguística Cognitiva? In: FERRARI, Lilian. **Introdução à Linguística Cognitiva**. São Paulo: Contexto, 2011.

FILLMORE, C. J. Frame Semantics. In: **Linguistics in the morning calm**, 111–138. Seoul, Korea: Hanshin, 1982.

\_\_\_\_\_. **Frames and the Semantics of Understanding**. Vol. 6, n.2, p. 222-254, 1985.

\_\_\_\_\_. Frame Semantics. In: GEERAERTS, D. (ed.) **Cognitive Linguistics: Basic Readings**. Berlin: Mouton de Gruyter: 2006, p. 373-400.

FONTENELLE, Thierry. A bilingual lexical database for Frame Semantics. BOAS, Hans C. **Multilingual FrameNets in computational lexicography: methods and applications**. Berlin: Mouton de Gruyter: 2009, p. 37-57.

**FrameNet**. Disponível em: <<https://framenet.icsi.berkeley.edu/fndrupal/>>. Acesso em: 2 fev. 2012.

**FrameNet Brasil**. Disponível em: <[www.framenetbr.ufjf.br](http://www.framenetbr.ufjf.br)>. Acesso em 2 fev. 2012.

GAWRON, Jean Mark. **Frame Semantics**. 2008. Disponível em: <[http://www.hf.uib.no/forskingskole/new\\_frames\\_intro.pdf](http://www.hf.uib.no/forskingskole/new_frames_intro.pdf)>. Acesso em: 21 set. 2012.

GEERAERTS, Dirk (ed.). **Cognitive Linguistics: Basic Readings**. Berlin / New York: Mouton de Gruyter, 2006.

HATIM, Basil; MUNDAY, Jeremy. **Translation: An advanced resource book**. Routledge. London & New York, 2004.

JAKOBSON, Roman. Aspectos linguísticos da tradução. In:\_\_\_\_\_. **Linguística e Comunicação**. São Paulo: Cultrix, p. 63-72, 1991.

KAY, Martin. Multilinguality. In Varile, G. and A. Zampolli (Eds). **Survey of the State of the Art in Human Language Technology**. Elsnet Publication, 1995.

**Kicktionary**. Disponível em: <[www.kicktionary.de/](http://www.kicktionary.de/)>. Acesso em: 9 fev. 2012.

**Kicktionary\_br**. Disponível em: <[projeto.unisinos.br/semantec/index.php/pt/projetos/13-projeto/54](http://projeto.unisinos.br/semantec/index.php/pt/projetos/13-projeto/54)>. Acesso em: 9 fev. 2012.

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

KRIFKA, Manfred. Semantic Fields and Componential Analysis. In: **Lexikalische Semantik**. Institut für deutsche Sprache und Linguistik, Humboldt - Universität zu Berlin, 2001. Disponível em: <[http://amor.cms.hu-berlin.de/~h2816i3x/Lehre/2001\\_HS\\_LexikalischeSemantik/LexSemantik2.pdf](http://amor.cms.hu-berlin.de/~h2816i3x/Lehre/2001_HS_LexikalischeSemantik/LexSemantik2.pdf)>. Acesso em 15 mar. 2013.

LAGE, Ludimila M. **Frames e construções: um estudo de caso da construção [SN V<sub>DEIXAR</sub> PARA/POR V<sub>INFINITIVO</sub>] e do frame de deixado\_por\_fazer**. Anais do SILEL. Volume 2, Número 2. Uberlândia: 2011.

LAKOFF, George. **Women, Fire, and Dangerous Things: What Categories Reveal about the Mind**. Chicago, IL: University of Chicago Press, 1987.

LANGACKER, Ronald W. **Foundations of Cognitive Grammar: theoretical prerequisites**. Vol. 1. Stanford, California: Stanford University Press, 1987.

LÖNNEKER-RODMAN, Birte. **Multilinguality and FrameNet**. Berkeley, CA: ICSI, 2007.

MUNDAY, Jeremy, ed. (2009): **The Routledge Companion to Translation Studies**. Routledge. London & New York, 2009.

OLIVEIRA, Alessandra Ramos de. **Equivalência: sinônimo de divergência**. Cadernos de Tradução. Florianópolis: 2007, v.1, n.19, p. 97-114.

PETRUCK, Miriam R. L. Frame Semantics. In J-O. Östman, J. Verschueren and J. Blommaert (eds.). **Handbook of Pragmatics**. Amsterdam & Philadelphia: John Benjamins, 1996.

PIOTROWSKI, Tadeusz. **Problems in bilingual lexicography**. Wrocław: Wydawn Uniwersytetu Wrocławskiego, 1994.

RUPPENHOFER, Josef et al. **FrameNet II. Extended Theory and Practice**. Berkeley, California. International Computer Science Institute: 2006. Disponível em: <[https://framenet.icsi.berkeley.edu/fndrupal/the\\_book](https://framenet.icsi.berkeley.edu/fndrupal/the_book)>. Acesso em: 27 jan. 2012.

SALOMÃO, Maria M. M.: FrameNet Brasil: Um trabalho em progresso. In: **Calidoscópico**. Vol. 7.3, 2009.

SAPIR, Edward. **Culture, Language and Personality**. Berkeley, Los Angeles, London: University of California Press, p. 65-77, 1949.

SCHANK, Roger C.; ABELSON Robert P. Scripts, plans and knowledge. Yale University. New Haven, Connecticut, USA, 1977. Disponível em: <<http://oak.conncoll.edu/parker/com316/progassign/scripts.pdf>> Acesso em 9 mar. 2013.

SCHMIDT, Thomas. The Kicktionary – A multilingual lexical resource of football language. In: BOAS, Hans C. **Multilingual FrameNets in computational lexicography: methods and applications**. Berlin: Mouton de Gruyter: 2009, p. 101-132.

SILVA, Augusto Soares da. Introdução: linguagem, cultura e cognição, ou a Linguística Cognitiva. In: SILVA, Augusto Soares da; TORRES, Amadeu; GONÇALVES, Miguel (orgs.). **Linguagem, Cultura e Cognição: Estudos de Linguística Cognitiva**. Vol. I, Coimbra: Almedina, 2004, pp. 1-18.

SNELL-HORNBY, Mary. The bilingual dictionary – help or hindrance?. In: R. R. K. Hartmann (ed.). **LEXeter '83 Proceedings**. Tübingen (Lexicographica. Series Maior ), 1984, p. 274-281. Disponível em: <[http://www.euralex.org/elx\\_proceedings/Euralex1983](http://www.euralex.org/elx_proceedings/Euralex1983)>. Acesso em: 7 mai. 2013.

\_\_\_\_\_, Mary. Of catfish and blue bananas: scenes-and-frames semantics as a contrastive “knowledge system” for translation. In: DAM, Helle V.; ENGBERD, Jan; GERZYMISCH-ARBOGAST, Heidrun (ed.). **Knowledge systems and translation**. Berlin/New York: Mouton de Gruyter, 2005, p. 193-206.

TAGNIN, Stella. E. O. **A Identificação de equivalentes tradutórios em corpora comparáveis**. Anais do I Congresso Internacional da ABRAPUI: Belo Horizonte, 3 a 6 de junho de 2007.

TANNEN, Deborah. What's in a Frame? Surface evidence for underlying expectations. In: FREEDLE, Roy. **New directions in discourse processing**. Norwood, NJ: Ablex, 1979.

ANEXO A  
50 PALAVRAS MAIS FREQUENTES NOS *CORPORA* EM INGLÊS E EM PORTUGUÊS

50 ULs mais frequentes do corpus em PT

	Palavra	Frequência
1	MINUTOS	5755
2	BOLA	5311
3	GOL	5184
4	ÁREA	4622
5	FOI	4260
6	JOGO	3423
7	TEMPO	3043
8	TIME	2718
9	CAMPO	2359
10	DEFESA	2339
11	GOLEIRO	2300
12	PARTIDA	2246
13	FALTA	2141
14	ATAQUE	1871
15	EQUIPE	1870
16	FEZ	1857
17	PLACAR	1528
18	ATACANTE	1433
19	ESCANTEIO	1399
20	CHUTOU	1327
21	VITÓRIA	1304
22	ETAPA	1278
23	SÃO	1265
24	PONTOS	1251
25	JOGADA	1248
26	TEVE	1201
27	CONSEGUIU	1187
28	RECEBEU	1170
29	BATEU	1156
30	ENTRADA	1109
31	CHUTE	1089
32	EMPATE	1055
33	CHANCE	1053
34	PERIGO	1050
35	PASSOU	1031
36	COPA	1030
37	GOLS	1023
38	COBRANÇA	996
39	CASA	948
40	CRUZOU	942
41	LADO	939
42	LANCE	932
43	CHEGOU	924
44	ZAGA	918
45	CAMPEONATO	903
46	RESULTADO	896
47	RODADA	894
48	SAIU	891
49	MINUTO	884
50	CABEÇA	882

50 ULs mais frequentes do corpus em EN

	Palavra	Frequência
1	WAS	9192
2	BALL	4379
3	HAD	4230
4	MINUTES	3924
5	GOAL	3906
6	HAVE	3313
7	IS	3284
8	BACK	3030
9	TIME	2904
10	WERE	2854
11	SHOT	2835
12	CORNER	2297
13	KICK	2185
14	CROSS	2141
15	GAME	2125
16	HOME	2080
17	SIDE	2042
18	MINUTE	1691
19	POST	1632
20	EFFORT	1605
21	BEEN	1592
22	BOX	1520
23	LEAGUE	1337
24	MADE	1300
25	HEADER	1275
26	ARE	1260
27	HAS	1210
28	CHANCE	1151
29	CAME	1149
30	WIN	1079
31	PENALTY	1069
32	YARDS	1052
33	BREAK	1029
34	AREA	1025
35	PAST	1015
36	END	1005
37	CITY	973
38	PASS	973
39	SEASON	934
40	CUP	929
41	FINAL	909
42	LEAD	899
43	MATCH	882
44	CLEAR	878
45	PLAY	878
46	WAY	870
47	LOOKED	859
48	WENT	842
49	SAVE	821
50	GET	817

ANEXO B  
ANÁLISE INICIAL DAS 50 PALAVRAS MAIS FREQUENTES NO *CORPUS* EM  
PORTUGUÊS

Palavra	Freq.	Possível equivalente	Observação	Possível tipo de equivalência
MINUTOS	5755	minutes	-	Equivalência exata de conteúdo.
BOLA	5311	ball	-	Equivalência exata de conteúdo.
GOL	5184	goal	-	Equivalência exata de conteúdo.
ÁREA	4622	area yard, inner, court range, scope, field	Área e Campo com equivalentes em comum	Superioridade.
FOI	4260	be / go	-	Intersecção de conteúdo.
JOGO	3423	game, match	Jogo e partida com equivalentes em comum	Superioridade.
TEMPO	3043	time	-	Equivalência exata de conteúdo.
TIME	2718	team	No caso da copa, uso de <i>Seleção (team)</i>	Superioridade.
CAMPO	2359	field, area countryside, interior	Área e Campo com equivalentes em comum	Superioridade.
DEFESA	2339	defence	Escreve-se <i>defence</i> em inglês britânico	Equivalência exata de conteúdo.
GOLEIRO	2300	goalkeeper	-	Equivalência exata de conteúdo.
PARTIDA	2246	match, game	Jogo e partida com equivalentes em comum	Superioridade.
FALTA	2141	lack, absence	Polissemia.	Intersecção de conteúdo.
ATAQUE	1871	attack, offensive, charge	Metáfora. Futebol é guerra	Equivalência exata de conteúdo.
EQUIPE	1870	team	Time e equipe com equivalente em comum	Superioridade.
FEZ	1857	do, make	-	Intersecção de conteúdo.
PLACAR	1528	score, plaque, placate, appease	Polissemia. Placar físico / Resultado.	Intersecção de conteúdo.
ATACANTE	1433	striker, forward	Metáfora. Futebol é guerra	Equivalência exata de conteúdo.
ESCANTEIO	1399	corner	-	Equivalência exata de conteúdo.
CHUTOU	1327	kick	-	Equivalência exata de conteúdo.
VITÓRIA	1304	victory	Polissemia. Time Vitória / Resultado	Intersecção de conteúdo.
ETAPA	1278	stage	Polissemia. Etapa do jogo / Estágio da competição	Intersecção de conteúdo.
SÃO	1265	are	-	Intersecção de conteúdo.
PONTOS	1251	points, dots	-	Equivalência exata de conteúdo.
JOGADA	1248	play, game; throw, cast, move, hit, shot, stroke	Polissemia. Fazer uma jogada, receber uma jogada,...	Intersecção de conteúdo.
TEVE	1201	have	-	Intersecção de conteúdo.
CONSEGUIU	1187	obtain, achieve, get	Polissemia. Conseguir marcar, conseguir defender,...	Intersecção de conteúdo.
RECEBEU	1170	receive	Polissemia. Receber um cruzamento, receber um passe,...	Intersecção de conteúdo.
BATEU	1156	beat, hit, strike	Polissemia. Bater a gol, bater falta, bater na zaga,...	Intersecção de conteúdo.
ENTRADA	1109	entrance	Polissemia. Entrada da área, entrada de um jogador,...	Intersecção de conteúdo.
CHUTE	1089	kick	-	Equivalência exata de conteúdo.
EMPATE	1055	tie, draw	-	Equivalência exata de conteúdo.
CHANCE	1053	chance, opportunity	Polissemia. Chance de gol, chance de finalizar,...	Intersecção de conteúdo.
PERIGO	1050	danger, peril, hazard	Polissemia. Sair com perigo, chutar com perigo,...	Intersecção de conteúdo.
PASSOU	1031	pass, nail	Polissemia. Passar a bola, passar perto do gol,...	Intersecção de conteúdo.
COPA	1030	cup, competition	-	Equivalência exata de conteúdo.
GOLS	1023	goals	Ver <i>gol</i>	Equivalência exata de conteúdo.
COBRANÇA	996	charge, collection	Metáfora.	Intersecção de conteúdo.
CASA	948	house	-	Equivalência exata de conteúdo.
CRUZOU	942	cross, middle	-	Equivalência exata de conteúdo.
LADO	939	side	-	Equivalência exata de conteúdo.
LANCE	932	move	-	Equivalência exata de conteúdo.
CHEGOU	924	arrive, come close, approach	Polissemia. Chegar ao gol, chegar ao empate,...	Intersecção de conteúdo.
ZAGA	918	-	-	Equivalência exata de conteúdo.
CAMPEONAT O	903	championship	-	Equivalência exata de conteúdo.
RESULTADO	896	score	-	Equivalência exata de conteúdo.
RODADA	894	round	-	Equivalência exata de conteúdo.
SAIU	891	leave, go out, quit	Polissemia. Sair do campo, sair na frente, sair vitorioso,...	Intersecção de conteúdo.
MINUTO	884	minute	Ver <i>minutes</i>	Equivalência exata de conteúdo.
CABEÇA	882	cabeça	-	Equivalência exata de conteúdo.

Obs.:

- 1) Não estão sendo trazidas aqui as equivalências que estivessem sendo indicadas como adequadas para outros domínios.
- 2) No caso de palavras que podem ser substantivos ou verbos, a separação entre as classes nos possíveis equivalentes é indicada por uma barra, como em *chute / chutar*.



ANEXO C  
ANÁLISE INICIAL DAS 50 PALAVRAS MAIS FREQUENTES NO *CORPUS* EM  
INGLÊS

Palavra	Freq.	Possível equivalente	Observação	Possível tipo de equivalência
WAS	9192	ser / estar	-	Intersecção de conteúdo.
BALL	4379	bola	-	Equivalência exata de conteúdo.
HAD	4230	ter	-	Intersecção de conteúdo.
MINUTES	3924	minutos	Palavras <i>minutes</i> e <i>minute</i> com alta frequência. Possivelmente <i>minute</i> seja usado como adjetivo.	Equivalência exata de conteúdo.
GOAL	3906	gol, baliza	Polissemia. Fazer gol ou meta física	Intersecção de conteúdo.
HAVE	3313	ter	-	Intersecção de conteúdo.
IS	3284	ser / estar	-	Intersecção de conteúdo.
BACK	3030	costas, verso, traseira	-	Intersecção de conteúdo.
TIME	2904	tempo, horas, época / cronometrar	Polissemia. Tempo ou horas	Intersecção de conteúdo.
WERE	2854	ser / estar	-	Intersecção de conteúdo.
SHOT	2835	tiro / chute a gol	Polissemia. Chute ou chute a gol	Intersecção de conteúdo.
CORNER	2297	esquina, quina, canto	A consulta por <i>escanteio</i> retorna a tradução " <i>corner</i> "	Equivalência exata de conteúdo.
KICK	2185	chute / chutar	-	Equivalência exata de conteúdo.
CROSS	2141	mistura / cruzar, atravessar	-	Equivalência exata de conteúdo.
GAME	2125	jogo, partida, brincadeira	Polissemia. <i>Jogo</i> ou <i>partida</i>	Superioridade.
HOME	2080	casa, lar	Metáfora	Intersecção de conteúdo.
SIDE	2042	lado	-	Intersecção de conteúdo.
MINUTE	1691	minuto	Palavras <i>minutes</i> e <i>minute</i> com alta frequência. Possivelmente <i>minute</i> seja usado como adjetivo.	Equivalência exata de conteúdo.
POST	1632	trave	-	Equivalência exata de conteúdo.
EFFORT	1605	esforço	-	Intersecção de conteúdo.
BEEN	1592	ser / estar	-	Intersecção de conteúdo.
BOX	1520	caixa	-	Intersecção de conteúdo.
LEAGUE	1337	liga	-	Equivalência exata de conteúdo.
MADE	1300	make	-	Intersecção de conteúdo.
HEADER	1275	cabeçalho	Nova consulta para obter a possível tradução de <i>cabecear</i>	Equivalência exata de conteúdo.
ARE	1260	ser / estar	-	Intersecção de conteúdo.
HAS	1210	have	-	Intersecção de conteúdo.
CHANCE	1151	chance, sorte	Polissemia. <i>Chance for a goal, chance to shoot,...</i>	Intersecção de conteúdo.
CAME	1149	vir, chegar	Polissemia. <i>Come to an end, come close,...</i>	Intersecção de conteúdo.
WIN	1079	ganhar, vencer	-	Equivalência exata de conteúdo.
PENALTY	1069	pênalti	-	Equivalência exata de conteúdo.
YARDS	1052	jarda (0.9144 metros)	Medida não é utilizada em português.	Intersecção de conteúdo.
BREAK	1029	quebrar	Polissemia. <i>The break, chance to break,...</i>	Intersecção de conteúdo.
AREA	1025	área	<i>Área</i> e <i>Campo</i> agora não aparecem ambos como equivalentes.	Superioridade.
END	1005	acabar, terminar	-	Equivalência exata de conteúdo.
CITY	973	cidade	-	Equivalência exata de conteúdo.
PASS	973	passar, passe	-	Equivalência exata de conteúdo.
SEASON	934	estação, temporada	-	Equivalência exata de conteúdo.
CUP	929	taça, caneco	-	Equivalência exata de conteúdo.
FINAL	909	final, finalíssima	-	Equivalência exata de conteúdo.
LEAD	899	passe ofensivo	Polissemia. <i>Take the lead, double the lead,...</i>	Equivalência exata de conteúdo.
MATCH	882	partida, jogo	Polissemia. <i>Jogo</i> ou <i>partida</i>	Superioridade.
CLEAR	878	remover, desobstruir, transpor,...	Polissemia. <i>Fail to clear, clear the ball,...</i>	Intersecção de conteúdo.
PLAY	878	jogar	-	Equivalência exata de conteúdo.
WAY	870	caminho, maneira, modo, via, jeito,...	-	Equivalência exata de conteúdo.
LOOKED	859	olhar, ver, parecer	Polissemia. <i>Looked like, looked dangerous on,...</i>	Intersecção de conteúdo.
WENT	842	ir	-	Intersecção de conteúdo.
SAVE	821	salvar	-	Equivalência exata de conteúdo.
GET	817	conseguir, obter, pegar, chegar, etc.	-	Intersecção de conteúdo.
VISITORS	810	visitante	-	Equivalência exata de conteúdo.

Obs.:

- 1) Não estão sendo trazidas aqui as equivalências que estivessem sendo indicadas como adequadas para outros domínios.
- 2) No caso de palavras que podem ser substantivos ou verbos, a separação entre as classes nos possíveis equivalentes é indicada por uma barra, como em *chute / chutar*.