

A terminologia do futebol: um estudo direcionado pelo *corpus*

Football terminology: a *corpus*-driven study

Sabrina Matuda*
Stella E. O. Tagnin**

RESUMO: Este artigo tem como objetivo estudar a terminologia do futebol em inglês e português. Para tanto, a fundamentação teórica embasa-se na Linguística de *Corpus*, na Terminologia Textual, na tradução técnica como condicionante cultural e no conceito forma-representação. O *corpus* de estudo possui aproximadamente um milhão de palavras em cada língua. Para etiquetar o *corpus*, utilizamos o etiquetador *Tree Tagger*, desenvolvido por Helmut Schmid, e, para explorá-lo, o *WordSmith Tools*, de Mike Scott.

PALAVRAS-CHAVE: Tradução. Linguística de *Corpus*. Terminologia. Futebol e cultura.

ABSTRACT: This article aims at investigating football terminology in English and Portuguese. To that aim, the study is based on the notions of *Corpus* Linguistics and Textual Terminology. To explain cultural differences, technical translation is viewed as a communicative act subject to cultural restraints and the concept of ‘form-representation’ is called upon to elucidate such differences. Our corpus consists of approximately two million words. To tag the corpus we used Helmut Schmid’s *Tree-Tagger* and to explore the corpus we used Mike Scott’s *WordSmith Tools*.

KEYWORDS: Translation. Corpus Linguistics. Terminology. Football and Culture.

1. Por que estudar a terminologia do futebol

O futebol é o esporte mais praticado no Brasil e no mundo. É reconhecido mundialmente como competição, manifestação cultural e até mesmo como um mercado na ordem econômica. A FIFA (Federação Internacional de Futebol Associado) congrega mais de 140 milhões de jogadores de 300 mil clubes, em 207 países afiliados.

O Brasil, conhecido mundialmente como o “País do Futebol”, tem aproximadamente 40 milhões de praticantes, entre atletas profissionais e amadores, cerca de 11 mil jogadores federados, 800 clubes, mais de dois mil atletas atuando em outros países e 580 estádios. Além de, pelo menos, 20 mil “campinhos” de pelada, nos bairros de classe mais baixa, escolinhas de futebol e milhares de torcedores.

* Doutoranda da Faculdade de Filosofia, Letras e Ciências Humanas (FFLCH) da Universidade de São Paulo (USP).

** Livre Docente pela Faculdade de Filosofia, Letras e Ciências Humanas (FFLCH) da Universidade de São Paulo (USP).

O esporte é praticado em todo o país por camadas sociais distintas e em diferentes espaços: campos de várzea, quadras, praias, ruas, escolas, clubes etc. Enfim, é parte do cotidiano de muitos brasileiros. DaMatta (1982) destaca que o futebol praticado, vivido e discutido no Brasil é um dos modos pelo qual a sociedade brasileira fala, se apresenta e se deixa descobrir. Nessa mesma linha, é também notório o reconhecimento do futebol, no Brasil, como objeto das ciências humanas ao longo dos últimos anos. Sobretudo a partir da década de 1990, ampliaram-se as pesquisas acadêmicas e publicações editoriais acerca do futebol (GIGLIO; SPAGGIARI, 2010).

Essa popularidade, tanto no Brasil quanto em outros países, tem aumentado a participação da indústria do futebol na economia mundial, movimentando cerca de 250 bilhões de dólares anuais (LEONCINI; SILVA, 2005).

É fato que as relações futebolísticas entre o Brasil e os países da Europa crescem cada vez mais (CRUZ, 2005), seja pelo intercâmbio de contratos de jogadores e técnicos, pelos direitos de transmissão de campeonatos, pelo patrocínio de jogadores por grandes marcas ou por qualquer outra negociação que envolva um produto relacionado ao futebol.

Para que todas essas relações se materializem, estabelece-se uma comunicação que, na grande maioria das vezes, se dá em língua inglesa. No entanto, cada nação tem a sua maneira de jogar, torcer e narrar, maneira, essa, expressa por meio de sua língua materna. O problema surge quando se quer expressar essas particularidades em uma língua estrangeira.

Embora seja, muitas vezes, relacionado somente ao lazer, o futebol não deixa de ser uma área de especialidade tanto no Brasil quanto em outros países. Sendo assim, possui uma linguagem própria utilizada para descrever o universo a ele relacionado. Essa linguagem é padronizada e, justamente por esse motivo, não pode ser utilizada de qualquer forma. Ao falar em padronização, não pretendemos, de forma alguma, tentar normatizar o léxico do futebol para favorecer a eficácia das comunicações especializadas em torno dessa temática. Ao contrário, pretendemos favorecer as peculiaridades de cada texto dentro de seu discurso (KRIEGER; FINATTO, 2004) levando em conta os aspectos históricos e socioculturais presentes em cada cultura.

A fim de melhor entender como essas particularidades históricas e socioculturais se manifestam na língua, utilizamos o conceito forma-representação proposto por Toledo (2002). Segundo o antropólogo, o conjunto de regras que define a atividade como esporte não delimita as maneiras de jogar. Na verdade, é a apropriação e a interpretação cultural que cada

região faz das regras que determinam as “formas de jogo”. Justapostas, as regras e as “formas de jogo” dão origem às “representações”, ou seja, o ajustamento da observação empírica das “formas de jogo” em um plano simbólico que, por sua vez, consolida as conhecidas “escolas”, “jeitos” e “estilos” de se jogar.

2. Objetivos

O principal objetivo deste artigo, recorte de nossa pesquisa de doutorado que está sendo desenvolvida no programa de Estudos da Tradução da Universidade de São Paulo, é demonstrar como a extração terminológica e a identificação de equivalentes tradutórios em *corpora* comparáveis contribui para uma investigação terminológica que transpõe a esfera linguística e que considera fatores culturais e históricos no processo de entendimento do léxico especializado. Para tanto, estabelecemos dois princípios, a saber:

- 1) a pesquisa foi feita do português para o inglês;
- 2) como se trata de um estudo exploratório, ilustramos os passos seguidos para a extração terminológica e a identificação de equivalentes tradutórios por meio de um estudo de caso com ‘gol’, segundo termo mais frequente no *corpus*, e a unidade fraseológica¹ mais frequente que comporta o termo: ‘fazer um/o gol’.

3. Tradução técnica como condicionante cultural

Entendemos tradução técnica como um ramo da tradução que se ocupa da tradução de textos de línguas de especialidade. Os tradutores, assim como os intérpretes, os redatores técnicos, os jornalistas e os documentalistas são usuários indiretos das terminologias, pois a eles interessa o uso adequado dos termos, das fraseologias e de expressões idiomáticas para que o texto esteja de acordo com as normas de convencionalidade que regem a produção do tipo e do gênero textual em questão na cultura em que é produzido.

Muito comum é a ideia de que a tradução técnica, diferente da literária, constitui um universo à parte, pois, para alguns, sua terminologia não se deixa contaminar por relações contextuais e pragmáticas, possuindo certa estabilidade que favorece e facilita o processo tradutório. Com base nesses preceitos, a tradução técnica é, muitas vezes, definida como uma

¹ Em nosso trabalho, consideramos unidades fraseológicas associações de, no mínimo, duas palavras, sendo uma delas obrigatoriamente uma palavra-chave no corpus e que apresentem frequência maior ou igual a três.

operação de transcodificação em que os conceitos “constituem uma [sic] amálgama indissolúvel e imune aos efeitos do tempo e do espaço, a fim de poderem resistir a uma série de condicionantes a que estão expostos.” (AZENHA, 1999, p. 10).

Adotamos a proposta de Azenha (1999, p. 60), a de uma tradução técnica que vai além dos limites do texto, uma tradução que define o ato tradutório tendo como ponto de partida o ato de comunicação. Em virtude disso, o texto técnico, assim como qualquer outra forma de comunicação, está atrelado a uma realidade sócio-histórico-cultural.

Para o autor, não podemos deixar de lado as relações que o texto técnico trava com o contexto em que é produzido. Ao contrário, os termos, as fraseologias, as definições, os equivalentes e as expressões devem ser empregados de maneira convencional² no texto de chegada, respeitando as variáveis ligadas ao emissor, receptor, situação e objetivo de comunicação.

Neste estudo, consideramos “equivalente” uma unidade fraseológica que funcione no texto de chegada como funciona no texto de partida (TAGNIN, 2007, p. 1). O conceito, bastante amplo, nos permite identificar equivalentes não só no nível da palavra, do texto ou da frase, mas também equivalentes pragmáticos, ou seja, equivalentes que, embora não reflitam uma tradução direta para a língua de chegada, são utilizados no mesmo contexto e com o mesmo objetivo comunicativo.

Ao considerarmos que os termos são empregados em dado cenário histórico-cultural e, portanto, condicionados por normas sociais e linguísticas sempre sujeitas a alterações (AZENHA, 1999, p. 22), não vemos outra possibilidade senão o uso de *corpora* comparáveis para a identificação dos equivalentes tradutórios.

4. Terminologia Textual

Neste estudo, adotamos a Terminologia Textual como aporte teórico para nossa pesquisa.

Seria inviável falar de Terminologia Textual sem discutir, ainda que brevemente, seu objeto de estudo: o texto especializado que, segundo Hoffmann (1998 *apud* KRIEGER; FINATTO, 2004, p. 113) , possui dois eixos básicos: um horizontal, relacionado ao critério temático, englobando diferentes disciplinas e suas eventuais subdivisões, e um vertical,

² Adotamos a definição de convencionalidade proposta por Tagnin (2005): “forma peculiar de expressão de uma dada língua ou comunidade lingüística” p.14.

relacionado ao grau de especialização do texto, ou seja, à densidade terminológica. A classificação de Hoffmann ressalta o fator pragmático no âmbito das comunicações especializadas e enfatiza o papel que as unidades lexicais assumem em diferentes contextos especializados de comunicação.

Alguns tipos de texto possuem maior densidade terminológica do que outros: em nosso caso, as notícias de resultados de partidas possuem densidade terminológica bem menor do que os relatos de partida minuto a minuto. Por outro lado, seria ingenuidade de nossa parte analisar a densidade terminológica de nosso *corpus* somente pela divisão de gêneros textuais. Por esse motivo, optamos por adotar um caráter gradual para observar a densidade terminológica dos textos, pois, por exemplo, um texto sobre resultado de partida publicado por um tabloide inglês pode apresentar uma terminologia bem distinta daquela expressa em um texto do mesmo gênero publicado por um jornal mais tradicional. Essa variação terminológica não se refere apenas à densidade terminológica de um texto, mas também às diferentes terminologias encontradas em periódicos distintos. Por esse motivo, mais do que o tema, o grau de densidade informativa, a forma de se comunicar, a situação em que se comunica e para quem se comunica constituem fatores determinantes do grau de especialização de um texto.

Krieger e Finatto (2004) propõem que a variação tipológica no âmbito da comunicação especializada se reflete, por exemplo, na distinção entre uma tese, um artigo de periódico altamente especializado em determinada área do conhecimento e um texto de jornal ou de revista informativa redigido com a finalidade de divulgar ao grande público um acontecimento científico. Essa distinção não pode ser aplicada ao nosso caso por vários motivos. Primeiramente, devido à grande repercussão que o futebol tem na sociedade moderna, a delimitação de “grande público” e de especialistas se torna um pouco controversa. Em um primeiro momento, tenderíamos a caracterizar os jornalistas esportivos, jogadores, técnicos e membros da comissão técnica como especialistas e os leitores e fãs de futebol em geral como “grande público”. No entanto, o que acontece é que, em se tratando de futebol, todo torcedor pode ser considerado um especialista na área.

Devido a essa grande dificuldade de delimitação de “grande público” e de especialistas, adotamos, aqui, a classificação de atores sociais do futebol proposta por Toledo (2002, p. 15):

Parto de três [atores sociais] dentro do campo esportivo: os profissionais (jogadores, técnicos, dirigentes, juízes, preparadores, médicos, etc.), os especialistas (a crônica esportiva) e o conjunto genérico de torcedores, “comuns” ou nomeados e reunidos em certas coletividades específicas.

Essa classificação é utilizada por Toledo para uma análise das lógicas do futebol, com o propósito de caracterizar os grupos sociais que se expressam por meio do futebolês. Dentro dessa ordem, os profissionais são todos aqueles que interferem diretamente no jogo, quer dentro do campo, como os jogadores, quer fora do campo, como os dirigentes e instituições como federações. Os especialistas são representados pela crônica esportiva e pelo jornalismo esportivo e, segundo Toledo, ocupam um lugar simbólico equidistante entre os profissionais e os torcedores, pois não participam efetivamente da partida, mas também não se comprometem em um nível de emoção partidária, pelo menos em teoria. Por fim, o “grande público” é caracterizado pelos torcedores.

É a classe dos especialistas que mais nos interessa neste trabalho, pois é através do produto do fazer jornalístico que observamos o futebol fora de seu *locus* de ritualização máxima, que é a partida. Ademais, é também por meio dos textos jornalísticos que o futebol alcança o “grande público”. Contudo, não podemos deixar de enfatizar a importância que o terceiro grupo, o conjunto genérico de torcedores, desempenha em nossa pesquisa, pois frequentemente recorremos a torcedores e amantes do esporte para melhor entender uma jogada ou drible.

Deve-se somar, ainda, à classificação de Toledo, o grau de densidade informativa da notícia, a forma de se comunicar, os traços de impessoalidade do jornalista, o contexto em que a notícia foi produzida³, o propósito da comunicação, que pode ser mais informativo, mais descritivo, mais técnico, para que, assim, seja possível caracterizar o texto técnico propriamente.

Nessa concepção de texto especializado, a presença de termos deixa de ser o elemento primordial que configura o caráter de uma comunicação especializada (KRIEGER; FINATTO, 2004). Ao contrário, os mecanismos linguísticos, textuais e pragmáticos dos quais um texto especializado faz uso também constituem elementos caracterizadores de dada língua

³ Uma notícia sobre o resultado de partida de um jogo do campeonato brasileiro provavelmente é produzida em um contexto mais específico e, portanto, voltada para um público mais específico do que uma notícia sobre uma partida de Copa do Mundo, pois durante esse acontecimento mais pessoas torcem, inclusive torcedores não habituais.

de especialidade. O uso de todos esses elementos em conjunto nos permite enxergar a complexidade do texto especializado.

Uma vez caracterizado seu principal objeto de estudo, passemos, agora, a uma breve discussão sobre as características da Terminologia Textual.

Para Krieger e Finatto (2004,), essa Terminologia, na qual o objeto ‘termo’ dá lugar ao objeto ‘texto’, possui duas características principais. A primeira refere-se ao reconhecimento do papel do cenário comunicativo e, conseqüentemente, do texto especializado para a descrição de uma terminologia. A segunda está relacionada ao estudo e caracterização do texto especializado.

Ao reconhecer o papel do cenário comunicativo, a Terminologia textual parte do pressuposto de que os termos são utilizados em situação de comunicação e que, portanto, não devem, ou melhor, não podem ser estudados à parte do contexto sociocultural em que ocorrem.

Os termos passaram a ser analisados em uso, ou seja, em textos especializados, permitindo, assim, a verificação de fenômenos até então ignorados, ou deixados em segundo plano pelos terminólogos. O contexto discursivo, antes considerado insignificante, passa a representar uma das principais características da Terminologia, contribuindo para um novo tratamento das Unidades Terminológicas (UTs), deixando de lado a ideia de que os termos constituem construtos teóricos idealizados em um sistema linguístico independente.

5. A Linguística de *Corpus* neste estudo

São inúmeras as definições para Linguística de *Corpus* (LC) e não nos cabe aqui, por questões de tempo e propósito de pesquisa, apresentar um panorama com todas as definições, que diferem consideravelmente umas das outras.

Os teóricos da LC frequentemente discordam quanto à sua caracterização. Autores como Berber Sardinha (2000), consideram a LC uma abordagem, outros, como é o caso de Rocha (2001), uma metodologia. Existem ainda os estudiosos que preferem ser mais neutros em suas definições e não tomam nenhum partido como, por exemplo, Aijmer e Altenberg (1991, p. 2), que definem a LC como “o estudo da língua por meio de *corpora*”. Há ainda os que adotam as duas definições como Bowker e Pearson (2002, p. 20): “uma abordagem ou metodologia para o estudo da língua em uso”.

Berber Sardinha chama atenção para o fato de que a LC não constitui somente um

metodologia, um instrumental, do qual outras áreas podem se valer para o estudo da linguagem. Para o autor, a LC apresenta também uma nova perspectiva de se chegar à linguagem possibilitando aos seus seguidores produzir conhecimento novo que muitas vezes coloca teorias linguísticas tradicionais em questão.

Neste trabalho, adotamos duas definições de LC que, até o presente momento, nos pareceram as mais abrangentes. Primeiramente, assim como Berber Sardinha (2004 p.32), acreditamos que a LC constitui uma abordagem empirista que toma a língua como sistema probabilístico, refletindo uma nova maneira de enxergar a linguagem que, conseqüente e futuramente, poderá dar origem a uma nova teoria linguística. Adotamos, também, a visão de McEnery e Hardie (2012:1), que definem LC como “[...] área que foca em um conjunto de procedimentos, ou métodos, para o estudo da língua [...]”⁴.

As duas definições acima se enquadrariam perfeitamente nos propósitos de nossa pesquisa se utilizássemos a LC unicamente para explorar um fenômeno linguístico. No entanto, o escopo de nossa pesquisa vai além da esfera linguística; estabelecemos, sempre que possível, um paralelo entre língua, cultura e história. Ademais, acreditamos que a LC pode e deve ser utilizada como metodologia, e, sim, unicamente como metodologia para pesquisas de outras áreas como, por exemplo, história, antropologia e jornalismo. Por esse motivo, acreditamos que as visões de Berber Sardinha e de McEnery e Hardie se mostram limitadas às áreas de estudos da linguagem. É fato que existe uma tendência mais atual de utilizar a LC em áreas afins; um bom exemplo dessa tendência foram os três cursos de verão oferecidos pela Lancaster University em julho de 2013: *UCREL Summer School in Corpus linguistics*, *ESRC Summer School in GIS in Corpus Approaches for Social Sciences* e *ERC Summer School in GIS for the Digital Humanities*. Como podemos observar, somente um dos cursos foi direcionado a linguistas.

Enfim, não nos cabe, aqui, criar uma nova definição de LC, apenas atentamos para o fato de que, embora utilizemos essas duas definições, acreditamos que a LC pode ser utilizada para outros fins, que não se restrinjam a pesquisas linguísticas.

Um *corpus* pode ser utilizado de diferentes maneiras para validar, exemplificar, contestar ou formular teorias linguísticas. Tognini-Bonelli (2001) distingue duas abordagens principais de pesquisa realizadas em *corpora*: abordagem baseada em *corpus* (*corpus-based*) e abordagem direcionada pelo *corpus* (*corpus-driven*).

⁴ “[...] area which focuses upon a set of procedures, or methods, for studying language [...]”

Na abordagem baseada em *corpus*, o linguista utiliza o *corpus* para explicitar, testar e exemplificar teorias e hipóteses pré-existentes e, principalmente, para extrair exemplos.

A vantagem dessa abordagem é que a extração de exemplos autênticos, seja para fins lexicográficos ou para a validação de hipóteses, confere mais autoridade à pesquisa. Por outro lado, utilizar o *corpus* somente para verificar dados limita a visão do linguista, que ignora novos fenômenos deixando de fazer novas descobertas e de desafiar teorias já existentes.

Na abordagem direcionada pelo *corpus*, o linguista analisa o *corpus* sem hipóteses pré-concebidas. O *corpus* mostra-lhe o caminho a ser percorrido. As descrições são feitas sempre com base nas evidências do *corpus*, possibilitando, assim, novas descobertas. Por isso, dizemos que nessa abordagem o linguista não busca evidências para classificá-las dentro de categorias pré-definidas. Ao contrário, se no decorrer da pesquisa não forem encontrados padrões linguísticos ou se os padrões encontrados não puderem ser classificados em alguma categoria, os achados constituirão argumentos de extrema relevância para a descrição da linguagem ou para a descoberta de novos fenômenos.

Nessa abordagem, o caminho metodológico percorrido pelo linguista é claro: a observação dos dados conduz à formulação de hipóteses que, conseqüentemente, leva à generalização dos resultados possibilitando, assim, a formulação de novas teorias (TOGNINI-BONELLI, 2001).

Apesar de as duas abordagens apresentarem características bem distintas, acreditamos que podem ser utilizadas em conjunto. Neste trabalho, utilizamos a abordagem direcionada pelo *corpus* para extrair os termos a serem estudados por meio das palavras-chave e de seus agrupamentos (*clusters*). Por outro lado, lançamos mão da abordagem baseada em *corpus* quando partimos de uma tradução *prima facie* para a busca dos equivalentes nas linhas de concordância.

6. O *corpus* de estudo

O *design* e a qualidade do *corpus* de estudo constituem o pilar de qualquer pesquisa em *corpus*. O quadro que segue mostra o *design* do *corpus* utilizado neste estudo⁵:

⁵ Chamamos atenção para o fato de que o *corpus* de estudo utilizado na presente pesquisa foi compilado para atender os objetivos de nossa pesquisa de doutorado, a saber: 1) verificar como os diferentes jeitos de jogar, a história do futebol em cada cultura, a apropriação cultural das regras na Inglaterra e no Brasil e outros fatores de ordem histórico-social contribuíram para a criação do léxico do futebol em português e inglês; 2) criar um glossário de futebol composto por verbetes que evidenciem diferenças culturais entre o Brasil e a Inglaterra.

Quadro 1: composição do *corpus* de estudo.

Conteúdo	especializado			
Assunto	futebol			
Autoria	de língua nativa			
Língua	português (BR) e inglês (ING); comparável			
Finalidade	de estudo			
Meio	eletrônico			
Modo	escrito			
tipo de texto	resultados de partidas, narrações minuto a minuto e narrações minuto a minuto com comentários de internautas.			
Período	2013 – 2014			
Seleção	de amostragem; balanceado			
Tamanho	aproximadamente 500 mil palavras em cada língua			
	Inglês		Português	
	No. palavras	No. textos	No. Palavras	No. textos
	600.079	612	469.765	864

O *corpus* utilizado na pesquisa é composto por textos de três periódicos ingleses sobre resultados de partidas da primeira divisão do campeonato inglês de 2013/2014⁶ e por textos de três periódicos brasileiros sobre resultados de partidas da primeira divisão do campeonato brasileiro de 2013⁷.

O *corpus* é caracterizado como *corpus* especializado, uma vez que é composto por textos de uma única área de especialidade: futebol. Coletamos somente textos escritos originalmente em português brasileiro e em inglês britânico.

No que se refere ao modo, nosso *corpus* é escrito, pois não trabalhamos com textos orais. Optamos por coletar três tipos de textos: resultados de partidas, narrações minuto a minuto, que são escritas durante a partida por um narrador *on-line*; e narrações minuto a minuto com comentários de internautas, também escritas durante a partida por um narrador e com contribuições de internautas que compartilham suas opiniões sobre os lances do jogo.

7. O *corpus* de referência

O *corpus* de referência é utilizado para contrastar com o *corpus* de estudo a fim de evidenciar as formas mais frequentes nesse último, filtrando os elementos mais genéricos. Em

⁶ Os periódicos selecionados foram: o jornal *The Guardian*, o tabloide *Daily Mail* e o site sobre futebol *Football.com*.

⁷ Os periódicos selecionados foram: o jornal *O Estado de São Paulo*, o jornal esportivo *Gazeta Esportiva* e a revista sobre futebol *Placar*.

geral, deve-se incluir vários gêneros textuais em um *corpus* de referência, de modo que proporcione uma escolha não-marcada das palavras-chave, pois suas características influenciam de forma direta os tipos de palavra que podem se tornar chave (BERBER SARDINHA, 2004).

O tamanho do *corpus* de referência pode influenciar o número de palavras-chave obtidas. Berber Sardinha (2004, p. 102) recomenda que o *corpus* de referência seja entre três e cinco vezes maior que o *corpus* de estudo.

Compilar um *corpus* muito maior do que o recomendado não retornará, necessariamente, maior número de palavras-chave (BERBER SARDINHA, 2004). No entanto, não existem restrições que limitem o tamanho do *corpus* de referência.

Em nossa pesquisa utilizamos dois *corpora* de referência: o BNC (British National Corpus)⁸ e o Banco de Português⁹. O BNC, *corpus* fechado (1990-1994), possui 100 milhões de palavras e foi desenvolvido com o objetivo de ser representativo das variantes escrita e falada do inglês britânico. O Banco de Português é um *corpus* monitor do português do Brasil, ou seja, está aberto e é constantemente atualizado. Conta com aproximadamente um bilhão de palavras.

8. Metodologia

Para nossa pesquisa, utilizamos o *software WordSmith Tools* versão 5.0, desenvolvido por Mike Scott e publicado pela *Oxford University Press*. O programa possui três ferramentas principais *WordList*, *KeyWords* e *Concord* e uma série de aplicativos extremamente úteis para a análise linguística. Ressaltamos que o *WordSmith Tools 5* é um *software* riquíssimo para a análise linguística, sendo que cada uma de suas ferramentas possui vários instrumentos de análise. Contudo, descreveremos aqui somente as ferramentas e aplicativos que estão sendo utilizados em nossa pesquisa.

Para extrair os candidatos a termo, geramos uma lista de palavras-chave para o *corpus* de português, já que a extração terminológica foi realizada na direção português-inglês. Para tanto, comparamos a *wordlist* do nosso *corpus* com a *wordlist* do *corpus* de referência para obter as *keywords*.

⁸ Disponível em: < <http://corpus.byu.edu/bnc/> > Acesso em: 15 ago 2014.

⁹ Disponível em: <http://www2.lael.pucsp.br/corpora/bp/conc/> Acesso em: 15 ago 2014.

Figura 1: 1: lista de palavras do *corpus* de estudo; 2: lista de palavras do *corpus* de referência; 3: lista de palavras-chave.

Após obter as palavras-chave, geramos linhas de concordância para as palavras-chave e examinamos seus agrupamentos (*clusters*) a fim de encontrar colocações e unidades fraseológicas. Para tanto, utilizamos a ferramenta *Concord*, que gera listas das ocorrências de um item específico, chamado de ‘palavra de busca’ ou ‘nódulo’. Esse item pode ser formado por uma ou mais palavras e é apresentado com o contexto ao seu redor (Figura 2).

N	Concordance	Set	Tag	Word #	t. #	os. #	. #	os. #	t. #	os. #	File	%
1	Assist by Michael McIndoe. 50:00 GOAL - Clinton MorrisonCoventry 2 - 1			867	70	7%	0	8%	0	8%	nar-bbc_003.txt	51%
2	the assist for the goal. 90:00+2:06 GOAL - Anderson De SilvaDerby 1 - 3			160	5	6%	0	1%	0	1%	nar-bbc_007.txt	16%
3	for the goal came from Lee Croft. 26:06 GOAL - Rob HulseDerby 1 - 0 Barnsley			1.101	75	3%	0	6%	0	6%	nar-bbc_007.txt	78%
4	against league opposition. 4:06pm: GOAL! Blackpool 2 (David Vaughan 48)			1.586	78	0%	0	6%	0	6%	nar-gua_008.txt	66%
5	provided the assist for the goal. 52:07 GOAL - Gary HooperScunthorpe 3 - 0			857	66	7%	0	5%	0	5%	nar-bbc_002.txt	48%
6	twice at Bloomfield Road. 4:07pm: GOAL! Birmingham 0 West Ham 1			1.610	81	0%	0	7%	0	7%	nar-gua_008.txt	67%
7	Mutch 3) Nottm Forest 0! 3:09pm: GOAL! Sunderland 1 (Danny Welbeck 9)			831	44	0%	0	5%	0	5%	nar-gua_008.txt	36%
8	official got that just about right. 21:10 GOAL Game back on. Zoran Tomic pings			491	23	0%	0	2%	0	2%	tso-bbc_009.txt	14%
9	for it, opened the scoring with his 100th goal for the club, and was set up for a			159	1	2%	0	9%	0	9%	jo-if-ind_023.txt	28%
10	Rooney came close to his 100th goal in the 83rd minute, but Owen should			766	24	5%	0	5%	0	5%	jo-if-bbc_039.txt	95%
11	both started looking for their 100th goal in the Premier League and both			216	4	1%	2	3%	0	2%	jo-if-eye_005.txt	62%
12	Premier League. Rooney got the 100th goal of his career with a header to break			152	2	5%	0	9%	0	9%	jo-if-foo_032.txt	41%
13	it in the back of the net. 10MINS Goal! Clinton Morrison scores for			231	11	0%	0	1%	0	1%	-nar-foo_049.txt	93%
14	according to Soccer Saturday. 3:10pm: GOAL! Blackpool 1 (Neal Eardley 10)			850	46	0%	0	5%	0	5%	nar-gua_008.txt	37%
15	Ham 1 (Frederic Piquionne 48) 4:10pm: GOAL! Blackpool 2 Everton 2 (Seamus			1.621	82	0%	0	7%	0	7%	nar-gua_008.txt	68%
16	<p>Moumouni Dagano scored his 10th goal of the 2010 FIFA World Cup South			91	0	2%	0	1%	0	3%	jo-if-fif_001.txt	71%
17	Milan striker Adriano scored his 10th goal of the season to give Flamengo a			297	8	3%	0	5%	0	5%	jo-if-ft_008.txt	72%
18	remaining.</p><p>A 10th-minute goal by Clint Dempsey off a cross by			222	12	1%	4	1%	0	4%	jo-if-lat_001.txt	43%
19	Mark Hughes' side with a 10th-minute goal. </p><p>Jonathan Pitroipa, who			277	7	0%	3	6%	0	6%	jo-if-bbc_003.txt	43%
20	Player of the Year scored his 11th goal of the season eight minutes into the			131	0	1%	0	8%	0	8%	jo-if-ft_016.txt	61%

Figura 2: linhas de concordância para a palavra “goal”.

A figura 2 mostra as linhas de concordância de ‘goal’ ordenadas alfabeticamente pela primeira palavra à esquerda.

O próximo passo foi examinar as linhas de concordância e, quando necessário, gerar

clusters para essas linhas. A figura 3 mostra os *clusters* de ‘gol’:

N	CLUSTER	FREQ.	N	CLUSTER	FREQ.
1	GOL FEITO POR	561	11	ESQUERDA NO GOL	169
2	NO GOL FEITO	561	12	DE ESQUERDA NO	169
3	DIREITA NO GOL	391	13	DO GOL E	165
4	DE DIREITA NO	391	14	O SEGUNDO GOL	149
5	PARA O GOL	318	15	DEPOIS DO GOL	146
6	POR CIMA DO	230	16	O GOL DA	126
7	CIMA DO GOL	221	17	SAÍDA DEPOIS DO	123
8	DO GOL DE	196	18	COM GOL DE	109
9	O PRIMEIRO GOL	173	19	LONGE DO GOL	108
10	O GOL DE	173	20	SAI DO GOL	106

Figura 3: primeiros 20 *clusters* de “gol”

No ajuste realizado acima para ‘gol’, o programa foi preparado para encontrar *clusters* de três palavras, com frequência mínima de três e em uma janela de cinco palavras à direita e cinco à esquerda.

O último passo para extrair os candidatos a termo foi expandir as linhas de concordância das palavras-chave e acessar o texto integral da ocorrência, sempre que necessário, para melhor entender seu funcionamento e validá-las como termos.

Utilizamos procedimentos diferentes para o estabelecimento de equivalentes tradutórios de cada termo. A não padronização de uma metodologia deu-se pelo fato de que alguns termos apresentam um equivalente *prima-facie* em inglês, sempre mais simples de ser encontrado, ao passo que outros não apresentam esse tipo de equivalente e alguns sequer possuem equivalente. Em suma, geramos linhas de concordâncias para os termos e unidades fraseológicas especializadas (UFEs) em português, para melhor entender seu funcionamento, ou seja, o tipo de texto, a situação em campo e o contexto em que ocorrem. Após entender o funcionamento dos termos e UFEs em português, geramos linhas de concordância para suas traduções *prima facie* em inglês. Por exemplo, para chegar ao equivalente de “gol”, geramos linhas de concordância para *goal*. A figura 4 mostra as linhas de concordância do termo *goal*:

N	Concordance	Set	Tag	Word #	t. #	os.	. #	os.	. #	os.	. #	os.	File	%
1	Assist by Michael McIndoe. 50:00 GOAL - Clinton MorrisonCoventry 2 - 1			867	70	7%	0	8%	0	8%	0	8%	nar-bbc_003.txt	51%
2	the assist for the goal. 90:00+2:06 GOAL - Anderson De SilvaDerby 1 - 3			160	5	6%	0	1%	0	1%	0	1%	nar-bbc_007.txt	16%
3	for the goal came from Lee Croft. 26:06 GOAL - Rob HulseDerby 1 - 0 Barnsley			1.101	75	3%	0	6%	0	6%	0	6%	nar-bbc_007.txt	78%
4	against league opposition. 4:06pm GOAL! Blackpool 2 (David Vaughan 48)			1.586	78	0%	0	6%	0	6%	0	6%	nar-gua_008.txt	66%
5	provided the assist for the goal. 52:07 GOAL - Gary HooperScunthorpe 3 - 0			857	66	7%	0	5%	0	5%	0	5%	nar-bbc_002.txt	48%
6	twice at Bloomfield Road. 4:07pm GOAL! Birmingham 0 West Ham 1			1.610	81	0%	0	7%	0	7%	0	7%	nar-gua_008.txt	67%
7	Mutch 3) Nottm Forest 0 3:09pm GOAL! Sunderland 1 (Danny Welbeck 9)			831	44	0%	0	5%	0	5%	0	5%	nar-gua_008.txt	36%
8	official got that just about right. 21:10 GOAL Game back on. Zoran Tasic pings			491	23	0%	0	2%	0	2%	0	2%	tso-bbc_009.txt	14%
9	for it, opened the scoring with his 100th goal for the club, and was set up for a			159	1	2%	0	9%	0	9%	0	9%	jo-if-ind_023.txt	28%
10	Rooney came close to his 100th goal in the 83rd minute, but Owen should			766	24	5%	0	5%	0	5%	0	5%	jo-if-bbc_039.txt	95%
11	both started looking for their 100th goal in the Premier League and both			216	4	1%	2	3%	0	2%	0	2%	jo-if-eye_005.txt	62%
12	Premier League. Rooney got the 100th goal of his career with a header to break			152	2	5%	0	9%	0	9%	0	9%	jo-if-foo_032.txt	41%
13	it in the back of the net. 10MINS Goal! Clinton Morrison scores for			231	11	0%	0	1%	0	1%	0	1%	-nar-foo_049.txt	93%
14	according to Soccer Saturday. 3:10pm GOAL! Blackpool 1 (Neal Eardley 10)			850	46	0%	0	5%	0	5%	0	5%	nar-gua_008.txt	37%
15	Ham 1 (Frederic Piquionne 48) 4:10pm GOAL! Blackpool 2 Everton 2 (Seamus			1.621	82	0%	0	7%	0	7%	0	7%	nar-gua_008.txt	68%
16	<p>Mourmouni Dagnano scored his 10th goal of the 2010 FIFA World Cup South			91	0	2%	0	1%	0	3%	0	3%	jo-if-fif_001.txt	71%
17	Milan striker Adriano scored his 10th goal of the season to give Flamengo a			297	8	3%	0	5%	0	5%	0	5%	jo-if-ft_008.txt	72%
18	remaining.</p><p>A 10th-minute goal by Clint Dempsey off a cross by			222	12	1%	4	1%	0	4%	0	4%	jo-if-lat_001.txt	43%
19	Mark Hughes' side with a 10th-minute goal.</p><p>Jonathan Pitroipa, who			277	7	0%	3	6%	0	6%	0	6%	jo-if-bbc_003.txt	43%
20	Player of the Year scored his 11th goal of the season eight minutes into the			131	0	1%	0	8%	0	8%	0	8%	jo-if-ft_016.txt	61%

Figura 4: linhas de concordância de *goal*

Esse tipo de pesquisa nos permitiu validar *goal* como equivalente de “gol” uma vez que observamos que o contexto em que as duas palavras ocorrem é o mesmo. No entanto, não se mostrou eficiente para a validação de alguns equivalentes como, por exemplo, o equivalente da unidade fraseológica “fazer um gol”, já que não encontramos um verbo que co-ocorre com *goal* com frequência relativamente alta. Por esse motivo, resolvemos gerar linhas de concordância para “scored”, tradução *prima facie* de ‘marcar’) a fim de verificar se o termo co-ocorre com ‘gol’. Observemos a figura 5:

N	Concordance	Set	Tag	Word #	t. #	os.	. #	os.	. #	os.	. #	os.	File	%
1	blocked, but it ran to Anderson, who scored with impunity. Redknapp			751	36	5%	0	4%	0	4%	0	4%	jo-if-day_049.txt	85%
2	then provided the assist for Anelka who scored Chelsea's second with a low			174	3	8%	0	8%	0	8%	0	8%	jo-if-bbc_038.txt	36%
3	the Togo striker signed from Arsenal, who scored a dazzling opener after three			287	8	7%	0	7%	0	7%	0	7%	jo-if-ft_002.txt	73%
4	because he was on a plane. Ballack, who scored with a header either side of			434	19	9%	0	2%	0	2%	0	2%	jo-if-gua_077.txt	68%
5	got anywhere near Sébastien Bassong, who scored the first goal of his career on			395	10	8%	0	7%	0	7%	0	7%	jo-if-ind_001.txt	43%
6	is working just as hard on Darren Bent, who scored with a header in the fifth			150	0	2%	0	6%	0	6%	0	6%	jo-if-tel_027.txt	61%
7	PENALTY SAVED! (58 mins) Cardozo, who scored in the shoot-out against			3.005	110	9%	0	0%	0	0%	0	0%	nar-gua_016.txt	72%
8	Amkar Perm. Martin Kushev (centre), who scored Amkar's goal, bursts through			129	1	6%	0	3%	0	3%	0	3%	jo-if-bbc_056.txt	32%
9	half-time break further ahead. Elliott, who scored Burnley's play-off final winner			452	14	9%	0	3%	0	3%	0	3%	jo-if-bbc_028.txt	67%
10	especially for Omar Bravo, the forward who scored Mexico's first two goals. The			201	4	5%	0	1%	0	1%	0	1%	jo-if-ny_035.txt	26%
11	who beat Mozambique 2-0.</p><p>Gabon, who scored an upset victory in Morocco			477	18	8%	10	7%	0	5%	0	5%	jo-if-gua_013.txt	96%
12	led by the prolific Theofanis Gekas, who scored ten goals during			1.321	65	1%	0	7%	0	7%	0	7%	-nar-foo_002.txt	97%
13	claim, picked out Steven Gerrard, who scored by heading back across the			374	16	6%	0	1%	0	1%	0	1%	jo-if-gua_096.txt	49%
14	can celebrate," defender Fabio Grosso, who scored the go-ahead goal against			619	16	8%	11	6%	0	7%	0	7%	jo-if-wp_045.txt	69%
15	first one to point out that it's Holland who scored, not Uruguay as it says at			2.026	83	4%	0	5%	0	5%	0	5%	nar-gua_017.txt	46%
16	start. The Poland international, who scored the first goal of his loan spell			320	10	1%	0	5%	0	5%	0	5%	jo-if-365_035.txt	52%
17	start. The Poland international, who scored the first goal of his loan spell			320	10	1%	0	5%	0	5%	0	5%	jo-if-365_037.txt	52%
18	will know, was the Haitian immigrant who scored the goal in the USA's			181	4	0%	0	4%	0	4%	0	4%	nar-gua_031.txt	6%
19	will know, was the Haitian immigrant who scored the goal in the USA's			211	4	0%	0	5%	0	5%	0	5%	nar-gua_003.txt	6%
20	will know, was the Haitian immigrant who scored the goal in the USA's			210	4	0%	0	4%	0	4%	0	4%	nar-gua_008.txt	6%

Figura 5: parte das linhas de concordância de *scored* ordenadas pela duas primeiras palavras à esquerda.

Em um primeiro momento, fomos levados a validar *score* ou *score a goal* como

equivalente de “fazer/marcar um gol”, o que não está errado. Entretanto, a diferença entre a frequência da UFE em português “marcar/fazer um gol” e da UFE em inglês “score a goal” era muito grande. Decidimos, então, conduzir buscas para todos os verbos do *corpus* utilizando as seguintes etiquetas morfossintáticas: VV (verbo no infinitivo), VVD (verbo no passado simples), VVG (verbo no gerúndio), VVN (verbo no particípio passado) e VVZ (verbo no presente – 3ª pessoa do singular) a fim de encontrar outros verbos que descrevessem o ato de fazer um gol.

A etiquetagem morfossintática, ou *part of speech tagging* (POS), consiste em colocar em cada palavra do *corpus* uma etiqueta que indique sua classe gramatical. A etiquetagem foi realizada com o etiquetador *Tree-Tagger*, disponibilizado pelo professor Tony Berber Sardinha, da Pontifícia Universidade Católica de São Paulo (PUC), no site do LAEL (Programa de Linguística Aplicada e Estudos da Linguagem)¹⁰.

Os resultados dessas buscas serão detalhados mais adiante no estudo de caso. Ressaltamos que priorizamos os *corpora* para a extração e validação dos termos e UFEs, bem como para a busca de equivalentes. No entanto, sempre que necessário, recorreremos à internet quando não encontramos o equivalente no *corpus* ou quando precisamos esclarecer a origem ou o significado de algum termo.

9. Estudo de caso: FAZER¹¹ um gol

Ao analisar as linhas de concordância de “gol” em nosso *corpus* de português, encontramos, ao total, 920 ocorrências de **FAZER|MARCAR [o|um] gol / MARCAR o tento / FINALIZAR / EMPATAR**, termos e unidades fraseológicas utilizados para descrever um gol.

Ao buscar pelos possíveis equivalentes em inglês, encontramos 259 ocorrências de:

- SCORE({a | the} goal)
- SNATCH [a goal]
- DELIVER a goal
- ADD {a|the} ORDINAL (goal)
- BLAST in ([a goal])

A primeira pergunta que nos ocorreu foi: por que temos mais ocorrências de “gol” em

¹⁰ disponível em: <<http://corpuslg.org/tools/etiquetagem/>>

¹¹ utilizamos letras maiúsculas para indicar a forma lematizada dos verbos

português do que de *goal* em inglês, já que o número de gols da Série A do Campeonato Brasileiro é quase o mesmo da *Premier League*? A figura 6 mostra o número de gols do Campeonato Brasileiro de 2006 a 2013 e o número de gols das temporadas 2011/2012, 2012/2013 e 2013/2014 da *Premier League*:

Confira a média e o total de gols do Brasileirão de 2006 para cá:				
2006 (2,71)	1.030 gols	58	1948/1949	1303
2007 (2,76)	1.047 gols	59	2011/2012	1066
2008 (2,72)	1.035 gols	60	1899/1900	856
2009 (2,88)	1.094 gols	61	2010/2011	1063
2010 (2,57)	978 gols	62	2012/2013	1063
2011 (2,68)	1.017gols	63	1900/1901	855
2012 (2,47)	940 gols	64	1999/2000	1060
2013 (2,46)	936 gols	65	1984/1985	1288
		66	1985/1986	1288
		67	1911/1912	1057
		68	2009/2010	1053
		69	2013/2014	1052
		70	1920/1921	1276

Figura 6: número de gols do Campeonato Brasileiro de 2006 a 2013¹² e número de gols da *Premier League* das temporadas de 2011/2012, 2012/2013 e 2013/2014¹³

Como podemos observar, o Campeonato Brasileiro de 2013 teve 936 gols e a *Premier League* de 2013/2014 teve 1052 gols. Nosso segundo questionamento foi: se *goal* não descreve os gols marcados, como esses gols estão sendo narrados? Nesse momento, nos demos conta de que o verbo *score*, presente na unidade fraseológica *SCORE* [*a | the* goal], muitas vezes ocorria sozinho, observemos os exemplos que seguem:

- a) The United States got out to an early lead as Clint Dempsey **scored** just 32 seconds into the game for the fastest goal in World Cup history for the U.S.
- b) Striker Asamoah Gyan also **scored** right before the first half ended.

Mais adiante, ao ler os textos que coletávamos para compilar o *corpus*, nos demos conta de que a língua inglesa utiliza outros verbos que, muito frequentemente, não ocorrem com o termo “gol”, ou seja, são usados como sinônimos de *SCORE* [*a | the* goal]. Sendo assim, geramos linhas de concordância para as etiquetas de verbo do *corpus* em inglês

¹² Disponível em: < <http://www.srgooool.com.br/Noticia/Brasileirao-2013-teve-o-menor-numero-de-gols-e-a-pior-media-de-tentos-dos-pontos-corridos> > Acesso: 16 jun 2014.

¹³ Disponível em: <<http://www.worldfootball.net/stats/eng-premier-league/1/>> Acesso: 16 jun 2014.

e analisamos todas as ocorrências à procura de verbos que descrevessem um gol. A tabela 1 mostra o número de linhas de concordância para cada etiqueta do *corpus*:

Tabela 1: Número de linhas de concordância para cada etiqueta gramatical

etiqueta	número de ocorrências
VB - Verb, base form	25.283
VBD - Verb, past tense	28.778
VBG - Verb, gerund or present participle	19.406
VBN – verb, past participle	20.457
VBZ - Verb, 3rd person singular present	13.853

Aplicamos o dispositivo *Re-sort*, que permite ordenar as linhas de concordância pelo nóculo de busca, e ajustamos as configurações para que as linhas fossem ordenadas pela etiqueta gramatical e pela primeira palavra à direita (forma canônica do verbo) e à esquerda.

Observamos todas as linhas de concordância e identificamos 43 verbos que poderiam fazer parte de UFEs utilizadas para relatar um gol. Depois, geramos concordâncias no *corpus* sem etiquetas gramaticais para esses verbos a fim verificar como ocorrem nos textos, ou seja, os padrões em que ocorrem.

O passo seguinte foi validar essas unidades fraseológicas por meio da análise das concordâncias e do contexto expandido. Ainda no caso de *fire*, ampliamos os contextos das linhas em que *fire home* era um possível equivalente de “fazer um gol”. Observemos, por exemplo, a ocorrência número 153:

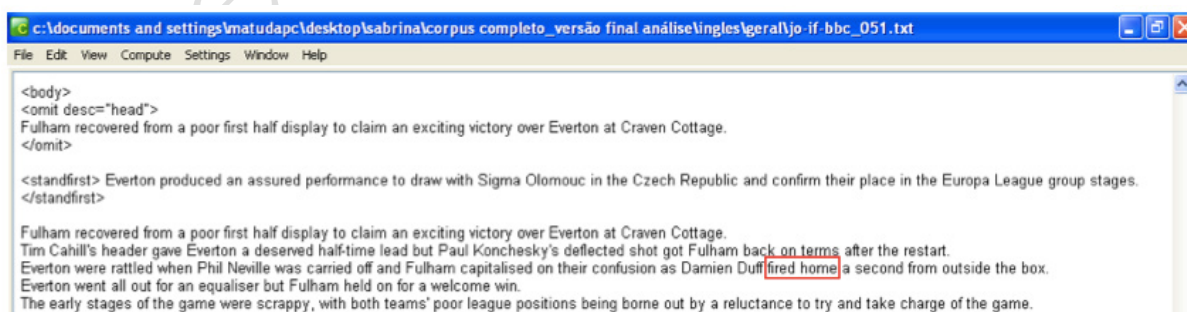


Figura 7: Parte do contexto expandido da linha de concordância 153 de *fire**

Após a validação das unidades fraseológicas, agrupamos as UFEs por categoria semântica tendo como base o *Roget's International Thesaurus*. Optamos por esse tipo de

classificação porque, durante a análise, notamos que o significado de algumas UFEs é bastante similar e que o sentido que carregam está relacionado ao sentido do verbo na linguagem geral..

Apresentaremos, nos parágrafos que seguem, os equivalentes por categoria semântica. Embora intercambiáveis, todos apresentam características particulares que denotam categoria semântica, prosódia semântica e *lexical priming*¹⁴ distintos, causando, dessa forma, reações diferentes no leitor.

9.1 Pontuação

As unidades fraseológicas **SCORE** (*{a|the} goal*) e **NOTCH** (*{his ORDINAL|CARDINAL}{goal}*) compõem a categoria semântica de pontuação. Ambas são utilizadas para pontuar os gols e seus significados recaem sobre o ato de adicionar gols ao placar de uma partida.

Os exemplos abaixo ilustram o uso de *notch* e *score* seguido do número de gols, enfatizando a ideia de soma de pontos:

Rooney **notched his fifth goal** of the season and showed that United are serious contenders to retain their Premier League crown despite the loss of Cristiano Ronaldo during the summer.

But Eduardo still had time **to score the sixth**, netting from close in after Andrey Arshavin's shot had come back off a post with two minutes to go.

9.2 Ação violenta

Durante o agrupamento semântico, identificamos 13 UFEs que transmitem a ideia de uma ação violenta. São elas:

- a) FIRE {home | in (a goal) | (the ball) past [goalkeeper]} into the net} ¹⁵
- b) HAMMER {in | home}

¹⁴ Teoria linguística desenvolvida por Michael Hoey (2005) que parte do princípio de que à medida que temos contato com uma nova unidade lexical, seja verbo, substantivo, adjetivo etc., essa nova unidade é adquirida juntamente com os contextos linguísticos, sociais, culturais em que frequentemente ocorre. Dessa forma, ao reproduzir a unidade lexical tendemos a utilizá-la nos mesmos contextos em que a encontramos. Toda a informação adquirida por meio da unidade lexical em questão é considerada o *lexical priming* da palavra.

¹⁵ {} elenco de opções; | combinações possíveis; () uso opcional; [] hiperônimo.

- c) SLAM {(the ball) past [goalkeeper] | in | into the roof of the net}
- d) BLAST IN (a goal)
- e) LASH home {(a goal) | into an empty net}
- f) SMASH home
- g) STAB the ball {home | past [goalkeeper]}
- h) POKE {theballhome|homepast[goalkeeper]}
- i) KNOCK {the ball {into the net | in | home}} / NUMBER goals
- j) THUMP {home | past [goalkeeper]}
- k) CLIP the ball into the net
- l) SNATCH [a goal]
- m) THUNDER a goal home

A alta ocorrência de verbos que transmitem a ideia de ação violenta em inglês nos levou à seguinte pergunta: será que o uso de verbos dessa categoria semântica também é comum em português?

Para responder a pergunta, geramos linhas de concordância para a etiqueta gramatical V¹⁶, que indica um verbo, no *corpus* de português e observamos as ocorrências.

As concordâncias nos mostram que em português não utilizamos verbos de ação violenta com a mesma frequência que em inglês. Encontramos 222 ocorrências do verbo “empurrar”, sendo 87 sinônimas de “fazer um gol”. No entanto, em grande parte das ocorrências, o verbo “empurrar” estava relacionado à facilidade com que o gol foi feito, e não à intensidade do chute dado. Observemos o exemplo abaixo:

Em desvantagem o Corinthians tentou pressionar, enquanto o Santos tentava concluir os bons contra-ataques que estava conseguindo. Mas aos 34 minutos Bill aproveitou bola rebatida na trave e **empurrou** para o gol, empatando o jogo no Pacaembu.

Também encontramos 6.000 ocorrências do verbo “bater”, das quais apenas 28 utilizadas para descrever um gol. Ao analisar as linhas de concordância, constatamos que das 28 UFEs com o verbo “bater”, somente 16 narram um gol feito. As outras representam tentativas que não deram certo:

Jorge Wagner domina e **bate** para o gol. Mas a bola vai fraca e sem direção. MUCHA SALVA!! Robben faz bela jogada pela esquerda e cruza rasteiro. Mathijsen **bate** para o gol e bola explode no rosto de Mucha.

¹⁶ As etiquetas gramaticais em português não fazem distinção entre os tempos verbais.

Após analisar cuidadosamente as linhas de concordância, expandindo o contexto e visitando o texto integral, concluímos que a diferença na frequência dos verbos de ação violenta é resultado da própria forma de jogar futebol no Brasil e na Inglaterra. Em outras palavras, se os *corpora* são realmente representativos, eles devem refletir, por meio do léxico, a forma como o futebol é jogado em diferentes culturas.

Se perguntarmos para qualquer pessoa que goste, acompanhe ou entenda, pelo menos um pouco, de futebol, sobre as diferenças entre o futebol brasileiro e o futebol inglês, as respostas serão muito parecidas: o futebol inglês é mais rápido e mais “pegado”, mais “forte”. Observemos o depoimento de Ramirez, volante do Chelsea, em entrevista ao site da FIFA em outubro de 2010 sobre o futebol inglês:

“O futebol inglês é muito físico e rápido. Aqui você não tem muito tempo para ficar com a bola no pé e tem que definir logo a jogada. Além de procurar as jogadas quando tenho a bola, eu sempre marquei forte. Por isso, não estou encontrando muitas dificuldades. Já estou me acostumando ao ritmo das partidas aqui.¹⁷”

Se tomarmos o depoimento acima, podemos caracterizar o futebol inglês como um futebol que tem a força como um de seus elementos centrais. Por esse motivo, tratamos a discrepância da frequência de verbos de ação violenta em inglês e português por meio do conceito forma-representação de Toledo (2002).

Retomando Toledo (2002), o conjunto de regras não determina nem influência as maneiras de jogar. Na verdade, a interpretação e apropriação cultural que cada região faz das regras é que revelam as “formas de jogo”. Justapostas, as regras e as “formas de jogo” dão origem às representações, ou seja, o ajustamento da observação empírica das “formas de jogo” em um plano simbólico que, por sua vez, consolidam as “escolas”, “jeitos” ou “estilos” próprios que se traduzem no código linguístico.

Investigar a apropriação cultural das regras no Brasil e na Inglaterra foge ao escopo deste artigo estudo. Contudo, reconhecer que essa interpretação influencia o “estilo” de jogo e, por conseguinte, a terminologia utilizada nas duas culturas é essencial para entendermos algumas discrepâncias linguísticas. Assim, o fato de o futebol inglês ser mais “pegado” transparece na grande quantidade de verbos de ação violenta.

¹⁷ Disponível em: <<http://www.guiame.com.br/v4/70368-1663-Destaque-no-site-da-Fifa-Ramires-diz-estar-adaptado-ao-futebol-ingl-s.html>>. Acesso em: 12 abr. 2011.

Para comprovar nossa hipótese, compilamos um *mini-corpus* de, aproximadamente, 1000 palavras em cada língua, composto por 15 textos sobre o estilo do futebol brasileiro e 15 textos sobre o estilo do futebol inglês. Utilizamos *football*, *Brazil*, *style* e *school* como palavras de busca no *Google* para encontrar textos sobre o estilo do futebol brasileiro e *football*, *England*, *UK*, *school* e *style* para encontrar textos sobre o estilo do futebol inglês. Esse *mini-corpus* é monolíngue, coletamos textos somente em inglês uma vez que o *corpus* foi compilado para comprovar a hipótese que seria apresentada em um congresso na Inglaterra.

A análise foi feita de forma bastante simples. Primeiramente, geramos listas de palavras dos *mini-corpora*. Em seguida, comparamos essas listas com listas de palavras de *corpora* de referência a fim de obter as palavras-chave. Como *corpus* de referência, utilizamos um *corpus* jornalístico, com aproximadamente de 100.000 palavras. Em seguida, analisamos as palavras-chave e selecionamos aquelas utilizadas para descrever as características dos dois futebolis. O quadro 2 mostra as *keywords* associadas às duas escolas:

Quadro 2: palavras-chave associadas ao estilo brasileiro e ao estilo inglês

Palavras-chave associadas ao estilo brasileiro	Palavras-chave associadas ao estilo inglês
<ul style="list-style-type: none"> - [good] dribblers - creative - spontaneous - improvisational - free-flowing - possession - samba [beat] - control [of the ball] - [highly] skilled - [flowing] passing [game] - beautiful - flair - invention 	<ul style="list-style-type: none"> - physical [prowess/strength] - fast - athleticism - rough - [long] shoots - quick - [direct] passes - [no] risk - courageous - hard - [break-neck] speed

Ao fim da análise, nos questionamos se a frequência de verbos que denotam ação violenta tende a aumentar em português, uma vez que a prática e o aprendizado do futebol mudou bastante no Brasil. Antigamente as crianças jogavam futebol na rua, no pátio da escola e em “campinhos”. Hoje em dia, as crianças frequentam, desde muito cedo, escolinhas de futebol, ambientes em que meninos e meninas são treinados, por meio de acompanhamento

tático, prática de jogadas e, muitas vezes, condicionamento físico. É comum ouvir de comentaristas e jornalistas esportivos que o “estilo” do futebol brasileiro, sempre caracterizado por jogadas de habilidade, está passando por um processo de transformação: a troca do “futebol arte” pelo “futebol força” ou, até mesmo, a junção desses dois futebolis. De qualquer forma, cremos que essas mudanças, futuramente, terão reflexo na constituição do léxico do futebol em português, incorporando neologismos. No entanto, embora as mudanças venham acontecendo há um tempo, seu reflexo no sistema linguístico, muito provavelmente, ocorrerá de forma mais lenta e, até o momento, o “futebol força” não é a “forma de jogo” que temos em nosso plano simbólico ao pensar no futebol brasileiro e, por esse motivo, não é traduzido em nosso código linguístico.

9.3 Facilidade

Outro campo semântico bastante comum em inglês, e que não foi encontrado em português, pelo menos não com frequência significativa, é o de “facilidade”. Observemos as seguintes unidades fraseológicas

- a) SLIDE {home his ORDINAL (goal) | the ball home};
- b) SWEEP {home | the ball into the net};
- c) DINK in [a goal];
- d) TUCK home (a goal);
- e) STROKE a shot past [goalkeeper].

Nelas, o sentido principal recai sobre a facilidade com que o gol foi feito. Na realidade, a facilidade não caracteriza o gol, mas sim a trajetória da bola, a posição do jogador ou o toque que o jogador dá na bola.

9.4 Velocidade

As UFEs **TAP** {in|home} e **FIZZ** {the ball past goalkeeper|a sidefooter at goal} compõem a categoria semântica de velocidade. Ambas enfatizam a velocidade do gol, não somente a velocidade com que o jogador recebe e chuta ao gol, mas também a rapidez dos toques até o momento do gol:

Scott Carson, oh Scott Carson. Drogba lines up the shot, **fizzes** a sidefooter at

goal, but it's straight at the keeper. Who proceeds to let the ball squirm from his grasp, drop to Mikel who dinks it square for Malouda to **tap in**.

Novamente, se tomarmos os depoimentos dos jogadores, jornalistas e técnicos sobre o futebol inglês, veremos que a alta ocorrência desses verbos de velocidade nos periódicos ingleses reflete o “estilo” do futebol inglês. Retomemos o depoimento de Ramirez: “O futebol inglês é muito físico e rápido. Aqui você não tem muito tempo para ficar com a bola no pé e tem que definir logo a jogada...”.

Sabemos que o português, assim como o inglês, utiliza palavras de outras categorias gramaticais para indicar a rapidez e velocidade de uma jogada como, por exemplo, os termos “rápido” e “de primeira”. No entanto, chamamos atenção para o fato de que, pelo menos em nosso *corpus*, o jogo rápido se mostra tão característico no futebol inglês que existe uma necessidade de expressá-lo por meio de verbos específicos.

Algumas UFEs, não descrevem nenhuma particularidade do estilo do futebol inglês, mas refletem, de certa forma, uma característica da cultura e, conseqüentemente, da língua inglesa: um alto grau de detalhamento nas informações.

Para entender essa particularidade cultural, recorreremos aos conceitos de *high-context culture* e *low-context culture* do antropólogo Edward Hall (1976). A teoria de Hall parte do pressuposto de que a quantidade de informação linguística e contextual necessária para transmitir o significado varia de acordo com a cultura e, para descrever as culturas, o autor cria dois grupos: *high-context culture*, formado por países como Brasil, Itália, Grécia e países da África, Ásia e América Latina; e *low-context culture*, formado por países como Estados Unidos, Alemanha, Suíça e outros países da Europa ocidental e do Reino Unido.

As culturas que pertencem ao primeiro grupo, as *high-context cultures*, são caracterizadas por recorrerem a uma grande quantidade de elementos contextuais, os quais auxiliam seus integrantes a entenderem as regras, ou seja, o modo como a cultura funciona. Por esse motivo, parte da informação aparece de forma implícita, é o dito pelo não dito, informações que estão subentendidas na fala ou, em nosso caso, em um texto, e podem contribuir muito mais para a transmissão do significado do que aparentam. Essas culturas possuem um forte senso de comunidade, fator que implica em diferentes estilos comportamentais (MANCA, 2010, p. 373), que são dados como subentendidos e, raramente, aparecem de forma explícita.

As culturas que se enquadram no segundo grupo, *low-context*, são caracterizadas por

explicitar os elementos contextuais. Quase nada é subentendido, os assuntos são debatidos exaustivamente de modo que não restem dúvidas ou margem para uma interpretação equivocada. A comunicação é feita da forma mais explícita possível e é realizada por meio da transmissão de fatos, sem a expressão de sentimentos.

A teoria de Hall é bastante complexa e detalhada. Contudo, pode ser utilizada para entender o processo de produção de textos sobre futebol nas duas culturas.

9.5 Detalhamento

As UFEs que apresentaremos nos próximos parágrafos transmitem um detalhamento que não encontramos nos verbos e UFEs utilizados para descrever um gol em português. As UFEs do primeiro grupo expressam o movimento que a bola faz para dentro do gol:

a) **SLOT {home([agoal])|(theball)into{an|the}emptynet|past[goalkeeper]| in}**
Robinho, who was the game's outstanding player, then provided one of the assists of the championship so far as he threaded a delightful ball to Elano in the 72nd minute before the midfielder **slotted home** a second.

b) **DRILL home**
The defender, 25 yesterday, settled a tight game and embarked on a celebration that would have looked more at home on Strictly Come Dancing than Match of the Day. He **drilled home** Swansea's 69th-minute winner after Leon Britton had rattled the home side's bar with a stunning overhead kick.

c) **BURY the ball past [goalkeeper]**
For good measure, Fabregas popped in a peach for his second and Arsenal's fifth as he ran up field, picked his spot and **buried the ball past** Howard.

No caso desses verbos, só entendemos o funcionamento das UFEs, que utilizam *drill*, *bury* e *slot* de forma figurativa, pelos seus significados na língua geral.

9.6 Realização

As UFEs **FINISH (off)** e **DELIVER a goal** expressam o sentido de ‘completar’, ‘realizar uma tarefa’, assim como o verbo “finalizar” em português:

a) And 43 minutes into his full debut, Stanislas **finished off** a great end-to-end move by tapping in Luis Boa Morte's cross.

b) It's been 25 years since a match between these two teams **delivered a goal** and it doesn't look like we're going to get one tonight.

c) O goleiro Felipe nem se mexeu. Novamente impaciente com o atacante Souza, a torcida do Corinthians passou a pedir a entrada de Dentinho. Só parou quando Elias, aos 42, **finalizou** de fora da área no ângulo: 2 a 1.

9.7 Movimento da bola

Também detectamos a presença de duas UFEs que descrevem o movimento da bola ou seu controle por um jogador:

a) **HOOK in ([a goal])**

When defender Marius Zaliukas **hooked in** the second shortly after the break, belief they were capable of a sensational result coursed through the home ranks.

b) **STEER the ball past [goalkeeper]**

His pass picks out the run of Osman who **steers the ball past** Schwarzer.

Analisamos as ocorrências de *HOOK in ([a goal])* no *corpus* e chegamos a duas possibilidades do uso. Primeiramente, pensamos que a UFE poderia descrever o tipo de jogada realizada para fazer o gol, já que *hook in* descreve uma jogada feita com o pé no ar, a meia altura, conhecida como “voleio” em português, fato que foi confirmado após a observação de outras ocorrências no Google.uk. Nossa segunda hipótese era que a UFE seria utilizada como o termo *hook* de uso corrente no *cricket* e no golfe, cujo significado indica um modo de bater na bola que faz com que ela desenhe uma curva em vez de ir reto. Confirmar essa segunda hipótese foi um pouco mais complicado do que a primeira, pois se o verbo *hook* designa, em seu significado, o movimento da bola, o texto, muito provavelmente, não usa nenhum outro recurso linguístico que possamos analisar para validar nossa hipótese. O detalhamento das notícias em inglês foi de grande valia para confirmar nossa ideia. Acessamos o site Football.co.uk, do qual um dos textos que contém a UFE foi coletado, vimos as fotos dos gols do jogo e clicamos em um link do YouTube que direcionava para um vídeo que mostrava os gols da partida.

A última UFE, ***STEER the ball past [goalkeeper]***, não descreve o movimento da bola, como a anterior. Seu significado recai sobre o controle da direção que a bola toma, como no exemplo em que o jogador direciona a bola e chuta para o gol:

His pass picks out the run of Osman who steers the ball past Schwarzer.

9.8 Gols de cabeça

Identificamos três UFEs utilizadas para narrar gols de cabeça:

- a) **NOD {home (into an empty goal) | in | into the net};**
- b) **HEAD (the ball) {home|past (goalkeeper)|into the (roof of the) net};**
- c) **{HAMMER | POUND} a header into the net.**

A alta frequência dos verbos que narram gols de cabeça em inglês nos chamou a atenção quando comparada à frequência de “cabecear”, “bater de cabeça” e “tentar de cabeça”, 1037 e 387 ocorrências em cada língua, respectivamente.

Recorremos novamente ao conceito forma-representação de Toledo para entender essa discrepância. Tomemos o depoimento de Robinho, em sua apresentação ao Milan, sobre o futebol inglês:

“Não tive nenhum problema com o Roberto Mancini. O que acontece é que o futebol inglês não é muito bom para jogador brasileiro, muita bola alta e a gente gosta de jogar com ela no chão¹⁸.”

Jogadas aéreas, principalmente de bola na área, definem um estilo bem inglês, uma vez que exigem a força, a impulsão e a altura características dos jogadores ingleses. Desse estilo, e também das condições climáticas da Inglaterra -- devido à chuva constante as equipes eram forçadas a mandar a bola para o alto, já que era quase impossível fazê-la correr na grama encharcada -- surgiu a expressão “chuveirinho”, para designar a grande quantidade de jogadas aéreas na área, que caracterizou, por muitos anos, a escola inglesa.

O cabeceio, que definia o “estilo” inglês, era considerado como a melhor forma de atacar e de concluir uma jogada, levando a seleção da Inglaterra à conquista da Copa do Mundo de 1966, mas perdendo sua força nos anos seguintes.

Não nos cabe, aqui, discutir a origem e a eficácia do “estilo” inglês, mas sim reconhecer que o papel das jogadas aéreas no futebol inglês é fundamental para entender a

¹⁸ Disponível em: <http://opiodopovo.wordpress.com/tab/chuveirinho/> Acesso em: 21 mai. 2011.

divergência entre a frequência de UFEs que narram gols de cabeça em inglês e em português.

9.9 Adição

A estrutura de *ADD {a|the} ORDINAL (goal)* pode ser enquadrada na categoria semântica de adição, pois a referência ao número do gol que foi marcado é parte integrante da UFE.

a) Rumbings of discontent were just beginning to emerge as Arsenal took the lead through Abou Diaby in the 18th minute and the midfielder **added a second** three minutes later, prompting Wenger to compare his midfielder with a Highbury great.

9.10 Empate

Identificamos os verbos *EQUALISE*, *DRAW* e *TIE* como equivalentes de EMPATAR. Entretanto, cada um apresenta suas particularidades. O verbo *EQUALISE* é o mais comum para narrar um gol de empate. Os trechos que seguem exemplificam seus principais usos:

Martinez`s lads were 13 minutes from victory until Clint Dempsey **equalised** with a diving header and Zamora had the last word with his bonce on 81 minutes in this gripping fifth-round replay.

Os verbos *TIE* e *DRAW*, outros possíveis equivalentes do verbo “empatar” em português, também ocorrem no *corpus*. Entretanto, o primeiro ocorre majoritariamente em jornais e revistas americanos e o segundo ocorre para narrar um jogo que empatou, e não um gol de empate.

The Cottagers beat Arsenal 1-0 in August, **drew** 2-2 with Chelsea over Christmas and stunned United 2-0 last month.

9.11 Posicionamento

A UFE *NESTLE {in | into} the (back of the) net* nos dá a idéia de acomodar a bola no gol:

Not a brilliant penalty-struck low and only a few feet to the left of the centre of the goal, but Julio Cesar went the other way and the ball **nestled in the back of the net**.

Como podemos observar, o exemplo ilustra a posição da bola após o gol.

9.12 Realização

Identificamos três UFEs que pertencem à categoria de realização, entendida aqui como consecução, ou seja, o ato de alcançar um objetivo. Observemos os exemplos:

a) **NET [a goal]**

Liverpool's season is firmly back on track after Yossi Benayoun **netted** a hat-trick in Saturday's 4-0 Premier League home win over Burnley. Substitute Stefan Maierhofer **netted** a late consolation for Wolves.

b) **BAG [a goal]**

Dave Kitson is finally beginning to find his feet at Stoke as he bagged the winner in a 1-0 victory over Sunderland at the Britannia Stadium. Drogba was a real menace throughout, heading just over from a Jose Bosingwa cross before **bagging** the all-important goal.

c) **NICK [a goal]**

Birmingham City had the early Premier League leaders begging for the final whistle before Aaron Lennon **nicked** a winner for Tottenham deep into stoppage time. Lions fans were already leaving the Den when Paul Shaw **nicked** a consolation goal on 79 minutes but the roar that went up when Tim Cahill (right) made it 3-2 within seconds brought many back.

Independentemente do significado de seus verbos na linguagem geral, as três UFEs descrevem um ato de realização. Quando utilizadas, enfatizam a conquista de um gol que é necessário ou importante. Se tomarmos como exemplos os colocados da UFE nos exemplos acima, *hat-trick* (três gols marcados por um jogador na mesma partida), *consolation* (gol de honra), *the winner* (gol da vitória) e *the-all-important goal* (gol importante, que pode ser decisivo para a vitória ou para a permanência de um time em um campeonato), veremos que além de descrever um gol feito, a UFE realça a importância do gol, tanto pelos seus colocados como pelo seu uso em detrimento de algum outro equivalente.

10. Considerações Finais

Ao final de nosso estudo de caso, verificamos que a LC é fundamental para expandir o escopo de pesquisas terminológicas e contribui, de forma significativa, para a identificação de aspectos culturais de uma área de especialidade.

A existência de UFEs que expressam o sentido de violência, explosão, rapidez, facilidade e gols de cabeça em inglês e a ausência do mesmo tipo de UFE em português pode ser explicada pelo conceito forma-representação, do mesmo modo que a maior variação de UFEs que descrevem a trajetória da bola, tanto em um gol quanto em uma tentativa, pode ser explicada pelo conceito de *high* e *low-context culture*. O torcedor brasileiro (*high-context*) espera um relato mais breve da partida e ênfase no produto final, o gol. Já o torcedor inglês (*low-context*) espera que os lances sejam narrados de forma precisa, com ênfase nos meios, ou seja, nas jogadas, que levaram ao gol.

Dessa forma, acreditamos que a Terminologia não é uma atividade prescritiva, na qual os termos devem ser normatizados a fim de garantir a eficácia de uma comunicação especializada. Ao contrário, os fatores culturais, o contexto e a situação e a finalidade de uso influenciam de forma direta o funcionamento das terminologias; por esse motivo, o fazer terminológico, principalmente o bilingue, deve considerar todos esses elementos na compilação de obras terminográficas.

Referências bibliográficas

- AIJMER, K.; BENGT, A. **English Corpus Linguistics**. New York: Longman, 1991.
- AZENHA JR., J. **Tradução técnica e condicionantes culturais**: primeiros passos para um estudo integrado. São Paulo: Humanitas/FFLCH-USP, 1999.
- BERBER SARDINHA, A. Lingüística de Corpus: Histórico e problemática. **Revista D.E.L.T.A.**, São Paulo, v. 16, n. 2, p. 323-367, 2000.
- _____. **Lingüística de Corpus**. Barueri: Manole, 2004.
- BOWKER, L.; PEARSON, J. **Working with Specialized Language**. A Practical Guide to Using Corpora. London/New York: Routledge, 2002. **crossref**
<http://dx.doi.org/10.4324/9780203469255>
- CRUZ, A. H. O. **A Nova Economia do Futebol**: Uma análise do processo de modernização de alguns estádios brasileiros. Dissertação (Mestrado em Antropologia Social) Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

DAMATTA, R. (Org.). Esporte na sociedade: Um ensaio sobre o futebol brasileiro. IN: DAMATTA, R. **Universo do Futebol**. Rio de Janeiro: Pinakotheke, 1982.

GIGLIO, S. S.; SPAGGIARI, E. A produção das ciências humanas sobre futebol no Brasil: um panorama (19980 – 2009). **Revista de História** (USP), v. 163, p. 293-350, 2010. **crossref** <http://dx.doi.org/10.11606/issn.2316-9141.v0i163p293-350>

GIL, G. A copa da cultura no futebol. **Jornal O Globo**, ed. 5 set. 2004. <<http://www.suapesquisa.com/futebol/>> Acesso em 16 nov. 2008.

KRIEGER, M. da G.; FINATTO, M. J. B. **Introdução à Terminologia**. Teoria & prática. São Paulo: Contexto, 2004.

LEONINCE, M. P.; SILVA, M. T. Entendendo o futebol como negócio: um estudo exploratório. **Gestão e Produção**, v.12, n.1, p.11-23, jan./abr. 2005. **crossref** <http://dx.doi.org/10.1590/S0104-530X2005000100003>

MANCA, E. From phraseology to culture: Qualifying adjectives in the language of tourism. In: RÖMER, U.; RAINER, S. **Patterns, Meaningful Units and Specialized Discourses**, 105–122, 2010.

MCENERY, T and HARDIE, A. **Corpus Linguistics: Method, Theory and Practice**. Cambridge: Cambridge University Press, 2012.

ROCHA, M. A. E. O uso de corpora computadorizados no ensino de língua portuguesa: metodologia e avaliação. In. GRIMM CABRAL, L. et all (orgs). **Linguística e ensino: novas tecnologias**, Blumenau: Nova Letra, p.137-55, 2001.

TAGNIN, S. E. O. A identificação de equivalentes tradutórios em corpora comparáveis. In: **I Congresso Internacional da ABRAPUI**, Minas Gerais, 2007.

TOGNINI-BONELLI, E. **Corpus Linguistics at Work**. Amsterdam/Philadelphia: John Benjamins, 2001. **crossref** <http://dx.doi.org/10.1075/scl.6>

TOLEDO, L. H. de. **Lógicas no Futebol**. São Paulo: Hucitec/Fapesp, 2002.

Artigo recebido em: 15.10.2014

Artigo aprovado em: 30.11.2014