

## Fraseologia no Futebol: Um estudo bilíngüe baseado em Corpus

Sabrina Matuda\*

**Resumo:** a presente pesquisa tem como objetivo estudar a terminologia do futebol em inglês e português por meio do estabelecimento de equivalentes fraseológicos. Para tanto, a fundamentação teórica embasa-se na Lingüística de Corpus e em uma Abordagem Textual da Terminologia. Partiremos do pressuposto de que um termo raramente ocorre sozinho e de que o significado estático que quase sempre lhe é atribuído muda de acordo com o contexto e com as palavras com as quais co-ocorre. O *corpus* de estudo, em fase de compilação, possui, aproximadamente, 285 mil palavras - 156.146 em inglês e 127.984 em português. Em vista da complexidade da compilação de um corpus representativo analisaremos somente alguns primeiros achados.

**Palavras-chave:** Lingüística de Corpus; futebol; fraseologia; equivalência.

**Abstract:** this research aims at investigating the football terminology in English and Portuguese establishing phraseological equivalents. In order to do so the theoretical foundations are that of Corpus Linguistics as well as Textual Terminology. The study is based on the assumptions that a term is not likely to be used apart from other lexical items and also based on the fact that the protected status that is often attributed to it changes according to the context and the words to which it co-occurs. The *corpus*, still being compiled, consists of approximately 285 thousand words – 156.146 in English and 127.984 in Portuguese. Due to the complexity of compiling a representative corpus just a preliminary account of the findings is presented.

**Keywords:** Corpus Linguistics; football; phraseology; equivalence.

### Introdução

O presente artigo, recorte de nossa pesquisa de mestrado que está sendo desenvolvida no Departamento de Letras Modernas da Universidade de São Paulo, reflete sobre o uso do termo *gol* partindo da premissa de que o texto especializado é o objeto de estudo da Terminologia e de que o reconhecimento do tipo textual, bem como a cultura em que está inserido, sua finalidade como texto, o público a que se dirige, seus traços de sistematicidade e o grau de densidade terminológica compõem a parte central do levantamento e da identificação terminológica (Krieger e Finatto 2004:81). Sendo assim, o texto técnico será visto como elemento integrante de uma cultura, como uma de suas formas de manifestação mais poderosas, e “não como um fenômeno isolado, suspenso num vácuo” (Snell-Hornby 1998:39 apud Azenha 1999:28).

---

\* Mestranda em Letras (Estudos Lingüísticos e Literários em Inglês) pela Universidade de São Paulo (USP/FFLCH)

## **Por que estudar as fraseologias do futebol?**

Embora seja, muitas vezes, relacionado somente ao lazer, o futebol é uma área técnica, pois possui um vocabulário específico organizado em torno de uma temática. Constitui uma área técnica tanto no Brasil quanto em outros países. Sendo assim, possui uma linguagem específica utilizada para descrever acontecimentos. 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 (Finatto 2004).

No que se refere aos estudos lingüísticos do Futebol já existentes, podemos mencionar o livro *“The Linguistics of Football”* (2008) que contém uma série de artigos sobre Terminologia, expressões idiomáticas, metáforas, discurso, ensino de línguas e mídia do futebol. Todos os artigos são escritos em inglês, mas os estudos são feitos em várias línguas, dentre elas: francês, inglês, alemão, croata, búlgaro, árabe, espanhol, polonês e malaio. Mencionamos também o *Kicktionary* (Schmidt 2008), dicionário online multilíngüe (inglês – francês – alemão) de futebol, desenvolvido no *FrameNet project* da universidade de Berkeley, que explora como teorias de semântica lexical, lingüística de corpus e processamento de línguas naturais contribuem para a elaboração de obras de referência destinadas à pesquisa lexical. Atualmente o dicionário conta com 1.900 unidades lexicais. No Brasil, temos vários trabalhos que retratam aspectos da linguagem futebolística, dentre eles: Feijó (1994), D’Ambrosio (2006), Maranhão (1998) e Capinssú (1998), dicionários monolíngües e ainda trabalhos como de Maria Cristina Andrade Santos na área de Lingüística da Universidade Federal de São Carlos que propõe a construção de uma ontologia do domínio futebol. No entanto todos esses trabalhos, que foram realizados com rigor metodológico e que descrevem particularidades da linguagem específica que o futebol constitui, são monolíngües.

Já existem alguns dicionários bilíngües inglês-português sobre futebol. Porém todos os que encontramos em nossa pesquisa, além de não terem sido elaborados com o rigor metodológico dos trabalhos citados acima, preocupam-se somente com a descrição do termo.

## Objetivos

O principal objetivo desse trabalho é fazer um levantamento semi-automático das fraseologias mais freqüentes em que o termo *gol* está inserido utilizando o programa *WordSimth Tools*. Para tanto, estabelecemos 3 princípios, a saber:

- 1) O levantamento será feito do português para o inglês;
- 2) Uma vez que para a identificação das possíveis fraseologias utilizaremos o dispositivo *Clusters*<sup>1</sup> que é acoplado à ferramenta *Concord*, trabalharemos somente com *clusters* de quatro palavras, com freqüência mínima de quatro e com o MI-score<sup>2</sup> maior ou igual a dois;
- 3) Como se trata de um estudo exploratório, escolheremos apenas uma das fraseologias levantadas para fazer uma análise mais detalhada com o objetivo de identificar padrões específicos;

## Fraseologia

A definição e delimitação das estruturas fraseológicas são objetos de grande complexidade. A problemática inicia-se na própria terminologia utilizada para nomear essas estruturas: *terminological phrase*, *LSP phrase*, *technical phrase*, *phraseme or phraseological unit*, *phraseological term*, *professionalism ou jargon Word*, *multi-word lexical item*, entre outros<sup>3</sup>.

Em nosso estudo observamos as fraseologias dentro de duas perspectivas presentes em Krieger & Finatto (2004):

- 1) formulações prototípicas de cada tipo de comunicação especializada;
- 2) estruturas sintagmáticas pluriverbais que comportam um termo;

Sendo assim, consideraremos fraseologias tanto cadeias de caracteres não especializados, que assumem status de termo por serem freqüentes em uma linguagem

---

<sup>1</sup> Dispositivo que nos permite visualizar grupos de palavras que são encontradas em seqüência em um texto.

<sup>2</sup> Medida estatística que permite calcular o grau de associação de duas ou mais palavras em um corpus tendo como base a freqüência independente das palavras.

<sup>3</sup> termos apresentados por Giuseppe Palumbo em *The use of phraseology for training and research in the translation of LSP texts* ao discorrer sobre a problemática da definição e da nomenclatura de unidades fraseológicas em linguagens especializadas.

especializada, como em *sem chance de + verbo*, quanto estruturas que comportam um termo, como em *fazer um gol contra*.

Ao considerar essas duas perspectivas, pretendemos tirar do termo o status protegido que geralmente lhe é atribuído (Tognini-Bonelli 2002) e reconhecer que os termos são, em sua grande maioria, vinculados a expressões maiores que caracterizam um modo de expressão específico de uma dada linguagem especializada.

A importância da utilização apropriada do termo é ressaltada por Bowker e Pearson em “Working with Specialized Language – A Practical Guide to Using Corpora”:

“... How many times have you found yourself armed with the correct terminology but unable to use it simply because you do not know which word goes around which?”

A consciência de que um termo não pode ser utilizado de qualquer maneira em um texto reflete uma visão probabilística da linguagem, que tem raízes na lingüística Hallydaiana – teoria lingüística que descreve a probabilidade dos sistemas lingüísticos, dados os contextos em que os falantes os empregam (Berber Sardinha 2004:30). Essa visão pressupõe que, embora muitas combinações e arranjos sejam possíveis gramaticalmente, não são comumente utilizados.

Dentro dessa mesma perspectiva, Sinclair (1991) chama a atenção para a tendência terminológica e para a tendência fraseológica, ou seja, a tendência de um termo ocorrer sozinho, sem estabelecer relações com outros itens, e a tendência de um termo ocorrer acompanhado de outros itens e, conseqüentemente, ser parte integrante de uma unidade de sentido maior, como em *fez o gol da vitória*.

Para o tipo de levantamento e de descrição que nos propomos a fazer, nos ancoramos a uma visão textual da Terminologia, que observa como os termos são usados em situação de comunicação, analisando, assim, as unidades terminológicas em seus contextos reais de ocorrência considerando todos os componentes constitutivos dos universos do discurso especializado (Krieger & Finatto 2004:38).

Em nossa pesquisa utilizamos uma definição pragmática de *termo* presente em Pearson (1998). De acordo com essa definição existem três categorias de termos: *subject specific vocabulary* (vocabulário específico de uma linguagem especializada), *non subject-specific specialized vocabulary* (vocabulário específico de qualquer linguagem especializada) e *general vocabulary* (vocabulário geral). O primeiro grupo é composto

por itens lexicais que fazem parte de um único domínio; são monorreferenciais. O segundo é composto por itens lexicais que fazem parte de mais de um domínio especializado. O terceiro é composto por itens lexicais de língua geral que adquirem *status* de termo simplesmente por serem utilizados em uma terminologia.

## O Corpus de Estudo

Com o objetivo de estudar as fraseologias típicas do futebol criamos o seguinte design para nosso corpus:

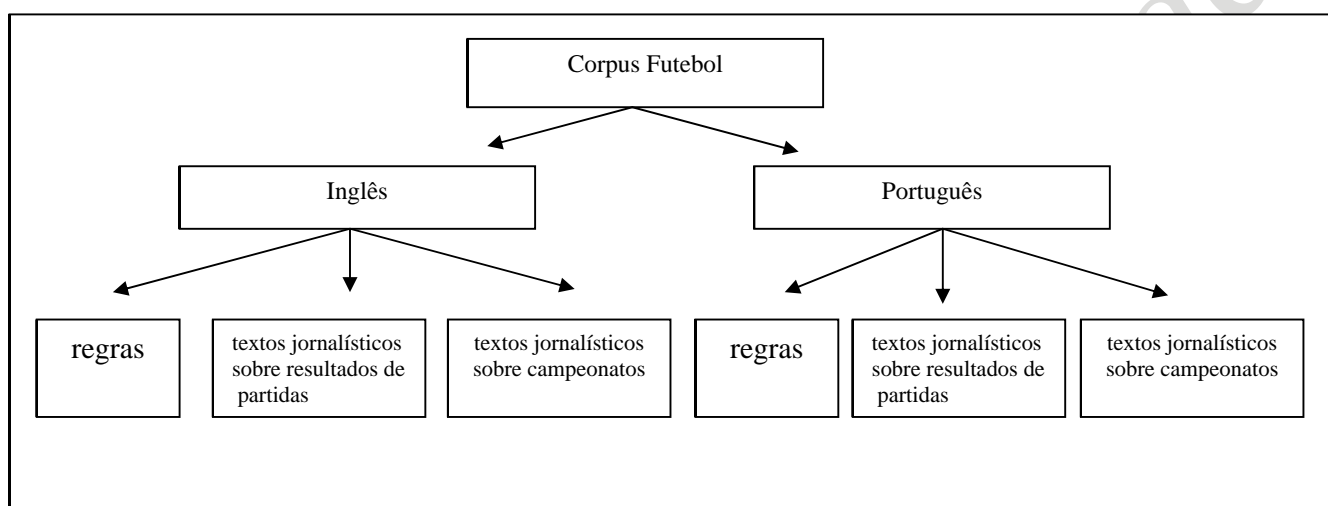


Figura 1: design do corpus

Como podemos observar, o *corpus* é dividido em Futebol Português (FP) e Futebol Inglês (FI). Cada corpus é subdividido em três *subcorpora*: regras do jogo, textos jornalísticos sobre resultados de partidas e textos jornalísticos sobre campeonatos.

Em *Corpora in Applied Linguistics*, Susan Hunston (2002) discute quatro aspectos que devem ser considerados para a compilação de um corpus, são eles: tamanho, conteúdo, balanceamento, representatividade e temporalidade.

### a) tamanho

O *corpus* de estudo, em fase de compilação, possui, aproximadamente 285 mil palavras - 156.146 em inglês e 127.984 em português.

Sabe-se (Capinussú 1998:15) que a linguagem do futebol é muito criativa, cheia de desvios e marcas culturais. Por esse motivo, decidimos compilar um *corpus* de 500 mil

palavras, um corpus de tamanho médio de acordo com a classificação de Berber Sardinha (2004:26).

b) conteúdo

O conteúdo do *corpus* é definido com base nos resultados que uma pesquisa pretende alcançar. Em nosso caso, a linguagem típica utilizada para falar de futebol em inglês e português, mais especificamente, as fraseologias utilizadas por especialistas para transmitir os acontecimentos para o grande público.

É com base nesse objetivo final que fizemos a divisão das tipologias textuais expostas na figura 1.

Escolhemos trabalhar com textos jornalísticos de resultados de partidas e campeonatos devido à alta concentração terminológica nesses tipos de notícias. Além disso, muitas das notícias sobre futebol, sejam elas publicadas em jornais ou em periódicos esportivos, tratam de temas com os quais não pretendemos trabalhar, como contratação de jogadores, organização de campeonatos, eleição dos melhores jogadores etc. Já o *subcorpus* de regras do jogo, que é um *corpus* paralelo, pois temos as regras da FIFA escritas originalmente em inglês e as regras da CBF que são a tradução das regras da FIFA para o português, foi compilado para que pudéssemos identificar a terminologia utilizada pelas associações internacionais.

As regras do jogo foram coletadas dos sites da FIFA e da CBF. Os textos jornalísticos estão sendo coletados de periódicos ingleses, americanos e brasileiros<sup>4</sup> publicados na internet.

c) representatividade e balanceamento

Em sua essência, um corpus, seja de que tipo for, é tido como representativo da linguagem, de um idioma, ou de uma variedade dele (Berber Sardinha 2004:22).

---

<sup>4</sup> periódicos ingleses: The Guardian; The Daily Telegraph; The Times; The Sun; Daily Mail; The Football.com;

periódicos americanos: ESPN; New York Times; The Washington Post;

periódicos brasileiros: Folha de São Paulo; O Estado de São Paulo; LANCENET; CBF News; Gazeta Esportiva; Globo Esporte; ESPN Brasil; SporTV; Sites de Clubes Esportivos;

Para que nosso corpus seja representativo da linguagem utilizada para descrever os acontecimentos de uma partida para o grande público, ou seja, para cobrir os vários aspectos de uma partida de futebol, selecionamos as três tipologias textuais mencionadas.

Durante a coleta e armazenamento dos textos, deparamos com alguns problemas como, por exemplo, a questão das regras: a FIFA é responsável pelas regras do jogo e as associações representantes dos países, em nosso caso a CBF, tem que utilizar as regras estabelecidas pela FIFA, ou seja, não existe um conjunto de regras da CBF ou de qualquer outra associação que tenha sido escrito originalmente em português; o que existe é o conjunto de regras da FIFA traduzido para o português. Mas, uma vez que as regras da CBF representam uma tradução “consagrada” e são utilizadas para a criação de regras de associações menores, serão aqui consideradas como texto original.

Deparamos também com o problema da grande diferença no número de palavras de um texto jornalístico sobre resultados de partidas em inglês e o mesmo tipo textual em português. Os textos em inglês tendem a ser mais explicativos e, conseqüentemente, possuem um número maior de palavras. Já os textos em português tendem a ser mais sucintos e menores. Até o presente momento, temos:

Língua		total	copa	geral	regras	Type/token ratio
Inglês	tokens	156.146	118.151	15.329	22.666	7,43
	nº de textos	147	133	13	1	
Português	tokens	127.984	61.449	41.821	24.714	8,13
	nº de textos	222	124	97	1	

Tabela 1: Corpus de Estudo

a) temporalidade

Uma vez que pretendemos estudar e descrever a linguagem futebolística contemporânea, compilamos um corpus sincrônico, ou seja, um corpus que compreende um período no tempo. Iniciamos a coleta dos textos durante a copa do mundo de 2006 e o corpus está *aberto*, pois ainda estamos coletando textos. No momento, nossa meta é chegar a 250 mil palavras em cada língua.

**Metodologia**

Nesta parte do trabalho, descreveremos a metodologia utilizada em nossa pesquisa. Primeiramente, mostraremos os procedimentos adotados para o levantamento

das possíveis fraseologias e, em seguida, apresentaremos os procedimentos utilizados para a validação das fraseologias encontradas e, ao final, descreveremos os passos seguidos para o estabelecimento dos equivalentes fraseológicos entre o par de línguas português-inglês.

### *Levantamento das Possíveis Fraseologias*

Inicialmente, geramos as linhas de concordância da palavra *gol* na ferramenta *Concord*. Em seguida, aplicamos o dispositivo *Re-sort* que nos permite ordenar as linhas de concordância de acordo com as palavras próximas à palavra de busca. No entanto, devido ao grande número de ocorrências que obtivemos (396), seria muito laborioso identificar as possíveis fraseologias somente pelas concordâncias. Resolvemos, então, utilizar o dispositivo *Cluster*. Ajustamos a configuração para identificar *clusters* de quatro palavras, com frequência mínima de quatro e com o MI-score maior ou igual a dois. Obtivemos, então, a seguinte tabela:

	Cluster	Freq.	Length	N	Cluster	Freq.	Length
1	O GOL DA VITÓRIA	7	4	9	SEM LEVAR UM ÚNICO	4	4
2	DE MARCAR UM GOL	7	4	10	CLARA DE MARCAR UM	4	4
3	UMA OPORTUNIDADE CLARA DE	7	4	11	LEVAR UM ÚNICO GOL	4	4
4	OPORTUNIDADE CLARA DE GOL	7	4	12	MELHOR CHANCE DE GOL	4	4
5	10 O GOL MARCADO	5	4	13	O GOL DE VIRADA	4	4
6	UM GOL PODE SER	5	4	14	OPORTUNIDADE CLARA DE MARCAR	4	4
7	UM GOL NÃO É	5	4	15	DE CHUTES A GOL	4	4
8	COM UM GOL DE	5	4	16	MARCOU UM GOL CONTRA	4	4

Figura 2: *clusters* de gol

Como podemos observar, o fato de algumas palavras serem encontradas repetitivamente em uma mesma seqüência não lhes garante o status de fraseologia.

### *Validação das Fraseologias*

O primeiro passo para validar as fraseologias foi agrupar os *clusters* que pareciam expressar uma mesma unidade de sentido, sendo assim, após o agrupamento chegamos à tabela abaixo:



N	Cluster
1	O GOL DA VITÓRIA DE MARCAR UM GOL CLARA DE MARCAR UM
2	OPORTUNIDADE CLARA DE MARCAR UMA OPORTUNIDADE CLARA DE
3	OPORTUNIDADE CLARA DE GOL UM GOL PODE SER
4	GOL PODE SER MARCADO
5	UM GOL NÃO É
6	COM UM GOL DE SEM LEVAR UM ÚNICO
7	LEVAR UM ÚNICO GOL
8	MELHOR CHANCE DE GOL
9	O GOL DE VIRADA
10	MARCOU UM GOL CONTRA

Figura 3 – *clusters* de gol agrupados

Após esse agrupamento conseguimos reduzir o número de *clusters* de 16 para 10. Mas alguns como *um gol pode ser*, *um gol não é* e *com um gol de* ainda pareciam estar “quebrados”. Para verificarmos o comportamento desses *clusters*, voltamos para as linhas de concordância onde pudemos constatar que todos faziam parte de uma unidade maior:

um gol pode ser	→	um gol pode ser marcado diretamente de um tiro de meta
um gol não é	→	quando um gol não é marcado de acordo com a regra X.
com um gol de	→	com um gol de (nome do jogador)

Classificamos os *clusters* acima como partes integrantes de uma unidade de sentido maior porque nenhum deles ocorreu separadamente das fraseologias acima, por exemplo, o *cluster um gol não pode ser* não ocorreu nenhuma vez desvinculado de *marcado diretamente de um tiro de meta* e assim por diante.

#### *Estabelecimento dos equivalentes fraseológicos*

Para identificar os equivalentes das dez fraseologias mais frequentes com a palavra *gol* em português no *corpus* de inglês fizemos duas concordâncias para a palavra *goal*, uma ordenada alfabeticamente pelas três primeiras palavras à esquerda, para busca de equivalentes como *o gol da vitória* e *o gol de virada*, que em inglês provavelmente

teriam a ordem invertida devido à posição do adjetivo, e outra ordenada alfabeticamente pelas três palavras à direita.

É importante ressaltar que, em alguns casos, os equivalentes das fraseologias aqui estudadas são bem diferentes nas duas línguas, seja pela estrutura ou pelo uso de uma outra fraseologia convencionalizada, em nosso caso no inglês, que não contém a palavra *goal*.

### Discussão e análise dos Dados

Dentre os 16 *clusters* iniciais, 10 foram classificados como fraseologias e, para melhor apresentá-las, optamos por dividi-las entre as utilizadas nas regras do jogo e as utilizadas em periódicos.

#### *Regras do Jogo*

Em nosso levantamento, encontramos quatro fraseologias específicas para descrever as regras do jogo, a saber:

fraseologia	freq. no corpus
1) oportunidade clara de gol	7
2) oportunidade clara de marcar um gol	4
3) um gol pode ser marcado diretamente de um tiro de meta	4
4) quando um gol não é marcado de acordo com a regra X.	4

Tabela 2: fraseologias utilizadas nas regras do jogo

A busca pela palavra *goal* ordenada alfabeticamente pela direita nos permitiu identificar os equivalentes para as fraseologias 3 e 4 da tabela acima: *a goal may be scored directly from a goal kick* e *when a goal is not scored in accordance with Law X*, respectivamente. No entanto não conseguimos identificar nenhum equivalente para as duas primeiras fraseologias mesmo com a busca ordenada pela esquerda. Decidimos, então, fazer uma busca para a palavra *opportunity* ordenada alfabeticamente pelas três primeiras palavras à esquerda:

N	
8	lty area) 5. denies <b>an obvious goalscoring opportunity</b> to an opponent moving towards
9	If a player denies <b>an obvious goal-scoring opportunity</b> playing in a dangerous manner,
10	denies an opponent <b>an obvious goalscoring opportunity</b> may be an offence that incurs
11	nt off if he denies <b>an obvious goalscoring opportunity</b> by holding an opponent · No fu
12	ls see the 32-nation tournament as a grand <b>opportunity</b> , they also acknowledge that, t
13	e. We understand that, but we have a great <b>opportunity</b> with young players like John T
14	deal with denying an oppo- nent an obvious <b>opportunity</b> to score a goal. It is not nec

15 's prowess may well be rooted in a lack of **opportunity** that has sired creativity. Whi  
 16 ermined to cash in on a once-in-a-lifetime **opportunity**. South Africans are preparing  
 17 an. The Black Antelopes had the first real **opportunity** in the 12th minute with Mateus  
 18 revents a goal or **an obvious goal- scoring opportunity** by deliberately handling the b  
 19 gs have changed a lot." Now comes a second **opportunity** to star in the World Cup. If o  
 20 elop -- for an Argentine midfielder to see **opportunity** where others saw nothing -- an  
 21 have it removed at the earliest pos- sible **opportunity**. 62 LAW 3 - THE NUMBER OF PLA  
 22 he player off unless there is a **clear subsequent opportunity** to score a goal. · The positio  
 23 ul play unless there is a **clear subsequent opportunity** to score a goal. The referee s  
 24 conduct unless there is a **clear subsequent opportunity** to score a goal. The referee s

Figura 4: seleção de linhas de concordância para *opportunity*

Essa busca não só nos permitiu identificar *an obvious goalscoring opportunity* como equivalente para *uma oportunidade clara de gol* como mostrou que *oportunidade clara de marcar um gol* não ocorreu nenhuma vez separadamente de uma fraseologia maior: *a menos que haja uma subsequente oportunidade clara de marcar um gol* para a qual identificamos *unless there is a clear subsequent opportunity to score a goal* como equivalente.

### Periódicos

Encontramos 6 fraseologias utilizadas em periódicos sobre futebol com a palavra gol, a saber:

fraseologia	freq. no corpus
1) o gol da vitória	7
2) com um gol de (jogador)	5
3) sem levar um único gol	4
4) melhor chance de gol	4
5) o gol da virada	4
6) marcou um gol contra	4

Tabela 3:

fraseologias encontradas nos periódicos

A busca pela palavra *goal* ordenada alfabeticamente pela esquerda nos permitiu identificar como equivalentes para as fraseologias 1 e 6 *the winning goal* e *scored an own goal*, respectivamente.

Ao observar as linhas de concordância encontramos quatro ocorrências de *go-ahead goal*:

N Concordance  
 187 1. The French were pushing for a **go-ahead goal** in the 104th minute Sunday, when Zid  
 188 d then set up Steven Gerrard's **go-ahead goal** in the 85th for England. But the Engli  
 189 der Fabio Grosso, who scored the **go-ahead goal** against Germany in the 119th minute,  
 190 has long defended, set up the **go-ahead goal** and scored the third. So even those Me

Figura 5: linhas de concordância de *go-ahead goal*

A princípio, não conhecíamos a colocação, mas depois de ler as linhas e ampliar os contextos para verificarmos as situações em que esse tipo de gol ocorre, identificamos *go-ahead goal* como equivalente para o *gol da virada*. Esse é um exemplo de como as concordâncias podem ser úteis para identificar equivalentes mesmo quando a palavra procurada, em nosso caso *virada*, não é a mesma nas duas línguas. Ao ampliarmos os contextos, verificamos que o placar de um jogo que estava empatado sempre virava depois do *go-ahead goal*, fato de extrema importância para a confirmação de nossa hipótese.

Como as concordâncias com *goal* não nos permitiram identificar equivalentes para as outras fraseologias, resolvemos gerar linhas de concordância com outras expressões de busca. Para *melhor chance de gol* fizemos concordâncias com a palavra *chance* e encontramos:

N	Concordance
60	al. It is far easier to build a <b>scoring chance</b> from 40 yards away than from 80. "I
61	len had the Soca Warriors' <b>best scoring chance</b> , a well-struck shot from the right s
62	bench. Saha set up France's final <b>scoring chance</b> , using his chest to drop the ball

Figura 6: seleção de linhas de concordância para *chance*.

Como a figura acima mostra, encontramos três ocorrências de *scoring chance* e apenas uma de *best scoring chance*. Resolvemos, então, verificar no Google.uk se a fraseologia é frequente em inglês ou se ocorreu esporadicamente em nosso corpus e, com o resultado do Google, que nos retornou 13.700 ocorrências, pudemos constatar que é uma fraseologia de uso corrente em periódicos britânicos que falam de futebol.

Devido à dificuldade de encontrar um equivalente para *com um gol de (nome do jogador)*, resolvemos ler as linhas de concordância da fraseologia em português para verificar a situação em que ocorre:

N	
2	a venceu o México por 2 a 1, neste sábado, <b>com um gol de Maxi Rodríguez</b> na prorrogaçã
3	te nas oitavas Time europeu bate a Tunísia <b>com um gol de Shevchenko</b> e avança na
4	virtudes. Perdeu para a Espanha de virada, <b>com um gol de Fernando Torres</b> no último mi
5	ia por 1 a 0 nesta terça-feira, em Berlim. <b>Com um gol de Kaká</b> ainda no primeiro tempo

Figura 7: seleção para linhas de concordância de *com um gol de* \*

As concordâncias revelaram que, no geral, a fraseologia ocorre para retratar uma mudança no jogo após o gol, ou seja, o foco não é o jogador, mas sim o efeito que o gol teve na partida. Tendo em vista a função da expressão de busca, fizemos uma concordância para o advérbio *when* ordenada alfabeticamente pela segunda palavra à

direita, já tendo em mente que a primeira palavra provavelmente seria o nome de algum jogador:

N	
1	strong German pressure early in the match <b>when Cristiano scored</b> in the fourth minute
2	Tunisians took the lead in the 23rd minute <b>when Jaziri scored</b> the second goal. Yass
3	rous." These plans for the match went awry <b>when Zidane scored</b> in the fourth minute a

Figura 8: seleção de linhas de concordância de \* *scored*

Encontramos 16 ocorrências para a expressão de busca, que por razões de espaço reduzimos a 3. Nesse caso, ampliamos as linhas para observar se o efeito da expressão era o mesmo. E, após constatar que sim, fizemos a busca na direção inversa, ou seja, geramos uma concordância para *quando* ordenada da mesma maneira que *when* e, para nossa surpresa, não encontramos nenhuma ocorrência de *quando* \* *marcou* em nosso corpus. Esse tipo de busca nos evidenciou que embora a expressão *quando* \* *marcou* seja gramaticalmente possível não é a expressão convencionalizada para descrever o efeito de um gol em uma partida<sup>5</sup>.

### Gol da vitória

Escolhemos a fraseologia mais freqüente em nosso corpus, *o gol da vitória*, para fazer uma análise mais detalhada e identificarmos possíveis padrões.

N	
1	s boas chances e de uma delas acabou <b>surgindo o gol da vitória, com</b> Fernando Torres, que
2	, depois de sair do banco de reservas, <b>marcou o gol da vitória contra</b> a Austrália. Ele pod
3	no aproveita chute errado de Gilberto e <b>marca o gol da vitória brasileira</b> A seleção brasil
4	O curioso é que foi justamente o meia que <b>fez o gol da vitória sobre</b> o Peixe por 3 a 2, no
5	cante Kléber Pereira está feliz por ter <b>feito o gol da vitória contra</b> o São Caetano, no úl
6	ado de "estrela em ascensão", que <b>conseguiu o gol da vitória brasileira</b> usando um pou
7	uperou o goleiro Pletikosa... ... e <b>comemorou o gol da vitória</b> . TUDO SOBRE A PARTIDA Ele

Figura 9: linhas de concordância para “o gol da vitória”

Ao observarmos as linhas, podemos identificar que a primeira palavra à esquerda é sempre um verbo e, na maior parte das vezes, no pretérito perfeito. À direita encontramos três padrões: 1) de um time; 2) com um jogador; 3) contra/sobre outro time;

<sup>5</sup>fizemos uma busca da expressão “quando \* marcou o gol” no Google que nos retornou 2.150 ocorrências. Geramos as linhas de concordância no KWIC Google (disponível em [lael.pucsp.br/corpora](http://lael.pucsp.br/corpora)) e, ao analisá-las, pudemos constatar que a expressão ocorre quase exclusivamente em blogs, fóruns sobre futebol e entrevistas, ou seja, em uma linguagem mais informal e não na linguagem utilizada pelo especialista para retratar uma partida ao grande público.

## Considerações Finais

Embora ainda em fase inicial, o estudo nos mostrou que a análise de unidades fraseológicas é uma abordagem promissora para uma extração terminológica que tenha como objetivo descrever a realidade dos usos terminológicos. As UFs levantadas compreendem sintagmas com elementos fixos, como em o *gol da vitória*, o *gol da virada*, *marcou um gol contra*, *melhor chance de gol* e *com um gol de* que representam o uso convencionalizado do termo, valorizando, assim, seu caráter discursivo.

Durante o levantamento, a validação e a identificação de possíveis equivalentes para as fraseologias, deparamos com alguns obstáculos, sendo assim, listamos algumas medidas que consideramos importantes para o aprimoramento da pesquisa:

- 1) expandir o corpus, que está em fase de compilação, para que tenhamos um corpus representativo da linguagem futebolística para, assim, validarmos os sintagmas encontrados como fraseologias sem precisar recorrer a fontes externas, como, por exemplo, o Google.
- 2) estabelecer melhor os critérios para o levantamento das fraseologias, uma vez que ao nos limitarmos aos *clusters* podemos deixar de observar associações entre palavras não-consecutivas<sup>6</sup>;

## Referências Bibliográficas

AZENHA JR., João. *Tradução Técnica e Condicionantes culturais*: primeiros passos para um estudo integrado. São Paulo: Humanitas/FFLCH-USP, 1999.

BERBER SARDINHA, Tony. *Linguística de Corpus*. Barueri: Manole, 2004.

BOWKER, Lynne. *Computer-aided Translation Technology: A Practical Introduction*. Ottawa: Ottawa Press, 2002.

BOWKER, Lynne & PEARSON, Jennifer. *Working with Specialized Language. A Practical Guide to Using Corpora*. London/New York: Routledge, 2002.

CAPINSSÚ, José Maurício. *A Linguagem Popular do Futebol*. São Paulo: Ibrasa, 1998.

---

<sup>6</sup> Embora não tenha sido mencionado na metodologia, utilizamos a ferramenta WConGram para validar algumas fraseologias.

D'AMBROSIO, Oscar. *Futebol na ponta da língua*. Jornal Unesp, Araraquara, a. 10, n. 212, 2006.

FEIJÓ, Luiz César Saraiva. *A Linguagem do Futebol no Brasil*. Disponível em: <[http://alinguagemdabola.com.br/fut\\_linguagem.html](http://alinguagemdabola.com.br/fut_linguagem.html)> acesso em: 23 mar. 2009.

FINATTO, Maria José Bocorny. *Termos, textos e textos com termos: novos enfoques dos estudos terminológicos de perspectiva lingüística*. In: ISQUERDO, A.N. & KRIEGER, M.G., (Org.). *Ciências do léxico* 2. 2 ed. Campo Grande: Editora da UFMS/PPG-Letras UFRGS, 2004, v. 01, p. 353-363.

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

KRIEGER, Maria da Graça; FINATTO, Maria José Bocorny. *Introdução à Terminologia*. Teoria & Prática. São Paulo: Contexto, 2004.

MARANHÃO, Haroldo. *Dicionário de futebol*. Rio de Janeiro/São Paulo: Record, 1998.

PALUMBO, Giuseppe. *The use of phraseology for training and research in the translation of LSP texts*. In Maia, Belinda, Johann Haller & Margherita Ulrych (eds.) *Training the Language Services Provider for the New Millenium - Proceedings of the III Encontros de Tradução de Astra-FLUP*, Porto: Faculdade de Letras Unviersidade do Porto, 199-211, 2003.

PEARSON, Jennifer. *Terms in Context*. Amsterdam: John Benjamins, 1998.

SANTOS, MARIA Cristina Andrade; VALE, Oto Araújo. *Abordagem linguística no desenvolvimento de ontologias: proposta de construção de uma ontologia do domínio futebol*. In: Seminário de pesquisa em ontologia no Brasil, 2008, Niterói?RJ. *Trabalhos aprovados - Seminário de pesquisa em ontologia no Brasil*, 2008.

SINCLAIR, John. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press, 1991.

TOGNINI-BONELLI, Elena. *Between Phraseology and Terminology in the Language of Economics*. In Nuccorini, S. (ed.) *Phrases and Phraseology: Data and Descriptions*, Bern: Peter Lang, 2002.

Sites:

Kicktionary

<http://www.kicktionary.de/>

WordSmith Tools

<http://www.lexically.net/wordsmith/>