

# A temática do futebol e a semântica de *frames* – um estudo exploratório

João Gabriel Padilha  
Bolsista IC Unisinos/CNPq  
joagabrielsl@hotmail.com



Orientadora  
Prof<sup>a</sup>. Dra. Rove Chishman  
rove@unisinos.br

## XXII Salão de Iniciação Científica – 2010

### Contextualização

#### O projeto FrameCorp

O projeto FrameCorp ocupa-se da construção de *corpora* anotados com informação semântica com base na Semântica de *Frames* (Fillmore, 1982, 1985) e do paradigma FrameNet (Fillmore; Johnson; Petruck, 2003).

#### Semântica de *Frames*

Os *frames* são “uma representação esquemática da estrutura conceitual e dos padrões de crenças, práticas, instituições, imagens, etc, que proporcionam uma fundação para uma interação plena de significado em uma certa comunidade de falantes”; (Fillmore, 2003, p. 235)

**Aporte teórico** – estudo dos *frames* futebolísticos através de 11 cenas (Schmidt, 2009);

**Recursos** – Ferramenta de anotação *Wordsmith* (Scott, 1999) e a base de dados da plataforma *kicktionary* (Schmidt, 2009);

#### Objetivo

Este trabalho tem como objetivo realizar um estudo exploratório do léxico do futebol em Português brasileiro, inicialmente investigando os conceitos de *cena* e *frame* propostos por Schmidt (2009).

Ele está inserido na segunda fase do projeto FrameCorp, a qual se chama *Kicktionary Br*.

O subprojeto *Kicktionary Br* (monolíngue) teve como inspiração os trabalhos de Schmidt (2009) presentes na home page homônima (<http://www.kicktionary.de/>) – uma ferramenta lexical multilíngue voltada à temática do futebol.

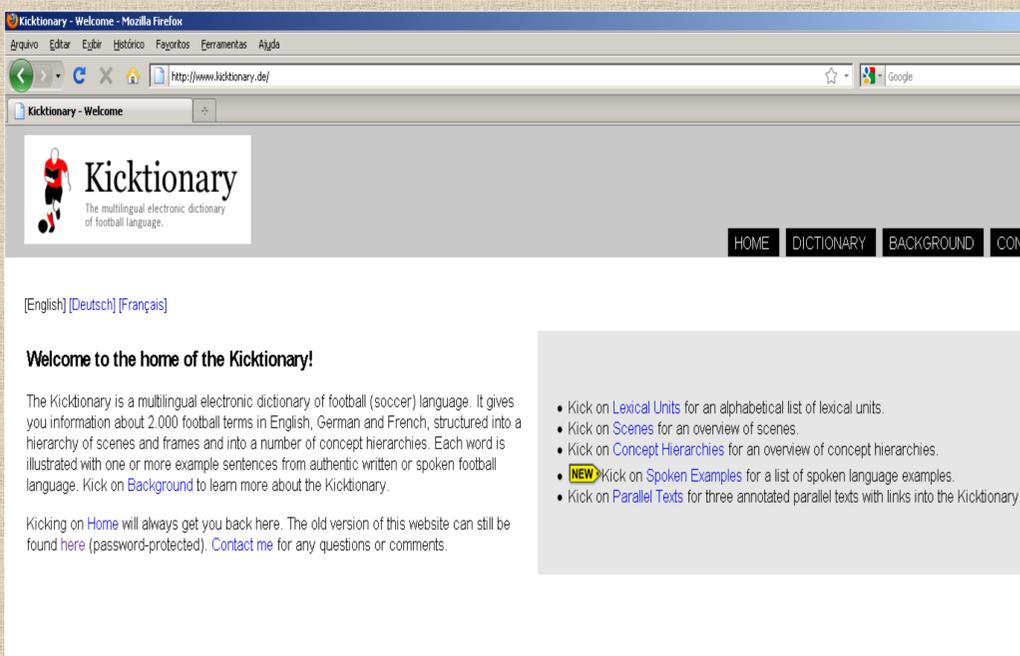


Figura 1 – Interface da ferramenta online *Kicktionary* ([www.kicktionary.de](http://www.kicktionary.de))

#### A plataforma *Kicktionary* (Schmidt, 2009)

✓ É uma ferramenta online voltada para a descrição do léxico do futebol que disponibiliza *corpora* semanticamente anotados nas línguas inglesa, francesa e alemã.

✓ Seu principal objetivo é prover descrição semântica para as unidades lexicais dessas três línguas com base na proposta do projeto FrameNet (Fillmore; Johnson; Petruck, 2003). de Berkeley, EUA.

✓ Encontrar uma forma para associar as informações lexicais do Inglês, Francês e Alemão foi um dos questionamentos iniciais de Schmidt.

✓ No *Kicktionary*, as unidades de descrição são a *cena*, o *frame* e as *unidades lexicais* – diferentemente do projeto FrameNet, que apresenta apenas o *frame* e as *unidades lexicais*.

✓ Schmidt afirma que as *cenais* são porções de conhecimento abstrato (não necessariamente linguístico), entidades *conceituais*, enquanto os *frames* são entidades linguísticas, e, como tal, englobam as unidades lexicais das três línguas.

### Scenários

#### On the pitch

##### 1 Shot

12 frames  
269 lexical units

##### 2 Ball\_And\_Goal (11 LUs)

Feign (6 LUs)  
Finish (5 LUs)  
Follow\_Up (7 LUs)  
Goal\_Kickoff (7 LUs)  
Hit (4 LUs)  
Intervene (23 LUs)  
Miss\_Goal (7 LUs)  
Save (46 LUs)  
Shoot\_At (5 LUs) 3  
Shot (123 LUs)  
Shot\_Supports (25 LUs)

Figura 2 – a cena *shot* (1) e os *frames* nela contidos (2). Ao lado deles, entre parênteses, o número de unidades lexicais que os evocam (3).

#### A pertinência dos conceitos *frame* e *cena* no *Kicktionary*

Para Schmidt, os falantes nativos do Inglês, Francês e Alemão partilham um conhecimento abstrato acerca dos eventos envolvidos em uma partida de futebol – espera-se o mesmo dos falantes do Português Br em relação à proposta *Kicktionary Br*.

#### O *corpus* do projeto *Kicktionary Br*

✓ O *corpus* piloto é constituído por 100 notícias sobre resultados de partidas de futebol extraídas de sites de 9 clubes brasileiros e totaliza 76.463 palavras;

✓ Trata-se de um *corpus* aberto (Berber Sardinha, 2004), uma vez que ainda está em fase de compilação.

#### Resultados preliminares e as próximas etapas deste estudo

1 – Os primeiros resultados desta segunda fase dizem respeito às diferenças entre os padrões de lexicalização entre a língua Inglesa – nossa principal referência no que tange aos equivalentes de tradução – e o PB.

✓ Embora este estudo não tenha caráter contrastivo, a consulta à base de dados da plataforma *Kicktionary* evidenciou essas peculiaridades sobre os padrões lexicais entre as duas línguas – uma questão a ser pensada nas próximas fases do projeto.

Considerem-se as seguintes sentenças:

*Chicão marcou o pênalti; Ronaldo marcou o segundo do Corinthians; Vágner Love marcou o adversário; o time marcou bem.*

Na língua inglesa, o verbo *marcar* tem 3 equivalentes distintos de tradução: *score*, *mark* e *defend* – o que evidencia, de um lado, a polissemia do verbo *marcar* em PB e, de outro, a inexistência de correspondentes diretos para os diversos sentidos desse verbo na língua inglesa.

2 – Outra constatação importante para este estudo é a alta incidência de nominais evocadores de *frame*: *jogo*, *final*, *competição*, *rodada*, *vitória*, *lance*, etc.

A alta incidência desses nominais nos remete à necessidade de expandir o tamanho do *corpus* a fim de atingir o conceito de *representatividade* postulado por Berber Sardinha (2009, p. 22) – esse conceito baseia-se na probabilidade e é fundamentado no fato de, para representar uma variedade da linguagem – neste caso, o léxico futebolístico – um *corpus* precisa ser grande o suficiente para que haja quantidade suficiente de dados a serem descritos.

Nosso objetivo é obter um *corpus* de tamanho médio segundo os critérios propostos por Berber Sardinha– 250.000 palavras (p. 26) – o que possibilitará maior variedade de dados e descrições semânticas mais precisas, conseqüentemente.

#### Referências

BERBER SARDINHA, A. P. *Linguística de Corpus*. São Paulo: Manole, 2004.

SCHMIDT, Thomas. *The Kicktionary – a multilingual lexical resource of football language*. In: BOAS, Hans. (Ed.) . *Multilingual FrameNets - Methods and Applications*. Berlin/New York: Mouton de Gruyter, 2009.