

## Equivalentes de Tradução entre Português, Espanhol e Inglês na Constituição de Dicionário Temático Trilíngue para A Copa 2014

Maucha Andrade Gamonal (PPG-Linguística/UFJF)

Simone Rodrigues Peron (PPG-Linguística/UFJF)

Tiago Timponi Torrent (orientador/UFJF)

Universidade Federal de Juiz de Fora – Linguística Cognitiva

Pesquisa iniciada em 03/2011 com previsão de término para 03/2013

O presente trabalho se insere no projeto COPA 2014 FrameNet Brasil (SALOMÃO et al., 2011), que tem por objetivo construir um dicionário temático trilíngue (Português, Espanhol, Inglês), em meio eletrônico, para os domínios do Futebol e do Turismo. O projeto é um desdobramento do Projeto FrameNet Brasil (<http://www.framenetbr.ufjf.br>) (SALOMÃO, 2009) que visa a construir para o Português do Brasil a contraparte linguística da rede semântica conhecida como FrameNet, sediada em Berkeley e liderada pelo linguista Charles Fillmore. O dicionário, baseado no aporte teórico da Semântica de *Frames* (FILLMORE, 1982), se caracteriza como um produto de inovação tecnológica, afastando-se dos dicionários tradicionais, tanto na organização dos verbetes quanto no sistema de consulta, isso porque a busca será feita por cenas relevantes dos domínios do Futebol e do Turismo. O corpus utilizado para sustentar nossas análises na construção do dicionário constitui-se de textos jornalísticos, guias de viagem, sites de órgãos governamentais de fomento à atividade turística e blogs de viajantes construídos preferencialmente nas três línguas. Os corpora utilizados para o trabalho ora apresentado foram retirados de matérias dos sites BrasilGov (<http://www.brasil.gov.br/>) e Braziltour (<http://www.braziltour.com/>), ambos trilíngues. O material compilado apresenta 65.415 *tokens*. Compilados os corpora, efetua-se a busca pelas Unidades Lexicais – pareamentos de uma forma lexical a um significado estruturado em termos de um frame – evocadoras dos frames relacionados à experiência turística. A partir das ULs, procede-se à anotação semântica das sentenças extraídas dos corpora, seguindo-se a metodologia definida para a FrameNet (RUPPENHOFER et al., 2010). As anotações são então contrastadas para as três línguas de modo a se determinarem os melhores equivalentes de tradução com base em critérios tanto semânticos – i.e. a capacidade de cada UL de perfilar os mesmos Elementos de Frame – quanto sintáticos – i.e. a forma como a UL organiza os Elementos de Frame em padrões de valência. A análise de um conjunto de dez unidades lexicais para cada língua demonstrou, em primeiro lugar, a fragilidade dos corpora utilizados até agora, apontando para a necessidade de utilização de corpora equivalentes, produzidos por falantes nativos de cada idioma, no lugar dos corpora traduzidos. Ademais, discutem-se as nuances de valência e perspectiva trazidas pelos equivalentes de tradução das ULs selecionadas para cada um dos três idiomas.

### Referências:

- FILLMORE, C. J. Frame Semantics. In: LSK. *Linguistics in The Morning Calm*. Seoul: Hanshin, 1982.
- RUPPENHOFER, J.; ELLSWORTH, M.; PETRUCK, M.; JOHNSON, C.; SCHEFFCZYK, J. **FrameNet II: Extended Theory and Practice**. Versão 14 set. 2010. Disponível em: <http://framenet.icsi.berkeley.edu/>.
- SALOMÃO, M.M.M. FrameNet Brasil: um trabalho em progresso. In: *Calidoscópico*, Vol. 7.2, 2009.
- SALOMÃO, M.M.M.; TORRENT, T.T.; CAMPOS, F.C.A.; BRAGA, R.M.M. & VIEIRA, M.B. *Copa 2014 Framenet Brasil*. Projeto de Pesquisa. Juiz de Fora, 2011.