

Elaboração de um Dicionário Trilíngue para COPA 2014: O *Frame* de Atração_Turística a partir das Unidades Lexicais *Atrair/ Atraer/ Attracts*

Simone Rodrigues Peron¹
Universidade Federal de Juiz de Fora (UFJF)
E-mail: speronjf@yahoo.com.br

RESUMO:

Neste artigo apresentaremos a anotação das Unidades Lexicais *atrain*, *atraer* e *attract* que evocam o *frame* de Atração_turística, e contrastaremos as sentenças, com base em critérios sintáticos e semânticos, de acordo com a metodologia definida pela FrameNet (RUPPENHOFER et al., 2006). Esse trabalho se insere no projeto COPA 2014 FrameNet BR, que tem por objetivo construir um dicionário temático trilíngue em meio eletrônico, para os domínios do Turismo e do Futebol.

PALAVRAS-CHAVE: Dicionário Temático Trilíngue; Unidades Lexicais *atrain / atraer / attract* ; Semântica de *Frames*; FrameNet BR.

ABSTRACT:

In this paper we present the annotation of the Lexical Units *atrain*, *atraer* and *attract*, which evoke the *Atração_turística* frame, and contrast sentences based on syntactic and semantic criteria, following the methodology defined for FrameNet (RUPPENHOFER et al., 2006). This work is part of the COPA 2014 FrameNet BR project, which aims to build a domain-specific trilingual electronic dictionary covering the Soccer and Tourism vocabularies.

KEYWORDS: Domain-Specific Trilingual Dictionary ; Lexical Units *atrain / atraer / attract*; Frame Semantics; FrameNet BR.

Introdução

O presente trabalho está inserido 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, que se constituirá como uma ferramenta que será utilizada no evento mundial a ser realizado no Brasil em 2014. O projeto está sendo desenvolvido na Universidade Federal de Juiz de Fora (MG) em associação com a UNISINOS (RS). O Projeto COPA 2014 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

¹ Mestranda em Linguística na Universidade Federal de Juiz de Fora.

como FrameNet, liderada por Charles Fillmore no *International Computer Science Institute* (EUA), sediado em Berkeley.

A construção do dicionário temático tem como base a Semântica de *Frames* (FILLMORE, 1977, 1982, 1985), seguindo o entendimento de que os domínios de conhecimento estão estruturados em *frames*.

Neste trabalho, analisaremos as Unidades Lexicais “atrair” em português, “atraer” em espanhol e “attracts” em inglês, tendo em vista que a UL “atrair” evoca o *frame* de *atração turística*. Para tanto, o artigo está organizado da seguinte forma: na primeira seção, discutiremos o aporte teórico que orienta o Projeto Copa 2014, o qual inclui a Linguística Cognitiva, a Semântica de *Frames* e a Linguística de Corpus; na segunda seção, faremos uma breve apresentação do Projeto FrameNet e de sua importância no cenário internacional. Na terceira seção, exporemos o *frame* de atração turística, já descrito pelo projeto FrameNet. Na quarta seção, apresentaremos os aspectos metodológicos para a constituição dos *corpora* e para a anotação lexicográfica. Na quinta seção, analisaremos as ULs “atrair”, “atraer” e “attracts” em quatro sentenças do português, espanhol e inglês, descrevendo-as para o uso no dicionário. Assim, temos o intuito de demonstrar como é realizado o trabalho do projeto brasileiro Copa 2014 de construção do dicionário temático trilingue, que se encontra, atualmente, em uma etapa inicial.

1. Aporte Teórico

Esse estudo possui como aporte teórico principal a Semântica de *Frames*, que está associada intrinsecamente à Linguística Cognitiva e à Linguística de Corpus. A seguir, abordaremos os aspectos relevantes de cada uma delas e como a rede semântica – FrameNet – pode ser utilizada na construção do dicionário.

A Semântica de *Frames* se insere no paradigma da Linguística Cognitiva, que, por sua vez, é compreendida dentro do cenário das “Ciências Cognitivas”.

As Ciências Cognitivas surgem na década de 50 do século passado no rastro do desenvolvimento do computador, chamado por muitos àquela altura de “cérebro eletrônico”. A denominada virada cognitiva da Linguística corresponde à revolução chomskyana que foi caracterizada como uma revolução contra a psicologia behaviorista, já que os processos computacionais de processamento da informação permitiam legitimar “estados internos” da mente emulada.

Daí se seguiu uma inversão de rumos investigativos contra as abordagens clássicas da mente e do cérebro, propondo uma abordagem científica interdisciplinar dos fenômenos relacionados. Para isso, tivemos a contribuição de diversas áreas, tais como da psicologia, da inteligência artificial, da linguística, da antropologia, das neurociências, das ciências da computação e da filosofia, constituindo um campo de investigação inteiramente novo.

Na Linguística, surge a contestação em ataque ao ascetismo gerativista com relação à significação, tanto em suas dimensões pragmáticas e sociais (Linguística Funcionalista) como em sua própria concepção de relação entre linguagem e cognição (Linguística Cognitiva).

Posteriormente, surge a abordagem da cognição e da linguagem como sistemas dinâmicos, visto que se desdobram na categoria de tempo, sendo este o principal fator diferencial, uma vez que leva em consideração o aspecto de mudança ao longo da vida, de aprendizagem social e histórica. Essa abordagem herda do connexionismo o foco em sistemas dinâmicos que se auto organizam, assim como o processamento linguístico (e o cognitivo) se dão pela emergência de relações probabilisticamente ponderadas. Nessa tendência teórica aparece o estudo da corporificação da linguagem e da cognição.

Com base nos postulados de Feldman (2006), podemos afirmar que os universais cognitivos acreditam que a aquisição da linguagem está intrinsicamente relacionada com a experiência corporificada pela qual cada sujeito vivencia em sua cultura. A conceptualização humana é corporificada: tanto conceitos elementares da experiência humana como ações sofisticadas envolvem imaginação e raciocínio, e, além disso, empregam capacidades sensório-motoras do sistema corpo-cérebro. Para esse autor, como temos uma complexa estrutura de redes neurais que comandam o nosso desenvolvimento tanto no sentido cognitivo como motor, há um elo fundamental das redes connexionistas entre o cérebro e o comportamento. Um argumento importante que endossa esse aspecto é que nenhum outro animal tem a capacidade de imitar como nós temos, o que se trata de uma façanha da capacidade neural (Meltzoff & Prinz 2002).

Feldman (2006) retoma o conceito de *frames* proposto por Fillmore (1989), que tem como finalidade mostrar que, para compreendermos o significado das palavras e as relações que estas estabelecem entre si, deveríamos entender a estrutura subjacente em relação ao conceito de *frame*.

Podemos então dizer que a Linguística Cognitiva trata a linguagem como uma habilidade cognitiva humana, que está ancorada tanto em princípios biológicos quanto socioculturais a partir das experiências que circundam o homem.

Dentro do paradigma teórico da Linguística Cognitiva, destacamos a Semântica de *Frames*.

Para Fillmore (1982), um *frame* é um sistema de categorias estruturadas de acordo com um determinado contexto motivador, sendo este um padrão de práticas vinculadas a uma instituição social, possibilitando que determinada categoria específica se torne inteligível na história de uma comunidade linguística. O autor ainda explica que:

algumas palavras existem para fornecer acesso ao conhecimento de tais *frames* aos participantes no processo comunicativo e, simultaneamente, servem para desempenhar uma categorização que pressupõe esses *frames*² (FILLMORE, 1982, p.381).

O exemplo mais conhecido é o *frame* de Transação Comercial (FILLMORE, 1982). Os elementos desta cena esquemática incluem uma pessoa com posse de dinheiro interessada em trocá-lo por uma mercadoria (comprador), o possuidor da mercadoria que quer trocá-la por dinheiro (vendedor), e as mercadorias que o Comprador pode adquirir ou adquiriu com o dinheiro (Mercadorias) e o dinheiro obtido pelo vendedor (Dinheiro). De acordo com a escolha lexical dos verbos que evocam esse *frame*, há um enfoque diferente. Se escolhermos o verbo “pagar”, a perspectiva enfoca tanto na ação do comprador em relação ao dinheiro, quanto do vendedor; já o verbo “vender” tem o foco na ação do vendedor, enquanto que o verbo “comprar” direciona para o comprador.

Segundo Fillmore (1982, p. 111) um *frame* também pode ser definido como “qualquer sistema de conceitos relacionados de tal forma que, para entender um deles, é necessário compreender toda a estrutura na qual ele se encaixa”. Assim, o *frame* é uma estrutura conceptual complexa, uma vez que uma palavra, ao ser mencionada, traz consigo a experiência socialmente compartilhada. Nessa direção, o autor considera que o termo *frame* abarca os conceitos sobre a compreensão da linguagem natural, que já são conhecidos na literatura como esquema, *script*, cenário, modelo cognitivo, entre outros.

Diante dessa nova abordagem, podemos verificar a importância da Semântica de *Frames* na construção do significado de modo mais aprofundado, já que leva em consideração

² “A *frame* (...) is a system of categories structured in accordance with some motivating context. Some words exist in order to provide access to knowledge of such *frames* to the participants in the communication process, and simultaneously serve to perform a categorization which takes such framing for granted.”

a experiência do falante e o conhecimento compartilhado entre os interlocutores. Para tanto, a Semântica de *Frames* dialoga com a Linguística de Corpus.

A Linguística de Corpus tem como finalidade explorar a linguagem através de dados empíricos de uma determinada língua, que são coletados e armazenados em meio eletrônico. A história dessa área está intimamente ligada e condicionada à tecnologia, que permite tanto a exploração como o armazenamento dos dados. Essa área vem sendo muito utilizada pelos campos da Lexicografia, Linguística Computacional e no Processamento da Linguagem Natural, entre outras áreas.

Segundo Sanchez (1995), a definição de corpus seria:

um conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso linguístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise' (SANCHEZ, 1995, p. 8-9)

Uma pesquisa baseada em *corpora* permite afirmar que os dados indicam o uso real dos falantes e não se tratam de sentenças criadas e moldadas para fins específicos, mas sim de evidencição do uso da linguagem em diversos contextos sociais, e que estarão sujeitos aos aspectos diacrônicos e sincrônicos.

Com base na Semântica de *frames* e na Linguística de Corpus, foi criado o Projeto FrameNet, que será detalhado na próxima seção.

2. O Projeto FrameNet

O projeto FrameNet (<http://FrameNet.icsi.berkeley.edu>) vem sendo desenvolvido por Charles Fillmore no *International Computer Science Institute* (ICSI), em Berkeley, na Califórnia desde 1997, com a finalidade de criar um recurso lexical *online* para o Inglês, com base na Semântica de *Frames* e com evidências em *corpus* (RUPPENHOFER et al. 2006).

De acordo com os dados disponíveis no site oficial do projeto, o banco de dados possui mais de 960 *frames* e mais de 10 mil unidades lexicais anotadas em mais de 170 mil sentenças. Esses dados são documentados a partir de uma gama de possibilidades combinatórias (valência) sintática e semântica de cada palavra em cada um dos seus sentidos (RUPPENHOFER et al. 2006).

Atualmente, o projeto tem se expandido para o espanhol, o alemão, o chinês, o sueco e o japonês. No Brasil, desde 2009, sob a liderança da professora Margarida Salomão, da

Universidade Federal de Juiz de Fora, está sendo realizada a contraparte linguística da rede semântica do português para a ampliação deste projeto multilíngue: a FrameNet Brasil (FN-Br)³.

Um dos projetos propulsores deste trabalho é o COPA 2014 FrameNet-Br (SALOMÃO et.al, 2011) que propõe:

(...) desenvolver, **em meio eletrônico, um dicionário temático trilíngue (Português – Inglês –Espanhol) abrangendo os domínios do Futebol e do Turismo.** Tal dicionário, com foco no uso humano, terá como audiência privilegiada a imprensa esportiva internacional, pessoas envolvidas na organização da Copa do Mundo FIFA 2014 e na recepção aos turistas estrangeiros, além dos próprios turistas. Em sintonia com as mudanças correntes na área das Tecnologias da Informação, o dicionário será um recurso lexical acessível eletronicamente de qualquer parte do mundo. (SALOMÃO et al. 2011, p. 2).

Nesse sentido, a busca, nesse dicionário, será feita por cenas conceptuais relevantes dos domínios pesquisados, o que distingue dos dicionários tradicionais, já que a busca não se dará pelos lexemas – conjunto de palavras relacionadas pelo sentido e pela sistematicidade de sua variação –, mas sim por Unidades Lexicais (ULs) – pareamentos de uma forma lexical a um significado estruturado em termos de um *frame* – visto que para compreender uma de suas partes, torna-se imprescindível o conhecimento do todo.

Para a elaboração do dicionário, há uma metodologia específica, na qual temos dois tipos de anotações: sendo uma denominada de anotação lexicográfica, que se caracteriza por se centrar na busca de uma unidade lexical em vários *corpora*, e que posteriormente, algumas serão selecionadas para serem anotadas; e a outra, a anotação de texto corrido, na qual as sentenças que constituem o texto escolhido é que são anotadas. A FR-Br adota a anotação lexicográfica.

Para cada anotação, escolhemos um *frame* e a palavra que será anotada como alvo evocadora deste *frame* em cada sentença. Anotamos os demais Elementos de *Frame* (EFs), que podem ser de três tipos: nucleares, periféricos e extratemáticos. Os elementos nucleares são aqueles fundamentais à composição do *frame*, uma vez que o particularizam e podem ser inferidos, mesmo que não estejam explicitamente lexicalizados. No *frame* de Transação_Comercial, por exemplo, os EFs nucleares são COMPRADOR, MERCADORIA, DINHEIRO, VENDEDOR. Os EFs periféricos, por sua vez, são informações acessórias a esse *frame*, que não o distinguem, mas atribuem ao *frame* algumas características. No mesmo

³ Para maiores informações acessar o site <http://www.framenetbr.ufjf.br>.

frame, MEIO, TARIFA e UNIDADE são EFs periféricos. Por fim, os EFs extratemáticos são aqueles que participam da cena descrita pelo *frame*, mas podem incluir um evento em um estado de coisas ou, até mesmo, evocar outros *frames*.

Quando os elementos de *frame* nucleares (EFs) não constam na sentença, eles devem ser indicados de alguma forma, pois são considerados relevantes para o delineamento do *frame*. Nesse caso, podem ocorrer três tipos de omissão que poderão ser focados de acordo com as seguintes categorias: Instanciação Nula Definida (IND)- omissão de um elemento que pode ser recuperado/ entendido no contexto linguístico ou discursivo; a Instanciação Nula Indefinida (INI)- omissão de um elemento que pode ser compreendido convencionalmente, sem precisar identificar o referente do discurso; e a Instanciação Nula Construcional (INC)- omissão permitida em determinadas construções gramaticais, como por exemplo, nas sentenças imperativas (Por favor, saia agora!) e em sentenças passivas (Nós fomos roubados.)

A anotação desses elementos é feita através de camadas, sendo que uma delas é constituída pela função semântica, ou elementos de frames, a outra é constituída pela função gramatical (FG) e por último pelos tipos de sintagmas (TS).

Adiante, estes aspectos poderão ser visualizados de uma maneira mais clara a partir da análise das ULs selecionadas para este trabalho.

3. *Frame* de Atração_turística

Como o nosso foco está centrado na anotação de ULs para a constituição do dicionário temático trilingue, escolhemos o *frame* de atração _turística que tem a seguinte definição:

<p>Atração_turística</p> <p>Definição:</p> <p>Um Lugar apresenta uma Atração_turística para chamar a atenção de um Turista.</p> <p>Elementos de Frame Nucleares:</p> <ul style="list-style-type: none">• Lugar: local em que se encontra a atração.• Atração_turística: espaço com valor cultural, importância histórica e/ou beleza natural ou artificial, que atrai turistas.• Turista: pessoa que visita o lugar em busca da atração.

Tabela 1. *Frame* de atração_turística

Podemos observar que, o *frame* de atração_turística se constituiu pelo interesse de um turista em visitar um determinado lugar, que oferece atrações. Uma das finalidades do turista

pode ser conhecer e aprender algo ou ter essa atividade como uma forma de lazer. Podemos ter, em alguns casos, informativos específicos como guias, panfletos e outros gêneros que se assemelham a estes para direcionarem melhor o turista.

Como EFs nucleares (centrais), temos o lugar, a atração turística e o turista. O primeiro é o local a ser visitado, enquanto o segundo é o que atrai o turista, e por último temos o elemento que está em busca de informação ou lazer no cenário específico para tais atividades.

Considerando o *frame* de atração_turística e os objetivos do projeto da FR-Br (COPA 2014), descreveremos, neste trabalho, as unidades lexicais “atrair”, “atraer”, “attracts” em suas línguas correspondentes.

4. Corpora e Método

O corpus utilizado até o momento 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.

Postularemos para análise da palavra alvo as camadas dos Elementos de *Frame* (EFs), da Função Gramatical (FG) e do Tipo de Sintagma (TS). A palavra alvo será grifada em caixa alta, marcada o seu fundo com a cor preta e a fonte da escrita é na cor branca. As unidades alvo jamais são etiquetadas em relação a sua função gramatical exercida na sentença analisada.

Na camada dos elementos de *frame*, tanto os elementos nucleares quanto os periféricos serão etiquetados, sendo que cada um deles será marcado por uma cor correspondente a sua função, e esta será a mesma em todas as análises posteriores.

Na camada da função gramatical os elementos marcados serão analisados de acordo com a função gramatical exercida por eles.⁴

Passaremos neste momento, para a análise das ULs “atrair”, “atraer” e “attracts” para contrastarmos as sentenças, que parecem ser bem equivalentes em sua tradução. É importante

⁴ É importante ressaltarmos, que as funções gramaticais e os tipos de sintagmas são analisados de acordo com as regras estabelecidas pela FrameNet de cada língua mencionada. Por isso, justificamos assim, as diferentes siglas nessas análises.

lembrarmos, que se trata de material com equivalentes já traduzidos pelos *corpora* pesquisados, conforme mencionamos anteriormente.

5. Anotação das Unidades Lexicais “atrair”, “atraer” e “attracts”

Nesta seção, faremos a anotação lexicográfica de acordo com o método especificado anteriormente, e para isso destacaremos os elementos de *frame*, a função gramatical e o tipo de sintagma de cada sentença que contenha as palavras alvo: *atrair/ atraer/ attracts*.

Iniciaremos com as sentenças descritas na nossa língua materna (português) que será sempre nomeada com a letra “a”, e em seguida, mencionaremos a mesma sentença na língua espanhola, que será nomeada com a letra “b” e por último anotaremos a língua inglesa que será intitulada com a letra “c”. Essa proximidade entre as sentenças nos possibilitará visualizar e contrastar melhor as traduções e a anotação lexicográfica realizada nas três línguas.

1(a) Já em Florianópolis, na costa catarinense, é a força das ondas que **ATRAI** surfistas de todo o mundo.

Camada com palavra alvo	em Florianópolis, na costa catarinense	a força das ondas	que	ATRAI	surfistas de todo o mundo.
Elemento de <i>Frame</i>	Lugar	Atração	Atração		Turista
Função Gramatical	DEP	EXT	EXT		ObjD
Tipo Sintagmático	NP	NP	NP		NP
Outros		Ant	Rel		

1(b) Ya en Florianópolis, en la costa catarinense, es la fuerza de las olas que **ATRAE** surfistas de todo el mundo.

Camadas	en Florianópolis, en la costa catarinense	la fuerza de las olas	que	ATRAE	surfistas de todo el mundo.
EF	Lugar	Atracción	Atracción		Turista
FG	Dep	EXT	EXT		Dobj
TS	NP	NP	NP		NP

1 (c) In Florianópolis, on the Catarinense coast, the wave strength ATTRACTS surfers from across the globe.

Camadas	In Florianópolis, on the Catarinense coast	the wave strength	ATTRACTS	surfers from across the globe.
EF	Place	Attraction		Tourist
FG	DEP	EXT		OBJ
TS	PP	NP		NP

2 (a) A capital catarinense, também conhecida como Floripa, ATRAI anualmente milhares de turistas que vão em busca de suas maravilhas naturais. (IND)

Camadas	A capital catarinense, também (...)	ATRAI	anualmente	milhares de turistas
EF	Lugar		Frequência	Turista
FG	EXT		DEP	ObjD
TS	NP		AVP	NP

2 (b) La capital catarinense, también conocida como Floripa, ATRAE anualmente a miles de turistas que van en búsqueda de sus maravillas naturales. (IND)

Camadas	La capital catarinense, también conocida (...)	ATRAE	anualmente	a miles de turistas
EF	Lugar		Frecuencia	Turista
FG	EXT		AdvObj	ObjD
TS	NP		AVP	NP

2 (c) The capital of Santa Catarina – also known as Floripa – ATTRACTS thousands tourists every year who are searching for its natural wonders. (IND)

Camadas	The capital of Santa Catarina – also known as Floripa	ATTRACTS	thousands tourists	every year
EF	Place		Tourist	Frequency
FG	EXT		OBJ	DEP
TS	NP		NP	AVP

3(a) Esse belo cenário ATRAI pessoas de todos os estados do Brasil. (Chapada dos Veadeiros- Goiás) (IND)

Camadas	Esse belo cenário	ATRAI	pessoas de todos os estados do Brasil
EF	Atração		Turista
FG	EXT		ObjD
TS	NP		NP

3 (b) Este bello cenário ATRAE a personas de todos los Estados de Brasil. (Chapada dos Veadeiros – Goiás) (INI)

Camadas	Este bello escenario	ATRAE	a personas de todos los Estados de Brasil
EF	Atracción		Turista
FG	EXT		PObj
TS	NP		PP

3(c) This wonderful scenery ATTRACTS people from all Brazilian states. (Chapada dos Veadeiros – Goiás) (INI)

Camadas	This wonderful scenery	ATTRACTS	people from all Brazilian states
EF	Attraction		Tourist
FG	EXT		OBJ
TS	NP		NP

4(a) Conhecida como capital do forró, a cidade ATRAI visitantes com seu centro de arte figurativa, seus artistas populares e a famosa Feira de Caruaru, Patrimônio Cultural Imaterial Brasileiro. (Caruaru, Pernambuco)

Camadas	a cidade	ATRAI	visitantes	com seu centro de arte figurativa, seus artistas populares e a famosa (...).
EF	Lugar		Turista	Atração
FG	EXT		ObjD	DEP
TS	NP		NP	PP

4 (b) Conocida como capital del forró la ciudad ATRAE visitantes con su centro de arte figurativa, sus artistas populares y la famosa Feria de Caruaru, Patrimonio Cultural Inmaterial Brasileño.

Camadas	la ciudad	ATRAE	Visitantes	con su centro de arte figurativa, sus artistas populares y la famosa (...)
EF	Lugar		Turista	Atracción
FG	EXT		DObj	Compl
TS	NP		NP	PP

4 (c) Known as the capital of forró, **the city** **ATTRACTS** **visitors** **with its figurative art centre, its popular artists and its famous Caruaru Market, Brazilian Cultural Heritage Site.**

Camadas	the city	ATTRACTS	Visitors	with its figurative art centre, its popular artists and (...)
EF	Place		Tourist	Attraction
FG	EXT		OBJ	DEP
TS	NP		NP	PP

Podemos observar que as siglas referentes à função gramatical e ao tipo sintagmático diferem nas três línguas, visto que, cada uma delas foi analisada de acordo com as especificidades de cada FrameNet (brasileira, espanhola e americana).

Ao verificarmos as realizações sintáticas, podemos observar que o verbo “atrair” é transitivo direto nas sentenças analisadas em português e em inglês, enquanto que na língua espanhola, este mesmo verbo pode ser transitivo direto (sentenças 1(b) e 4 (b)), e transitivo direto preposicionado, conforme mostra as sentenças (2 (b), 3(b)).

Outro aspecto importante é a presença dos elementos de *frames* nucleares em quase todas as sentenças. Apenas na sentença 2 (em todas as três línguas) pudemos observar a ausência do EF “atração”, que foi marcado como Instanciação Nula Definida (IND), por ser recuperável pelo contexto seguinte ao da oração na qual se encontra a UL alvo.

Por se tratar de uma tradução muito próxima, uma vez que temos uma substituição de palavras sinônimas ou correspondentes nas línguas, podemos observar que os elementos de *frame* tiveram as mesmas correspondências nas três línguas, e em quase todas as sentenças tivemos a mesma estrutura canônica (sujeito – verbo – complemento verbal) evidenciando assim, as mesmas categorias semânticas.

Pretendemos em trabalhos posteriores, direcionar nosso olhar de forma mais específica para esses *corpora* traduzidos, para verificarmos a viabilidade de serem adequados ou não diante do seu uso na língua alvo. Também temos procurado pesquisar sentenças que são produzidas dentro de um contexto de uso de cada língua.

Vale ressaltar que a pesquisa realizada pelas FrameNets nos demais idiomas não buscam seus *corpora* por meio de traduções, mas a partir das ULs em *corpora* específicos de cada língua. O presente trabalho partiu de sentenças traduzidas por ter o objetivo de apresentar o papel da sintaxe em uma mesma sentença nas três línguas, já que estas serão

utilizadas como exemplos ilustrativos no dicionário. A determinação de um *frame* carece da identificação da estrutura sintática em corpus.

6. Considerações Finais

Podemos sinalizar a partir deste trabalho uma ideia da estrutura gramatical e metodológica que está permeando a construção do dicionário. Esses aspectos abordados e analisados não serão visualizados no dicionário, uma vez não são relevantes para os interlocutores que utilizarão essa ferramenta, ademais, estas informações estarão disponíveis no banco de dados da FN-Br.

No entanto, como vimos no presente trabalho é a partir da anotação lexicográfica que poderemos visualizar a construção da rede semântica que fará conexão entre os Elementos de *Frame* e os *frames* descritos pelo dicionário.

Essas questões dos *corpora* traduzidos ou não, das possibilidades combinatórias (sintáticas e semânticas) têm direcionado nossas buscas e reflexões, para que o projeto de criação e inovação tecnológica possa ser concretizado. Além disso, esperamos que esse projeto contribua para a ampliação metodológica desenvolvida a partir da Semântica de *Frames*.

REFERÊNCIAS BIBLIOGRÁFICAS

CHOMSKY, N. 1955. *The logical structure of linguistic theory*. Manuscript. Published in 1975 by Plenum Press, New York.

CHOMSKY, N. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

FELDMAN, J. *Firts words, Conceptual schemas and cultural frames, learnig spatial relacion words, Embodied Knowledge of actions, learning about words*. Cambridge, Mass: MIT Press, 2006

FILLMORE, C. J. *Scenes-and-frames semantics*, *Linguistic Structures Processing*. In Zampolli, Antonio (Ed.): *Fundamental Studies in Computer Science, No. 59*, North Holland Publishing, 55-88, 1977a.

_____. The need for a *frame* semantics in linguistics, In Karlgren, Hans (Ed.): *Statistical Methods in Linguistics* 12: 5-29, 1977b.

_____. *Frame semantics*. In *Linguistics in the Morning Calm*, Seoul, Hanshin Publishing Co., 111-137, 1982.

- _____. *Semântica de Frames*. Tradução: Galeno Fae da Silva: revisão Danilo Nogueira Marra. Caderno de Tradução. Porto Alegre, nº 25, jul –dez, 2009 p. 25-54.
- _____. *Frames and the semantics of understanding*. In *Quaderni di Semantica*, Vol. 6.2: 222-254, 1985a.
- _____. *Fillmore's cases grammar: a reader*. In: DIRVEN, R. & RADDEN, G. (Ed). Heidelberg, 1987.
- JAKOBSON, R. *Linguística e Comunicação*. Editora Cultrix. São Paulo. 1975
- MELTZOFF, A. N. *Imitation as a mechanism of social cognition: Origins of empathy, theory of mind, and the representation of action*. In: U. Goswami (Ed.), *Handbook of childhood cognitive development*. Oxford: Blackwell Publishers. (2002b).
- SANCHEZ, A. *Definicion e historia de los corpus*. In: A. SANCHEZ et al (org.). *CUMBRE – Corpus Linguístico de Espanol Contemporaneo*. Madrid: SGEL. 1995
- RUPPENHOFER, Josef; ELLSWORTH, Michael; PETRUCK, Miriam R. L.; JOHNSON, Christopher R.; SCHEFFCZYK, Jan. *FrameNet II: extended theory and practice*, 2006. Disponível em: www.FrameNet.icsi.berkeley.edu. Acesso em: 20 abr. 2012.
- SALOMÃO, M. M. M. *FrameNet Brasil: um trabalho em progresso*. *Calidoscópico*, São Leopoldo: UNISINOS, vol. 7 n. 3, p. 171-182, set/dez 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.